Journal of Geophysics Zeitschrift für Geophysik

Volume 48 1980

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 and J. Untiedt from Vol. 27 (1961). 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.

Springer-Verlag Berlin Heidelberg New York

Printed in Germany by Universitätsdruckerei H. Stürtz AG Würzburg

Author Index

Subject Index

Journal of Geophysics Zeitschrift für Geophysik

Volume 48 Number 1 1980

Springer International

J. Geophys. — ISSN 0340-062X — JGEOD4 48(1) 1–56 (1980) — March 1980

Journal of Geophysics – Zeitschrift für Geophysik

Edited for the **Deutsche Geophysikalische Gesellschaft** by **W. Dieminger 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 Vol. 1–18 (1944). It reappeared in 1954 edited by B. Brockamp from Vol. 19–26 (1960), and edited by W. Dieminger and J. Untiedt from Vol. 27 (1961). 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.

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.

Subscription Information

Volumes 47+48 [3 issues each totalling about 246 pages (due to a larger format)] will appear in 1980. The price of a volume is DM 143,– or $80.00. Prices for backvolumes are available on request. In addition, we also offer microform editions in 16 mm and 35 mm microform and microfiche. Correspondence concerning subscriptions should be addressed to the publisher.

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

Members. Members of the Deutsche Geophysikalische Gesellschaft are entitled to purchase the Journal for their own use at a privilege price of DM 98,— 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. Subscription rate: $169.00, including postage and handling. Subscriptions are entered with prepayment only. Orders should be addressed to: Springer-Verlag New York Inc., Service Center Secaucus, 44 Hartz-Way, Secaucus, N.J. 07094, USA, Tel. (201) 348-4033, Telex 0023-125994.

Japan. Subscription rate: DM 314,— including postage (surface airmail lifted) and handling. Orders can either be placed with your bookdealer or sent directly to: Springer-Verlag, Heidelberger Platz 3, D-1000 Berlin 33, Tel. (030) 8207-1, Telex 01-83319.

All Other Countries. Subscription rate: DM 286,—, plus postage and handling. Orders can either be placed with your bookdealer or sent directly to: Springer-Verlag, Heidelberger Platz 3, D-1000 Berlin 33, Tel. (030) 8207-1, Telex 01-83319.

Offices

Springer-Verlag, Heidelberger Platz 3, D-1000 Berlin 33, Tel. (030) 8207-1, Telex 01-83319.

Springer-Verlag, Journal Production Department II, Postfach 105280, D-6900 Heidelberg 1, Tel. (06221) 487-1, Telex 04-61690.

Springer-Verlag New York Inc., 175 Fifth Avenue, New York, N.Y. 10010, USA, Tel. (212) 477-8200, Telex 0023-23223.

Responsible for Advertisements

G. Sternberg, Kurfürstendamm 237, D-1000 Berlin 15, Tel. (030) 8821031, Telex 01-85411.

Printed in Germany by Universitätsdruckerei H. Stürtz AG, Würzburg

Das Heft enthält eine Beilage des Springer-Verlages Berlin Heidelberg New York

Springer International

J. Geophys. 48, 1–6 1980

Journal of Geophysics

Original Investigations

The Analysis of Simultaneous Observations of Nighttime Pi Pulsations on an East-West Profile

L.N. Baranskiy[1], V.A. Troitskaya[1], I.V. Sterlikova[1], M.B. Gokhberg[1], N.A. Ivanov[2], I.P. Khartchenko[3], J.W. Münch[4], and K. Wilhelm[5]

[1] Institute of the Physics of the Earth, Academy of Sciences, Moscow, USSR
[2] Geophysical Institute of the Ural Scientific Centre, Sverdlovsk, USSR
[3] Siberian Institute of Terrestrial Magnetism, Academy of Sciences, Irkutsk, USSR
[4] University of Siegen, D-5900 Siegen 21, Ferderal Republic of Germany
[5] Max-Planck-Institut für Aeronomie, D-3411 Katlenburg-Lindau 3, Federal Republic of Germany

Abstract. Magnetic pulsations of the types Pi 1 and Pi 2 were observed on a profile consisting of five stations between Lindau and Novosibirsk (87.5 to 155.0 deg geomagnetic longitude and 49.7 to 54.0 deg geomagnetic latitude). The Pi events were analysed with respect to their propagation characteristics. The frequency-time analysis demonstrated that individual Pi pulsation events attained their highest frequencies near local magnetic midnight. In addition, it was found that the east-west component of the horizontal disturbance vector exhibited a larger frequency shift than the north-south component. The phase variations observed were indicative of a propagation along the profile. The polarization of the pulsations generally was elliptical with a left-hand rotation of the disturbance vector and local time dependent direction of the major axis. All findings support the conclusion that the propagation of the two components is governed by different conditions. The phase velocities of the events ranged from 100 to 500 km/s.

Key words: Magnetic Pi pulsations – Hydromagnetic wave propagation.

1. Introduction

Irregular geomagnetic pulsations of the type Pi are known to occur in close relation to magnetospheric substorm onsets. They predominantly are night-time phenomena and exhibit a pronounced/latitudinal dependence of their amplitude variations with an absolute maximum near the auroral zone and a relative maximum near the plasmapause location. On the basis of this evidence, it was concluded that the primary excitation process is operating on auroral magnetic field lines and that the plasma density gradient at the plasmapause gives rise to a secondary resonance effect (Jacobs, 1970; Kuwashima, 1978). The observations on the polarisation properties of the Pi pulsations are by far more controversial than those on the amplitude characteristics. It is likely that phase reversals across boundary lines in longitude and in latitude, such as the auroral zone or the plasmapause, substantially complicate the issue, as the interpretation of a given measurement depends on the accurate spatial determination of the appropriate boundary with respect to the observation site. A detailed description of the properties of Pi 2 pulsations and the corresponding theoretical interpretations was recently given by Kuwashima (1978) supported by a compilation of the relevant literature.

In order to further the understanding of mid-latitude Pi pulsations and their propagation characteristics, an observational programme was performed at five sites with geomagnetic latitudes between 49.8 and 54.0 deg and geomagnetic longitudes reaching from 87.5 to 155.0 deg. The corresponding McIlwain parameters L were smaller than 2.9. Consequently, all stations were located south of the plasmapause for magnetically quiet and moderately disturbed time intervals, thus eliminating any complications by the expected polarisation reversals across this boundary.

2. Methods of Observation and Data Analysis

The observations used in this study were made by compensation-type pulsation magnetometers described elsewhere (Wilhelm, 1966; Wilhelm et al., 1977). In summary, both pulsations and variations of the geomagnetic field were recorded in both components H and D of the horizontal disturbance vector. The amplitude and phase response curves of all recording systems as a function of the pulsation period were nearly identical, showing 3-dB points at periods of 1 and 150 s.

Data were recorded both on paper charts and digital magnetic tapes. Special features of the magnetic tape recordings were sampling times of 0.4 and 3.2 s for pulsation and variation channels, respectively, and a resolution of the digitization of 0.03 nT in the frequency range corresponding to the plateau of the response function. Internal crystal clocks and frequent comparisons with radio time signals maintained a very good time accurancy at the widely separated stations. The time adjustments required typically were of the order of 40 ms and never exceeded 100 ms giving an indication of the synchronism achieved.

Five identical instruments were operated on a profile reaching from Lindau, FRG, to Novosibirsk, USSR. The stations and their geomagnetic coordinates are listed in Table 1.

The experiment was conducted from August 1–October 25 1974 as a joint venture of the Max-Planck-Institut für Aeronomie and the Institute of the Physics of the Earth. A more detailed description of the project was given by Münch et al. (1975) and first results were published by Troitskaya et al. (1976a and b).

0340-062X/80/0048/0001/$01.20

Table 1. List of stations

Station	Nomenclature	Geomagnetic latitude, deg	Geomagnetic longitude, deg	L-parameter	Local magnetic time at 00:00
Lindau	LIN	49.8	87.5	2.40	02:00
Kaliningrad	KNG	52.1	98.0	2.65	02:42
Borok	BOR	54.0	114.0	2.89	03:48
Sverdlovsk	SVE	52.5	135.5	2.70	05:12
Novosibirsk	NOV	49.7	155.0	2.39	06:30

The pulsation recordings of the H and D components designated in this paper by P_H and P_D, respectively, stored on magnetic tape in a digital data format were processed using a computer programme by Grusdeva et al. (1974), which carried out an analysis of the spectral contents of the pulsation signals with respect to time. The principle of the mathematical formalism can be summarized as follows:

Let $U(t)$ denote the time dependent signal of the pulsation activity, an estimate of the spectrum of the signal $s(\omega)$ as a function of the angular frequency averaged over a time interval τ can then be obtained by

$$s(\omega)=\int_0^\tau U(\vartheta)\exp(-i\omega\vartheta)\,d\vartheta. \tag{1}$$

The interval τ has to be chosen in accordance with the accuracy of the estimate and the time resolution required. Using narrow-band filters, the time dependent change of the spectrum can be computed by

$$Y(t,\omega_\nu)=\frac{1}{\pi}\int_0^\infty\left[\int_0^\tau U(\vartheta)\exp(-i\omega\vartheta)\,d\vartheta\right]\cdot\exp[i\omega t-\alpha(\omega-\omega_\nu)^2/\omega_\nu^2]\,d\omega \tag{2}$$

with discrete values of $\nu=1\ldots N$, where ω_ν is the central frequency of the ν-th filter. The factor α determines the slope of the filter response curve. In the calculation presented here, α and N were chosen to be equal to 25 and 30, respectively. Discrete time series of the P_H and P_D components and of the horizontal magnetic disturbance $P=(P_H^2+P_D^2)^{1}/_{2}$ were separately treated according to Eq. (2). The magnitude $r=|Y(t,\omega_\nu)|$ of $Y(t,\omega_\nu)$ solely depends on the amplitude of the filtered signal and, similarly, the argument of $Y(t,\omega_\nu)$ is related to the phase φ. The quantities r_{P_H}, r_{P_D} and r_P as well as φ_{P_H} and φ_{P_D} were thus obtained, by which the pulsation activity could be characterized. In addition, the phase difference $\varphi_{P_H}-\varphi_{P_D}$ is of interest for a comparison of the horizontal components. The magnitudes r were plotted in (t, T)-planes with a time resolution corresponding to $\tau=6.4$ s. The pulsation period T instead of the frequency was used here for reasons of simplicity. In these plots, contour lines of constant pulsation activity could be drawn thus delineating wave packets of the Pi pulsations. Furthermore, the largest amplitude for each individual frequency band could be determined.

3. Observations and Results of the Data Analysis

Two examples of the dynamic spectra of Pi events recorded simultaneously at all five stations have been compiled in Figs. 1 and 2. Inspection of auroral zone magnetograms showed that the centre of the substorm activity related to Fig. 1 was located near northern Scandinavia (geomagn. long. 120 deg) starting at 18:45 UT. The amplitude signal of the pulsation activity between 18:45 and 18:46 UT was largest at Borok. The event in Fig. 2 resulted from a substorm of more than 500 nT negative depression in the H component over Iceland and Greenland. In addition, four other cases of simultaneous Pi events were analysed that could clearly be identified at all observation sites in the time interval from October 18–22, 1974. In Fig. 3 some of these events are depicted as functions of time. The two features that are most apparent from Figs. 1 and 2 are the different positions of the wave packets in the (t,T)-domain for different stations and the displacement between the P_H and P_D components relative to each other. In

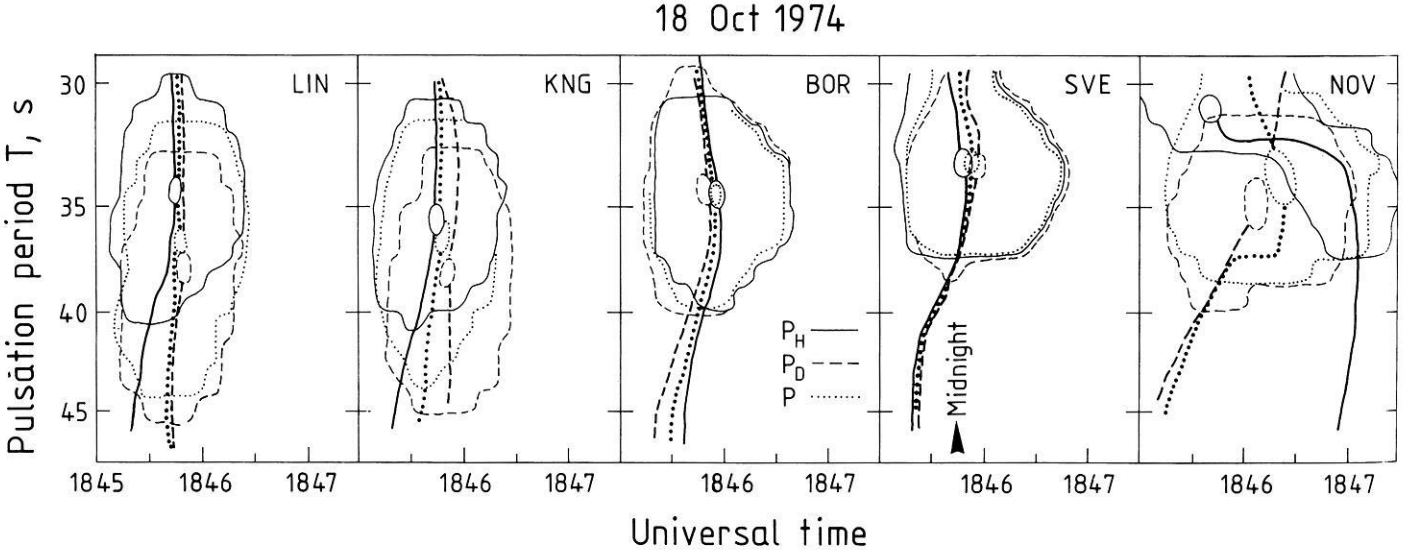


Fig. 1. Dynamic spectra of a Pi pulsation event on October 18, 1974 at 18:46 UT. Amplitude contours of the pulsation components (for details see text) are given in a universal time-pulsation period co-ordinate system together with the location of the amplitude maxima

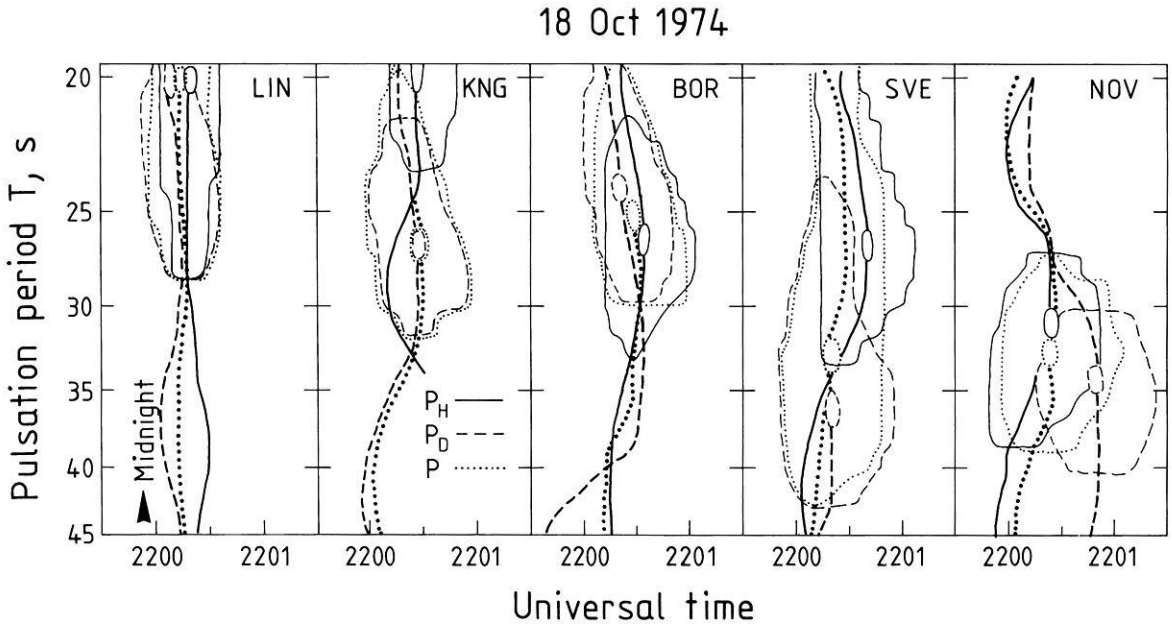


Fig. 2. Dynamic spectra of a pulsation event on October 18, 1974 at 22:00 UT

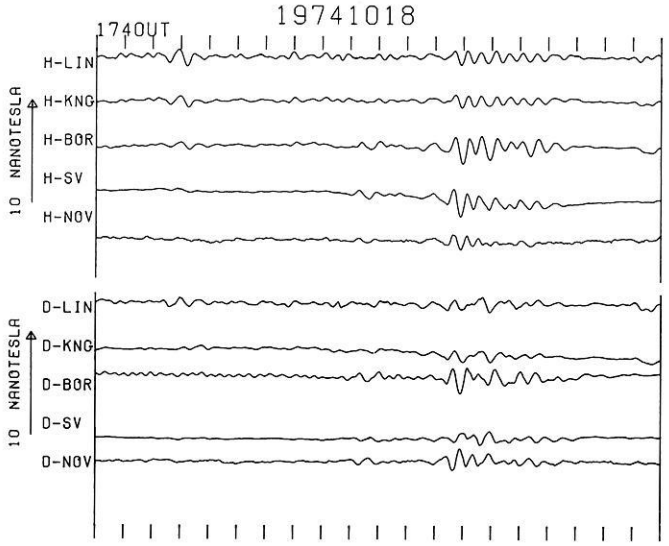


Fig. 3. Amplitude-time displays of Pi pulsation events in both the H and D components as observed at the five observational sites. The *arrows* point in the positive direction and have a length corresponding to 10 nT. The time span shown is 20 min starting at 17:40 UT on October 18, 1974

particular, there is a tendency towards a shift of the position of the wave packets to longer periods with increasing distance from magnetic midnight. Beginning at stations near local magnetic midnight where the positions of the wave packets of the P_H and P_D components almost coincide, this shift of the main period of the Pi pulsations increases with distance from the midnight meridian and is significantly stronger for the P_D than for the P_H component, causing the large displacement of the wave packets in the (t,T)-domain of the P_D relative to the P_H component at stations located sufficiently far from the magnetic midnight meridian.

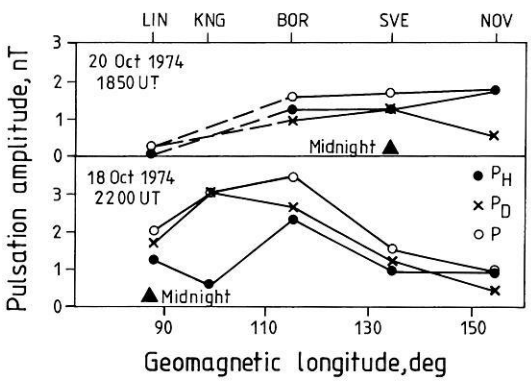
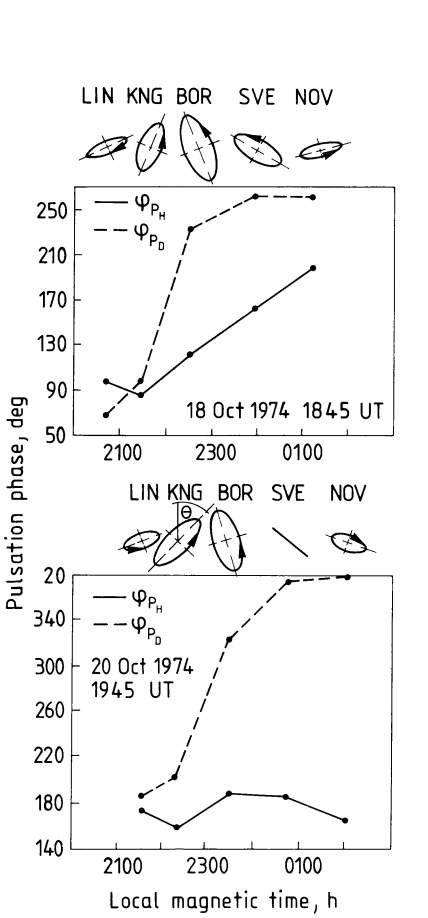


Fig. 4. Mean peak pulsation amplitude of two Pi events as a function of geomagnetic longitude

In order to establish the distribution of the intensity of a Pi event along the profile the following method was applied:
For each event, an area in the (t,T)-plane was defined that contained the peak values of the amplitudes at all five stations. Over this area, the average amplitude values were then calculated for different stations. The amplitude information thus obtained could directly be used in comparing the different observation points. This was done in Fig. 4, showing two examples of the distribution of the average amplitude peaks of the quantities P_H, P_D, and P along the profile. The maxima of these distributions exhibit a weak tendency to follow local magnetic midnight. In addition, it could be noticed in these and other examples not subject to the restriction made in this paper of simultaneous occurrence at all five stations that Borok normally showed the largest amplitudes. This seems to reflect the fact that the stations were not precisely located along a constant latitude profile but that the central stations lay at higher latitudes (see Table 1). Taking into account the increase of Pi 2 amplitudes with increasing latitude of typically 0.2–0.9 nT/degree (Baranskiy et al., 1974), which is a large gradient as compared to the weak longitudinal dependence, we can presume



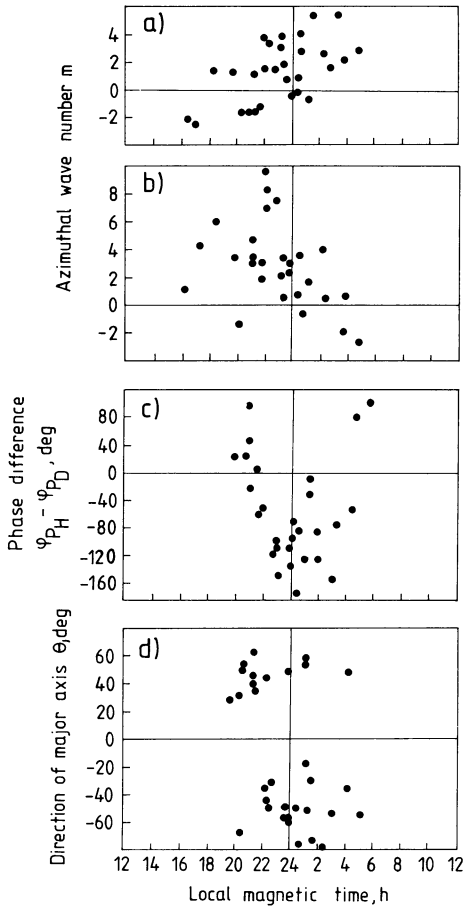


Fig. 5. Variation of the pulsation phase for both components along the observation profile for two events. Also given are schematic polarization ellipses and the sense of the rotation of the disturbance vector

Fig. 6. Dependence of Pi wave parameters on local magnetic time

that the Pi amplitudes at the stations BOR, KNG, and SVE are usually larger than those at the stations LIN and NOV mainly because of their geomagnetic location.

In order to investigate the shape of the Pi waves and their variations in space along the chain of stations considered here, it was necessary to study the phase relations of the P_H and P_D components. For each individual wave packet the phase relationship turned out to be rather stable. This allowed the determination of the average phase values at all stations. For two pulsation events the resulting phase curves as a function of local magnetic time are given in Fig. 5.

Also indicated in this figure are the polarisation ellipses for each station. When assessing the phase dependence of Pi pulsations on local time, it should be remembered that the phase of sinusoidal waves propagating in one direction along the profile would linearly change in this direction. It can be seen that this ideal case is only a very crude approximation for some portions of the propagation path. The observed nonlinear changes of the phases point to more complicated wave structures and changes or even reversals of the phase velocity. The polarisation ellipses support earlier investigations (Baranskiy et al., 1970; Björnsson et al., 1971) that Pi 2 pulsations are characterized by a dominant left-hand rotation of the disturbance vector in the horizontal plane and the change of the sign of the orientation angle θ around midnight. In order to reveal some of the statistical properties of Pi pulsation events, the information obtained for all events studied here have been compiled in Fig. 6 as a function of local magnetic time. An important parameter is the phase difference between pairs of stations. It was taken positive when the phase angle of the signal at the eastern station was greater than that at the western station. The resulting phase difference could only be determined modulo π and was selected to lie between $\pm\pi$. Any ambiguity could be excluded by the assumption that the change of the phase along the profile of stations was continuous. The azimuthal wave number m could then be defined by $m = \Delta\varphi/\Delta\lambda$ where $\Delta\varphi$ is the difference of the phase of the signal at two neighbouring stations and $\Delta\lambda$ is the difference of their longitudes. The scatter plots of m exhibit little systematic behaviour both for the H (Fig. 6a) and D (Fig. 6b) component. Most of the values are positive and seem to have a tendency to change sign around midnight for both components. Should this indication be supported by further investigations, it would mean that the phase velocity of the P_H signal is directed towards local magnetic midnight whereas the P_D signal is directed away from it.

The dependence of the phase difference between the P_H and P_D signals on the local magnetic time is displayed in Fig. 6c. This difference characterizes the direction of the rotation of the disturbance vector of the pulsations in the horizontal plane. A right-hand rotation will lead to positive values of $\varphi_{P_H} - \varphi_{P_D}$ and a left-hand rotation correspondingly results in negative values.

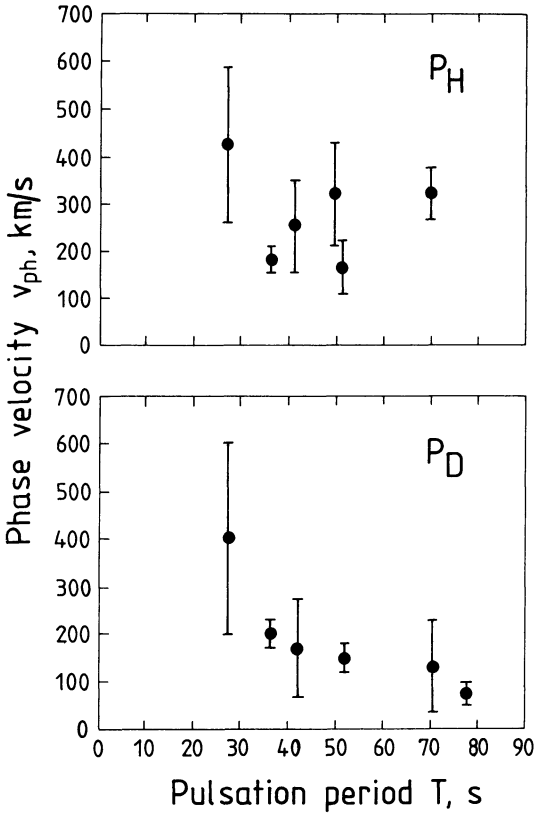


Fig. 7. Phase velocities of the two horizontal wave components as a function of the pulsation period

It follows that the rotation of the horizontal vector of Pi pulsations is predominantly negative in the 21:00 to 04:00 local magnetic time interval and positive both before and after this interval.

Finally, the dependence of the angle of the major axis of the polarisation ellipses with the north direction on local magnetic time is presented in Fig. 6d. The angle is positive, if the axis of the ellipse is pointing into the north-east quadrant and negative, if it is deflected to the west. The scatter diagram does not indicate a clear local time dependence. However, the mean local times of the positive and negative clusters have significantly different values of 22.3 and 0.6 LMT, respectively, confirming earlier findings of Baranskiy et al. (1970) that a change of the sign of the angle θ in the vicinity of the midnight meridian is a common feature for Pi pulsation events.

Figure 7 summarizes the findings on the phase relationships of Pi pulsations in mid-latitudes by presenting the phase velocities separately both for the *H* and *D* components. A decrease of the phase velocity with increasing pulsation period seems to be indicated at least for the D-component.

4. Discussion

Observations of Pi pulsations along a latitudinal profile extending over 60 deg in longitude were carried out for the first time. Of special interest is the value and the direction of the Pi phase velocity. The figures obtained are in good agreement with calculations of the azimuthal wave number by Lanzerotti and Fukunishi (1974) and with a Pi 2 generation model suggested by Chen and Hasegawa (1974). However, the interpretation cannot be unambiguous before a moving source, possibly related to the westward travelling surge of a magnetospheric substorm event (Pytte et al., 1976; Kisabeth and Rostoker, 1973), can definitely be excluded. Further investigations taking into account the auroral Pi activity will be required to solve this problem.

Other results of this investigation of Pi pulsations are the discoveries of the pulsation period increase on either side of the midnight meridian and the polarisation splitting. The different spectral composition and the dependence of the azimuthal wave number on local magnetic time as well as the differences of the intensity distribution of the *H* and *D* components of the Pi pulsations can all be considered as evidence that the primary hydromagnetic disturbance experiences different conditions of propagation for the *H* and *D* components.

In conclusion, it may be useful to compare these findings with results obtained for pulsations in the 8–25 mHz band (Mier-Jedrzejowicz and Southwood, 1979) and for dayside pulsations Pc 3 and Pc 4 (Green, 1976). The values of the azimuthal wave numbers obtained for nighttime events were comparable to the ones reported here, whereas Pc 3 and Pc 4 pulsations had slightly larger wave numbers indicating shorter wave lengths for Pc 3 and Pc 4 than for Pi waves.

Acknowledgements. The observations discussed in this paper were made during a compaign jointly sponsored by the Deutsche Forschungsgemeinschaft and the Academy of Sciences of the USSR. The significant contributions of B.N. Kazak and R. Schmidt during the field operations are greatly acknowledged.

References

Baranskiy, L.N., Shchepetnov, R.V., Afanas'Yeva, L.T., Zybin, K.Yu., Hillebrand, O., Sanker Harayan, P.V.: Intensity distribution of Pi 2 pulsations along the geomagnetic meridian and on the night side of the earth. Geomagn. Aeron. **14**, 743–746, 1974

Baranskiy, L.N., Vinogradov, P.A., Raspopov, O.M.: Polarization of geomagnetic pulsations of the Pi 2 type. Geomagn. Aeron. **10**, 743–745, 1970

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

Chen, L., Hasegawa, A.: A theory of long-period magnetic pulsations 1. Steady state excitation of field line resonance. J. Geophys. Res. **79**, 1024–1032, 1974

Green, C.A.: The longitudinal phase variation of mid-latitude Pc 3–4 micropulsations. Planet. Space Sci. **24**, 79–85, 1976

Grusdeva, N.P., Levshin, A.L., Pisarenko, V.F., Pruchkina, F.J.: Spectral time analysis of seismic waves. Theor. Comput. Geophys. Nauka 5–15, Moscow, 1974

Jacobs, J.A.: Geomagnetic micropulsations. In: Physics and Chemistry in Space, Vol. 1, J.G. Roederer and J. Zähringer, eds.. Berlin, Heidelberg, New York: Springer 1970

Kisabeth, J.L., Rostoker, G.: Current flow in auroral loops and surges inferred from ground-based magnetic observations. J. Geophys. Res. **78**, 5573–5584, 1973

Kuwashima, M.: Wave characteristics of magnetic Pi 2 pulsations in the auroral region – Spectral and polarization studies. Mem. Nat. Inst. Polar Res., Series A, No. 15, Tokyo, 1978

Lanzerotti, L.J., Fukunishi, H.: Relationships of the characteristics of magnetohydrodynamic waves to plasma density gradients in the vincinity of the plasmapause. J. Geophys. Res. **80**, 4627–4634, 1975

Mier-Jedrzejowicz, W.A.C., Southwood, D.J.: The East-West structure of mid-latitude geomagnetic pulsations in the 8–25 mHz band. Planet. Space Sci. **27**, 617–630, 1979

Münch, J.W., Wilhelm, K., Troitskaya, V.A., Kazak, B.N., Baranskiy, L.N.: Observations of geomagnetic micropulsations on a mid-latitude east-west-profile. Kleinheubacher Berichte **18**, 123–132, Darmstadt, 1975

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

Troitskaya, V.A., Baranskiy, L.N., Gokhberg, M.B., Münch, J.W., Wilhelm, K., Voelker, H., Siebert, M., Hillebrand, O., Ivanov, N.A., Vinogradov, D.A.: Preliminary results of the german-soviet experiment on simultaneous registrations of geomagnetic pulsations on meridional and latitudinal profiles. Kleinheubacher Berichte **19**, 545–552, Darmstadt, 1976a

Troitskaya, V.A., Baranskiy, L.N., Gokhberg, M.B., Sterlikova, L.V., Belen'kaya, B.N., Münch, J., Wilhelm, K., Voelker, H., Siebert, M., Hillebrand, O., Kharchenko, I.P., Ivanov, N.A., Kopytenko, Yu.A.: Preliminary results of a soviet-german experiment on synchronous recording of geomagnetic pulsations on meridional and latitudinal station profiles. Geomagn. Aeron. **16**, 558–561, 1976b

Wilhelm, K.: Registrierung and Analyse erdmagnetischer Pulsationen der Polarlichtzone sowie ein Vergleich mit Bremsstrahlungsmessungen. Mitt. a.d. MPI f. Aeronomie, Nr. 27. Berlin, Heidelberg, New York: Springer 1966

Wilhelm, K., Münch, J.W., Kremser, G.. Fluctuations of the auroral zone current system and geomagnetic pulsations. J. Geophys. Res. **82**, 2705–2716, 1977

Received September 7, 1979; Revised Version December 19, 1979

J. Geophys. 48, 7–17, 1980

Journal of Geophysics

Coordinated Magnetic Observations of Morning Sector Auroral Zone Currents With Triad and the Scandinavian Magnetometer Array: A Case Study

H. Sulzbacher[1], W. Baumjohann[1], and T.A. Potemra[2]

[1] Institut für Geophysik, Universität Münster, Gievenbecker Weg 61, D-4400 Münster, Federal Republic of Germany

[2] Applied Physics Laboratory, Johns Hopkins University, Laurel, Maryland 20810, USA

Abstract. By using coordinated two-dimensional ground-based and satellite magnetic measurements obtained during a morning sector pass of the Triad satellite over the Scandinavian Magnetometer Array we were able to derive current densities of a large-scale westward electrojet system as well as those of a small-scale arc-associated system. The broad westward electrojet constituted a Hall current flowing in the same region where balanced field-aligned current sheets were observed. The field-aligned currents were directed downward in the north and upward in the south and were closed in the ionosphere by southward Pedersen currents. South of the maximum of the broad electrojet, in the region of grossly upward field-aligned current and near to a quiet auroral arc, another small-scale current system was found with essentially the same configuration as the electrojet system. The current densities of this arc-associated system were slightly higher or comparable to those calculated for the electrojet system.

Computation of the Σ_H/Σ_P ratio gave an indication of the latitudinal distribution of energetic particle precipitation. In contrast to the evening sector, the energetic particles precipitate in the southern half of the auroral oval, where Σ_H/Σ_P ratios of 2, and up to 4 near the auroral arc, have been found, while ratios close to 1 in the northern half indicate lower energetic electron precipitation. In addition we have also paid attention to the comparatively smaller magnetic disturbances observed by Triad perpendicular to the field-aligned current sheets and along the vertical axis. The perpendicular horizontal variation may be explained by the leakage of the toroidal magnetic field due to east-west gradients of the field-aligned sheet current density. The variations in the vertical component are due to the same source as the north-south component and can be seen because of the inclination of the field lines.

Key words: Triad satellite – Scandinavian Magnetometer Array – Morning sector auroral electrojets – Field-aligned currents – Auroral arc currents.

1. Introduction

It is a well-known fact that magnetic measurements on the earth's surface are not sufficient in order to determine the *real* three-dimensional current distributions responsible for geomagnetic disturbances. Instead, there are always various *equivalent* current systems that may explain a magnetic disturbance field configuration observed on the ground. A good example for this basic problem in the field of geomagnetism was the different way in which on one side Birkeland and Alfvén (Birkeland 1908, 1913; Alfvén 1939, 1940) and on the other hand Chapman and Vestine (Chapman 1935; Vestine and Chapman 1938) explained the main features of high-latitude magnetic variations. Birkeland and Alfvén proposed a three-dimensional ionospheric and field-aligned current system whereas Chapman and Vestine preferred a two-dimensional system confined to the ionosphere. At that time, when only ground-based magnetic observations were possible, the debate could not be settled since both current systems have equivalent ground magnetic effects (Fukushima 1969).

On the other hand, magnetic measurements still play an important rôle in the exploration of ionospheric and magnetospheric currents, because the more direct methods as, for example, the measurement of plasma parameters by means of incoherent scatter, sounding rocket or satellite experiments, also have inherent difficulties. For example, one can underestimate the field-aligned current flow by counting only particles in a limited energy range with rocket and satellite spectrometers, as described by Klumpar et al. (1976) and Evans et al. (1977). Furthermore, the 'equivalence problem' can be overcome by including measurements of additional relevant quantities like conductivity and electric fields in the analysis (see Baumjohann et al. (1980) for a broader discussion of this topic).

Another possible solution would be magnetic observations made on a three-dimensional grid. Up to now nobody has built a 'three-dimensional magnetometer array', but one magnetometer configuration which for certain magnetic disturbances may come close to this ideal case consists of the two-dimensional Scandinavian Magnetometer Array (Küppers et al. 1979) and the Triad satellite (Armstrong and Zmuda 1973). Provided that the current flows are stable during the satellite pass, the magnetic measurements along the trajectory can partially add the third dimension.

For a first case study in using coordinated Scandinavian Magnetometer Array and Triad observations for analysing auroral zone currents we have chosen a morning sector pass of the Triad satellite over a westward electrojet since here the ground magnetic fields were rather stable and indicated two-dimensionality, i.e., independency of the east-west coordinate. There have been extensive studies of ionospheric and field-aligned currents based on data from the Alaska or Alberta meridian chains of magnetometers and the geomagnetically east-west aligned component of the Triad magnetometer (e.g., Armstrong et al. 1975; Rostoker et al. 1975; Kamide and Akasofu 1976; Kamide et al. 1976b; Kamide and Rostoker 1977). As compared to these earlier studies, we have incorporated three

0340-062X/80/0048/0007/$02.20

improvements in our analysis: The first improvement is that with the array measurements we were able to show that the magnetic fields were indeed nearly two-dimensional which verified the validity of a two-dimensional analysis. The second important difference is that our analysis is quantitative instead of qualitative, as were the earlier ones. The last improvement is an additional discussion and interpretation of the comparatively small magnetic variations observed by the Triad satellite in the north-south and vertical directions.

By means of separation and subsequent upward continuation of the external part of the ground magnetic fields to ionospheric heights with a method similar to that recently applied by Mersmann et al. (1979) we were able to compare directly the latitudinal distribution of magnetic fields at 100 km (just below the ionospheric current layer) and at the Triad altitude of 800 km (above the ionosphere and in the region of field-aligned currents). Using this data set we calculated current densities for Hall, Pedersen and field-aligned currents. The high spatial resolution of our results permitted us to discuss even a small-scale structure in the westward electrojet system associated with an auroral arc. In addition, computation of the ratio between Hall and Pedersen conductivity allowed us to draw some conclusions on the latitudinal distribution of the energy of precipitating particles. Finally, we were also able to explain the magnetic observations made by Triad in the north-south and vertical directions in the framework of the whole three-dimensional current system.

2. Instrumentation and Data

On February 22, 1978, between 2237 and 2239 UT (around 0130 MLT) the Triad satellite traversed the postmidnight auroral oval over Scandinavia. Figure 1 displays the temporal variations of the geomagnetic components observed at the observatory of Kiruna, Sweden (67.8° N, 20.4° E). The magnetogram indicates that the pass occurred between two negative magnetic bays and

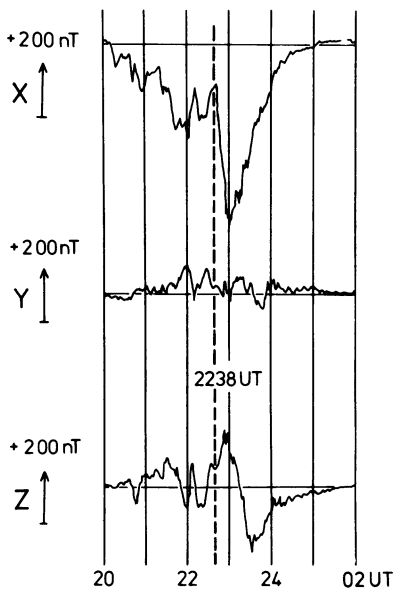


Fig. 1. Standard magnetogram from the geomagnetic observatory at Kiruna for February 22, 1978. The vertical dashed line gives the time when Triad passed approximately over Kiruna

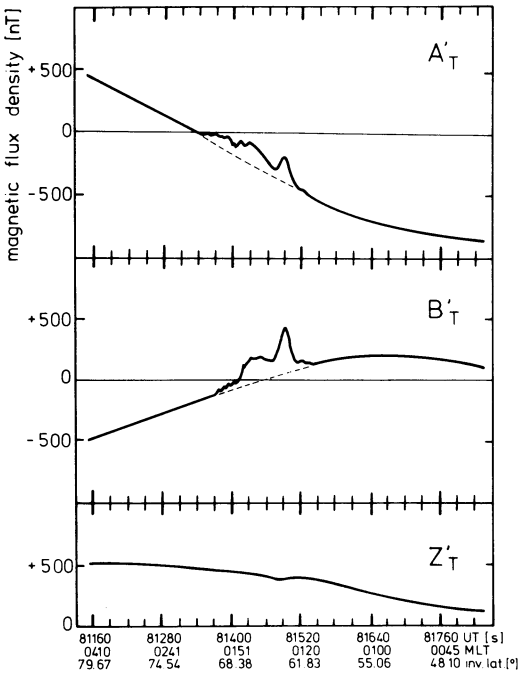


Fig. 2. Magnetic deflections along the satellite trajectory observed on February 22, 1978 around 2238 UT. The A'_T, B'_T, and Z'_T components are more commonly known as A, B, and Z components, but have been renamed to omit confusion with the Kiruna system A, B, and Z components (see Fig. 3 and text). The *dashed lines* in the A'_T and B'_T diagrams give the baseline relative to which the disturbances investigated in our analysis have been defined

during a time when the magnetic field was rather stable intermittently. This is confirmed by our other ground observations. Accordingly, all disturbances observed with the Triad magnetometer along the satellite track could be attributed to spatial variations.

A detailed description of the triaxial magnetometer onboard Triad has been given by Armstrong and Zmuda (1973). Figure 2 shows the magnetic variations observed at a rate of 2.25 samples per axis and per second during the above mentioned time interval. When passing southbound over Scandinavia the A'_T Triad magnetometer axis is directed approximately 30° west of geomagnetic north, and the B'_T axis is directed about 30° south of geomagnetic west. Accordingly, opposite to Alaskan passes, where the A'_T sensor is aligned nearly in the magnetic east-west direction, both horizontal sensors show significant variations. The Z'_T sensor shows only very weak variations. This sensor is directed vertically upwards, and is therefore not antiparallel to the main magnetic field vector.

A complete description of the Scandinavian Magnetometer Array has been given by Küppers et al. (1979). We therefore only display the locations of the magnetometers used in this study in Fig. 3, by giving the spatial distribution of 2 min averaged (2237–2239 UT) equivalent current vectors on the earth's surface (in nT). These vectors have their origin at the station where the corresponding magnetic disturbance has been observed. The 10 s averaged horizontal magnetic field vectors along the satellite track are also drawn into the same figure, with the satellite trajectory projected down to 100 km height along the magnetic field lines. The equivalent current vectors are given relative to

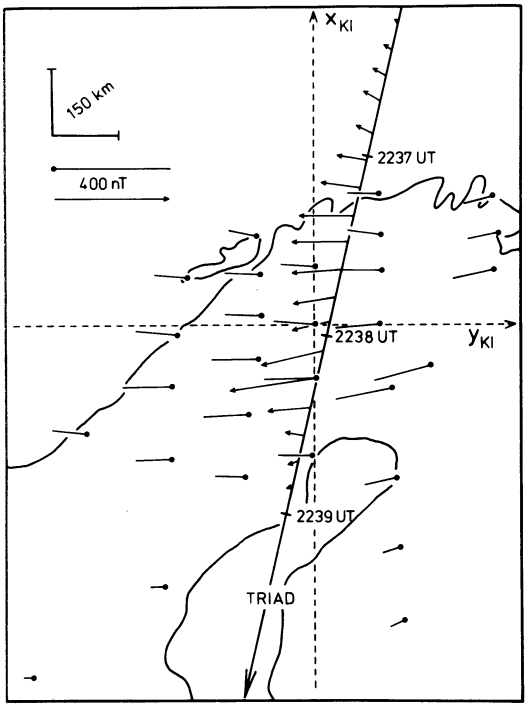


Fig. 3. Two minutes averaged equivalent current vectors on the ground and 10 s averaged horizontal magnetic disturbance vectors along the satellite track (projected down along the fieldlines to 100 km). The equivalent current vectors have their origin (*solid dot*) at the station where the corresponding magnetic disturbance has been observed. Also indicated are the axes of the Kiruna system which is explained in the text

the quiet night level, while the Triad magnetic field vectors are given relative to the baselines shown in Fig. 2. These baselines have been constructed by fitting cubic splines to the undisturbed magnetic components north of 70° and south of 60° invariant latitudes.

It can be seen that the main equivalent current flow constitutes a westward electrojet and that both equivalent current arrows and satellite magnetic fields are nearly everywhere westward directed. The constant westward direction of the equivalent current arrows indicates that the electrojet was flowing uniformly in longitude. The maximum westward magnetic field at 800 km height coincides (as projected down along the field line) with the maximum westward equivalent current arrows on the ground, but a secondary maximum can be seen in the Triad data near the northern coast of Norway. The ground-based magnetometers show no corresponding maximum of equivalent currents, but a slower decrease of current intensity north of the maximum westward current flow than south of it.

The coordinate system indicated in Fig. 3 was described by Küppers et al. (1980) and was named the Kiruna system. It is a Cartesian coordinate system obtained by a stereographic projection of the globe onto a tangential plane centered at Kiruna, Sweden. Cartesian coordinates are very suitable for analysing local ionospheric-magnetospheric current systems as done within this paper. The y_{KI} axis of the system whose origin is situated at Kiruna has been chosen as the tangent to the projection of the line $\phi_c = \phi_c$ (KIR) = 64.8° with ϕ_c denoting the revised corrected geomagnetic latitude as given by Gustafsson (1970). The x_{KI} axis points about 12° west of geographic north at Kiruna. The horizontal components of the magnetic disturbances observed with the satellite and on the ground have been mapped into this system and have been denoted (A_T, B_T) and (A_M, B_M), respectively. The A components are aligned parallel to the x_{KI} axis and the B components parallel y_{KI}. Z_M and Z_T components and z axis are aligned perpendicular to this plane and are positive when vertically downward directed.

3. Separation and Upward Continuation of Two-Dimensional Ground Magnetic Fields

The rather constant westward direction of equivalent current arrows and Triad magnetic fields suggest that it is possible to apply methods of two-dimensional potential theory for computing the external part of the southward magnetic field distribution along the x_{KI} axis at ionospheric heights. Two-dimensionality in this respect means that all quantities are independent of one coordinate which in our case turns out to be y_{KI}.

The latter fact is illustrated by Fig. 4 where we have plotted the observed A_M and Z_M components (average values between 2237 and 2239 UT, see above) versus x_{KI}. The figure shows that there is indeed little y_{KI} dependence especially in the A_M component, while the larger scatter in the Z_M components is very probably due to local induction anomalies or coastal effects (see

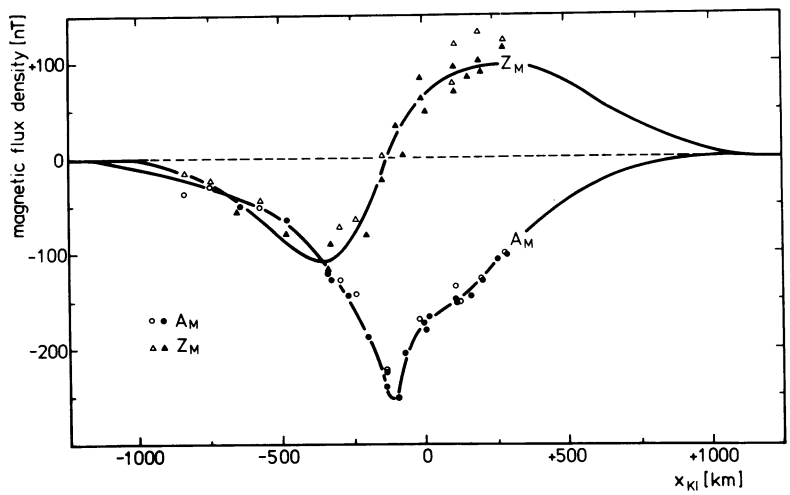


Fig. 4. Average latitude profiles of A_M and Z_M components (*solid lines*) observed on the ground. The *solid dots* and *triangles* denote A_M and Z_M field values observed at stations within $-150\,\text{km} < y_{KI} < +150\,\text{km}$, i.e., close to the Triad trajectory (cf. Fig. 3), while *open dots* and *triangles* denote components observed at more distant locations

Küppers et al. 1979, for a discussion of this topic). Accordingly a two-dimensional analysis can be applied to the (A_M, Z_M) components in a (x_{KI}, z) coordinate system. For that purpose we have constructed an average latitudinal profile of the observed A_M and Z_M component each along the x_{KI} axis, with an extrapolation towards zero on both sides (solid lines in Fig. 4). It should be noted that the A_M profile clearly shows the aforementioned asymmetry around the sharp peak. The Z_M profile crosses the zero level at the minimum of A_M and also shows an asymmetric distribution.

If the A_M and Z_M components are assumed to be periodic along the x_{KI} axis, with $2\pi k_0^{-1}$ defining the basic spatial wavelength (large as compared to the length of our profiles), the following Fourier expansions of the external and internal parts of these components are valid between the ionosphere and the conducting layers in the earth (in these and the following formulas we use x and y instead x_{KI} and y_{KI} for the sake of simplicity)

$$
\begin{aligned}
A_{Me}(x,z) &= -j\sum_{n=0}^{\infty} a_n\, e^{jkx-kz},\\
A_{Mi}(x,z) &= -j\sum_{n=0}^{\infty} b_n\, e^{jkx+kz},\\
Z_{Me}(x,z) &= +\sum_{n=0}^{\infty} a_n\, e^{jkx-kz},\\
Z_{Mi}(x,z) &= -\sum_{n=0}^{\infty} b_n\, e^{jkx+kz},
\end{aligned}
\tag{1}
$$

where $k = n \cdot k_0$ denotes the wavenumber and j the imaginary unit. The A_{Me} and Z_{Me} components are caused by external sources, e.g., in the ionosphere and magnetosphere, and vanish for $z \to \infty$, while the A_{Mi} and Z_{Mi} vanish for $z \to -\infty$ and are caused by internal sources, i.e., induced currents below the earth's surface ($z=0$). The observed total magnetic components on the ground are given by the superposition of the external and internal parts

$$
\begin{aligned}
A_M(x,0) &= -j\sum_{n=0}^{\infty} (a_n+b_n)\, e^{jkx},\\
Z_M(x,0) &= \sum_{n=0}^{\infty} (a_n-b_n)\, e^{jkx}.
\end{aligned}
\tag{2}
$$

A comparison of Eqs. (1) and (2) shows that $F[A_{Me}(x,0)]$, the Fourier transform of the external part of A_M at the ground, i.e., the set of coefficients $-ja_n$, can be determined by a superposition of the Fourier transforms of the observed values of $A_M(x,0)$ and $Z_M(x,0)$ according to

$$
F[A_{Me}(x,0)] = \tfrac{1}{2}(F[A_M(x,0)] + jF[Z_M(x,0)]). \tag{3}
$$

Accordingly, Fourier analysis of the observed A_M and Z_M latitudinal profiles on the ground (solid lines in Fig. 4), combination of the A_M and Z_M Fourier transforms according to Eq. (3), multiplication of the resultant Fourier coefficients with e^{+kh} (h is the assumed height of the base of the ionospheric current layer) and subsequent Fourier synthesis yields the external magnetic north-south component which is due to ionospheric and magnetospheric currents just below the ionospheric current layer. We have chosen $h = 100$ km according to the observed average height distribution of westward electrojets (Kamide and Brekke 1977).

This method is similar to that recently applied by Mersmann et al. (1979). The basic difference is that they separated external and internal parts in the spatial domain by using the Kertz operator, which is basically a Hilbert transform (Kertz 1954; Siebert and Kertz 1957; Weaver 1964), while we separated in the wavenumber domain. In order to avoid the problem of unstable solutions during continuation towards the sources (large factors e^{+kh} greatly enhance small scatter in short wavelength parts; see Mersmann et al. 1979) we have computed the Fourier spectra by a harmonic analysis with a sufficiently large spacing of $\Delta x_{KI} = 120$ km between neighbouring field values which were taken from the solid curves shown in Fig. 4. The average upward continued spectrum has then been synthesized with a 60 km spacing to be comparable with the 10 s averages of the westward B_T component observed by Triad.

In Fig. 5 we display the resultant A_{Me} profile at 100 km height together with the profile of the east-west disturbances B_T observed by Triad. The asymmetry of the A_M and Z_M profiles observed on the ground gives two clearly separated minima in the A_{Me} profile at ionospheric height. Both $A_{Me}(x_{KI}, -h)$ and $B_T(x_{KI})$ have nearly the same shape, the same small-scale structure, an almost identical amplitude (for both the minimum is -400 nT), and they both are confined to nearly the same latitudinal extent (about 1200 km). The wavy structures at the ends of

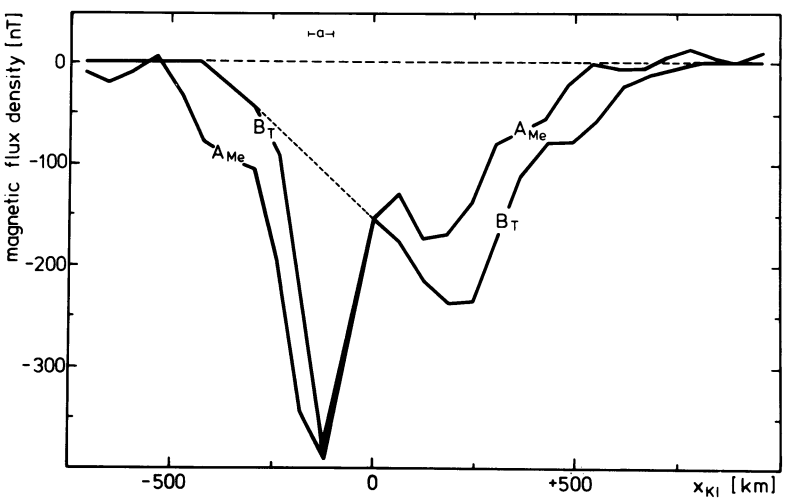


Fig. 5. Latitudinal profiles of the approximately northward external magnetic horizontal component A_{Me} at the height of 100 km (just below the assumed ionospheric current layer) and of the approximately eastward horizontal disturbance B_T observed by Triad and projected down to 100 km along the fieldlines. The A_{Me} and B_T curves can be regarded as giving westward and southward ionospheric height-integrated current density if one substitutes 100 nT by 160 mAm^{-1} for A_{Me} or 100 nT by 80 mAm^{-1} for B_T, respectively. The *horizontal dashed line* gives the zero-level, the other dashed line in the region of the small-scale structure gives the approximate magnetic flux density B_T associated with the large-scale electrojet system in this area and the bar around the small 'a' denotes the approximate position of the observed auroral arc (see text)

the A_{Me} profile belong to a spatial wavelength of 240 km, i.e., the smallest wavelength resolved in our analysis, and are very probably related to the above mentioned instability problem. The factor e^{+kh} amounts to about 15 for this wavelength, and accordingly the observed amplitude of ± 15 nT of this wavelength at $h=100$ km seems to be due to height-continued errors in A and Z that are of the order of 1 nT at ground.

4. Ionospheric and Field-Aligned Currents and the Hall to Pedersen Conductivity Ratio

For a two-dimensional situation like the present one we can relate the magnetic field distributions displayed in Fig. 5 to the *real* (not merely *equivalent*) auroral zone current system in the following way:

As Boström (1964) first clearly pointed out this system (if two-dimensional and if confined to the auroral latitudes) consists of the electrojet whose magnetic field is observed on the ground, and of meridional currents that are field-aligned above the ionosphere and are closed within the lower ionosphere (and in the magnetosphere). The meridional currents possess a toroidal magnetic field that may be parallel or antiparallel to the electrojet and that is unobservable on the ground.

Since in our case the vertical thickness of the electrojet, that flows in the negative y direction is small as compared to its latitudinal extent, its large-scale features may be described by a surface current density J_y (negative in our case). This quantity is related to the height-continued northward magnetic flux density according to

$$J_y(x)=\frac{2}{\mu_0} A_{Me}(x, -h) \tag{4}$$

if h is the height of the base of the electrojet layer (note that z denotes depth). Accordingly, the curve denoted by A_{Me} in Fig. 5 gives directly the electrojet height-integrated current density J_y if 100 nT are replaced by 160 mAm^{-1}. The maximum westward current density is 620 mAm^{-1} for the southern and 280 mAm^{-1} for the northern peak of the J_y distribution. Between these two extrema the surface current density decreases to about 210 mAm^{-1}, and the total current is approximately $2.4 \cdot 10^5$ A.

The magnetic component B_T as observed by the Triad satellite (Fig. 5) must be interpreted as the toroidal magnetic flux density generated by the poloidal (meridional) current system, as it has been done for similar disturbances since the first analyses of Triad data (e.g. Armstrong and Zmuda 1970; Zmuda and Armstrong 1974). This system may be described by a spatial field-aligned current density $j_\parallel(x)$ (positive if parallel to the earth's magnetic field) above the ionospheric current layer, and by a surface current density $J_x(x)$ that connects the field-aligned currents within the ionosphere. Because B_T is zero outside this current system we get from Maxwell's first equation

$$j_\parallel(x)=\frac{1}{\mu_0}\frac{dB_T}{dx}. \tag{5}$$

Current continuity within the poloidal system may be expressed by

$$\frac{dJ_x}{dx}=j_\parallel(x). \tag{6}$$

The combination of these two equations yields

$$J_x(x)=\frac{1}{\mu_0} B_T(x). \tag{7}$$

This equation shows that the second curve in Fig. 5 may be interpreted as the height-integrated density of the northward current (negative, i.e., southward in this case) in the electrojet region if 100 nT are replaced by 80 mAm^{-1}. Accordingly, we recognize a maximum southward current density of 310 mAm^{-1}, about 190 mAm^{-1} for the northern secondary extremum, and 120 mAm^{-1} in between.

The field-aligned current density has been calculated by differentiating the B_T profile (Fig. 5) according to Eq. (5). The result is displayed in the upper part of Fig. 6. The $j_\parallel(x)$ curve shows relatively pronounced structure that will be discussed below. The maximum current densities are 2.5 μAm^{-2} with upward and 1.5 μAm^{-2} with downward flow. The total field-aligned surface current density is 370 mAm^{-1} for both up- and downward flowing current.

Since the large-scale field-aligned currents have a magnetospheric source (e.g. Boström 1975; Rostoker and Boström 1976) the associated ionospheric electric field must be southward directed along the x_{KI} axis for the given gross pattern of upward field-aligned current flow in the south and downward in the north. Accordingly, we can interpret $J_x=\Sigma_P E_x$ as Pedersen and $J_y=\Sigma_H E_x$ as Hall current and get from Eqs. (4) and (7)

$$A_{Me}(x, -h)=\frac{1}{2}\cdot\frac{\Sigma_H}{\Sigma_P}(x)\cdot B_T(x). \tag{8}$$

The latitudinal profile of the ratio Σ_H/Σ_P which results from the profiles of A_{Me} and B_T (Fig. 5) according to this relation is given in the lower panel of Fig. 6. It shows a decrease with increasing latitude from values close to 4 near the southern boundary of the current system and in the region of intense upward field-aligned current flow to values close to 1 in the north where a broad region of downward field-aligned current is observed.

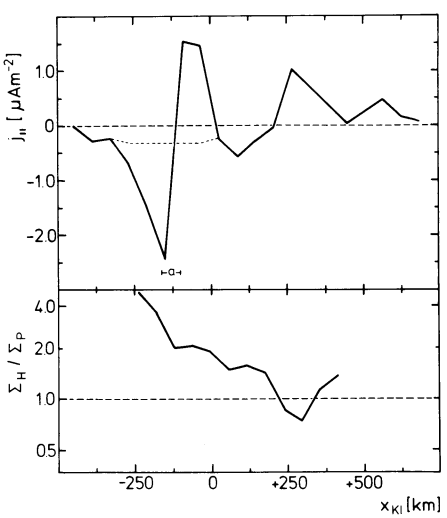


Fig. 6. The upper diagram displays the latitudinal distribution of field-aligned current density calculated by differentiating the B_T profile given in Fig. 5. Positive values denote downward currents. *Dashed line* and bar around '*a*' have the same notation as in Fig. 5. The lower panel gives the derived (cf. text) latitudinal profile of the ratio between Hall and Pedersen conductivity on a logarithmic scale

5. Magnetic Variations in the North-South and Vertical Component at Triad Altitude

In the previous sections we were concerned with the strong magnetic variations observed by Triad in the B_T (approximately geomagnetic east-west) direction and related to the meridional current system of balanced field-aligned currents and Pedersen currents. In this section we will try to interpret the comparatively smaller variations in the A_T component (approximately geomagnetic north-south; maximum amplitude about 50 nT) and the even smaller ones in the Z_T component (vertically downward; maximum amplitude about 20 nT). One possible source of magnetic variations in the A_T and Z_T components at 800 km altitude is the magnetic field associated with the electrojet. We have calculated this contribution from $J_y(x)$ as given by the curve $A_{Me}(x)$ in Fig. 5 in combination with Eq. (4). The result is given by the curves A_I and Z_I in the upper part of Fig. 7. One can see that they are changing only slowly with latitude and that they have maximum amplitudes of 50 and 20 nT for A_I and Z_I, respectively. Magnetic fields of more distant sources like the ring current should change even more slowly and may hardly be distinguishable from the quiet level. If we subtract A_I and Z_I from A_T and Z_T, respectively, we should get the residual magnetic disturbances due to probably more local sources. Since A_I and Z_I have a very gradual slope south of 60° and north of 70° invariant latitude, i.e., in the regions where we have fitted the cubic splines to construct the baseline, no decision can be made between these variations and the well-known slow variations due to the spacecraft attitude uncertainty. Accordingly, we have here first subtracted the fields of ionospheric origin and then constructed a new baseline by fitting cubic splines to the northern and southern 'undisturbed' region. The results are given in the lower part of Fig. 7 and the variations indeed reflect a quite local character.

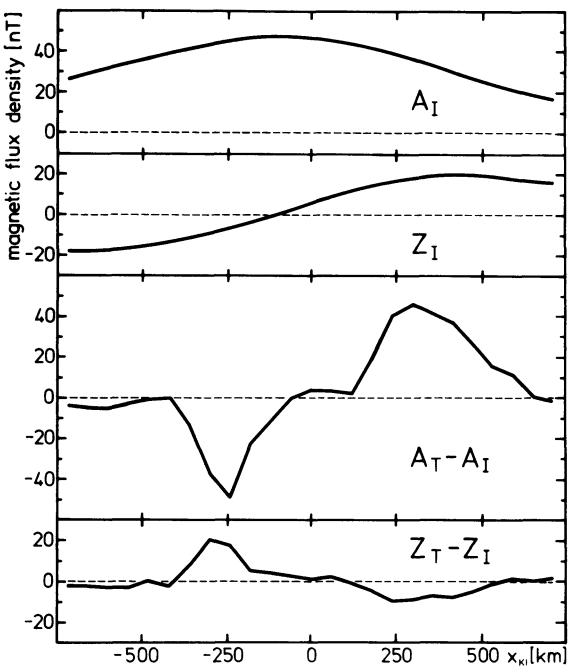


Fig. 7. The two upper panels display the magnetic flux density of the ionospheric westward electrojet (cf. Fig. 5) at a Triad altitude. The two lower panels give the magnetic disturbances observed by Triad at 800 km altitude in the northsouth and vertical component after subtraction of the above-mentioned fields of ionospheric origin

The (A_T-A_I) variations show an almost antisymmetric shape with respect to $x_{KI}\approx 0$. If one projects this location along the magnetic field lines down to 100 km altitude in order to compare it with the location where the large-scale field-aligned currents change their direction as given in the upper part of Fig. 6 (around $x_{KI}=200$ km), one finds that both locations approximately coincide. One also finds that the northern maximum (A_T-A_I) is located at the northern edge of the field-aligned region and that the southern minimum is close to the southern edge. Both these facts hint to the explanation that the (A_T-A_I) magnetic variations constitute a leakage of toroidal magnetic flux $\int B_T$ due to a westward decrease in the field-aligned current intensity. This decrease is consistent with a slight decrease of the westward electrojet in the same direction which can be noted when carefully examining Fig. 3.

The actual leakage magnetic flux density depends on both the gradient of the westward decrease and on the longitudinal location with respect to the central meridian and the western termination of the electrojet. With the data set available we cannot determine these parameters, but we have calculated the A_T component in 800 km altitude by assuming that the Pedersen-field-aligned current system given in Figs. 5 and 6 decreases linearly by 10% per 100 km in the westward direction and that the western termination was 1000 km and the central meridian 2000 km apart. The resultant curve is displayed in Fig. 8 and one can see that there is a reasonable agreement with the observed (A_T-A_I) curve of Fig. 7 regarding shape, zero-crossover, location of the extrema, and amplitude in the central region between the two arrows.

The observed (A_T-A_I) profile cannot show the very gradual increase and decrease of the calculated A_T south and north of the locations indicated by the arrows, since these are the regions

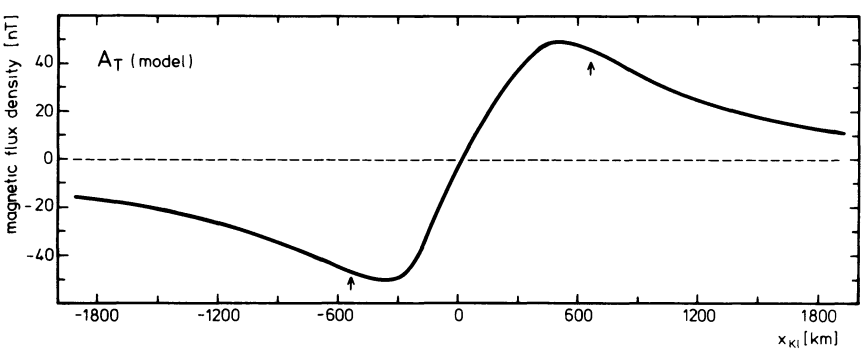


Fig. 8. Calculated northward component of magnetic flux density due to a leakage of the toroidal magnetic flux (the model current system is explained in the text). Outside the region between the *two arrows* cubic splines have been fitted to the observed data for baseline determination and no agreement between calculated and observed A_T components (Fig. 7) can be expected

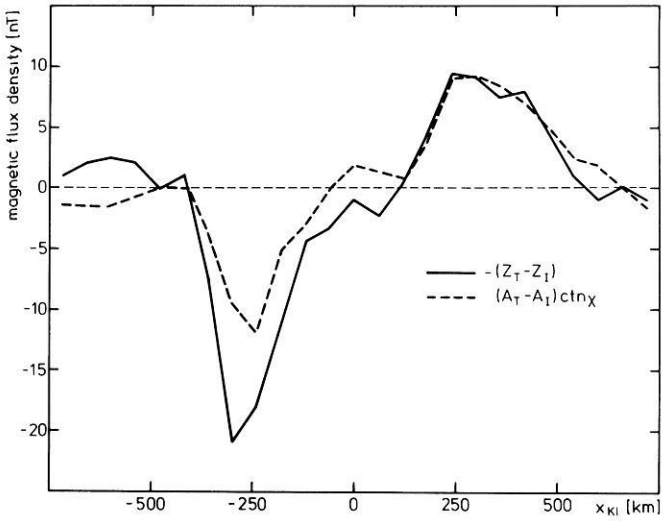


Fig. 9. Comparison of the two quantities related in Eq. (9) (χ denotes the inclination of the earth's magnetic field)

where cubic splines have been fitted to the variations observed by Triad for the baseline determination and consequently the $(A_T - A_I)$ curve must be close to this baseline. We should also note that a westward decrease of 10% per 100 km is about twice as much as the westward decrease of the southward magnetic field observed on the ground, but we feel that our calculations give at least good evidence that the observed $(A_T - A_I)$ variations are due to the leakage to toroidal magnetic flux.

If the Triad Z sensor would be aligned along the magnetic field direction one would expect to observe no effect of field-aligned currents in this sensor. But the sensor is actually aligned gravitationally and therefore the following relation holds (with $\chi(x)$ denoting the latitude dependent inclination of the fieldlines as given by geomagnetic charts):

$$-[Z_T(x) - Z_I(x)] = [A_T(x) - A_I(x)]\ \mathrm{ctn}\ \chi(x). \tag{9}$$

We have calculated the above two quantities and display them in Fig. 9. It can be seen that both curves agree rather well and it can therefore be assumed that the $(Z_T - Z_I)$ variations are due to the same sources as the $(A_T - A_I)$ variations, i.e., very probably due to the leakage of toroidal magnetic flux.

6. Discussion

Our analysis has provided results for the distribution of ionospheric Hall and Pedersen currents and field-aligned currents associated with a westward electrojet in the morning sector. The large-scale features of such current systems are already rather well known and generally in agreement with our results. Small-scale structures in the latitudinal current density distribution are still a much debated topic (Kamide 1979) and the relationship of the sharp southern peak in ionospheric current density and the intense field-aligned current sheets with an auroral arc will be discussed as the second topic. Another important result is the latitudinal dependence of the conductivity ratio, and from this we will draw some conclusions about the particle precipitation in the morning sector during the present event.

Large-Scale Currents

The electrojet constitutes a westward Hall current of about $2.4 \cdot 10^5$ A flowing in nearly exactly the same region where grossly balanced field-aligned current sheets of 370 mAm^{-1} are observed flowing downward in the north and upward in the south. These currents are linked in the ionosphere by southward Pedersen currents, and the toroidal magnetic field caused by this meridional current system was observed by Triad above the ionosphere and was of the order of 200–400 nT. In contrast, the meridional magnetic field of the electrojet was about 30–50 nT at the Triad altitude.

The values given above are within the range found in earlier studies with the Chatanika radar and the Triad satellite for the same local time sector (e.g. Brekke et al. 1974; Kamide and Brekke 1975; Iijima and Potemra 1976, 1978; Horwitz et al. 1978a, b; Kamide and Horwitz 1978). Also our observation that balanced field-aligned currents close along one meridian is in agreement with the statistical field-aligned current flow pattern given by other authors (Yasuhara et al. 1975; Sugiura and Potemra 1976; Iijima and Potemra 1976, 1978; Hughes and Rostoker 1977), who all note balanced field-aligned currents during 2 to 4 h around magnetic midnight.

We should note that our assumption of a purely meridional southward directed ionospheric electric field caused by the balanced field-aligned currents (and consequently the westward current to be a pure Hall current) is not without contradiction. For example Hughes and Rostoker (1977, 1979), using observations of westward electric fields around midnight by Mozer and Lucht (1974), proposed that the quiet time westward electrojet in the 22–02 MLT sector is mainly a Pedersen current while the Hall current is directed more northward. After examining a great amount of published electric field patterns (Heppner 1973, 1977; Holzworth et al. 1977; Horwitz et al. 1978a, b; Maynard 1974; Wedde et al. 1977), we have found that westward electric fields are confined mostly to the Harang-discontinuity region (where indeed northward Hall currents may be found; e.g. Kamide and Horwitz 1978; Baumjohann 1979; Baumjohann et al. 1980) in the late evening sector. Furthermore southward directed fields appear to predominate in the postmidnight sector in support of our earlier assumptions. This apparent contradiction may be resolved by the fact that Hughes and Rostoker (1977, 1979) discussed the quiet time current configuration and used quiet time electric field data while our data and most of the above cited electric field patterns are representative for more disturbed times, when the westward Pedersen current is concentrated behind the surge in the evening sector (Rostoker and Hughes 1979).

Therefore we suggest that during disturbed times the zonal westward closure current of the net field-aligned currents around magnetic midnight (Kamide et al. 1976b, c) is the Hall current whose magnetic signatures have been observed on the ground. This closure is in agreement with a model already proposed by Heppner et al. (1971a, b), and has been verified recently by Baumjohann et al. (1980) for the eastward electrojet in the dusk sector by comparing two-dimensional distributions of ionospheric electric and ground magnetic fields.

Small-Scale Current System Associated With an Auroral Arc

Mersmann et al. (1979) have found an electrojet current density that was not a smooth latitudinal distribution from their study

of an eastward electrojet with essentially the same method as applied here. In their case a small decrease in current density was found, and evidence was given that this decrease was possibly due to a local westward current inside an auroral arc. The existence of this arc could not be verified due to cloudiness and some ambiguity remained. In our case a similar structure is found both in the ionospheric and field-aligned current densities, but it is much stronger in amplitude, and is of a greater spatial wavelength (about 500 km). Again the weather in Scandinavia was rather cloudy but one of the Scandinavian all-sky-cameras produced analysable data. This camera was located at Hankasalmi in southern Finland and during the time of the Triad pass a rather stable quiet arc was seen near the northern horizon of the camera. This arc was located approximately between $x_{KI} = -160$ and -100 km (H.J. Opgenoorth private communication, 1979), which means in the region where we observed intense upward field-aligned current and strong westward ionospheric current flow (cf. Figs. 5 and 6).

Accordingly, we interpret the peaks in the observed A_{Me} and B_T curves (Fig. 5) at $x_{KI} = -120$ km as enhanced westward Hall and southward Pedersen current flow between a pair of relatively thin sheets (each about 100 km wide) of intense field-aligned currents (cf. Fig. 6 upper panel) that is associated with the observed auroral arc. The center of this system is situated about 300 km south of the secondary peaks which now may be contributed to the maximum of a broad westward electrojet current density distribution. Of the total 620 mAm^{-1} Hall current density inside or near the arc, about 450 mAm^{-1} may be attributed solely to the small-scale arc system. The arc-associated westward Hall current density is therefore greater than the maximum current density of the broad westward electrojet (270 mAm^{-1}), and the total arc-associated Hall current contributes about 20% to the total westward current. Of the maximum southward Pedersen current density of 310 mAm^{-1} about 220 mAm^{-1} may be arc-associated and are therefore of about the same magnitude as the current density near the maximum of the westward electrojet. The maximum field-aligned current densities (cf. Fig. 6, upper part) can be nearly solely attributed to the arc and we get 2 μAm^{-2} associated with the arc compared with a maximum of 1 μAm^{-2} associated with the main electrojet. In terms of sheet current densities, this means that a current of about 200 mAm^{-1} flows downward to the north of the arc and flows upward above or south of the arc. This field-aligned current closes via the southward Pedersen currents. Compared to the electrojet-associated field-aligned current, whose sheet current densities may be estimated to be also about 200 mAm^{-1}, the total arc-associated vertical currents have the same magnitude.

It should be noted that the current densities given are more minimum current densities for the arc system, since the latitudinal resolution of our analysis is restricted to spatial wavelength greater than 200 km. Auroral arcs and most probably also the associated current systems are much smaller (e.g. Davis 1978) and our data may display smoother curves and lower peaks than were actually there.

We have also compared our observations with results of rather rare earlier studies on morning sector auroral arcs. Kamide and Rostoker (1977) noted that discrete arcs in the morning sector are found in the region of intense upward field-aligned currents in the southern half of the westward electrojet. Beaujardiere et al. (1977) found, by means of incoherent scatter observations, an arc-associated enhancement of the westward current flow of about 160 mAm^{-1}. Because of their fine resolution rocket-borne experiments are best suited to study phenomena associated with auroral arcs. But to the best of our knowledge, only one rocket flight into postmidnight aurora has been made (Sesiano and Cloutier 1976). Since this rocket was flown through a highly structured multiple arc system the results are hardly comparable with ours. The field-aligned sheet current densities determined in our study are within the range (160 to 1200 mAm^{-1}) obtained from the Rice University rocket flights into auroral arcs in the premidnight sector (Anderson and Vondrak 1975).

We should note once more that our interpretation of the westward current as a Hall current, and the southward current as a Pedersen current is based on the assumption that the field-aligned currents are driven by a magnetospheric source with the ionosphere acting as a load. As described earlier, this mechanism is commonly accepted for the large-scale currents, but it is still a matter of debate if the field-aligned current pairs associated with auroral arcs are driven by the magnetosphere (e.g. Mallinckrodt and Carlson 1978) or by a polarization electric field in the ionosphere (e.g. Carlson and Kelley 1977). The second model implies an ambient zonal westward electric field which, as described earlier, is rarely observed in this MLT sector. Furthermore the non-existence of this electric field can be concluded from the rather stable location of the auroral arc (arcs drift southward under $E \times B$ motion of a westward field; see for example Kelley et al. 1971). Therefore we tend to believe that a magnetospheric source can indeed be assumed, but this question can only be decided by simultaneous electric field measurements and we hope that a future study incorporating Triad, the Scandinavian Magnetometer Array, and the STARE-radars (Greenwald et al. 1978) will help in this respect.

Hall to Pedersen Conductivity Ratio

South of $x_{KI} = 0$ km the Σ_H/Σ_P ratio varies between 2 and 4 and is within the range given by Brekke et al. (1974) based on their Chatanika radar observations. North of $x_{KI} = 0$ this ratio reduces to values around and slightly below 1, and is therefore smaller than the minimum value given by Brekke et al. (1974). But Evans et al. (1977), who computed ionospheric conductivities based on auroral electron data obtained during a sounding rocket flight, determined values between 0.8 and 1.4. Both groups used a different atmospheric model and the difference in the ratios found may very probably be attributed to uncertainties in these models, while our calculations are free from these ambiguities.

The ratio between the two conductivities allows some conclusions on the energy of precipitating particles, since energetic auroral electrons penetrating the atmosphere reach different altitude levels depending on their energy (Rees 1963). Experimental evidence for this can be found for example in Brekke et al. (1974) and Evans et al. (1977) who both noted that maximum enhancements of the Σ_H/Σ_P ratio are found to be coincident with energetic particle precipitation.

For our case this means that particles with higher energies precipitate in the southern third of the westward electrojet region, where the most intense Hall and upward field-aligned current densities and the auroral arc have been found. The average energy of the precipitating particles gradually decreases towards north, where the field-aligned currents flow downward. This latitudinal distribution of particle precipitation is oppositely to that in the evening sector (see, for example, Lui et al. 1977) but agrees with other results found in the morning sector

by direct measurements with the Isis (McDiarmid et al. 1975; Klumpar et al. 1976) and Injun (Frank and Ackerson 1971) satellites, by monochromatic all-sky observations (Mende and Eather 1976) and by studying the relative location of discrete aurora in the electrojet region (Kamide et al. 1976b; Kamide and Rostoker 1977).

7. Summary and Conclusions

In the present paper, for the case of a morning sector pass, we have demonstrated the usefulness of coordinated two-dimensional ground-based and Triad satellite magnetic observations. We have shown that, for the two-dimensional case studied, it is possible to derive quantitatively the real ionospheric-field-aligned current system and the latitudinal distribution of the Σ_H/Σ_P ratio and that it is possible to explain also variations observed by Triad in the north-south and vertical component. This has been done by using only two very basic and well accepted assumptions: that the balanced field-aligned currents are of magnetospheric origin thereby causing a purely southward electric field, and that the altitude of the ionospheric current layer is 100 km.

For the event which we investigated the more detailed results include the following:

1. A broad westward electrojet (about 1200 km wide) was observed at about 0130 MLT. It constituted a Hall current of about $2.4 \cdot 10^5$ A flowing in the same region where balanced upward and downward flowing field-aligned current sheets of about 200 mAm^{-1} were observed. These currents where closed in the auroral ionosphere along the same meridian by southward Pedersen currents of about the same maximum current density as the field-aligned currents.

2. South of the maximum of the broad electrojet and near the location of a quiet auroral arc another latitudinally small but very intense peak in the ionospheric current density was observed. This small-scale arc-associated current system had essentially the same configuration as the large-scale electrojet system. The about 250 km wide westward Hall current had a maximum density of 450 mAm^{-1}, greater than the maximum current density of the main electrojet, and contributed 20% of the total westward current. The field-aligned currents flew downward in the north and upward over or south of the arc and their maximum current densities of about 2 μAm^{-2} were twice as high as those of the electrojet associated pair. The total field-aligned current flow was closed by southward Pedersen currents of 220 mAm^{-1}.

3. While we found Σ_H/Σ_P ratios close to 1 in the northern half of the westward electrojet and in the region of downward field-aligned currents, this ratio increased to values between 2 and 4 in the southern half with the highest values near the location of the auroral arc and the intense upward field-aligned currents. Relating these ratios to the energy of precipitating particles we have concluded that in contrast to the evening sector energetic particles precipitate in the southern half of the auroral oval (as defined by the electrojet borders) where the discrete arc has been observed. Lower energetic particle precipitation appears to occur in the northern half.

4. The disturbances observed by the geomagnetic northsouth component of the Triad magnetometer (perpendicular to the field-aligned current sheets) may be explained by a leakage of the toroidal magnetic flux due to a probable decrease of the field-aligned current intensity in the westward direction. Due to the inclination of the earth's magnetic field these variations can also be seen in the gravitationally aligned vertical component.

It should be mentioned, that the present conclusions are derived on the basis of only one case studied. Future work has to show if the conclusions may be generalized.

Acknowledgements. We are greatly indebted to those past and present members of the magnetometer group at the University of Münster, who were involved in collecting the magnetic data. The magnetic observations were performed in cooperation with the Royal Institute of Technology at Stockholm, the Finnish Meteorological Institute at Helsinki, the University at Tromsö, the University of Bergen, the Geophysical Observatory Sodankylä of the Finnish Academy of Science and Letters, the Kiruna Geophysical Institute, the University of Oulu, and the Aarhus University. To these institutions our sincere thanks are due for permanent support. Invaluable assistance in the TRIAD data collection and processing was provided by S. Favin, J. Du Brul, and J. Nearvy, and the entire project was made possible by the Space Department of APL/JHU. We also wish to thank H.J. Opgenoorth for information on the auroral condition during the Triad pass and Ch. Sucksdorff and G. Gustafsson for additional magnetic data. We also thank G. Gustafsson and the Kiruna Geophysical Institute for assistance in the acquisition of the Triad data. This work has been supported financially by the Deutsche Forschungsgemeinschaft, the National Science Foundation, and the Office of Naval Research. We are grateful to Y. Kamide and G. Rostoker for their instructive referees' comments.

References

Alfvén, H.: A theory of magnetic storms and of the aurorae I. R. Swedish Acad. Sci., Proc. 3 Ser.: **18**, No. 3, 1939

Alfvén, H.: A theory of magnetic storms and of the aurorae II and III. R. Swedish Acad. Sci., Proc. 3 Ser.: **18**, No. 9, 1940

Anderson, H.R., Vondrak, R.R.: Observations of Birkeland currents at auroral latitudes. Rev. Geophys. Space Phys. **13**, 243–262, 1975

Armstrong, J.C., Akasofu, S.-I., Rostoker, G.: A comparison of satellite observations of Birkeland currents with ground observations of visible aurora and ionospheric currents. J. Geophys. Res. **80**, 575–586, 1975

Armstrong, J.C., Zmuda, A.J.: Field-aligned current at 1100 km in the auroral region measured by satellite. J. Geophys. Res. **75**, 7122–7127, 1970

Armstrong, J.C., Zmuda, A.J.: Triaxial magnetic measurements of field-aligned currents at 800 km in the auroral region: initial results. J. Geophys. Res. **78**, 6802–6807, 1973

Baumjohann, W.: Spatially inhomogeneous current configurations as seen by the Scandinavian Magnetometer Array. In: Proceedings of the International Workshop on Selected Topics of Magnetospheric Physics, Japanese IMS Committee, ed., pp. 35–40, Tokyo, 1979

Baumjohann, W., Untiedt, J., Greenwald, R.A.: Joint two-dimensional observations of ground magnetic and ionospheric electric fields associated with auroral zone currents. 1. Three-dimensional current flows associated with a substorm-intensified eastward electrojet. J. Geophys. Res. in press, 1980

Beaujardiere, O. de la, Vondrak, R., Baron, M.: Radar observations of electric fields and currents associated with auroral arcs. J. Geophys. Res. **82**, 5051–5062, 1977

Birkeland, K.: The Norwegian Aurora Polaris expedition, 1902–1903, Vol. 1: On the cause of magnetic storms and the origin of terrestrial magnetism, sect. 1. Christiania: H. Aschehoug 1908

Birkeland, K.: The Norwegian Aurora Polaris expedition 1902–1903, Vol. 1: On the cause of magnetic storms and the origin of terrestrial magnetism, sect. 2. Christiania: H. Aschehoug 1913

Boström, R.: A model of the auroral electrojets. J. Geophys. Res. **69**, 4983–5000, 1964

Boström, R.: Mechanisms for driving Birkeland currents. In: Physics of the hot plasma in the magnetosphere, B. Hultqvist, L. Stenflo, eds.: pp. 341–351. New York: Plenum Press 1975

Brekke, A., Doupnik, J.R., Banks, P.M.: Incoherent scatter measurements of *E* region conductivities and currents in the auroral zone. J. Geophys. Res. **79**, 3773–3790, 1974

Carlson, C.W., Kelley, M.C.: Observation and interpretation of particle and electric field measurements inside and adjacent to an active auroral arc. J. Geophys. Res. **82**, 2349–2360, 1977

Chapman, S.: The electric current system of magnetic storm. Terr. Magn. Atmos. Electr. **40**, 349–370, 1935

Davis, T.N.: Observed characteristics of auroral forms. Space Sci. Rev. **22**, 77–113, 1978

Evans, D.S., Maynard, N.C., Trøim, J., Jacobsen, T., Egeland, A.: Auroral vector electric field and particle comparisons. 2. Electrodynamics of an arc. J. Geophys. Res. **82**, 2235–2249, 1977

Frank, L.A., Ackerson, K.L.: Observations of charged particle precipitation into the auroral zone. J. Geophys. Res. **76**, 3612–3643, 1971

Fukushima, N.: Equivalence in ground geomagnetic effect of Chapman-Vestine's and Birkeland-Alfvén's electric current systems for polar magnetic storms. Rep. Ionos. Space Res. Jap. **23**, 219–227, 1969

Greenwald, R.A., Weiss, W., Nielsen, E., Thomson, N.R.: STARE: A new radar auroral backscatter experiment in northern Scandinavia. Radio Sci. **13**, 1021–1039, 1978

Gustafsson, G.: A revised corrected geomagnetic coordinate system. Ark. Geofys. **5**, 595–617, 1970

Heppner, J.P.: High latitude electric fields and the modulations related to interplanetary magnetic field parameters. Radio Sci. **8**, 933–948, 1973

Heppner, J.P.: Empirical models of high-latitude electric fields. J. Geophys. Res. **82**, 1115–1125, 1977

Heppner, J.P., Stolarik, J.D., Wescott, E.M.: Electric-field measurements and the identification of currents causing magnetic disturbances in the polar cap. J. Geophys. Res. **76**, 6028–6053, 1971a

Heppner, J.P., Stolarik, J.D., Wescott, E.M.: Field aligned continuity of Hall current electrojets and other consequences of density gradients in the auroral ionosphere. In: The radiating atmosphere, B.M. McCormac, ed.: pp. 407–426. Dordrecht: D. Reidel 1971b

Holzworth, R.H., Berthelier, J.-J., Cullers, D.K. Fahleson, U.V., Fälthammer, C.-G., Hudson, M.K., Jalonen, L., Kelley, M.C., Kellog, P.J., Tanskanen, P., Temerin, M., Mozer, F.S.: The large-scale ionospheric electric field: its variation with magnetic activity and relation to terrestrial kilometric radiation. J. Geophys. Res. **82**, 2735–2742, 1977

Horwitz, J.L., Doupnik, J.R., Banks, P.M.: Chatanika radar observations of the latitudinal distributions of auroral zone electric fields, conductivities, and currents. J. Geophys. Res. **83**, 1463–1481, 1978a

Horwitz, J.L., Doupnik, J.R., Banks, P.M., Kamide, Y., Akasofu, S.-I.: The latitudinal distributions of auroral zone electric fields and ground magnetic perturbations and their response to variations in the interplanetary magnetic field. J. Geophys. Res. **83**, 2071–2084, 1978b

Hughes, T.J., Rostoker, G.: Current flow in the magnetosphere and ionosphere during periods of moderate activity. J. Geophys. Res. **82**, 2271–2282, 1977

Hughes, T.J., Rostoker, G.: A comprehensive model current system for high-latitude magnetic activity – I. The steady state system. Geophys. J. R. Astron. Soc. **58**, 525–569, 1979

Iijima, T., Potemra, T.A.: The amplitude distribution of field-aligned currents at northern high latitudes observed by Triad. J. Geophys. Res. **81**, 2165–2174, 1976

Iijima, T., Potemra, T.A.: Large-scale characteristics of field-aligned currents associated with substorms. J. Geophys. Res. **83**, 599–615, 1978

Kamide, Y.: Recent progress in observational studies of electric fields and currents in the polar ionosphere: A review. Antarct. Rec. **63**, 61–231, 1979

Kamide, Y., Akasofu, S.-I.: The auroral electrojet and field-aligned current. Planet. Space Sci. **24**, 203–213, 1976

Kamide, Y., Akasofu, S.-I., Brekke, A.: Ionospheric currents obtained from the Chatanika radar and ground magnetic perturbations at the auroral latitude. Planet. Space Sci. **24**. 193–201, 1976a

Kamide, Y., Akasofu, S.-I., Rostoker, G.: Field-aligned currents and the auroral electrojet in the morning sector. J. Geophys. Res. **81**, 6141–6147, 1976b

Kamide, Y., Yasuhara, F., Akasofu, S.-I.: A model current system for the magnetospheric substorm. Planet. Space Sci. **24**, 215–222, 1976c

Kamide, Y., Brekke, A.: Auroral electrojet current density deduced from the Chatanika radar and from the Alaska meridian chain of magnetic observatories. J. Geophys. Res. **80**. 587–594, 1975

Kamide, Y., Brekke, A.: Altitude of the eastward and westward auroral electrojets. J. Geophys. Res. **82**, 2851–2853, 1977

Kamide, Y., Horowitz, J.L.: Chatanika radar observations of ionospheric and field-aligned currents. J. Geophys. Res. **83**. 1063–1070, 1978

Kamide, Y., Rostoker, G.: The spatial relationship of field-aligned currents and auroral electrojets to the distribution of nightside auroras. J. Geophys. Res. **82**, 5589–5608, 1977

Kelley, M.C., Starr, J.A., Mozer, F.S.: Relationship between magnetospheric electric fields and the motion of auroral forms. J. Geophys. Res. **76**, 5269–5277, 1971

Kertz, W.: Modelle für erdmagnetisch induzierte elektrische Ströme im Untergrund. Nachr. Akad. Wiss. Göttingen. Math. Phys. Kl. 101–110, 1954

Klumpar, D.M., Burrows, J.R., Wilson, M.D.: Simultaneous observations of field-aligned currents and particle fluxes in the post-midnight sector. Geophys. Res. Lett. **3**, 395–398. 1976

Küppers, F., Untiedt, J., Baumjohann, W., Lange, K., Jones. A.G.: A two-dimensional magnetometer array for ground-based observations of auroral zone electric currents during the International Magnetospheric Study (IMS). J. Geophys. **46**, 429–450, 1979

Lui, A.T.Y., Venkatesan, D., Anger, C.D., Akasofu, S.-I., Heikkila, W.J., Winningham, J.D., Burrows, J.R.: Simultaneous

observations of particle precipitations and auroral emissions by the Isis 2 satellite in the 19–24 MLT sector. J. Geophys. Res. **82**, 2210–2226, 1977

Mallinckrodt, A.J., Carlson, C.W.: Relations between transverse electric fields and the field-aligned currents. J. Geophys. Res. **83**, 1426–1432, 1978

Maynard, N.C.: Electric field measurements across the Harang discontinuity. J. Geophys. Res. **79**, 4620–4631, 1974

McDiarmid, I.B., Burrows, J.R., Budzinski, E.E.: Average characteristics of magnetospheric electrons (150 eV to 200 eV) at 1400 km. J. Geophys. Res. **80**, 73–79, 1975

Mende, S.B., Eather, R.H.: Monochromatic all-sky-observations and auroral precipitation patterns. J. Geophys. Res. **81**, 3771–3780, 1976

Mersmann, U., Baumjohann, W., Küppers, F., Lange, K.: Analysis of an eastward electrojet by means of upward continuation of ground-based magnetometer data. J. Geophys. **45**, 281–298, 1979

Mozer, F.S., Lucht, P.: The average auroral zone electric field. J. Geophys. Res. **79**, 1001–1006, 1974

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. **80**, 3571–3579, 1975

Rostoker, G., Boström, R.: A mechanism for driving the gross Birkeland current configuration in the auroral oval. J. Geophys. Res. **81**, 235–244, 1976

Rostoker, G., Hughes, T.J.: A comprehensive model current system for high-latitude magnetic activity – II. The substorm component. Geophys. J. R. Astron. Soc. **58**, 571–581, 1979

Sesiano, J., Cloutier, P.A.: Measurements of field-aligned currents in a multiple auroral arc system. J. Geophys. Res. **81**, 116–122, 1976

Siebert, M., Kertz, W.: Zur Zerlegung eines lokalen erdmagnetischen Feldes in äußeren und inneren Anteil. Nachr. Akad. Wiss. Göttingen, Math. Phys. Kl. 87–112, 1957

Sugiura, M., Potemra, T.A.: Net field-aligned currents observed by Triad. J. Geophys. Res. **81**, 2155–2164, 1976

Vestine, E.H., Chapman, S.: The electric current-system of geomagnetic disturbance. Terr. Magn. Atmos. Electr. **43**, 351–382, 1938

Weaver, J.T.: On the separation of local geomagnetic fields into external and internal parts. Z. Geophys. **30**, 29–36, 1964

Wedde, T., Doupnik, J.R., Banks, P.M.: Chatanika observations of the latitudinal structure of electric fields and particle precipitation on November 21, 1975. J. Geophys. Res. **82**, 2743–2751, 1977

Yasuhara, F., Kamide, Y., Akasofu, S.-I.: Field-aligned and ionospheric currents. Planet. Space Sci. **23**, 1355–1368, 1975

Zmuda, A.J., Armstrong, J.C.: The dirunal flow pattern of field-aligned currents. J. Geophys. Res. **79**, 4611–4619, 1974

Received November 20, 1979; Revised Version January 22, 1980

J. Geophys. 48, 18–27, 1980

Journal of Geophysics

K–Ar Age Determinations on Tertiary Volcanic Rocks*: V. Siebengebirge, Siebengebirge-Graben

W. Todt** and H.J. Lippolt

Laboratorium für Geochronologie der Universität, D-6900 Heidelberg, Im Neuenheimer Feld 234, Federal Republic of Germany

Abstract. Sanidines from 16 tuffs of the first volcanic phase in the Siebengebirge and the Siebengebirge-Graben were dated by the K/Ar-method. Ages range from 24.1 (Nachtigallental) to 22.9 M.y. (core Rott).

Three mineral ages from surface outcrops in the Central Siebengebirge (23.9 ± 0.5 M.y.) are statistically indistinguishable from those of the sub-surface samples Stieldorf-1, Stieldorf-2 and Rott (23.0 ± 0.5 M.y.). These ages confirm the Uppermost Oligocene biostratigraphic age of the 'Blätterkohle Rott' which is interlayered with the trachyte tuff.

Sanidine and biotite mineral ages from five trachyte samples give similar apparent ages from 26.4 to 24.6 M.y. These results show the trachytes to be older than the tuffs, in disagreement with the geologically established sequence. The sequence of eruptions in the Siebengebirge area however (trachyte-tuff, trachyte, latite and alkaline-basalt) was confirmed wherever outcrops allow observations. This discrepancy is discussed.

The apparent ages of three dated latites are in the same range (26.2 to 25.1 M.y.), in agreement with the geological sequence. The latite ages show that the time span between the eruptions of the trachytes and the latites must have been very small. Isolated basalts belonging to a fourth phase yield K/Ar ages from 25 to 19 M.y. Pliocene volcanism could not be verified in the Siebengebirge.

In addition, stratigraphically well-dated glauconites from the Tertiary of the Niederrhein area were used to correlate our isotopic age date with the Cenozoic time scale.

Key words: Potassium-Argon dating – Tertiary – Miocene – Oligocene – Sanidine – volcanic rocks – paleomagnetism – Siebengebirge – Germany.

1. Introduction

The area of interest for our study is the well known mountainous region of the Siebengebirge in western Germany and adjacent areas (Fig. 2). The age of the volcanism in the Siebengebirge is not well defined by stratigraphical methods. The Siebengebirge volcanics are considered to be approximately of the same age as the analogous rocks in the Westerwald. Ahrens (1957) placed the trachyte tuffs of the Westerwald into the Upper Oligocene (Middle Chattian). The basaltic eruptions in the Westerwald continued through the Lower Miocene and perhaps through the Pliocene. Lippolt and Todt (1978) demonstrated that the volcanic action in the Westerwald had several phases and was still active in the Quaternary. The most intensive phase was between Upper Oligocene and Lower Miocene (22–25 M.y.). One object of this study was to check the alleged synchronism of volcanic activity in the Siebengebirge and the Westerwald and to define the chronological sequence in the Siebengebirge.

The stratigraphic classification of the volcanic products in the Siebengebirge-Graben, the most southerly point of the Niederrheinische Bucht, has also been unsatisfactory. The volcanic action there was placed in the Oligocene, since the trachyte tuff is older than the 'Blätterkohle' of Rott, which Stehlin (1932) dated as Upper Oligocene. The second object of this study was to date several trachyte tuffs, which are intercalated with Tertiary sediments. In addition, samples of glauconite from well dated sediments in the Niederrhein area offered another opportunity for correlations.

The third object of this study was to date all the volcanic rocks for which paleomagnetic data have been published, in order to obtain more information about the chronologic evolution of the magnetic field at this time.

2. Geological Setting

The area of the Niederrheinische Bucht formed a sedimentary basin in the Tertiary. Marine transgressions are noticeable in the Oligocene and Miocene sections near the Ruhr-river. Deposition of fine clastic sediments occurred contemporaneously. These sediments now appear as large seams of brown coal. A movement of the crust took place in the course of the so-called Savian tectonic phase at the boundary of the Oligocene and Miocene. Volcanic eruptions in the Siebengebirge (Hesemann 1970, 1975) may have been triggered by a rapid subsidence of the graben-system in the Niederrheinische Bucht. This volcanic activity in the Siebengebirge belongs to the large arc of volcanic centers from the Eifel in the West to the Oberpfalz in the East (Fig. 1).

A period with several phases of eruptions in the area of the Siebengebirge produced, in succession, trachyte tuffs, trachytes, latites and finally alkaline basalts. The trachyte tuffs are interlayered with the sediments of the Niederrheinische Bucht in the Siebengebirge-Graben, so that it is possible to correlate isotopic ages with stratigraphic ages.

The only available K/Ar age published is the age of sanidine separated from the Drachenfels trachyte (22.8 M.y. Lippolt 1961,

* Former project title 'K–Ar age determinations on volcanics with known paleomagnetic field parameters'

** Now at Max Planck Institut für Chemie, Mainz

0340-062X/80/0048/0018/$02.00

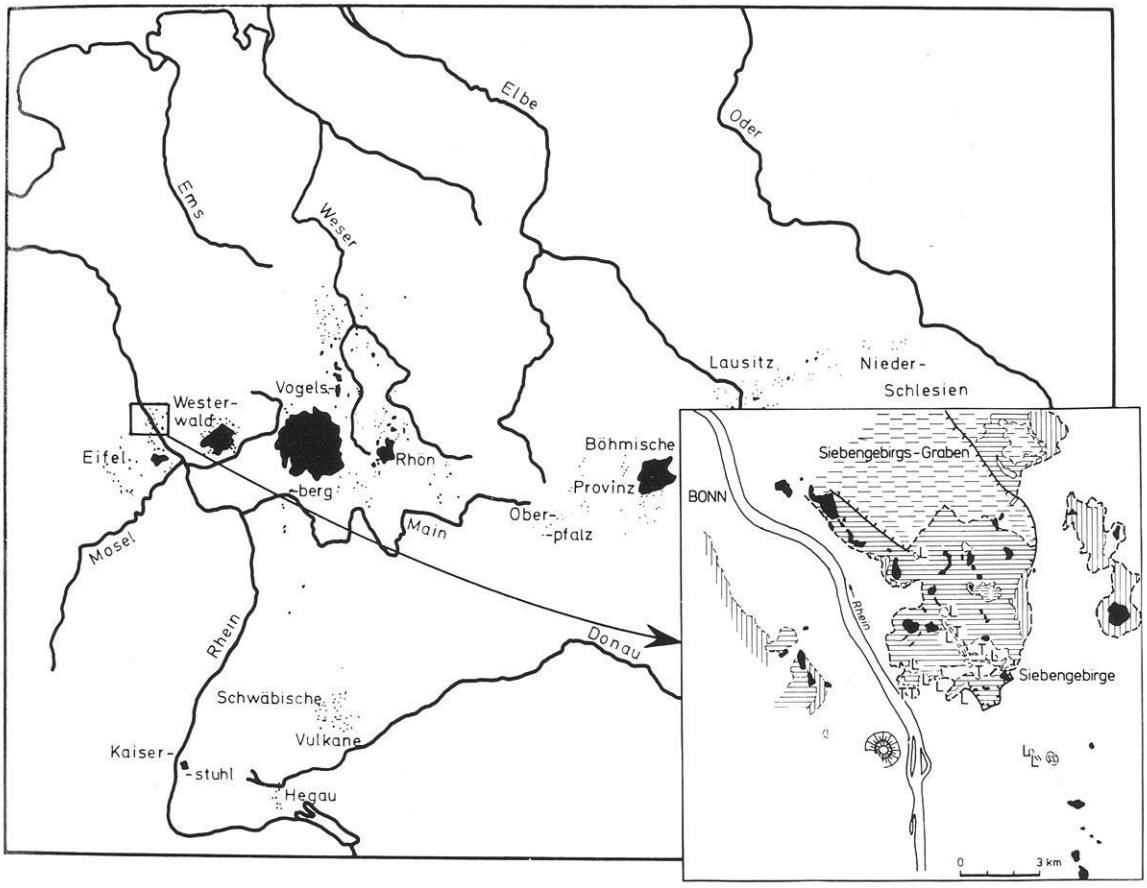


Fig. 1. Map of volcanic centers in Central Europe with area under investigation. Inset: Geological Map of Siebengebirge area

Frechen and Lippolt 1965, 23.1 M.y. calculated with the constants recommended by Steiger and Jäger 1977). The trachyte tuff from the Drachenfels area into which this trachyte intruded should be older than this intrusive rock.

The volcanism in the Siebengebirge has been studied extensively. Detailed studies of the geology and petrography and in regard of the deposition sequence were published by Laspeyres (1900). The chemistry of the rocks was published by Berg (1933). Further mineralogical and chemical analyses were carried out and interpreted by Frechen (1961, 1976), Vieten (1961, 1972) and Frechen and Vieten (1970).

According to the present geological concept, the volcanism in the Siebengebirge occurred in several stages, whereby the pressure of the Mesozoic sedimentary cover (>1,000 m) was relieved during the first phase of explosive volcanism. In this phase trachyte tuffs were deposited over a large area (>10 km^2) with an average thickness of about 100 m. Succesive trachyte-, latite-, and alkaline-basaltic magma intruded into these tuffs, forming domes, funnels, pipes, and dikes. The sequence of eruptions is the same throughout the area. It is particularly well established where the intrusive relationships are seen. Differences in the petrography of samples of different localities, however, suggest that it was not a contemporary eruption with a uniform sequence. It is more likely that the volcanism in different areas occurred with a time delay depending on the tectonic situation. This appears to have resulted in a similar, but often incomplete sequence. Volcanism in the Siebengebirge followed different pathways of eruptions as well as complex chemical compositions. The nearby volcanic areas of Eifel and Westerwald (compare Fig. 1) are separated from it by a zone of few eruptions. Figure 2 shows a simplified map of the Siebengebirge area and the majority of the sample localities which are described in the following section and in the appendix.

Further north in the Siebengebirge-Graben the 'Blätterkohle' of Rott is imbedded in trachyte tuff. This 'Blätterkohle' is considered to be a typical example of the brown coal of the Upper Oligocene and contains numerous fossils. In this 'Rott' locality, a *Microbunodon minimum Cuvier* (=Anthracotherium breviceps Troschel) has been dated as Upper Oligocene by Stehlin (1932). According to the Phanerozoic Time Scale (PTS), compiled at the 'Holmes Symposium' 1964 (Funnel 1964) and supplemented by Rast (1971), the boundary between Oligocene and Miocene lies at 25 ± 2 M.y. Since this time new K/Ar data have placed this boundary at 22.5 M.y. (Berggren 1972, Van Eysinga 1975) or even lower at 21–22 M.y. (Odin et al. 1975).

The well defined stratigraphical classification of the 'Blätterkohle' of Rott and the intercalation of the trachyte tuffs in the sediments of the Siebengebirge-Graben made it possible to date this coal seam absolutely and to assign hereby a further isotopic age to the Oligocene/Miocene boundary. Further calibration points for the PTS come from three glauconite bearing horizons in the Niederrheinische Bucht, where fauna relicts have been very well dated stratigraphically. They are:

(a) Marlstones of the Upper Paleocene in the borehole Hamsfeld near Issum. This horizon is dated as Thanetian (Anderson 1958; Indans 1958). Samples HAM Al and HAM Bl.

(b) Marlstones of the uppermost part of Oligocene in the Rossenray 1 and Hoerstgen mines. The fauna with *Chlamys bitica acuticostata* classified this layer as Lower Eochattian (Anderson

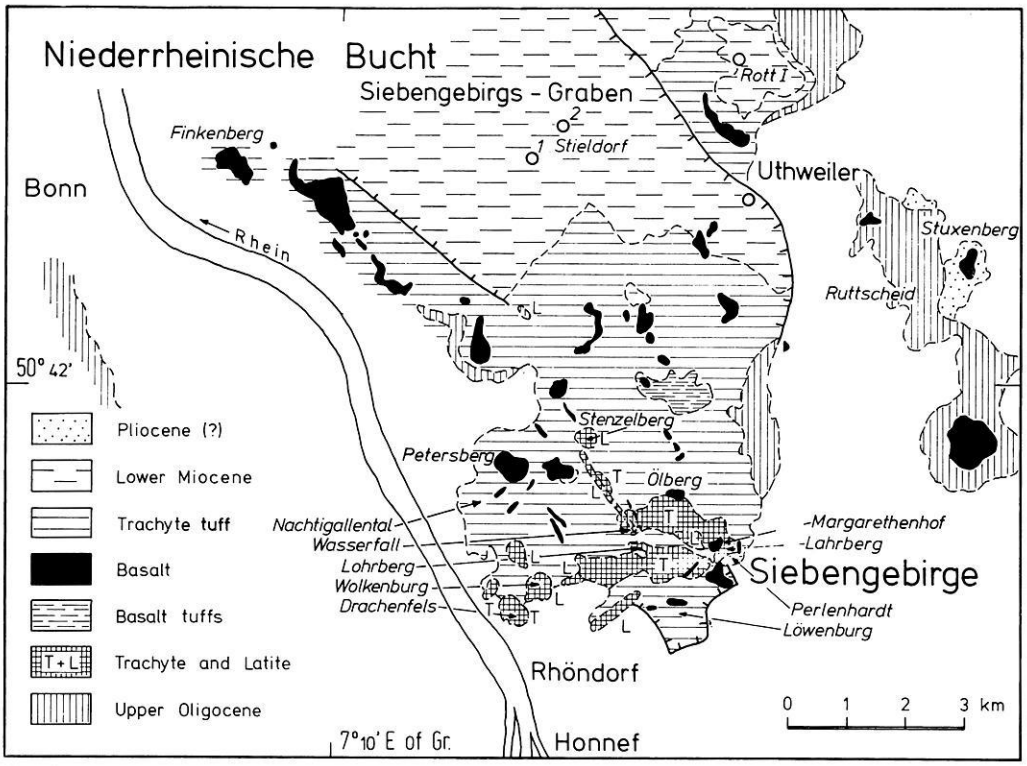


Fig. 2. Geological sketch map of Siebengebirge and Siebengebirge-Graben and sample localities (Samples E 33, E 34, E 35, E 39 in the South and E 36 in the Northeast are not included). Map after Teichmüller (1974) in Burghardt (1979)

1958; Indans 1958; Ellermann 1958). Samples ROS and HOER.

(c) Very fine pelitic sand from the Miocene of the Hoerstgen mine. This sand is also rich in fauna, with large numbers of tropical forms. Therefore it very likely belongs to the Hemmoor formation (Anderson 1958; Indans 1958). Sample 261A.

Glauconite separates from these layers prepared by the Geological Survey of Nordrhein-Westfalen served to correlate our data on the volcanism with the Geological Time Scale.

3. Analysed Samples

Cores from boreholes drilled by the Geologisches Landesamt Nordrhein-Westfalen (Krefeld) in the Siebengebirge-Graben, supplied samples of the graben area. Minerals from the tuff horizon in boreholes Stieldorf-1 and Stieldorf-2 were used to date this trachyte tuff. The cores of the two drill holes allowed Grünhagen (1970) to describe, layer by layer, the petrographic composition of the 100-m-thick graben section composed of trachyte tuff. Separation of sanidine allows us to determine their K/Ar ages. Grünhagen (1970) was able to recognize similar layers in the core of the drillhole Rott-1 but not to correlate them with the cores of Stieldorf. Sanidine has also been separated for K/Ar analyses from these layers. Tuff layers with sanidine had been found close to the 'Blätterkohle' near Uthweiler (Schmidt 1951). This sanidine was also dated by K/Ar analyses.

Sanidines from three layers in the central Siebengebirge (Nachtigallental, Heisterberg and Ruttscheid) were separated in order to evaluate the correlation between the Siebengebirge-Graben and the Siebengebirge trachyte tuff. Representative suites were chosen from the igneous rocks in the Siebengebirge covering the different phases of volcanism. Wherever mineral separations were possible we analysed sanidine- and biotite-samples, in the other cases we resorted to total rock analyses (TR). All samples were as fresh as possible, and classified by the criteria of Horn et al. (1972) as fresh volcanic rocks with the exception of Finkenberg basalt (No. 40) which was altered. Descriptions of the samples are given in the appendix.

4. Analytical Methods

Potassium was analysed by flame photometry with a 1σ-error of $\sim 0.5\%$. The argon was determined by isotope dilution in a 180°-5 cm-mass-spectrometer (Varian MAT GD150). The samples (~ 0.5 g) were wrapped in nickel foil and carefully predegassed in vacuum prior to the analyses to minimize the atmospheric argon contribution. The argon blank was about $1 \cdot 10^{-8}$ cm^3 STP argon. A detailed description of the analytical technique is published in Horn et al. (1972). The following decay constants were used (Steiger and Jäger 1977): $\lambda(^{40}K_{\beta^-}) = 4.962 \cdot 10^{-10}$/yr; $\lambda(^{40}K_e) + \lambda'(^{40}K_e) = 0.581 \cdot 10^{-10}$/yr; $^{40}K = 0.01167$ atom%. These constants yield ages about 1% higher than those calculated with the former set of constants. The error of the argon determination was calculated after Todt (1971) basing on the statistical error of the measurements and the atmospheric argon correction. The error of the age is the sum of the K- and Ar-errors. On the standard sample LP6-Bio of the US Geological Survey we determined $4.335 \cdot 10^{-5}$ cm^3 STP $^{40}Ar_{rad}$ per gram.

5. Results and Discussion

The results of the analyses are presented in Table 1. The discussion of the results begins with our glauconite data (Table 1A) since

glauconites are the basis for the correlation of the stratigraphy in the Niederrheinische Bucht (Kreuzer et al. 1973).

A. Ages of Stratigraphically Controlled Samples

The apparent age of the iron-rich glauconite fraction from the Thanetian sand in the Hamsfeld core is 57.8 ± 1.3 M.y. According to the PTS, the time base of the Upper Paleocene is 58–60 M.y. (Table 2). The age of the less-magnetic fraction of this sample is too high (67.2 ± 1.5 M.y.). X-ray analyses and mineralogical tests by Grünhagen (unpublished) suggest that impurities (detrital illite) might be the reason. These illites were not completely converted into glauconites.

Glauconites of the Asterigerina-horizon (Ellermann 1958) at the base of the Chattian in the Hoerstgen mine yield an age of 28.3 ± 1.0 M.y. in agreement with the PTS. The age of the glauconites (26.4 ± 0.9 M.y.) in a comparable layer (Lower Eochattian) from the Rossenray-1 mine is 2 M.y. lower and corresponds with the Uppermost Oligocene.

The results for the glauconite samples of Hemmoor in the Hoerstgen mine yield an age of 18.2 ± 0.9 M.y. and are somewhat younger than glauconites separated from the drillchips from Oldenburg Hemmoor (19.2 and 20.4 M.y.). Kreuzer et al. (1973) have measured ages of glauconites from Lattorfian sediments (Nannoplankton zone NP 21) close to Helmstedt (37.3 to 39.3 M.y.) and from corresponding glauconites close to Lehrte (38.4 and 38.9 M.y.). Four glauconite samples from the Neochat-

Table 1. Analytical results of the K/Ar measurements. For sample notations see chap. Appendix

Sample		Potassium (%)	Ar run No.	$^{40}Ar_{rad}$ (10^{-6})	^{40}Ar (atm) (%)	K-Ar age M.y. ± 1σ	Mean value M.y. ± 1σ
(A) Glauconite samples							
1. 261 A	Miocene (Hemmoor)	5.85	1260	4.146	51	18.2 ± 0.8	
2. ROS	Oligocene (Chatt)	6.40	595 790	6.609	6	26.4 ± 0.9	
3. HOER	Oligocene (Chatt)	5.83	599 601	6.333	28	28.3 ± 1.0	
4. HAM B1 (more magnetic)	Paleocene (Thanet)	5.42	596 598	12.305	23	57.8 ± 1.3	
5. HAM A1 (less magnetic)		5.12	597	13.870	20	67.2 ± 1.5	
(B) Tuff Samples from drill core Stieldorf-1							
6. G2S	Sanidine	8.16	715/1410	7.169	10	23.5 ± 0.7	
7. G3S	Sanidine	6.95	694/731/1824	6.186	14	22.5 ± 0.5	
8. G7S	Sanidine	7.85	718/819	5.986	11	(19.6 ± 0.6)	23.1 ± 0.5
9. G7H	Hornblende	1.46	602/787	1.334	43	23.3 ± 0.8	
10. Basalt	Totalrock	1.04	932/957	1.024	76	25.7 ± 2.3	
Tuff Samples from drill core Stieldorf-2							
11. G5S	Sanidine	7.53	588/688 817/1413	6.715	19	22.5 ± 1.5	
12. G1S	Sanidine	7.85	585/695/734 815/1828	7.371	13	24.0 ± 0.4	
13. G6S	Sanidine	8.57	589/693/1832	7.485	8	23.2 ± 0.5	
14. G4S	Sanidine	5.72	586/689/722	4.909	12	22.0 ± 0.4	23.0 ± 0.8
15. G8S	Sanidine	6.61	691/729	6.079	17	23.4 ± 0.8	
Tuff Samples from drill core Rott 1							
16. (A) 13m	Sanidine	8.09	1563/1585	7.292	20	23.0 ± 0.5	
17. (B) 25m	Sanidine	7.04	1043/1117	6.300	5	22.8 ± 0.5	23.0 ± 0.3
18. (C) 43m	Sanidine	8.16	1044/1120	7.306	6	23.1 ± 0.5	
Tuff from artificial outcrop							
19. Uthweiler	Sanidine	6.87	1408/1530	6.420	9	23.9 ± 0.7	

Table 1 (Continued)

	Bonn No.	Sample		Potassium (%)	Ar run No.	$^{40}Ar_{rad}$ (10^{-6})	^{40}Ar (atm) (%)	K-Ar-age M.y. $\pm 1\sigma$	Mean value M.y. $\pm 1\sigma$
(C) Hard rock samples from the Siebengebirge									
20.	BN (2)	Drachenfels	Sanidine	8.03	1588/1650	8.15	13	26.1 ± 1.5	
		Quartz trachyte	Biotite	7.23	1083/1049	7.14	22.5	25.4 ± 0.6	
21.	BN (4)	Perlenhardt	Sanidine	8.58	1235	8.60	6	25.8 ± 0.8	
		Quartz trachyte	Biotite	7.39	1076/1047	7.35	20	25.6 ± 0.6	
22.	BN (7)	Lohrberg	Sanidine	8.80	1236	8.79	5	25.7 ± 0.8	
		Trachyte	Biotite	7.02	1045/1081	6.93	23	25.4 ± 0.6	
23.	BN (8)	Wolkenburg	Totalrock	3.71	558/623	3.71	14	25.7 ± 0.6	
		Quartz latite							
24.	BN (11)	Stenzelberg	Totalrock	3.33	557/556	3.25	17	25.1 ± 0.6	
		Quartz latite							
25.	BN (21)	Löwenburg	Totalrock	3.02	561/605/619	3.04	43	25.9 ± 1.3	
		Foidic trachyte							
26a.	BN (37)	Ölberg-Südfuß	Sanidine	8.76	1247/1353	8.58	5	25.2 ± 0.6	
		Trachyte	Biotite	7.33	1048/1079	7.21	15	25.3 ± 0.5	
26b.	BN (54)	Alcaline basalt	Totalrock	1.14	555/615	1.02	35	23.0 ± 0.7	
27a.	BN (38)	Wasserfall	Sanidine	8.83	1250	9.06	6	26.4 ± 0.8	
		Trachyte	Biotite	7.46	1050/1085	7.13	15	24.6 ± 0.6	
27b.	BN (42)	Latite	Biotite	6.08	1893	6.19	53	26.2 ± 1.3	
28.	BN (39)	Lahrberg	Sanidine	8.67	1261/1309	8.49	7	25.2 ± 0.6	
		Trachyte							
29.	BN (36)	Am Margarethenhof	Totalrock	1.12	559/612	1.07	52	24.6 ± 1.1	
		Alc. Basalt							
(D) Trachytic tuffs from the Siebengebirge									
30.	BN (61)	Nachtigallental	Sanidine	5.60	1565/1590/1674	5.25	13	24.1 ± 1.5	
31.	BN (62)	Heisterbg/ Petersbg	Sanidine	5.39	1567/1596	4.99	14	23.8 ± 0.6	23.9 ± 0.5
32.	BN (63)	Ruttscheid	Sanidine	7.34	1568/1598	6.79	8	23.8 ± 0.6	
(E) Isolated basalt occurrences neighbouring the Siebengebirge; Totalrock measurements									
33.	BN (51)	Dächelsberg/ Niederbachem		1.22	566/614	1.19	60	25.1 ± 1.3	
34.	BN (57)	Kahlenberg/ Burgbrohl		0.968	565/613	0.918	56	24.4 ± 1.2	
35.	BN (58)	Asberg/ Kalenborn		1.15	564/611	1.073	56	24.0 ± 1.1	
36.	BN (60)	Stein/Eitorf		1.07	450/627	0.786	61	18.9 ± 1.1	
37.		Stieldorf	compare sample B-10						
38.		Stuxenberg/ Oberpleis		1.32	1594	1.277	29	24.9 ± 0.9	
39.	BN (59)	Steinsbergkopf/ Niederlützingen		1.31	452/626	1.135	43	22.3 ± 1.0	
40.		Finkenberg Bonn-Beuel		1.35	451/608/625	1.443	92	27.5 ± 5.4	

Table 2. Stratigraphic division of the Tertiary in the Niederrhein basin and pertinent K/Ar-data by Kreuzer et al. (1973) and from this work

		Data for Time scale boundaries		Data from Kreuzer et al. (1973)		Data from this work	
		Rast (1971)	Berggren (1972)	K-Ar Age M.y.	Stratigraphic Age	K-Ar-Age M.y. $\pm 1\sigma$	Stratigraphic Age
		10 ± 3	5				
Miocene	Upper Miocene Reinbek Hemmoor Vierland			19.2–20.4	Hemmoor	18.2 ± 0.9	Hemmoor
		25 ± 2	22.5	23.3–23.6	Olig/Miocene	23.0 ± 0.3	Upper Oligocene
Oligocene	Chattian Rupelian Lattorfian			23.8–24.8	Neo-Chattian	26.4 ± 0.9 28.3 ± 1.0	Lower Eo-Chattian
		37 ± 2	37.5	37.3–39.3	Lattorfian		Asterigerina Horizon
Eocene		58 ± 4	53.5				
Paleocene	Thanetian Danian-Montian					57.8 ± 1.3 (67.2 ± 1.5)[a]	Thanetian
		67 ± 3	65				
Cretaceous							

[a] Discussion in section 5A

tian sediments below the Oligocene/Miocene transition zone are dated at 23.3 and 23.6 M.y. We conclude from these dates that K/Ar determinations on glauconites in the Niederrhein-Tertiary largely confirm the PTS (Table 2), with the only exception that the Oligocene/Miocene time boundary might be too high.

Core Rott-1: From the trachyte tuffs of the Rott-1 core three sanidine samples from the depths of 13, 25 and 43 m were analysed. All results agree perfectly (23.0 ± 0.3 M.y.). Because of the intercalation of the 'Blätterkohle'-seam with these trachyte tuffs, the result also dates this seam. The PTS puts the Oligocene/Miocene transition at 25 ± 2 M.y. This implies that the 'Blätterkohle' would belong to the Lower Miocene, in disagreement with the presumed stratigraphic age; conversely if we believe the stratigraphic age to be right, the Oligocene/Miocene transition must be younger than 23.0 ± 0.3 M.y.

These data on the sanidines and the data on the glauconites do support the suggestions by Berggren (1972) and Odin et al. (1975) that the Miocene-Oligocene boundary is younger than that given in the PTS.

B. Results From Sanidines From Stieldorf-1 and Stieldorf-2

Stieldorf-1. Three sanidines, one hornblende and the whole rock basalt were analysed from this drill sample. The sanidines from the two upper samples give K/Ar ages of 23.5 ± 0.7 and 22.5 ± 0.5 M.y. Sanidine and hornblende analyses of sample G7 yield 19.6 ± 0.6 and 23.3 ± 0.8 M.y. ages respectively. This sanidine was the only sample with isotropic grains – perhaps glass. Glass is known for its low argon retentivity (compare references in Dalrymple and Lanphere 1969, p. 176). An additional possibility to explain the argon loss is the concordant age of the hornblende with the two other samples of this core. The age of the TR-sample of the basalt (25.7 ± 2.3 M.y.) was unexpected, because it was assumed to be a sill (Teichmüller, personal communication). This result means that the basalt was already present when the trachyte-tuff erupted, and that the basalt was filled in by the tuff.

Stieldorf-2. Five sanidine samples from this core resulted in ages between 22.0 and 24.0 M.y. (similar to Stieldorf-1). The youngest sample G4 shows patches with microcline cross-hatching, which could be a reason for argon loss (compare references in Dalrymple and Lanphere 1969, p. 168). The results show that these tuffs were formed before or shortly after the Oligocene-Miocene transition.

C. Results on the Samples From the Siebengebirge

Three sanidine samples from the trachyte tuff of the central Siebengebirge were analysed in order to evaluate the isotopic age position of the trachyte cores of the area and to confirm the results on the trachyte tuff in the Siebengebirge-Graben. These three samples (30, 31, and 32) yield ages of 23.9 ± 0.5 M.y. thereby supporting our previous result that the trachyte tuffs have younger apparent ages than the trachytic igneous rocks. As will be discussed in the next sub-section, all igneous rocks (except some basalts) were older than 24 M.y.

The analytical results of the minerals and total rock samples for the volcanics from the Siebengebirge are presented in Table 1 C. We measured sanidines and biotites from all trachytes which have been collected. All sanidine- and biotite-ages agree within error-limits. The Wasserfall trachyte (27) is an exception since the biotite appears to be younger (24.6 ± 0.6 M.y.) than the sanidine (26.4 ± 0.8 M.y.). The reason for this is not yet understood. The average ages of all sanidines are 25.7 ± 0.5 M.y. The corresponding average age of all biotites is 25.3 ± 0.4 M.y. Hence the ages of both minerals agree.

The average age of the three analysed latities is 25.7 ± 0.6 M.y. which suggests the eruptions of the trachytes and latites occurred within a very short time span at about 25 M.y.

The alkaline-basalts, which are considered to be products of the last eruptive phase in the Siebengebirge area, yield K/Ar ages ranging from 25 to 19 M.y., indicating they were produced from the beginning of the volcanic activity and continual eruptions over a time span of several million years.

D. Paleomagnetism

Paleomagnetic data from eight analysed samples in the Siebengebirge showed reversed magnetization for five alkaline-basalts (26, 29, 33, 34, and 35) while two quartz-latites (23 and 24) and the feldspathoidal-trachyte (25) are normally magnetised. The age range for the normally magnetised samples is 25.9–25.1 M.y. and for the other samples 25.8–23.0 M.y. These results therefore show an older group with normal magnetization (latites and trachytes) and a younger group with reverse magnetization (alkaline-basalts). The event for the reversal from normal to reverse polarity occurred ~ 25 M.y. ago. This corresponds to the 25 ± 2 M.y. reversal, dated in the basaltic province of the Oberpfalz by Todt and Lippolt (1975).

6. General Discussion

Geologic studies of the area suggest the Siebengebirge volcanism took place in five phases (compare Hesemann 1970). It appears from local geologic observations that the first phase produced the trachytic tuffs as a thick layer over the whole area which was then intruded by trachytic magmas during a second volcanic phase. This relationship can be seen at the Drachenfels and in the Lohrberg quarry. Unfortunately these locations have not been sampled for the present study.

Our K/Ar results conflict with this chronology. Here the tuffs appear to be younger than the lavas. All our analysed ages for the tuff layers are in the range 24.1 to 22.0 M.y. while the youngest trachyte has an age of about 25 M.y. as do the latities. From this, one would conclude that all dated tuffs must have been deposited within a relative short time span (2 M.y.) and that they erupted after nearly all other magmas had been erupted. On the other hand our results for the suite of hard rocks show the expected sequence of eruption.

Since the same mineral (sanidine) was analysed in the tuff and in the hard rocks our K/Ar results should be comparable. Each set of data (hard rock/tuff) seems to be consistent. The hard rock sanidine ages are supported by coexistent biotite ages. The tuff sanidines give consistent values for all the boreholes of the graben and for the outcrops in the mountain range. They have all the same age within the errorlimits.

On the other hand there is no reason to expect an inferior retentivity behaviour of sanidines from tuffaceous origin. Baadsgaard et al. (1961) reported that fresh sanidine obtained from various volcanic bentonite horizons retained the radiogenic argon sufficiently well to yield reliable $^{40}Ar/^{40}K$ ages. Their short-term (days) tests of argon leakage from pure sanidines showed that the radiogenic argon is quantitatively retained at temperatures below 400° C. Moreover Baadsgaard et al. (1961) reported one case (Crowsnest volcanics, Alberta, Canada) were sanidine from a trachyte flow yielded significantly lower ages than the sanidine from a correlated bentonite. One could however argue that the tuffs and the flows of the Siebengebirge and the Siebengebirge-Graben might have experienced different geochemical histories. The tuffs are diagenetically altered and might have had more hydrothermal alteration.

In our suite of samples there is unfortunately no pair of neighboring tuff and trachyte where the age relationship is clearly discernible in the field. We are therefore faced with two alternatives:

1. Production of tuff is not confined to a short phase of activity but lasted a rather long time. Sanidines from the tuff production of the starting phase were not observed in the field and therefore could not be analysed. This would imply that we have to look for samples where the field relationships are clear (Lippolt and Vieten in preparation).

2. There is a hitherto unknown process observed which influenced the results on the sanidines from the tuffs. Earlier studies have demonstrated the reliability of sanidine in K/Ar dating. There is however a grain size and a chemical difference between the sanidines from the two rock types which may affect Ar retentivity. The sanidines of the tuffs are smaller (mm) than those from the trachytes (cm) and the potassium contents of the tuff-sanidines are lower than the contents of the others and they have a much wider spread (5.4–8.6% compared to 8.0–8.8%).

At the moment we cannot distinguish between these two alternatives. To solve the question we intend to begin a more detailed study of two pairs of sanidines from tuff and trachyte in neighboring position. By determining the K/Ar ages using conventional and neutron activation techniques, and studying the chemical compositions of the minerals by XRF technique we hope to resolve this problem (Lippolt and Vieten in preparation).

7. Conclusion

Our results are illustrated in a simplified crossection through the Niederrheinische Bucht in Fig. 3. The right side of the figure shows the results on the igneous rocks of the Siebengebirge. We consider these to be the most reliable ages because of the concordance of sanidine and biotite ages. These analyses fix the main phase of the Siebengebirge igneous rock production at about 26–24 M.y. The left side of the figure shows the age determinations of glauconites, implying that the age levels of the Tertiary part of the PTS (after Funnel 1964 and Rast 1971) are also valid in the Niederrheinische Bucht.

At their face value, the results on the minerals from the tuffs (right central part of the figure) may be interpreted as follows:

The ages of the sanidines in the Siebengebirge-Graben, particularly in the Rott core, add new evidence to the age of the Oligocene/Miocene transition. If we take the correct stratigraphical age of these layers as Upper Oligocene, this boundary must be younger than 23.0 ± 0.3 M.y. Hence the results on the tuffs favor the 22.5 M.y. transition age as published by Van Eysinga (1975). This would also be the correct age for the Oligocene/Miocene transition in the Siebengebirge-Graben. But the discrepancy between results from tuffs and igneous rocks does not allow us to resolve this question at present.

In Fig. 4 our new results are compared to K/Ar age distributions of other areas in the Rhenish Shield. Approximately 24 M.y. ago all volcanic areas were active, which conflicts with the idea of a hotspot moving from E to W or a plate moving from W to E over the hotspot (Duncan et al. 1972; Burke et al. 1973). Two periods of volcanism prevail in the Eifel area: the Tertiary phase from 44 to about 24 M.y. and the well known Quaternary phase. The Tertiary Eifel volcanism mostly precedes the Wester-

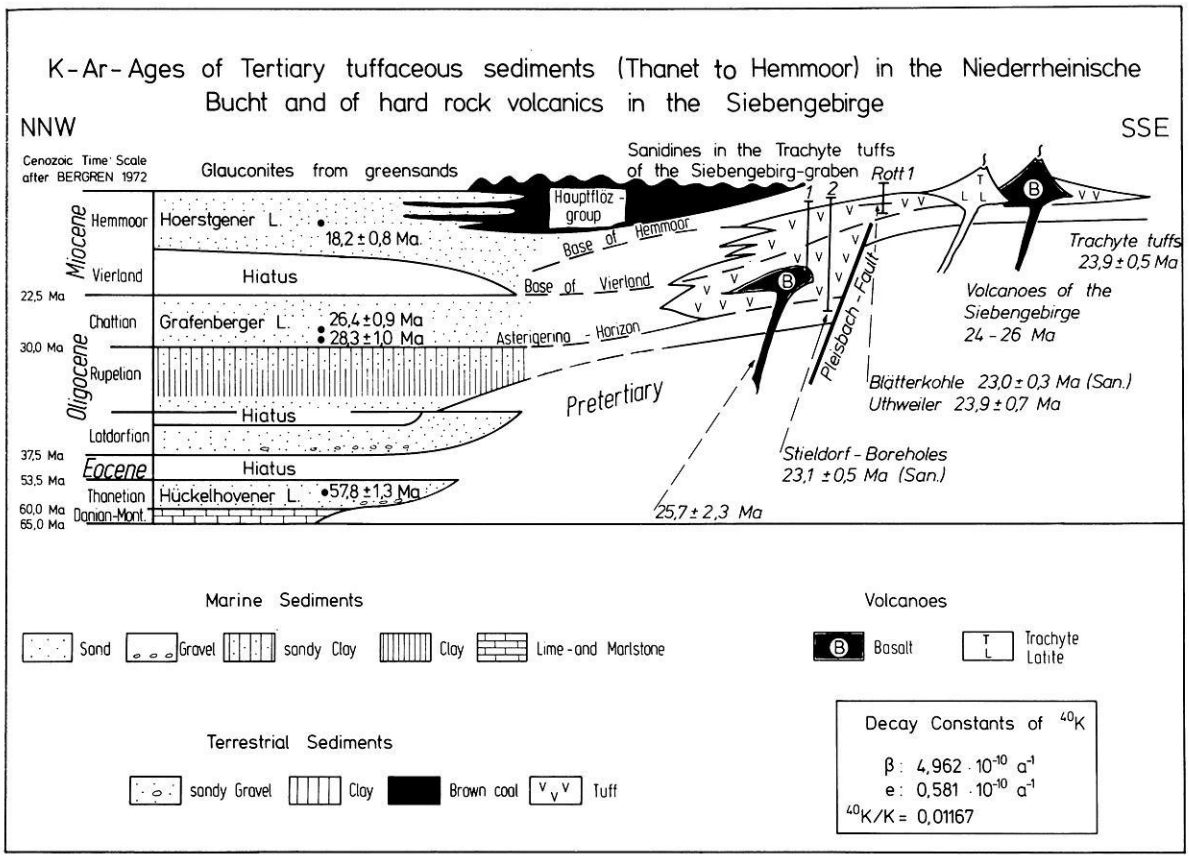


Fig. 3. Synopsis of the results as a simplified cross section (NNW-SSE) through the Niederrhein basin and the Siebengebirge. (After Teichmüller 1974)

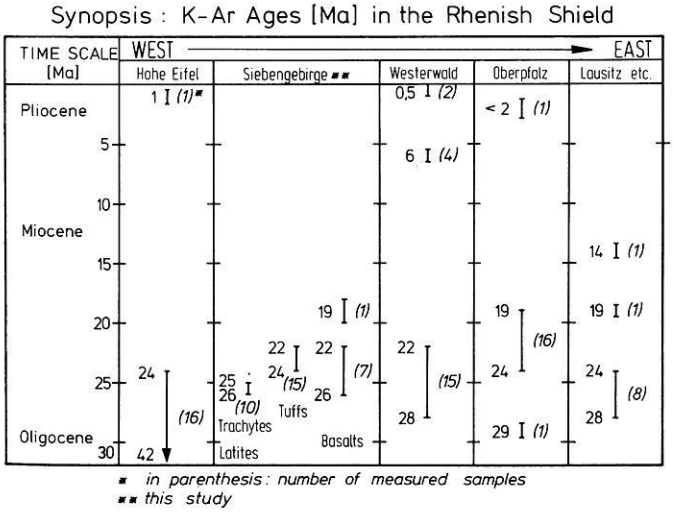


Fig. 4. Age distribution of volcanic activity in the Rhenish Shield from Hocheifel (*W*) to Saxony (*E*) after the results of papers published earlier and of this study (Siebengebirge area)

wald-Siebengebirge volcanism. However the contemporaneity of the volcanic activity in these two latter areas is still correct based on our data. After a longer time span volcanism began again during the Pliocene/Quaternary in the Eifel, Westerwald and Oberpfalz. In the Siebengebirge area proper, however, it has not been possible to find any Pliocene volcanism.

Acknowledgements. We are indebted to Drs. H. Grünhagen and R. Teichmüller (Geologisches Landesamt Nordrhein-Westfalen in Krefeld) for the preparation of the samples 1 to 19 and 37 to 38 and helpful discussions about the Tertiary in the Niederrheinische Bucht. We are also indebted to Prof. Drs. J. Frechen and K. Vieten (Universität Bonn) for separating minerals from crystalline rocks in the Siebengebirge (samples 20 to 29), and preparation of the sanidine samples 30 to 32. Professor J. Frechen also guided us when we collected samples 33 to 36 and 39. With Professor K. Vieten we had fruitful discussions concerning the tuff-igneous rock sequence.

We are grateful to H.-J. Funke and A. Lutz for their careful technical assistance.

We further wish to thank R. Hurst for taking the trouble to improve the English text.

This work was partly supported by the Deutsche Forschungsgemeinschaft (Bonn) under the priority research program 'Geologische Korrelationsforschung'.

8. Appendix: Sample Description

A. Glauconite-Samples

These samples are prepared in the Geologisches Landesamt Nordrhein-Westfalen in Krefeld: After wet sieving, the 200–500 µm-fraction was separated on a Frantz-magnetic separator. Handpicking under a stereo-microscope produced about 100% clean separates (Grünhagen 1970).

1. *Sample 261 A:* shaft IV of the Friedrich-Heinrich AG at Hoerstgen. TK 25, sheet 4404 Issum, $r=32\ 826$, $h=09\ 320$; depth 35.8 m, stratigraphic classification: Hemmoor.
2. *Sample ROS:* shaft I Rossenray TK 25, sheet 4405 Rheinberg, $r=38818$, $h=09395$, depth 91.1 m, stratigraphical classification: Lower Eochattian.
3. *Sample HOER:* Shaft IV of the Friedrich-Heinrich AG at Hoerstgen. cp sample 261 A, depth 128.2 m, stratigraphic classification: Lower Eochattian.
4. *Sample HAM B 1:* Borehole Hamsfeld. TK 25, sheet 4404 Issum. $r=29437$, $h=10069$, depth 309.65–309.75 m, stratigraphic classification: Upper Paleocene, strong manetic fraction.
5. *Sample HAM A 1:* The same sample as HAM B 1, but a less magnetic fraction with impurities of illite.

B. Tuff-Samples From the Siebengebirge-Graben

These samples are also prepared in the Geologisches Landesamt Nordrhein-Westfalen in Krefeld. The slightly solidified tuffs are ground wet in a porcelain-mortar, then elutriated. The feldspars are separated from the 200–400 µm and 400–600 µm sieve-fractions with a Frantz-magentic separator and heavy liquids. The sanidine-concentrates were handpicked under a stereo-microscope and checked for purity with refracting liquids. With one exception (G7S: 85% sanidine) the concentrations are higher than 98% (Grünhagen 1970; Todt 1971).

Borehole Stieldorf-1: TK 25, sheet 5209 Siegburg, $r=25\ 85575$, $h=56\ 2273$

6. *G 2 S:* Sanidine from a fine greenish trachyte tuff, depth: 45.5 m.
7. *G 3 S:* Sanidine from a lightgreen trachyte tuff, fine grained, depth: 58.5 m.
8. *G 7 S:* Sanidine from a grey, solid trachyte tuff, depth 67.5 m.
9. *G 7 H:* Hornblende from a grey, solid trachyte tuff, depth 67.5 m.
10. *Basalt:* depth 81.2 m, cp. E-37.

Borehole Stieldorf-2: TK 25 sheet 5209 Siegburg, $r=25\ 68100$, $h=56\ 23320$.

11. *G 5 S:* Sanidine from a green to grey trachyte tuff, depth 82.6–83.7 m.
12. *G 1 S:* Sanidine from a green to grey trachyte tuff, depth: 92–94 m.
13. *G 6 S:* Sanidine from a grey, finegrained trachyte tuff, depth 124.0–125.4 m.
14. *G 4 S:* Sanidine from a green, mediumgrained solid trachyte tuff, depth 134.1–137.0 m.
15. *G 8 S:* Sanidine from a brown clay, downwards more and more carbonaceous, depth 148.0–148.7 m.

Borehole Rott-1: $r=25\ 89200$, $h=56\ 25100$.

16. (A) Sanidine from a trachyte tuff, depth 12.70–12.95 m.
17. (B) Sanidine from a trachytic lapilli tuff, argillaceous with rock fragments, depth 25.0–25.5 m.
18. (C) Sanidine from a trachytic tuff and tuffite, the tuff is partly rich on sanidine, partly with feldspar and biotite, depth 43.25–43.4 m.
19. Uthweiler: Streamscarp at Frechwinkel. Sanidine from the lower tuff-bank in the 'Blätterkohle', about 2 m below the hanging boundary of the 'Blätterkohle'. TK 25, sheet 5209 Siegburg, $r=88\ 890$, $h=22\ 600$.

20.–32. *C, D. Igneous-Rock- and Tuff-Samples From the Siebengebirge*

These samples are prepared in the Mineralogisch-Petrologisches Institut der Universität Bonn. The purity of all mineral-separates was higher than 98%. In Table 1 we also list the lab-numbers of the University of Bonn in addition to our lab-numbers. Detailed description of all these samples may be found in the following publications: Vieten (1961), Vieten (1972), Frechen (1976), Frechen and Vieten (1970).

E. Isolated Basalts

33. Dächelsberg/Niederbachem. Alkaline-Ol-Basalt, unfresh.
Phenocrysts: mostly Px, some with reabsorptions, Ol only marginal transformed into Iddingsite.
Matrix: fine-crystalline, Plg more than Px, in some parts strong calcitization and zeolitization.
34. Kahlenberg/Burgbrohl. Alkaline-Ol-Basalt, unfresh.
Phenocrysts: mostly Ol, with transitions to completely transformed bowlingite, minor Px.
Matrix: coarse-crystalline, carbonates mostly concentrated in spots, partly strong zeolitization.
35. Asberg/Kalenborn. Alkaline-Ol-Basalt, fresh.
Phenocrysts: Ol with no or little transformation to bowlingite, dominating over Px.
Matrix: very fine-crystalline, Plg dominating over Px, minor amount of calcite.
36. Stein/Eitorf. Alkaline-Ol-Basalt, fresh.
Phenocrysts: Ol dominating with beginning bowlingitization, minor Px with resorption-features.
Matrix: coarse-crystalline, rich of ore-minerals, minor calcite and zeolites, Plg dominating.
37. Stieldorf (B-10). Alkaline-Ol-Basalt, unfresh.
Phenocrysts: rare, mostly Ol (strongly to completely bowlingitized), minor Px.
Matrix: mostly Plg, strong calcitization and zeolitization.
38. Stuxenberg. Nepheline-Basanite, fresh.
Phenocrysts: abundant, mostly Px, also Ol (−1 mm), marginal transformed to bowlingite.
Matrix: poor in Plg, partly Neph, fine-crystalline, mostly prismatic Px, ore-minerals, minor amount of glass, beginning calcitization.
39. Steinsbergkopf/Niederlützingen. Alkaline-Ol-Basalt, fresh.
Phenocrysts: mostly Px, also fresh Ol, minor amount of Px, xenocrysts with strong reaction-rims.
Matrix: localy: calcite and zeolites, otherwise fresh, fine-crystalline, rich in ore-minerals, Plg dominating over Px and Ol.
40. Finkenberg/Bonn-Beuel. Alkaline-Ol-Basalt, unfresh.
Phenocrysts: abundant, mostly Px locally concentrated, minor amount of Ol heavy to completely bowlingitized, also very fresh.
Matrix: rich in calcite and zeolites (in spots and along cracks) Plg dominating over Px.

References

Ahrens, W.: Überblick über den Aufbau des Westerwälder Tertiärs, mit besonderer Berücksichtigung der stratigraphischen Stellung der vulkanischen Gesteine. – Fortschr. Mineral. **35**, 109–116, 1957

Anderson, H.J.: Zur Stratigraphie and Paläogeographie des marinen Oberoligozäns und Miozäns am Niederrhein aufgrund der Mollusken-Faunen. – Fortschr. Geol. Rheinl. Westfalen **1**, 277–295, 1958

Baadsgaard, H., Lipson, J., Folinsbee, R.E.: The leakage of radiogenic argon from sanidine. Geochim. Cosmochim. Acta **25**, 147–157, 1961

Berg, G.: Geologische Beobachtungen im Siebengebirge. – Verh. nat-hist. Ver. Rheinl. Westfalen **91**, 99–134, 1933

Berggren, W.A.: A Cenozoic time-scale – some implications for regional geology and paleobiography. – Lethaia **5**, 195–215, 1972

Burghardt, O.: Siebengebirge – Landschaft im Wandel – Geologisches Landesamt Nordrhein-Westfalen 64 p, 1979

Burke, K., Kidd, W.S.F., Wilson, J.T.: Plumes and concentric plume traces of the Eurasian Plate. – Nature 241 Phys. Sci. pp. 128–129, 1973

Collinson, D.W., Creer, K.M., Irving, E., Runcorn, S.K.: Palaeomagnetic investigations in Great Britain I–IV. – Proc. R. Soc. London, Ser. A: **250**, 71–156, 1957

Dalrymple G.B., Lanphere, M.A.: Potassium-Argon Dating. pp. 166–184, San Francisco: W.H., Freeman and Co. 1969

Duncan, R.A., Peterson, N., Hargraves, R.B.: Mantle plumes, Movement of the European Plate, and Polar Wandering. – Nature **239**, 82–86, 1972

Ellermann, C.: Die mikrofaunistische Gliederung des Oligozäns im Schacht Kapellen bei Moers. – Fortschr. Geol. Rheinl. Westfalen **1**, 205–214, 1958

Eysinga, F.W.B. Van (Compiler): Geological Time Table 3rd edn. – Amsterdam: Elsevier 1975

Frechen, J.: Einführung zu den Exkursionen der Deutschen Mineralogischen Gesellschaft. – Deutsche Mineral. Ges. 38. Jahrestagung in Bonn/Rl. Fortschr. Mineral. **39**, 161–185, 1961

Frechen, J.: Siebengebirge am Rhein – Laacher Vulkangebiet – Maargebiet der Westeifel. – Slg. Geol. Führer **56**, 4–27, – Borntraeger, Berlin-Stuttgart, 1976

Frechen, J., Lippolt, H.J.: Kalium-Argon-Daten zum Alter des Laacher Vulkanismus, der Rheinterrassen und der Eiszeiten. – Eiszeitalter Ggw. **16**, 5–30, 1965

Frechen, J., Vieten, K.: Petrographie der Vulkanite des Siebengebirges. I. Die subalkalische Gesteinsreihe Quarztrachyt-Latitbasalt. II. Die peralkalische Gesteinsreihe Alkalitrachyt-Sanidinbasanit. Decheniana **122**, 337–377, 1970

Funnel, B.M.: The Tertiary period. The Phanerozoic time-scale. A Symposium dedicated to Arthur Holmes. – Q. J. Geol. Soc. London **120**, [Suppl.], 179–191, 1964

Grünhagen, H.: Petrographie der Bohrungen Stieldorf-1 und –2. – Geol. Landesamt Nordrhein Westfalen, Krefeld 1970

Hesemann, J.: Zum Vulkanismus des Siebengebirges. – Decheniana **122**, 199–206, 1970

Hesemann, J.: Geologie Nordrhein-Westfalens. pp. 268–278, – Paderborn: F. Schöningh, 1975

Horn, P., Lippolt, H.J., Todt, W.: Kalium-Argon-Altersbestimmungen an tertiären Vulkaniten des Oberrheingrabens. I. Gesamtgesteinsalter. – Eclogae Geol. Helv. **65**, 131–156, 1972

Indans, J.: Mikrofaunistische Korrelationen im marinen Teritär der Niederrheinischen Bucht. – Fortschr. Geol. Rheinl. Westfalen **1**, 223–238, 1958

Kreuzer, H., Daniels, C.H., v., Gramann, F., Harre, W., Mattiat, B.: K/Ar-Dates of some Glauconites of the North West Germany Tertiary Basin. – Fortschr. Mineral. **50**, (Beih. 3), 94f., 1973

Laspeyres, H.: Das Siebengebirge am Rhein. – Verh. nat.-hist. Ver. Rheinl. Westfalen **57**, 119–591, 1900

Lippolt, H.J.: Alterbestimmungen nach der K-Ar-Methode bei kleinen Argon- und Kaliumkonzentrationen. – Heidelberg, Nat.-Math. Fak. Dissertation, 82p, 1961

Lippolt, H.J., Todt, W.: Isotopische Altersbestimmungen an Vulkaniten des Westerwaldes. – Neues Jahrb. Geol. Palaeontol. Monatsh. **6**, 332–352, 1978

Nairn, A.E.M., Kürsten, M.: Mitteilungen über paläomagnetische Untersuchungen an Basalten des Siebengebirges und seiner Umgebung. – Neues Jahrb. Geol. Palaeontol. pp. 348–352, 1959

Odin, G.S., Hunziker, J.C., Lorenz, C.R.: L'âge radiométrique du Miocène inférieur en Europe Occidentale et Centrale. – Geol. Rundsch. **64**, 570–592, 1975

Rast, N.: Isotope dating in the USSR – an essay review. – In: The Phanerozoic Time scale, W.B. Harland, E.H. Francis eds.: pp. 39–50. Geol. Soc. London Spec. Publ. No. 5, 1971

Schmidt, W.: Ein aufgeschürftes Tertiärprofil bei Uthweiler im Siebengebirge. – Geol. Jahrb. **65**, 603–610, 1951

Stehlin, H.G.: Über die Säugetierfauna der Westerwälder Braunkohlen. – Eclogae Geol. Helv. **25**, 314–319, 1932

Steiger, R.H., Jäger, E.: Subcommission on Geochronology: Convention on the Use of decay constants in Geo- and Cosmochronology. – Earth Planet. Sci. L. **36**, 359–362, 1977

Teichmüller, R.: Die tektonische Entwicklung der Niederrheinischen Bucht. Approaches to Taphrogenesis, Int. Union Comm. Geophys. Sci. Rep. **8**, 269–285, 1974

Todt, W.: Kalium-Argon-Altersbestimmungen an mitteleuropäischen miozänen Vulkaniten bekannter paläomagnetischer Feldrichtung. – Heidelberg, Naturwiss. Gesamtfakultät Dissertation, 74–106, 1971

Todt, W., Lippolt, H.J.: K-Ar-Alterbestimmungen an Vulkaniten bekannter paläomagnetischer Feldrichtung. I. Oberpfalz und Oberfranken. – J. Geophys. **41**, 43–61, 1975

Vieten, K.: Die Trachyt-Latit-Alkalibasalt-Assozation des Siebengebirge am Rhein. – Bonn, Nat.- Math. Fak. Dissertation 112 p. 1961

Vieten, K.: Über die Heteromorphie-Beziehungen in der Vulkanit-Assoziation des Siebengebirges und ihre petrogenetische Bedeutung. – Neues Jahrb. Mineral. Abh. **117**, 282–323, 1972

Received April 30, 1979; Revised Version October 25, 1979

J. Geophys. 48, 28–33, 1980

Journal of Geophysics

The Geothermal Anomaly of Landau/Pfalz: An Attempt of Interpretation

D. Werner[1] and M. Parini[1,2]

[1] Institut für Geophysik, ETH Zürich, CH-8093 Zürich, Switzerland
[2] Geophysikalisches Institut, Universität Karlsruhe, D-7500 Karlsruhe, Hertzstr. 16, Federal Republic of Germany

Abstract. The geothermal anomaly in the oil-field Landau (Rhinegraben, Germany) is modeled postulating the rise of deep groundwater along a fault in the basement. Relations are investigated between (a) the temperature in the centre of the anomaly, (b) its lateral extension, (c) the water flux, (d) the original depth of the water, and (e) the age of the anomaly. The numerical approach takes into account two processes: (a) Thermal conduction; (b) convection in a porous medium. The estimated age of the anomaly (about 100,000 y) does not compare with the age of the Rhinegraben (about 50 my). The required water flux is very small, and the source depth of the water flow is estimated to be about 6,000 m. The presented hydrothermal model possibly may be used to interpret the shallow temperature anomalies in the whole Rhinegraben.

Key words: Rhinegraben – Geothermal anomalies – Uprising deep waters – Hydrothermal processes – Heat transfer in porous media.

Introduction

The Rhinegraben between Frankfurt and Basel is one of the best studied continental rift systems. A nearly complete coverage of the results can be found in several proceedings of Rhinegraben symposia, such as edited by Rothe and Sauer (1967), Illies and Müller (1970) and Illies and Fuchs (1974). This area is noteworthy from the geothermal aspect. In the Rhinegraben we have to distinguish between the temperature distribution in the depth range of the lithosphere/astenosphere (about 200 km), corresponding to the history of the rift system, and the near surface temperature distribution within the sediments caused by hydrothermal processes. The aim of this paper is to find an explanation for the geothermal anomalies in the sedimentary cover of the Rhinegraben. A fairly well-known and striking anomaly is located in the oil-field of Landau/Pfalz (Fig. 1a), and is described by several authors (Doebl 1970; Hänel 1974; Werner and Doebl 1974; Werner et al. 1978). Within the German as well as French zone of the Rhinegraben, which is characterized by high subsurface temperatures, there are other significant anomalies (Delattre et al. 1970; Lauer 1976).

We start from the assumption that the observed thermal anomalies in the sediments of the Rhinegraben are caused by water circulation systems down to great depth. It has been shown that other attempts of interpretation (e.g., differences in the heat conductivity, existence of very young magma bodies, or the assumption of a very high density of radiogene heat sources, (see Werner 1975) lead to unrealistic results.

Our observations are limited to the uppermost part of the earth's crust. The existence of a deep geothermal anomaly in the upper mantle under the Rhinegraben (Werner and Kahle 1980) is not the subject of this paper.

Our considerations are based on new continuous temperature logs in oil wells of the Landau field. The logged depths range from 800 to 1,300 m. Further details about these measurements are described in an earlier paper (Werner et al. 1978). The results of the measurements are summarized in Fig. 1b. Apparently, significantly high temperatures exist in this area, with values of about 100° C at 1,000 m depth in the centre of the anomaly. This corresponds to a heat-flow maximum of about 120 mW/m^2

The Concept of the Model

The Rhinegraben is a rift zone characterized by numerous faults in its sedimentary cover (Illies 1974). Some major faults continue into the crystalline basement. In the Landau field a deep reaching fault system (the so called ω-fault and γ-fault, see Fig. 1b) is well known (Schad 1962; Doebl et al. 1974). The average thickness of the sedimentary cover in the Landau field amounts to about 2,000 m.

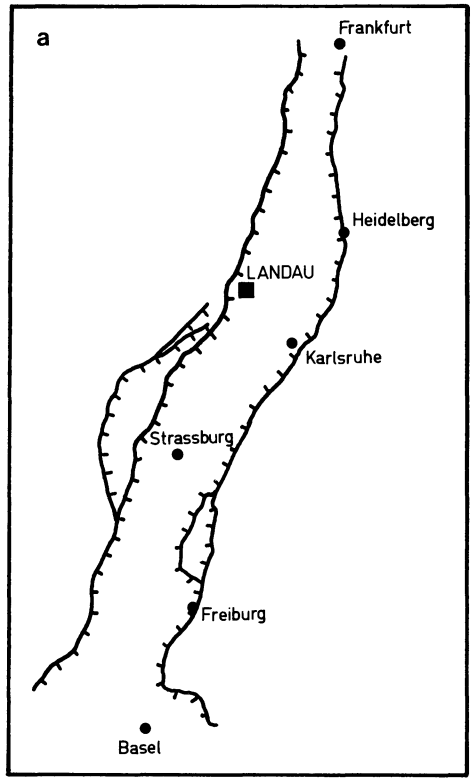
The question arises wether a deep-reaching fault can be considered as origin for a thermal anomaly. We do not expect that each major fault plays a geothermally active role. Our model is based on the assumption that some deep faults are permeable, at least locally, and accordingly thermal waters are able to rise up from a depth of some thousand meters. For this model it is necessary that a continuous narrow zone with a vertical permeability exists.

The basic principle is that rising groundwater transmits its heat content to the surroundings and builds up a thermal anomaly. Although we assume the existence of rising thermal waters, our model does not explain its hydromechanical cause. Starting with a given distribution of flow the model calculation leads to a corresponding temperature distribution. The aim of the model calculation is to find quantitative relations between:

(a) The temperature in the centre of the anomaly, e.g., at a depth of 1,000 m; this allows a comparison with the field results;

Contribution no. 249 of the Institut für Geophysik, ETH Zürich
Contribution no. 193 of the Geophysikalisches Institut, Universität Karlsruhe

0340-062X/80/0048/0028/$01.20



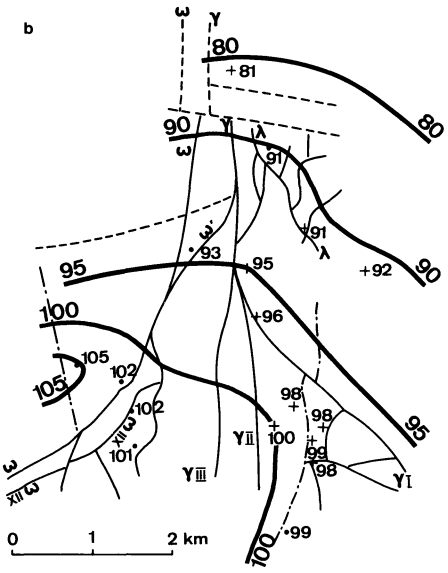


Fig. 1.
a Location of Landau.
b Isotherms and fault pattern of the Landau field at 900 m below sea level (about 1,050 m below earth surface). The *numbers* indicate temperature values in °C. The prominent fault in this region is the ω-fault

(b) the spatial extension of the anomaly, also to compare with field results;

(c) the age of the anomaly;

(d) the original depth of the thermal water;

(e) the thermal water flow, i.e., the quantity of water transported per unit time.

These quantities allow a discussion about the applicability of the model to the real situation.

The way of the rising deep groundwater consists of two paths: Path 1, from the original depth of the water up to the top of the crystalline basement; Path 2, its subsequent way through the overlying sediments. It will be shown that Path 1 is the critical one for the thermal model, i.e., the original depth of the water strongly influences the results. On the other hand, the distribution of the flow within the graben fill (Path 2) is of interest in determining the lateral extension of the thermal anomaly.

The Model Calculation

We consider a half-space in which heat transport takes place. The boundary condition at its surface is isothermal (mean annual temperature). We assume an undisturbed, 'normal' temperature field before the forming of the anomaly. Within the half-space heat is transported by:

(a) Thermal conduction;

(b) convection, caused by moving waters in a porous medium. In the heat conducting medium there are permeable zones, where movement of fluid can take place. The velocity field is given and is assumed to be constant, starting at an initial time $t=0$.

Furthermore, we distinguish two regions within the half-space: The sedimentary cover, and below this, the crystalline basement. We assume that permeable zones exist in both regions. In the permeable zones of the graben fill the flow of water is governed by the Darcy velocity, however in the faulted crystalline basement this definition is only approximately valid. No accurate description on the special mechanism of the permeability of a fault is available. We avoid this difficulty by assuming that within the fault zone the definition of a Darcy velocity is still approximately valid. Using this assumption. the two regions (sediments and crystalline) can be handled with the same mathematical formulae.

We use a basic equation, whose applicability to shallow aquifers has already been successfully tested (Werner and Kley 1977). In the permeable zones, the equation is

$$\gamma \mathbf{u} \operatorname{grad} T + \frac{\partial T}{\partial t} = D \nabla^2 T. \tag{1}$$

In zones with pure heat conduction Eq. (1) reduces to:

$$\frac{\partial T}{\partial t} = \kappa \nabla^2 T \tag{2}$$

where

T = temperature
t = time
$\mathbf{u}$ = vector of the Darcy velocity (approach velocity, specific discharge)
$\gamma = c_w \rho_w / c \rho$ = ratio of the heat capacities of water and rock
ρ = density of rock
$\kappa = K / c \rho$ = thermal diffusivity of rock in the zones of pure heat conduction
K = thermal conductivity
D = 'effective thermal diffusivity' in water permeable zones, which includes the effect of thermal dispersion.

The Darcy velocity **u** can be calculated from the mean flow velocity **v** (advance velocity) by

$$\mathbf{u} = p\mathbf{v}$$

where p indicates the effective porosity.

For the heat transport the Darcy velocity has to be used, because this quantity (specific discharge) describes the liquid flow as mass per unit time (in the strict sense **u** should not be called a velocity).

The parameter D, which we call 'effective thermal diffusivity', is exactly a tensor quantity influenced by diffusion and dispersion. We assume here that D can be considered as a scalar constant (see e.g., Werner and Kley 1977).

As it is evident from (1) and (2) radiogenic heat sources are not taken into account. They are omitted because their effect on our hydrothermal model is negligible.

We consider a two-dimensional problem. With x as horizontal and z as the positive downward direction coordinates (1) and (2) can be written as follows:

$$\frac{\partial T}{\partial t} = D\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial z^2}\right) - f_x \frac{\partial T}{\partial x} - f_z \frac{\partial T}{\partial z} \tag{3}$$

for the permeable zones, where $\mathbf{f} = (f_x, f_z) = \gamma \mathbf{u} = \gamma(u_x, u_z)$, and

$$\frac{\partial T}{\partial t} = \kappa\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial z^2}\right) \tag{4}$$

for zones with pure conduction.

The field **f** is a given quantity, i.e., it is an 'input parameter' of the calculation. In our model it is assumed that **f** is constant with time.

The equation of mass conservation is div **f** = div **u** = 0. This means that thermally induced density changes are not considered.

The problem is solved with a finite difference method, i.e., the differential Eqs. (3) and (4) are converted into equivalent difference equations. The temperature at every point of a grid with point distance $\Delta x = \Delta z = 200$ m is calculated after every time step Δt. It is advantageous to work with two different time steps, according to the different heat transport rates by either pure conduction or by convection. We select $\Delta t' = 60$ years for the permeable zones and $\Delta t = 300$ years for the zones with pure conduction. The choice of the ratio between Δx, Δz and Δt, $\Delta t'$ is limited by the stability condition of the numerical method.

The fault in the crystalline basement (more precisely, the narrow zone in which the thermal water can rise) is assumed as perpendicular to the earth's surface and is chosen to be 200 m wide, corresponding to one distance step Δx. Our two-dimensional calculation is limited to a symmetrical case where the axis of symmetry at $x = 0$ coincides with the z-axis.

Models for the Landau Anomaly

Figures 2, 4, and 5 show different model versions of the geothermal anomaly of Landau, which are distinguished by different assumptions on the path of the thermal water in the graben fill. In the first model (Fig. 2) it is assumed for simplicity that the water flows in a narrow, horizontal aquifer. In this case no water movement takes place within the sediments apart from the aquifer.

In our calculation the sediment cover has a laterally constant thickness of 2,000 m. The thickness of the aquifer is assumed to be 200 m, corresponding to the distance step Δz. The width of the vertical permeable zone in the basement is also assumed to be 200 m.

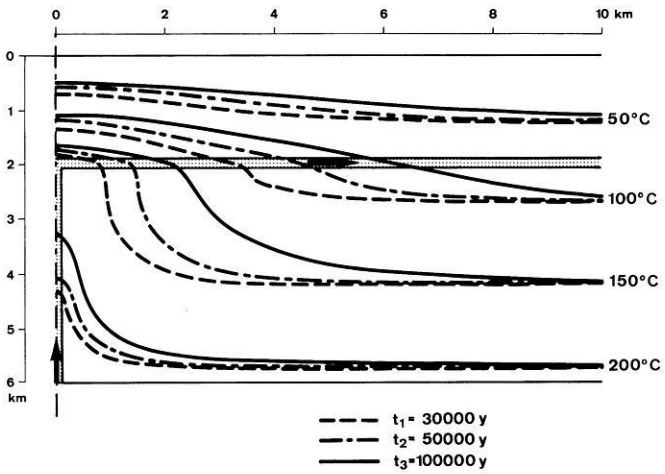


Fig. 2. Hydrothermal model 1 (two-dimensional, symmetrical) Deep groundwater rises along a fault zone and builds up a thermal anomaly. The dotted zones are assumed to be permeable, the arrows indicate the direction of the water flow. The isotherms show the evolution of the anomaly in space and time. $dT/dz =$ 33 deg/km is the temperature gradient of the undisturbed field. The total amount of the deep groundwater flow is $S = 100$ m^2/year ($= 100$ m^3/year for 1 m length perpendicular to the figure plane). In the centre of the anomaly at a depth $z = 1{,}000$ m a temperature of about 100° C is obtained after 100,000 years, corresponding to the observations in the Landau field

The quantity S (flux of deep groundwater) is assumed to be 100 m^2/y (i.e., 100 m^3/y along 1 m perpendicular to the plane of the figure).

A decisive parameter in our calculation is the original depth z_o of the thermal water. To build up a thermal anomaly as it is observed near Landau, we must assume an original water depth of about 6,000 m, if the initial temperature distribution is assumed as normal (33 deg/km).

The result of the calculation is shown in Fig. 2 in form of isotherms. The shape of the anomaly and its evolution with time can now be recognized. After 100,000 years a state is reached which is very similar to the observed temperature field in Landau. For a comparison of the model results with the observations, we can consider on the one hand the temperature at $z = 1{,}000$ m in the centre of the anomaly, on the other hand the decrease of the temperature with increasing distance from the centre of the anomaly. Figure 3 shows such a comparison: It can be seen that the observed temperature decrease (dots) agrees with the model results. There are, however, two reasons which allow only a qualitative comparison between model and observation:

(1) The model is two-dimensional and therefore not comparable in detail with three-dimensional data (see Fig. 1b);

(2) in our model the conductivity (resp. the diffusivity) of the sediment body is assumed as constant, whereas newest research work (Sattel 1979; see also Werner and Fuchs 1977) shows an increase of the thermal conductivity with depth. In order to demonstrate that the distribution of water flow within the sedimentary cover does not influence strongly the calculated temperature field we constructed two further models. In each model the quan-

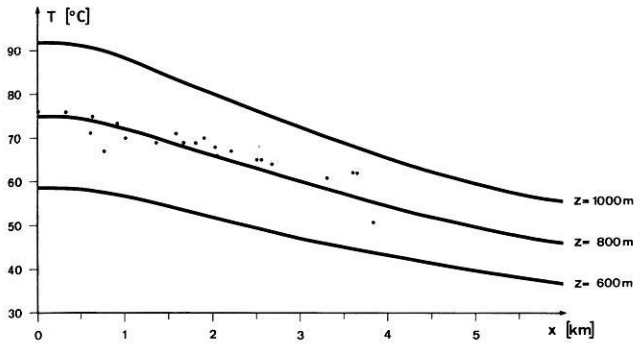


Fig. 3. The lateral temperature distribution at different depths z after 100,000 years according to model 1. For comparison measured values are shown plotted as function of the distance from the centre of the anomaly

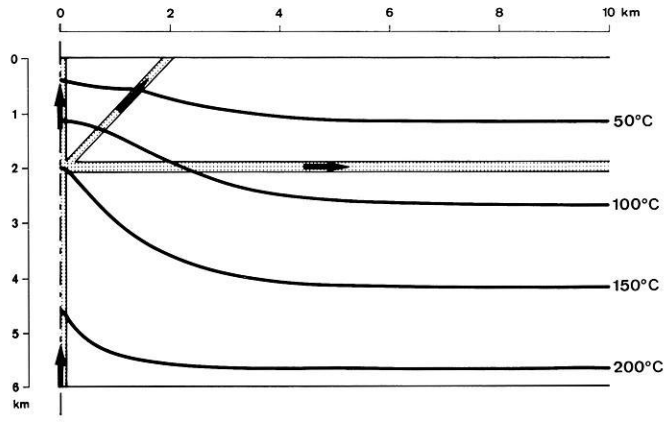


Fig. 4. Hydrothermal model 2. As distinguished from model 1 water-bearing faults within the sediment body are assumed. With this flow pattern a quantity $S=40\ m^2/y$ is sufficient to build up a similar temperature anomaly after 100,000 years

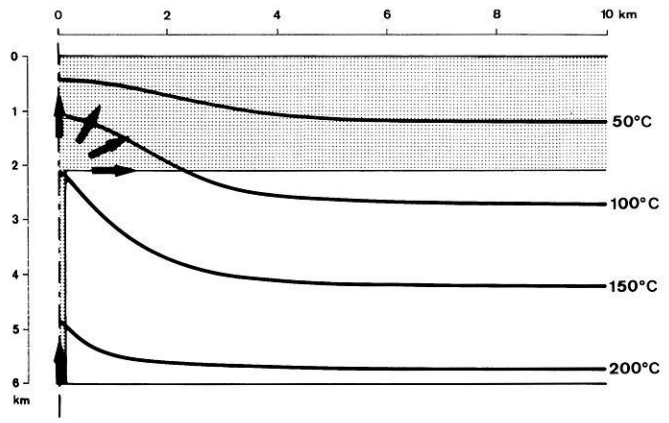


Fig. 5. Hydrothermal model 3. In this case the whole sediment body is assumed to be permeable with a radial flow distribution (*arrows*). The temperature field is shown for a flow $S=33\ m^2/y$ after 100,000 years. As in the models 1 and 2 the original depth z_0 of the thermal water must be chosen to be 6,000 m. With a smaller depth z_0 the Landau anomaly could not be explained without changing the basic conditions of the model

tity S was chosen so that the maximum temperature is nearly the same in the three models. The second model (Fig. 4) is an attempt to divide the groundwater flow into different paths. For this model the existence of permeable zones in the graben fill is required. In this case a flux of water $S=40\ m^2/y$ was assumed and the temperature distribution after 100,000 years is shown. The original depth z_0 is still 6,000 m. A comparison with the model in Fig. 2 shows that the flow pattern within the sediments is a secondary question.

This can be seen in the third model too (Fig. 5). Here we assume that the whole graben fill is permeable and that the flow of water is radially distributed. The situation in the basement is the same as in the other models (Figs. 2 and 4), and for this example $S=33\ m^2/y$. Figure 5 shows the isotherms for an age of the anomaly of 100,000 years.

The sequence of these model examples could be continued, changing the flow pattern in the sediments. The original depth z_0 of 6,000 m must be considered as a minimum value. Calculations with $z_0<6{,}000$ m cannot fit the observed anomaly if the same order of magnitude of the flux S is supposed. This means that our model is particularly suited for obtaining indications about the quantities z_0 and S.

The Origin of the Deep Groundwater

In our model the quantity of the water is prescribed, i.e., the hydromechanical causes are not explained. We assume that these deep waters are part of a wider circulation pattern. They accumulated in the basement by percolation from the earth's surface and are therefore not interpreted as juvenile waters.

For the geothermal anomaly of Landau large parts of the Pfälzer Wald come into question as source drainage area. If we assume a drainage area 100 km wide, the deep groundwater flux S amounts to some 0.1% of the annual rainfall. This value is sufficiently small to be considered as realistic. The original depth of the water at 6,000 m leads to the hypothesis of a 'crystalline aquifer' in a similar depth. Seismic investigations indicated a zone of about 10 km thickness with low seismic velocity (Müller et al. 1973). Possibly there is a relation between this zone and the postulated aquifer.

A support to our model is the salinity distribution in the Landau field. According to Schad (1962), a strong increase in the salt content of pore water is observed in the field from west to east; within a few kilometers it increases from 0 to 130 g/l. It is remarkable that the zone of fresh water almost coincides with the centre of the anomaly. This fact can be well explained with the flow distribution of the model in Fig. 5. Assuming a deep water flow of fresh water, which displaces the saline pore water, an explanation of the salt distribution in the field is obtained.

The Age of the Anomaly

It appears from the models that the Landau anomaly is about 100,000 years old. Figure 6 shows the evolution with time of the anomaly for the model in Figure 4, described by the increase of the temperature in the centre of the anomaly for different water fluxes S.

It follows from Fig. 6 that an estimate of the age can be different depending on the choice of S. For example an age of 100,000 years can be considered as a minimum age for the case $S=50\ m^2/y$.

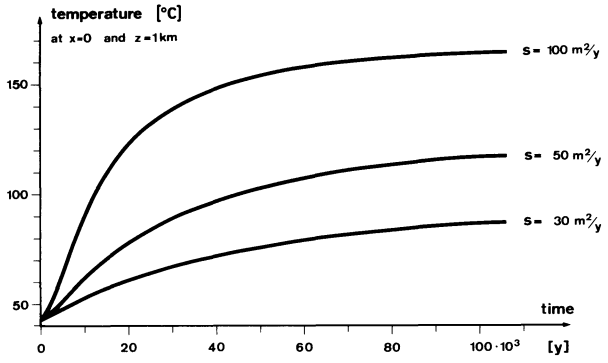


Fig. 6. The time depending temperature in the centre of the anomaly at $z = 1{,}000$ m calculated for different water flows S according to model 2. This diagram can be used for estimating the age of the anomaly

But it is also possible that the anomaly is in a phase of progressive evolution, hence in the area where the temperature curves of Fig. 6 show a marked increase. It must be noted that the quantity S can vary with time; in our model S can be considered as a mean value. With these assumptions the age of the anomaly can only be correct to within an order of magnitude.

This time period of about 100,000 years is much smaller than the age of the Rhinegraben (about 50 my). Accordingly, the presented model does not include assumptions connected with the rift building of the graben system. The process of rifting also leads to a prominent geothermal anomaly, but in a depth range which includes the lithosphere and the astenosphere (Werner and Kahle 1980). We must therefore make a distinction between two geothermal problems:

(a) the problem of shallow temperature anomalies (depth range: a few kilometers; time: about 100,000 years);

(b) the problem of the thermal anomaly of the upper mantle (depths down to 200 km; characteristic time about 50 my)

The very different orders of magnitude in space and time allow us to consider them as independent. This means that our model is not affected by the heat anomaly of the upper mantle. The age of the shallow anomaly considered here is to compare with the time period in which a permeable fault in the basement exists.

As the age of the anomaly is very small compared with the age of the graben, the existence of a water-bearing fault can be considered as a local 'episode' in the evolution of the graben. This suggests the assumption that during the graben evolution many such thermal events happened at different times and different places.

Our model is clearly valid for a temporally and spatially restricted situation within the graben tectonics. The last 100,000 years are indicated by Illies (1978) as a period of increased tectonic activity. This time span is well in accordance with the age of our model anomaly.

A further support to the young age of the Landau anomaly is the study of the degree of coalification of organic inclusions in the sedimentary rocks of the Rhinegraben (Teichmüller 1970; Doebl et al. 1974; Teichmüller and Teichmüller 1977; Buntebarth 1978). The degree of coalification of this organic material depends on the quantity temperature × time. Based on measurements of the degree of coalification the age of a thermal anomaly in the graben fill can be estimated. Investigations of Buntebarth (personal communication) support the idea that the Landau anomaly is a relatively young phenomenon, according with our model.

Outlook

The presented model is oriented to the special case of the thermal anomaly in Landau, but it could be applicated to other prominent anomalies within the whole Rhinegraben (e.g., Soultz/Pechelbronn, Stockstadt). Less pronounced local anomalies could be correspondingly interpreted with a minor source depth z_0 or smaller fluxes S. In any case a thermal anomaly would be connected with a water-bearing fault in the basement. With this we generalize the model for the case of Landau and assume that hydrothermal processes of this type are active in the whole Rhinegraben.

This supposition leads to the problem of the hydraulic system and of its causes. Such an hydraulic system includes a region which exceeds the proper graben zone. The considerations of the present paper are limited to the special part of the circulation system characterized by ascending deep groundwaters. It would be interesting to work out a hydromechanical model for the whole water circulation system in the underground of the wider surroundings of the graben, testing then the reality of our hydrothermal model.

Acknowledgments. This study was carried out under contract with the Commission of the European Communities (Project No. 321/79/2 EGD). We thank Wintershall A.G. for granting permission to carry out the measurements and especially Dr. F. Doebl for providing all necessary information. Thanks are also due to Dr. H. Scriba and Dr. J. Channel for critical reading of the manuscript.

References

Buntebarth, G.: The degree of metamorphism of organic matter in sedimentary rocks as a paleogeothermometer, applied to the Upper Rhinegraben. Pure Appl. Geophys. **117**, 83–91, 1978/1979

Delattre, J.N., Hentiger, R., Lauer, J.P.: A provisional geothermal map of the Rhinegraben (Alsatian part). In: Graben problems, J.H. Illies, St. Müller, eds.: pp. 107–110. Stuttgart: Schweizerbart 1970

Doebl, F.: Die geothermischen Verhältnisse des Ölfeldes Landau/Pfalz. In: Graben problems, J.H. Illies, St. Müller, eds. pp. 110–116. Stuttgart: Schweizerbart 1970

Doebl, F., Heling, D., Homann, W., Karweil, J., Teichmüller, M., Welte, D.: Diagenesis of tertiary clayey sediments and included dispersed organic matter in relationship to geothermics in the Upper Rhinegraben. In: Approaches to taphrogenesis, J.H. Illies, K. Fuchs, eds.: pp. 192–207. Stuttgart: Schweizerbart 1974

Hänel, R.: Bericht über Temperaturmessungen bei Landau/Pfalz. Report of the Niedersächsisches Landesamt für Bodenforschung Hannover 1974

Illies, J.H.: Taphrogenesis and plate tectonics. In: Approaches to taphrogenesis, J.H. Illies, K. Fuchs, eds.: pp. 433–460. Stuttgart: Schweizerbart 1974

Illies, J.H.: An extinct tensional rift valley, reactivated by shear. Proceedings International Symposium on the Rio Grande Rift, Santa Fé, New Mexico pp. 42–46, 1978

Illies, J.H., Fuchs, K. (eds): Approaches to taphrogenesis. Stuttgart: Schweizerbart 1974

Illies, J.H., Müller, St. (eds): Graben problems. Stuttgart: Schweizerbart 1970

Lauer, J.P.: Isothermes profondes et 'effects de socle' Bull. Soc. Fr., **18**, 1139–1149, 1976

Müller, St., Peterschmitt, E., Fuchs, K., Ansorge, J.: Crustal structure of the Rhinegraben area. Tectonophysics **20**, 381–392, 1973

Rothe, J.P., Sauer, K.: The Rhinegraben progress report 1967. Abh. Geol. Landesamtes Baden-Württenb. **6**, 1967

Sattel, G.: Aufbau einer Meßapparatur zur Bestimmung thermischer Parameter von Gesteinen und Messungen an tertiären Sedimenten, Dipl. Arbeit, Geophysik. Inst., Univ. Karlsruhe 1979

Schad, A.: Das Erdölfeld Landau., Abh. Geol. Landesamtes Baden-Württenb. **4**, 81–101, 1962

Teichmüller, M.: Bestimmung des Inkohlungsgrades von kohligen Einschlüssen in Sedimenten des Oberrheingrabens – ein Hilfsmittel bei der Klärung geothermischer Fragen. In: Graben problems, J.H. Illies, St. Müller, eds.: pp. 124–142. Stuttgart: Schweizerbart 1970

Teichmüller, M., Teichmüller, R.: Zur geothermischen Geschichte des Oberrhein-Grabens. Zusammenfassung und Auswertung eines Symposiums. Fortschr. Geol. Rheinl. Westfalen **27**, 109–120, 1979

Werner, D.: Probleme der Geothermik am Beispiel des Rheingrabens. Thesis, Univ. Karlsruhe 1975

Werner, D., Doebl, F.: Eine geothermische Karte des Rheingrabenuntergrundes. In: Approaches to taphrogenesis. J.H. Illies, K. Fuchs, eds.: pp. 182–191. Stuttgart: Schweizerbart 1974

Werner, D., Fuchs, K.: Exploration and interpretation of the geothermal anomaly in the Rhinegraben. In: Seminar on geothermal energy, Vol. I, pp. 189–206. Commission of the European Communities, Brussels 1977

Werner, D., Fuchs, K., Parini, M., Sattel, G., Schroth, G., Doebl, F.: Exploration and interpretation of the strongly anomalous temperature field underneath the Rhinegraben. Report of the European Communities 1978

Werner, D., Kahle, H.G.: A geophysical study of the Rhinegraben, part I: kinematics and geothermics, Geophys. J. in press, 1980

Werner, D., Kley, W.: Problems of heat storage in aquifers. J. Hydrol. **34**, 35–43, 1977

Received October 25, 1979; Revised Version December 20, 1979

J. Geophys. 48, 34–39, 1980

Journal of Geophysics

Magnetic and Gravity Investigations of the Dead Sea Rift and Adjacent Areas in Northern Israel

Y. Folkman
The Institute for Petroleum Research and Geophysics, 1 Hamashbir Street, Holon, Israel

Abstract. Magnetic and gravity data covering a wide strip across the rift, with dimensions suitable for crustal and upper mantle investigations are presented and analysed. They suggest that the thickness of the crust remains almost unchanged across the rift and that the upper mantle is also normal. On the other hand, the rift probably delineates the boundary of an upper crustal lithological transition zone, characterized by an increasing mafic component from east to west.

Local magnetic anomalies within the rift zone are believed to be related to basalt flows instead of the commonly interpreted intrusions. Based on this suggestion, fault patterns are delineated. They support the sinistral strike-slip hypothesis for the origin of the Dead Sea Rift.

Key words: Dead sea rift – Magnetic anomalies – Gravity anomalies – Crust – Upper Mantle.

Introduction

The structure of the Dead Sea Rift has been intensively investigated for almost one hundred years. However, in the field of geophysical crustal and upper mantle investigations, little has been published based on continuous data measurements across the rift (Knopoff and Belshe 1967).

This is the first study to present an analysis of gravity and magnetic data which cover a wide strip across a segment of the Dead Sea Rift. The dimensions of the strip are sufficient to allow interpretation of deep crustal and upper mantle characteristics as well as analyses of near surface features.

The investigated portion of the rift extends from Lake Kinneret (Sea of Galilee) to the town of Kiryat Shemona (Fig. 1). It forms an elongate valley, five to eleven kilometers wide, and is bounded by mountainous terrain – the Golan Heights in the east and the Galilee mountains in the west.

The geology in the northern rift valley has been described in detail by Picard (1965), Horowitz (1973), Michaelson (1973), Neev (1979), and Schulman (1962, 1978), and can be briefly summarized as follows: until Late Eocene or Early Oligocene times, the depositional environment in the area presently occupied by the valley was marine and similar to neighboring areas. However, since the Miocene and up to recent times, several phases of faulting and block tectonics which were associated with the creation of the rift resulted in the formation of inland lakes covering different parts of the area. These followed rapid accumulation of thick sequences of clastic deposits with minor chalks within deep grabens. In addition, two major phases of basalt eruption covered the area. The earlier one occurred in Miocene times and is represented in neighboring areas by a thickness of up to 600 m of the 'Lower Basalt' (Schulman 1962). The second phase, namely the 'Cover Basalt', followed intensive faulting activity during Late Pliocene and Early Pleistocene times (Freund et al. 1965). It covers the entire Golan Heights and parts of eastern Galilee. This 'Cover Basalt' is also found in the rift valley where major flows have descended from the mountains.

According to the structure and morphology, three zones are clearly distinguished in the valley. They are, from south to north (Fig. 1).:

(a) Lake Kinneret (Sea of Galilee) depression with the lake surface at 210 m below sea level.

(b) The Korazim block, a structural high rising to 500 m above sea level and covered by up to 200 m of the 'Cover Basalt,' and

(c) The Hula basin, a topographic and structural depression filled with a thick accumulation (more than 700 m) of clastic sediments.

The elevated areas on both sides of the rift differ considerably in their structure. In the east, the Golan Heights form an uplifted syncline covered by hundreds of meters of basalts (Michaelson 1973; Mor 1973). In the west, the Galilee Mountains exhibit a complex structure composed of transverse faults (i.e., normal to the rift valley), tilted blocks and faulted anticlines. Outcrops are composed mainly of carbonate rocks ranging in age from Early Cretaceous to Early Tertiary (Picard 1943; Picard and Golani 1965).

The Data

The magnetic data consist of previous aeromagnetic surveys (Domzalski 1967; Folkman 1971; Folkman and Yuval 1976) flown at an altitude of 1,000 m above sea level with a line spacing of 2 km. The gravity data were compiled from previous surveys and investigations (Amitai 1962; Yuval 1966; Ginzburg 1960, 1968) with additional recent measurements. The Bouguer anomalies refer to mean sea level datum.

The aeromagnetic map (Fig. 1) clearly demonstrates considerable differences in the magnetic field across the rift. As one goes from east to west, three zones, each characterized by distinctive anomalies, can be distinguished. Over the Golan Heights, elongated short wavelength high amplitude anomalies are arranged in a sublinear pattern trending NE (Folkman 1978). Near the eastern boundary faults of the rift valley, this pattern changes

0340-062X/80/0048/0034/$01.20

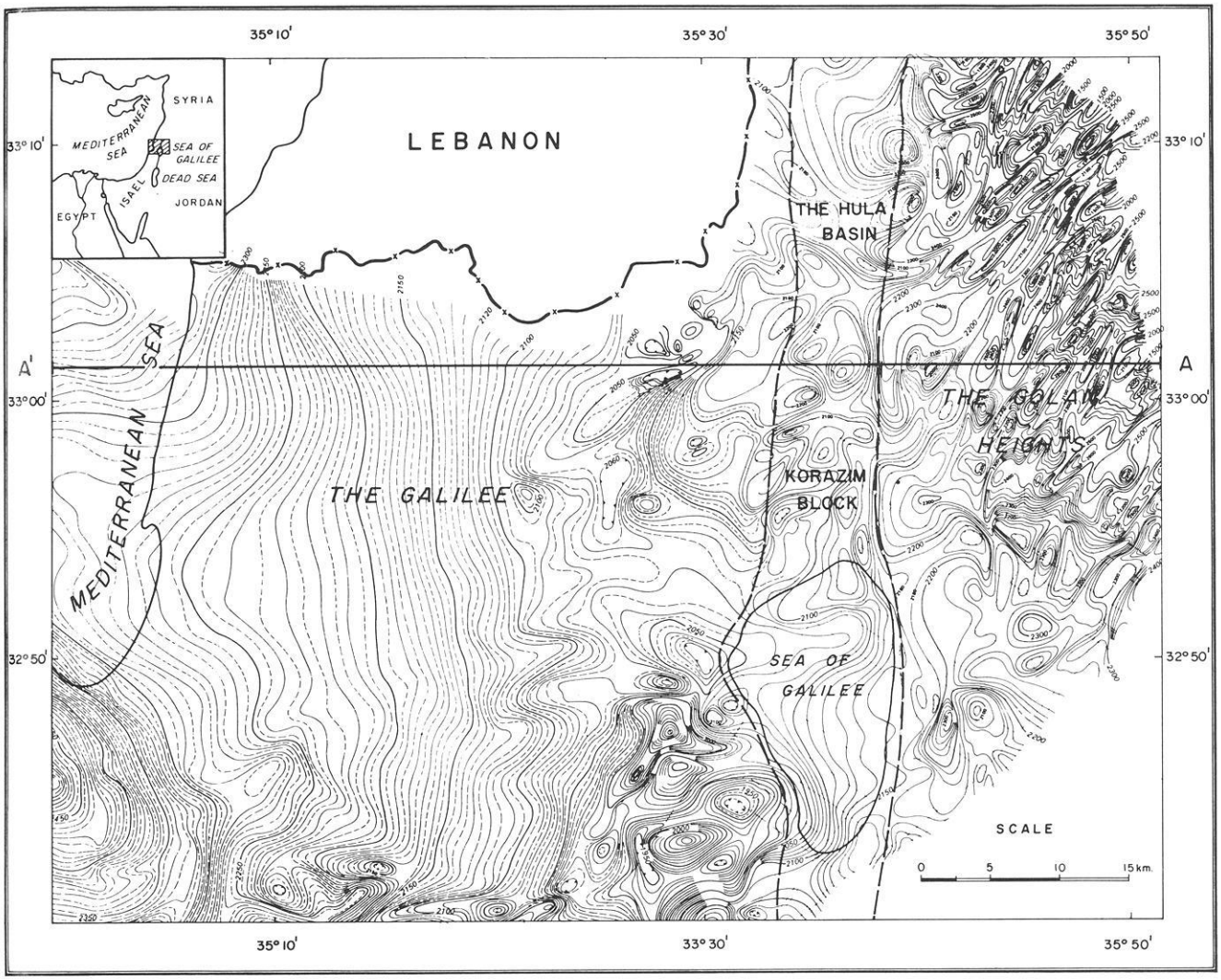


Fig. 1. Aeromagnetic map, after Folkman and Yuval, 1976 Flight altitude – 1,000 m above M.S.L. *Dashed lines* delineate the rift zone

abruptly into N-S trending contours which clearly delineate the eastern boundary fault system. Further to the west, the anomaly pattern changes again into well defined local anomalies of various orientations covering the rift valley and eastern Galilee. The western boundary faults of the rift are undetectable on the magnetic map.

The third zone having a characteristic magnetic field extends over the central and western Galilee, the coastal plain, and the continental shelf. Here, rarely affected by the complicated geological structure, the magnetic field increases continuously westward.

The detailed compiled Bouguer anomaly map is not authorized for publication. However, the main gravity patterns are demonstrated on a generalized map (Fig. 2) which shows that a good correlation usually exists with the main magnetic patterns. Again, the same three zones have typical and yet different characteristic patterns of Bouguer anomalies. Over the Golan Heights, an axis of a Bouguer low trending NE is located in the south with a rising gradient towards Mount Hermon in the north. Within the rift valley, two distinct negative anomalies cover the Hula basin and the eastern portion of Lake Kinneret. Apart from these anomalies, the disturbance caused by the rift is almost unnoticeable.

Further to the west, over the Galilee, the Bouguer anomaly map becomes very similar to the magnetic map. The gradient rises westwards with a dominant pattern of N-S trending contours that cannot be correlated with the known major structural features. Only over the coastal plain and the continental shelf can local anomalies be separated from the regional field and be correlated with structural features.

Interpreted Crustal Structure

The gravity and magnetic fields over the Galilee, which as described above, are not correlatable with local structures, resemble that of a regional component with only minor superimposed local effects. It may, therefore, be justified to assume causative structures which lie deeper in the crystalline crust or upper mantle. However, the magnetic field excludes a mantle solution because the Curie point is most probably reached near the Moho, under local conditions of a normal thermal gradient (Ben Avraham et al. 1978; Eckstein 1976) and a crustal thickness of approximately 30 Km (Ginzburg et al. 1979).

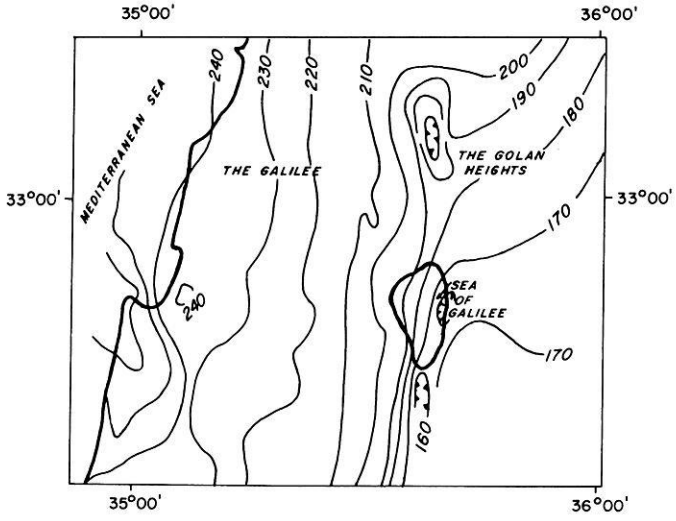


Fig. 2. Generalized relative Bouguer anomaly map. Contour interval 10 milligals

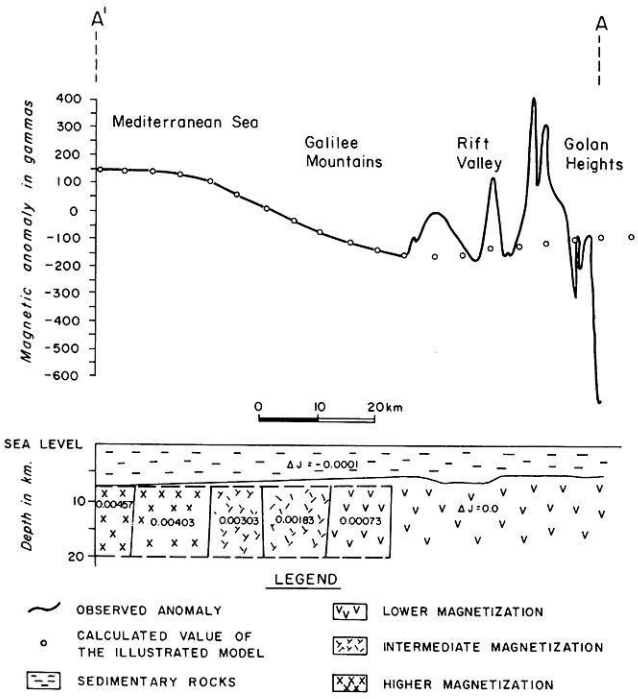


Fig. 3. A possible model interpretation of magnetic profile *A–A'* across northern Israel. *Δj* denotes the assumed magnetization contrast in e.m.u.

The pattern of the magnetic contours, as demonstrated in Fig. 1, allows a two dimensional interpretation along E-W oriented profiles across the entire area. The short wavelength high amplitude anomalies in the east are obviously caused by near surface basalts. They are superimposed on the main regional component and may be filtered out by choosing the best fitted zero curve. The interpretation of this curve yields a crustal model with N-S trending discontinuities across which lateral variations in rock magnetization occur (Fig. 3). The increasing magnetization westwards probably indicates an increase in the mafic content of the rock type. Thus, according to this interpretation, the rift and its eastern side are underlain by a crust which is different in composition from the crust underlying the western side of the rift.

This hypothesis is supported by the interpretation of the Bouguer gravity anomaly which is again based on two dimensional modeling along the same profile (Fig. 4). The Bouguer anomaly profile shows a steep gradient over the Galilee which becomes more moderate over the Mediterranean coastal plain and continental shelf, where low density marls fill a large sedimentary basin of Neogene age (Gvirtzman 1969). The gravity effect of the basin has been calculated and stripped off the anomaly, thus causing a continuation of the steep gradient also over the coastal plain and the continental shelf, as demonstrated in Fig. 4.

The resultant regional component of the Bouguer anomaly may be explained on the basis of the following three models:

(a) Lateral density variations in the upper mantle, which form an anomalous low density lens under the rift zone

(b) thinning of the crust towards the Mediterranean Sea

(c) lateral lithological variations in the crystalline crust across N-S trending discontinuities, with the density increasing westwards.

The first explanation is unacceptable because the Bouguer anomaly map of Syria (Bureau Gravimetrique International 1971; Tiberghien 1974) shows that the Bouguer values in the east continue to decrease towards a regional low over southeastern Syria. This means that if an anomalous low density lens exists in the upper mantle, then its axis should be sought after under southeastern Syria, while over the coastal plain of northern Israel, only minor gravity effects would be noticed.

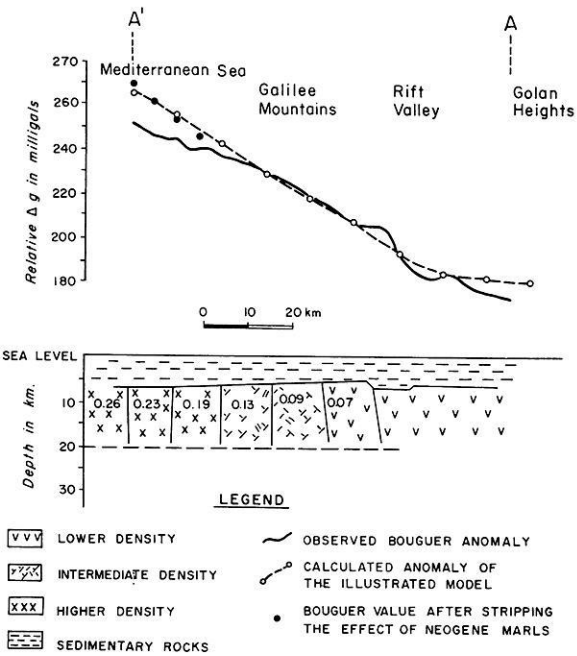


Fig. 4. A possible model interpretation of magnetic profile *A–A'*; *Δρ* is the assumed density contrast in g/cm^3 units

The second explanation is probably the most acceptable one for central and southern Israel (Folkman 1976; Ginzburg et al. 1979). However, across northern Israel the magnetic anomaly cannot be explained by crustal thinning and, therefore, the third possibility is more attractive; it presents an interpretation model, shown

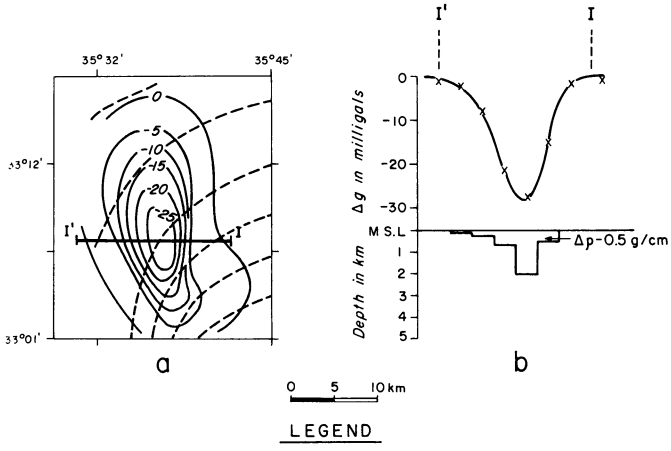


Fig. 5a and b. Gravity interpretation of the Hula basin. **a** Relative Bouguer anomaly over the basin. *Dashed contours* are the assumed regional component with a 5 mgal. interval. **b** Model interpretation of a profile across the basin after subtraction of the assumed regional background. Crosses denote the calculated anomaly of the illustrated model

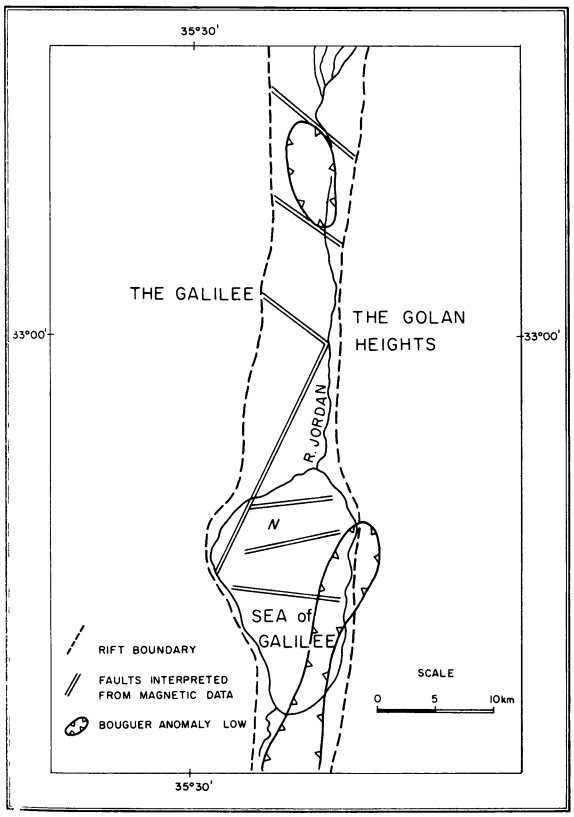


Fig. 6. Interpreted transverse and oblique fault patterns in the rift zone. *N* denotes a block interpreted to be free of basalt cover

in Fig. 4, which is almost identical with the one derived from the magnetic data.

It can be concluded, therefore, that the magnetic and gravity data suggest variations in the lithological composition of the crust across the northern segment of the Dead Sea Rift. Under the rift and its eastern side, magnetization and density of the rock type are relatively lower, whereas under the region west of the rift, the magnetization and density increase towards the continental shelf, probably indicating an increase in the proportion of mafic minerals in the rock type.

Local Anomalies in the Rift Valley

Gravity

Two distinct negative anomalies are located over the Hula basin and the southern and eastern portions of Lake Kinneret (Fig. 2). They clearly define deep depressions filled with thick accumulations of low density sediments and separated by a structural high.

Assuming a density contrast of -0.5 g/cm^3 between the dense limestone and dolomite of Cretaceous age and the soft clastic fill, the depth to the bottom of the depressions has been calculated by an inversion technique (Bott 1960), after removal of the assumed regional component of the Bouguer anomaly as shown in Fig. 5. A depth of 2,000 m was calculated to the Cretaceous carbonate rocks in the deepest part of the Hula basin and also under the Jordan River, just south of Lake Kinneret. The inversion models also indicate normal step fault boundaries east and west of the interpreted depressions.

Magnetics

Unlike the Gulf of Elat (Allan 1970; Allan and Morelli 1970) and the southern portion of the Dead Sea Rift (Folkman and Yuval 1976; Neev and Hall 1976) whose typical magnetic field is smooth, the northern area is characterized by short wavelength, well defined local anomalies, some of which probably indicate reversely magnetized basalts (Fig. 1). Those anomalies which delineate causative bodies of a three dimensional type, could well be interpreted as being caused by igneous intrusions, the existence of which has been speculated (Picard 1953; Vroman 1958).

However, the same anomalies can be more simply explained by bodies of extrusive material formed during several periods of basalt eruption, particularly in Late Pliocene and Early Pleistocene times. In this area, the most expected causative bodies may belong to the following three groups:

(a) local accumulation of basalts in a pre-existing topographic relief,

(b) vertical displacement across a faulted sheet of basalt, and

(c) successive crystallization of normal and reversed magnetized flows.

In those places where basalts crop out, the magnetic anomalies usually correlate reasonably with the outcrops. It seems, therefore, justified to suggest an interpretative approach based on extrusive material only. In this case the existence of dykes is obviously not supported by any direct evidence. On the other hand, magnetic anomalies can be used for the delineation of major faults and the estimation of vertical movements of blocks before and after the main periods of basalt eruption.

Let us first consider the Hula basin in the north (Fig. 6). Here, the deep depression based on the gravity interpretation is

bounded in the north and south by a pattern of NW-SE trending elongated magnetic anomalies which most probably delineate 'step models' or faults in a sheet of basalts, provided that one assumes dominant normal magnetization.

A fault pattern of this kind has been previously suggested by Freund et al. (1970) to account for the interpreted rhomb-shaped graben of the Hula. However, its existence has never been supported by direct observations because of the thick cover of young sediments. Thus the combined magnetic and gravity study is the first to detect and delineate the major rhomb-shaped faults of the Hula graben which have a downthrow of up to 2,000 m.

A transverse fault pattern has probably also dominated the tectonic history of the Lake Kinneret depression. Here, the magnetic field is very smooth in contrast to the surrounding areas either because the causative basalts are deeply burried under the lake, or because most of the area occupied by the lake formed an elevated ridge during the periods of eruption and therefore was not covered by basalts.

The dominant magnetic features which have also been detected on a follow-up ground survey (Ben Avrahan, personal communication) comprise a transverse low separating two distinct highs across the lake (Fig. 1). Again assuming dominant normal magnetization, a straightforward solution would relate this pattern to a block structure divided by transverse faults as shown on Fig. 6. The positive anomalies represent blocks which are covered by basalts, whereas the negative anomaly is interpreted as being caused by the absence of a basalt cover. If this is the case, then the positive anomaly near the northern coast of the lake probably describes an accumulation of basalts which descended from the north and were dammed by a barrier formed by an elevated transverse block. Consequently, the positive anomaly occupying the central portion of the lake probably describes a block tilted to the east with a maximum thickness of basalts covering the lower eastern side.

The magnetic evidence strongly supports Neev's (1979) suggested structural solution in which the northern portion of the lake is shown to have been an elevated continental block until 13,000 years ago. Moreover, the magnetic study clearly delineates the faulted boundaries of the block and may thus help to detect brine springs which contaminate the lake. These springs are known from onshore observations to ascend along fault planes.

The western and higher portion of the tilted block, suggested as occupying the central portion of the lake, may be a part of the NNE trending anticlinal ridge postulated by Michaelson (1973) to be buried under the western margins of the lake.

Structural Implications

The magnetic and gravity anomalies clearly indicate that unlike some other portions of the East African Rift system, which are associated with a regional thinning of the lithosphere and an extreme thinning beneath the rift floor (Girdler 1978), the northern segment of the Dead Sea Rift is probably underlain by a normal crust and upper mantle.

Unlike the Red Sea magnetic anomalies which are generally accepted to indicate sea floor spreading, here an explanation based on basalt flows is more convincing, hence the Red Sea structural model is not considered applicable for the Dead Sea rift even in the northern segment where well defined local magnetic anomalies are shown.

Some crustal thinning under the rift between the Dead Sea and the Gulf of Elat was suggested by Ginzburg et al. (1979) based on explosion seismology. Further investigations of the same seismic data (Perathoner et al. 1979; Ginzburg et al. in press) have indicated vertical lithological variations near the base of the crust under the rift, and a normal lower crust under the western side of the rift. Unfortunately, the seismic experiments did not cover the eastern side of the rift and, therefore, the possibility still remains that the crustal type under the rift also extends to the eastern side. This implies that the rift zone is adjacent to some kind of a lateral lithological transition in the crust. This lateral transition zone is strongly supported by the magnetic and gravity anomalies over the northern segment of the Dead Sea rift and adjacent areas. Under the rift and its eastern side, the crust is interpreted to be characterized by low magnetization and relatively low density. West of the rift, however, these properties gradually increase, probably indicating an increase in the mafic content of the crustal rock type. A similar model has been previously suggested by Knopoff and Belsche (1967) for the portion of the rift between the Dead Sea and Lake Kinneret.

Although the investigated portion of the rift does not coincide with the interpreted boundary of the crustal lithological transition, its vicinity supports the classic view of movements along a complicated transform fault system (Quennell 1958, 1965; Freund et al., 1970; Freund and Garfunkel 1976). This view is also supported by the detection of transverse and oblique fault patterns which divide the rift floor into a complicated block structure.

Acknowledgements. The author wishes to thank Professor R. Shagam and Dr. Z. Garfunkel for their critical reading and useful comments of this paper.

References

Allan, T.D.: Magnetic and gravity fields over the Red Sea. Philos. Trans. R. Soc. London, Ser. A: **267**, 153–180, 1970

Allan, T.D., Morelli, C.: The Red Sea. In: The Sea, E. Maxwell, ed.: **4** (part 2) 493–542, 1970

Amitai, U.: A gravity survey west of Lake Kinneret. Isr. Inst. Petrol. Res. Geophys., Rep. GR/120/60, 1962

Ben Avraham, Z., Hanel, R., Villinger, H.: Heat flow through the Dead Sea rift. Mar. Geol. **28**, 253–269, 1978

Bott, M.H.P.: The use of rapid digital computing methods for direct gravity interpretation of sedimentary basins. Geophys. J.R. Astron. Soc. **3**, 63–67, 1960

Bureau gravimetrique international: Anomalies de Bouguer Europe-Afrique. 3rd. edn. Paris 1971

Domzalski, W.: Aeromagnetic survey of Israel. Isr. Inst. Petrol. Res. Geophys. Rep. SMA/482/67, 1967

Eckstein, Y.: The measurement and interpretation of terrestrial heat flow in Israel (in Hebrew). Ph. D. thesis, Hebrew Univ., Jerusalem 1976

Folkman, Y : Aeromagnetic survey in the areas of the Golan, Shomeron, Judea and Gaza offshore. Isr. Inst. Petrol. Res. Geophys., Rep. SMA/760/70, 1971

Folkman, Y : Magnetic and gravity investigations of the crustal structure in Israel (in Hebrew). Ph. D. thesis, Tel Aviv Univ. 1976

Folkman, Y.. The minimal scanning density/feature size ratio in contour maps, a reply. Isr. J. Earth Sci. **26**, 41–43, 1978

Folkman, Y., Yuval, Z.: Aeromagnetic map of Israel, 1: 250,000. Isr. Inst. Petrol. Res. Geophys. Holon 1976

Freund, R., Garfunkel, Z.: Guidebook to excursion along the Dead Sea rift. Dept. of Geology, Hebrew Univ., Jerusalem 1976

Freund, R., Garfunkel, Z., Zak, I., Goldberg, M., Weissbrod, T., Derin, B.: The shear along the Dead Sea rift. Philos. Trans. R. Soc. London Ser. A: **267**, 107–130, 1970

Freund, R., Oppenheim, M.J., Schulman, N.: Direction of magnetization of some basalts in the Jordan valley and lower Galilee (Israel). Isr. J. Earth Sci. **19**, 200–227, 1965

Ginzburg, A.: Geophysical studies in the central and northern coastal plain and the western emeq. (in Hebrew). Ph. D. thesis, Hebrew Univ., Jerusalem 1960

Ginzburg, A.: Gravity map of northern Israel (note on the relative Bouguer anomalies). Isr. Inst. Petrol. Res. Geophys., Rep. GR/468/66, 1968

Ginzburg, A., Markis, J., Fuchs, K., Perathoner, B., Prodehl, C.: Detailed structure of the crust and upper mantle along the Jordan-Dead Sea rift. J. Geophys. Res. 1979 (in press)

Ginzburg, A., Markis, J., Fuchs, K., Prodehl, C., Kaminski, W., Amitai, U.: A seismic study of the crust and upper mantle of the Jordan-Dead Sea rift and their transition towards the Mediterranean. J. Geophys. Res. **84**, 1569–1582, 1979

Girdler, R.W.: Comparison of east African rift system and Permian oslo rift. In: Tectonics and Geophysics of Continental Rifts, I.B. Ramberg, E.R. Neumann, eds. Amsterdam: D. Reidel 1978

Gvirtzman, G.: The Saqiye Group (late Eocene to early Pleistocene) in the coastal plain and Hashephela regions, Israel. Bull. Geol. Surv. Isr. **51**, 1969

Horowitz, A.: Development of the Hula basin, Israel. Isr. J. Earth Sci. **22**, 107–139, 1973

Knopoff, L., Belshe, J.C.: Gravity observations of the Dead Sea rift. Univ. of California, Inst. Geophys. Publ. 476, 1967

Michaelson, H.: Geology of the Kinneret area. In: Lake Kinneret general background, T. Berman, ed! National Counc. Res. and Dev., Israel. **12–73**, 15–19, 1973

Mor, D.: The volcanism in the central Golan Heights. Tahal Ltd., Tel Aviv. Rep. 01/73/27, 1973

Neev, D.: The Geology of the Kinneret. In: The Kinneret, I. Gal, ed. The Kinneret Authority, Zemah (in Hebrew), pp. 15–26 1979

Neev, D., Hall, J.: Dead Sea geophysical survey, final report no. 2. Isr. Geol. Surv. Rep. MG/1/78, 1976

Perathoner, B., Fuchs, K., Prodehl, C.: Detailed interpretation of deep seismic sounding observation from the Jordan-Dead Sea rift and adjacent areas. Abst. Inter. symp. on Rift Zones of the Earth, the Dead Sea Rift, Jerusalem, Israel, 1979

Picard, L.: Structure and evolution of Palestine. Bull. Geol. Dept., Hebrew Univ., Jerusalem 1943

Picard, L.: Disharmonic faulting, a tectonic concept. Bull. Res. Counc. Israel. **3**, 132–135, 1953

Picard, L.: The geological evolution of the Quaternary in the central-northern Jordan graben. Israel. In: International studies of the Quaternary. Wright, H.E. Jr., Frey, D.G., eds. Geol. Soc. Am. **84**, 337–366, 1965

Picard, L., Golani, U.. Geological map of Israel, 1:250,000, northern sheet. Survey of Israel, Tel Aviv. 1965

Quennell, A.M.: The structural and geomorphic evolution of the Dead Sea rift. Q. J. Geol. Soc. London, CXIV. **1**, 1–24, 1958

Quennell, A.A.: Tectonics of the Dead Sea rift. Twentieth Int. Geol. Cong., Mexico, pp. 385–405, 1965

Schulman, N.: The geology of the central Jordan valley (in Hebrew). Ph. D. thesis, Hebrew Univ. Jerusalem 1962

Schulman, N.: The Jordan rift valley, In: Guidbook to field excursions. 10th. International Congress on Sedimentology, Jerusalem, pp. 57–94, 1978

Tiberghien, M.V.: Anomalies de Bouguer et structure profonde du Liban. C.R. Acad. Sci. Paris. T Ser. D.: **287**, 2401–2404, 1974

Vroman, A.: Geological map, 1:50,000, Series I Galilee. Isr. Geol. Surv. Jerusalem 1958

Yuval, Z.: Jordan valley gravity survey. Isr. Inst. Petrol. Res. Geophys. Rep. GR/427/66, 1966

Received June 5, 1979; Revised Version October 15, 1979

J. Geophys. 48, 40–46, 1980

Journal of Geophysics

Changes of Structure, Phase Composition and Electric Conductivity Under High-Temperature Oxidation of Titanomagnetites

V. Kropáček and M. Laštovičková

Geophysical Institute, Czechoslovak Academy of Sciences, Boční II/1401, 141 31 Praha 4, Czechoslovakia

Abstract. Temperature dependence of electrical conductivity in the temperature range from 200 to 900° C has been measured on a set of 17 magnetic fractions separated from Tertiary and Quaternary basaltic rocks of the Bohemian Massif. Under partial oxygen pressure of 0.1–1.0 Pa and temperatures over 600° C, conditions under which the sample can oxidize, and phase and structural changes occur in the specimens. Generally the electrical conductivity σ was found to decrease with increasing degree of oxidation z. A possible mechanism of electrical conductivity and the corresponding distribution of cations is proposed to explain the results. For $z<0.4$, the mechanism of valence transfer predominates, whereas for $z>0.4$ the mechanism of controlled valency is most important.

Key words: Titanomagnetic – Thermal treatment – Electric conductivity.

Introduction

The process of oxidation of titanomagnetites $Fe_{3-x}Ti_xO_4$, changes in the distribution of cations, and phase changes in the representation of individual spinel and other phases connected with this process, have been studied by a number of authors (Readman and O'Reilly 1970, 1972; Bleil 1971, 1975; O'Donovan and O'Reilly 1978). In most of these papers magnetic measurements were used to obtain an idea of the distribution of the cations into sublattices. However, direct use of magnetic measurements in restricted by the highest Curie temperature of titanomagnetite, about 500° C. Therefore, in this paper an attempt has been made to use another structurally sensitive physical parameter, electrical conductivity, to characterize the process of oxidation at temperatures up to 900° C.

Experimental Methods

Seventeen titanomagnetite specimens separated from Tertiary and Quaternary basaltic rocks of the Bohemian Massif (Kropáček 1974, 1977) were used. Specimens of the magnetic fractions were obtained by repeated magnetic separation in an A.C. field and by cleaning with the aid of heavy fluids. The resultant purity of the magnetic fractions was between 90 and 95%. These powdered fractions were used to prepare specimens for measuring electrical conductivity. The specimens were of cylindrical shape, 8 mm i.d. and 6 mm high. They were compressed under a pressure of 300 MPa in a fixture of Ti-bronze. A 5% admixture of talc was used as a plasticizer. The effect of the low-conductivity talc on the electrical conductivity σ and on the activation energy E_0 was checked in a series of measurements using specimens with different concentrations of talc (from 1.0 to 75%). These measurements (Laštovičková and Kropáček 1979) were used to determine corrections ($\Delta \log \sigma(200° C) = -0.25$ and $\Delta E_0 = -0.12$ eV) which are practically constant in the temperature interval 200°–900° C, for the content of talc 2–5%. Electrical contacts were of 'silver pain' paste which was not contaminated by the specimens. A two-electrode connection with a source of stabilized voltage and a continuous record of the electrical conductivity and temperature were used. Specimens were heated in a flow-through Ar-atmosphere with a partial oxygen pressure $P_{O_2} \sim 0.1$–1.0 Pa.

Structures and phase composition of the specimens were determined by means of a X-ray difractometer DRON 2.0 with Co radiation (Fe filter) and in some cases with a graphite or quartz monochromator. Calibration was carried out with the aid of NaCl p.a. Diffraction intensities were recorded over an angular range of $2\vartheta = 20°$–$80°$, for selected sections using a scanning interval of $2\vartheta = 0.05°$ and an integration time of 40–200 s. High temperature X-ray measurements were made in a UPVT-1500 diffractometric chamber in a vacuum of 10^{-5} Pa. Reflections were evaluated by means of a three-point parabola with correction for the Lorentz polarization factor, and lattice parameters were determined by the least-squares method (Azároff and Buerger 1958).

The degree of oxidation of titanomagnetites z is defined by the following relation (O'Reilly and Banerjee 1966):

$$Fe^{2+} + 1/2\,z\,O \rightarrow z\,Fe^{3+} + (1-z)\,Fe^{2+} + 1/2\,z\,O^{2-}.$$

First, z was determined by conventional chemical analysis from the relation

$$z = [2(1+x)(M+3L+2K) - 8M] \times [(1+x)(3M+6L+4K)]^{-1},$$

where x is the content of ulvöspinel in the titanomagnetite, $(1-x)\,Fe_3O_4 \times Fe_2TiO_4$ M is the content (mol%) of FeO, L the content of Fe_2O_3, K the content of TiO_2 (Ozima et al. 1974); and secondly by X-ray analysis using the intensities of reflection 111 (Readman and O'Reilly 1972) and from the intensities of reflections 110, 210, 211, 310, 321 (Gendler and Kropáček 1979).

0340-062X/80/0048/0040/$01.40

Table 1. Composition parameters of magnetic fractions and its changes during the thermal treatment; MT titanomagnetite, HI hemoilmenite, PB pseudobrookite, R + A rutile and anatase

Group	Specimen no.	Natural state											After heating up to 900° C					
		Fe^{2+}	Fe^{3+}	Ti^{4+}	□	x	z	a [nm]	MT	HI	PB	R + A	z	a [nm]	MT	HI	PB	R
I	201	1.38	0.85	0.68	0.09	0.68	0.15	0.8465	92	–	–	–	0.10	0.8470	80	10	–	–
	206	1.11	1.13	0.60	0.16	0.62	0.28	0.8447	85	10	–	–	0.25	0.8434	80	10	5	–
	227	1.04	1.38	0.43	0.13	0.44	0.25	0.8427	96	–	–	–	0.05	0.8459	80	6	–	10
	249	0.13	1.96	0.46	0.45	0.48	0.90	0.8376	92	–	–	–	0.40	0.8427	90	–	–	–
	256	0.97	1.90	0.09	0.04	0.09	0.10	0.8398	83	–	–	10	0.06	0.8423	85	–	–	5
	257	0.96	1.91	0.09	0.04	0.09	0.11	0.8433	93	–	–	–	0.08	0.8431	93	–	–	–
	259	1.38	0.88	0.65	0.09	0.65	0.15	0.8468	93	–	–	–	0.12	0.8472	90	5	–	–
	298	0.46	1.70	0.50	0.34	0.50	0.66	0.8373	92	–	–	–	0.34	0.8405	92	–	–	–
II	202	1.58	0.69	0.69	0.04	0.70	0.06	0.8467	94	–	–	–	0.70	0.8378	40	15	20	20
	218	0.93	1.35	0.52	0.20	–	–	0.8436 0.8425	90	–	–	–	1.00	0.8342	28	42	23	–
	220	0.15	2.40	0.12	0.33	0.14	0.85	0.8349	80	–	8	7	0.80	0.8377	45	20	20	5
	230	0.94	1.92	0.09	0.05	0.10	0.13	0.8398	85	5	3	–	0.30	0.8386	52	30	6	7
	235	0.20	1.88	0.49	0.43	0.50	0.85	0.8366	90	3	–	–	0.80	0.8377	55	20	12	5
	238	0.40	2.07	0.24	0.29	0.25	0.65	0.8336	90	5	–	–	0.66	0.8376	42	20	30	–
	263	1.22	1.10	0.60	0.08	0.60	0.15	0.8443	85	7	–	–	0.65	0.8357	72	15	–	5
	288	1.23	1.14	0.53	0.10	–	–	0.8445 0.8388	90	–	4	–	1.00	0.8345	50	15	23	5
	297	0.10	1.90	0.58	0.42	0.60	0.90	0.8348	80	12	2	–	0.93	0.8351	30	58	–	6

The number f of vacancies per formula unit of titanomagnetite was determined from known values of the parameter x and z as follows:

$$f=[3z(1+x)][8+z(1+x)]^{-1},$$

assuming the formula of titanomagnetite to be

$$Fe_b^{2+}Fe_c^{3+}Ti_d^{4+}\square_f O_4^{2-}.$$

Amounts of the individual phases (titanomagnetite, hemoilmenite, pseudobrookite, rutile, anatase) were determined by X-ray quantitative phase analysis with an accuracy of ±2.5%. For low contents (2–5%) the existence of these phases was determined only qualitatively. From conventional chemical analysis only FeO, Fe_2O_3, TiO_2 were determined.

The temperature dependence of specific heat c_p was measured using a DSM-2 differential scanning microcalorimeter in the temperature range of 20°–500° C. Differential thermic analysis (DTA) and thermogravimetric analysis (TG) were obtained with a Derivatograph instrument in air in the temperature range of 20°–1,000° C and a rate of 10° C/min.

Experimental Results

X-ray measurements indicated that titanomagnetite at various stages of oxidation constituted the largest part of the magnetic fractions beside other phases such as hemoilmenite, pseudobrookite, and rutile, as indicated in Table 1.

Composition of titanomagnetites $Fe_{3-x}Ti_xO_4$ varies over the range $x=0.08$–0.70, and the oxidation parameter z varies between 0.06–0.90. The lattice constant of the cubic phase in our samples was within the interval $a=0.8336$–0.8468 [nm] (Table 1).

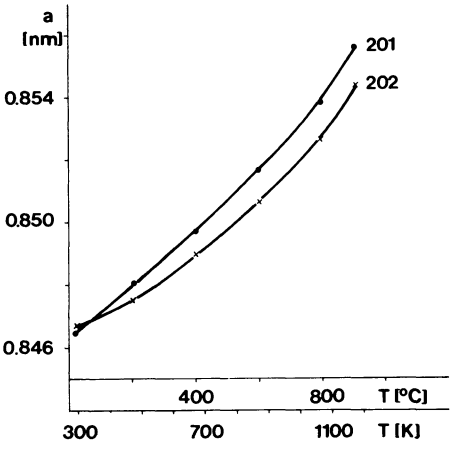


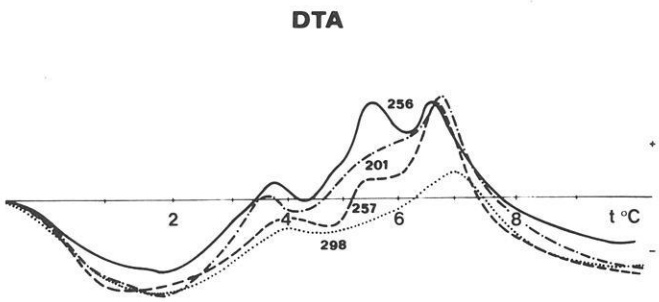
Fig. 1. Temperature dependence of lattice parameter a [nm] for the specimens 201 and 202 in vacuum

For two specimens (Nos. 201 and 202) the dependence $a(t)$ of the lattice constant a on temperature t [°C] was measured in the range 20°–900° C (Fig. 1). $a(t)$ can be described by the relations

$$a=0.8465+6.082\times10^{-6}t+4.283\times10^{-9}t^2 \quad \text{for No. 201}$$

$$a=0.8467+2.733\times10^{-6}t+6.035\times10^{-9}t^2 \quad \text{for No. 202.}$$

From high-temperature X-ray measurements we determined that no phase changes took place in specimen No. 201, whereas in specimen No. 202 3–5% hemoilmenite and rutile appeared above 700° C. The working vacuum was $\sim10^{-4}$ Pa. Curves from differential thermoanalysis (DTA) and thermogravimetry (TG) have been used for determining titanomagnetite oxidation states during heating in air (Figs. 2 and 3). From the DTA curves it is



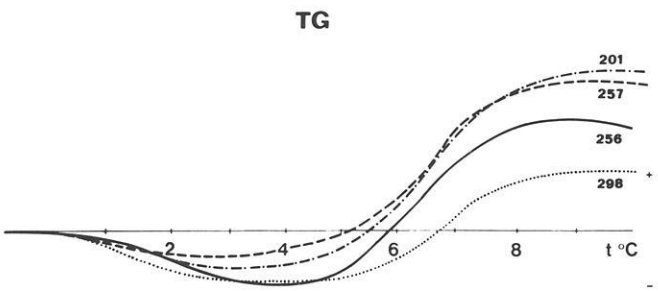
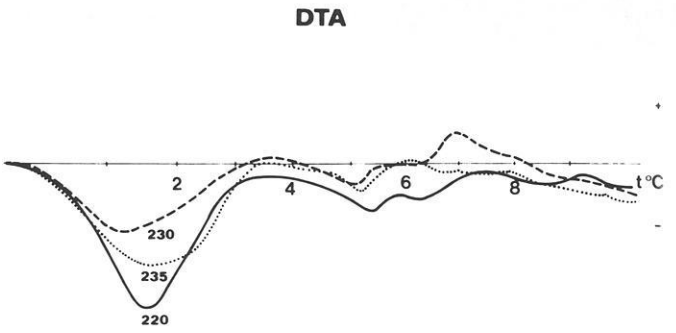


Fig. 2. Thermogravimetric (*TG*) and differential thermic curves for the specimens Nos. 201, 256, 257, 298. Heating in air, temperatures are in units of 100° C, *TG* and *DTA* curves in arbitrary units



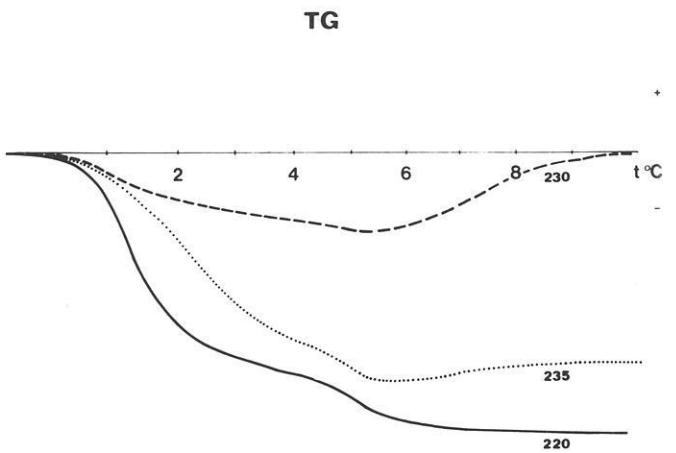


Fig. 3. *TG* and *DTA* curves for specimens Nos. 220, 230, and 235. Heating in air, temperatures are in units of 100° C

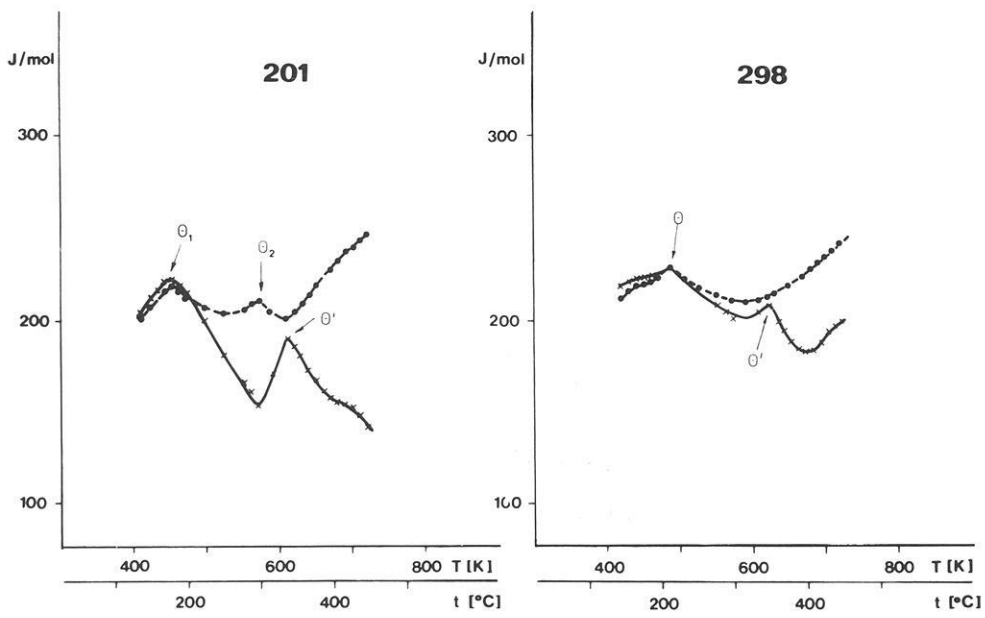


Fig. 4. Temperature dependence of specific heat $c_p = c_p(t)$. ●: Heating in nitrogen; ×: heating in air. θ, θ_1 are Curie temperature of the primary titanomagnetites and θ_2, θ' are Curie temperatures of secondary titanomagnetites produced during the heating

apparent that the first stage of oxidation produces exothermic effects in the temperature range 450° to 800° C with the greatest changes occuring between 600° and 700° C. At higher temperatures of 800° to 1,000° C the changes are endothermic. Similar results have been found from the thermogravimetric analysis (TG), where the oxidation processes also resulted in an increase of sample mass. The maximum rate of oxidation has again been found between 600° to 700° C, whereas above 800° C the mass of the samples has remained constant. X-ray phase analysis yielded the results that after heating the samples to 1,000° C in air the cubic phases had transformed into hemoilmenite, rutile, pseudobrookite and anatase.

The temperature dependence of the specific heat c_p (Fig. 4), determined in nitrogen and air, suggests formation of states with lower enthalpy $\Delta H = \int_{T_1}^{T_2} c_p \, dT$ during oxidation of titanomagnetites, that is formation of more stable phases at normal oxidizing condition. After heating to 900° C at a rate of 10° C/min and cooling in a Ar-atmosphere with $P_{O_2} \sim 0.1$–1.0 Pa (for the conductivity measurements) marked changes occurred in the structure and phase composition of the specimens. These changes can be divided into two categories and the specimens, depending on these changes, into two groups.

The first group of specimens (Nos. 201, 206, 227, 249, 256, 259, and 298) is characterized by a moderate decrease of the values of the oxidation parameter z (i.e., the stoichiometry of the specimens increases), and no other detectable phases are created. The average value of $\bar{z}$ is 0.32 in the natural state and $\bar{z} = 0.18$ after thermal treatment. Accordingly the value of the lattice constant $\bar{a}$ is found to have increased slightly, from 0.8423 to 0.8443 nm, after the thermal cycle. The content of any newly generated phase in this group is minimal and does not exceed 10 % on the whole (Table 1).

As regards the second group of specimens (Nos. 202, 218, 220, 230, 235, 238, 263, 288, and 297) very marked changes take place under thermal treatment. A whole series of other phases is generated (hemoilmenite, pseudobrookite, rutile, and anatase) and the content of the spinel phase decreases as much as to 28 % of the original value. The lattice constant also decreases from $\bar{a} = 0.8495$ nm in the initial state to $\bar{a} = 0.8365$ nm after thermal treatment. Correspondingly, the oxidation parameter z increases markedly from the original $\bar{z} = 0.48$ to about 0.83. Therefore, intensive oxidation associated with the generation of new phases takes place in the second group of specimens under thermal treatment with the following average proportion of phases: non-stoichiometric titanomagnetite 40–50 %, hemoilmenite 20–25 %, pseudobrookite 15–20 %, rutile and anatase 7–15 %.

Average values of the electrical conductivity at $t = 200°$ C of the specimens of the first group ($\log \sigma_I = -6.343\,[\Omega m]^{-1}$) are distincly higher the the average values of the second group ($\log \sigma_{II} = -7.942\,[\Omega m]^{-1}$); a smaller difference is preserved at a temperature of 600° C ($\log \sigma_I = -2.956\,[\Omega m]^{-1}$, $\log \sigma_{II} = -3.546\,[\Omega m]^{-1}$) (Table 2). Statistically significant relations, for which the correlation coefficient $r > r_{crit} = 0.71$ (for $N = 8$ and a 95 % probability) at a temperature of 200° C, are only observed in the first group:

$$\log \sigma(200°\,C) = -9.45 + 2.935\,b \quad (r = +0.74), \tag{1}$$

$$\log \sigma(200°\,C) = -5.02 - 7.960\,f \quad (r = -0.78), \tag{2}$$

$$\log \sigma(200°\,C) = -7.35 + 1.77\,b - 4.49\,f \quad (r = 0.84),$$

if we express titanomagnetite formula unit as

$$Fe_b^{2+} Fe_c^{3+} Ti_d^{4+} \square_f O_4^{2-}.$$

At a temperature of 600° C in the first group of the specimens attain statisticaly significant values of the coefficients of correlation for other relations:

$$\log \sigma(600°\,C) = -1.76 - 7.16\,f \quad (r = -0.80), \tag{3}$$

$$\log \sigma(600°\,C) = -5.45 + 2.50\,b \quad (r = +0.75), \tag{4}$$

$$\log \sigma(600°\,C) = -3.29 + 1.16\,b - 4.88\,f \quad (r = 0.78).$$

Table 2. Electrical conductivity σ at temperatures 200° C and 600° C, activation energy E_0 in the temperature intervals 200°–400° C, 600°–700° C and 800°–900° C

Group	Specimen no.	$\log \sigma$ (200° C) $[\Omega m]^{-1}$	$\log \sigma$ (600° C) $[\Omega m]^{-1}$	E_0 [eV] 200°–400° C	E_0 [eV] 600°–700° C	E_0 [eV] 800°–900° C
I	201	−5.717	−2.550	0.79	0.76	1.58
	206	−6.745	−2.594	1.08	1.51	1.70
	227	−7.092	−4.250	0.45	1.20	1.15
	249	−9.007	−5.264	0.58	1.25	1.32
	256	−6.430	−2.695	1.52	0.53	0.11
	257	−4.614	−1.303	0.75	1.30	–
	259	−4.309	−1.441	0.69	0.48	1.72
	298	−6.230	−3.523	0.98	1.39	0.85
II	202	−8.030	−3.625	1.41	2.01	1.88
	218	−7.114	−1.354	1.16	1.34	0.46
	220	−8.283	−5.143	0.48	0.86	2.44
	230	−8.827	−4.701	0.50	1.57	–
	235	−7.991	−3.251	0.88	0.71	3.00
	238	−8.551	−4.229	1.01	1.12	2.50
	263	−7.577	−4.126	0.55	1.16	1.91
	288	−6.538	−1.399	1.28	0.80	0.42
	297	−6.230	−3.523	0.98	1.39	0.95

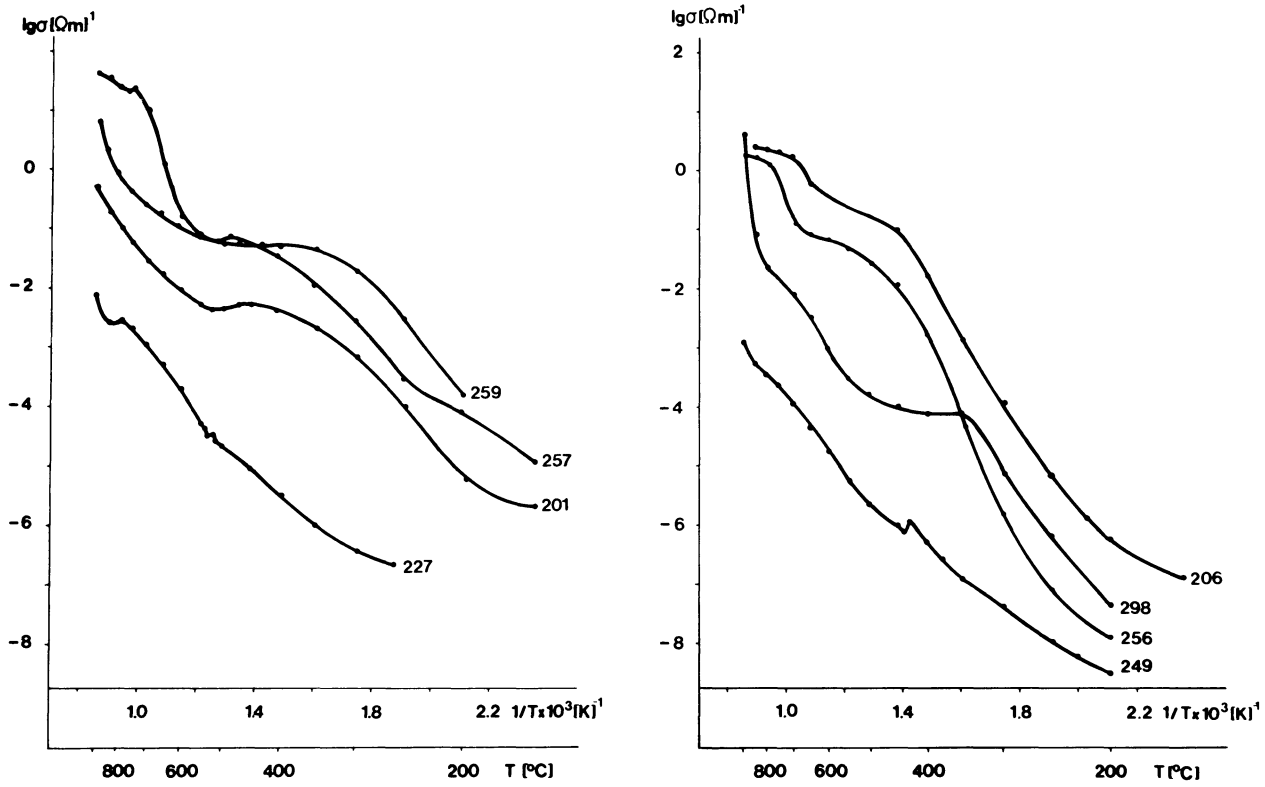


Fig. 5. Electrical conductivity $\sigma\,[\Omega m]^{-1}$ as a function of temperature for the first group of samples (see text)

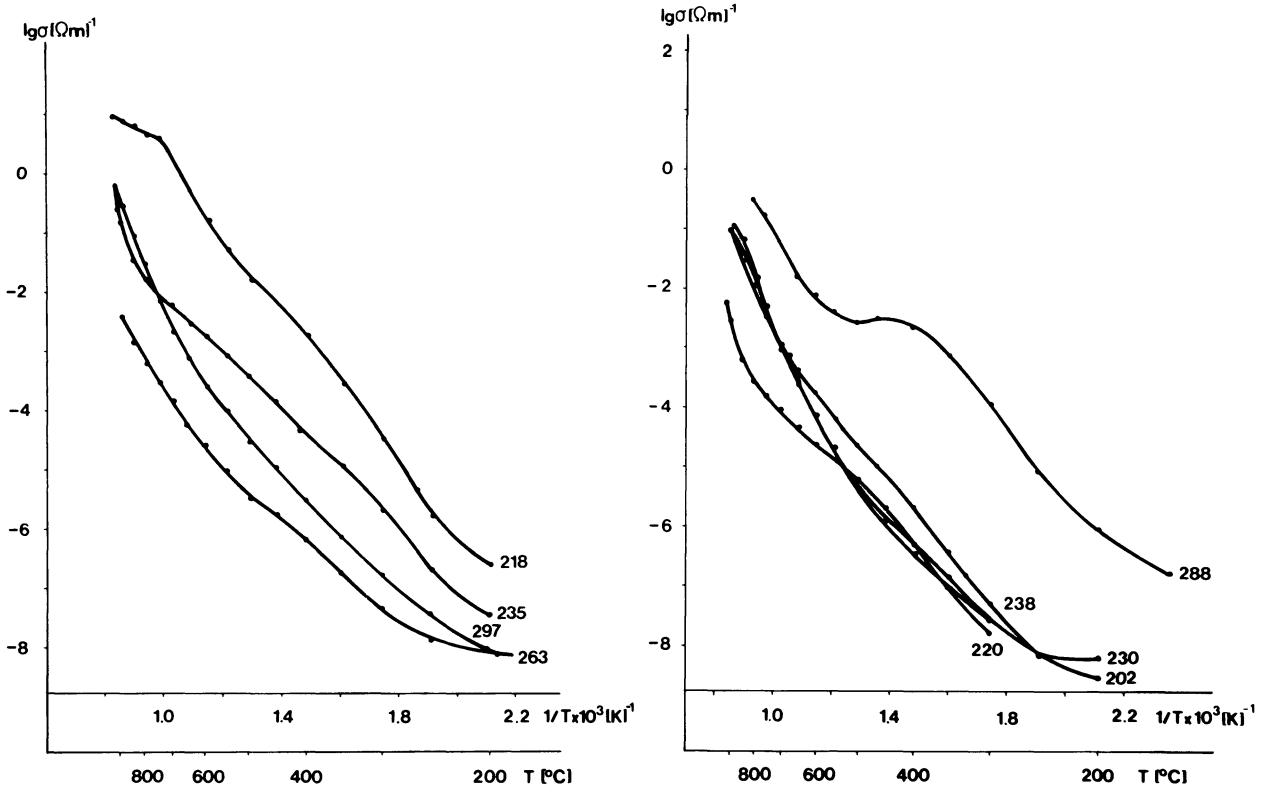


Fig. 6. Electrical conductivity $\sigma\,[\Omega m]^{-1}$ as a function of temperature for the second group of samples

For the second group there is no statistically significant correlation between $\log\sigma$ and any of the other parameters considered; higher values were only observed for $r_{\log\sigma(200),b} = 0.56$ and $r_{\log\sigma(200),f} = -0.40$.

Temperature dependence of the electrical conductivity for the specimens of the first group is shown in Fig. 5 and for those of the second group in Fig. 6. The comparison of Figs. 5 and 6 indicates that specimens of the first group display a relatively low increase of σ with temperature. The second group displays a high increase of σ with temperature, particularly above 600° C. To describe the dependence of the electrical conductivity on temperature we used

$$\sigma = \sigma_0 \exp(-E_0/kT),$$

where E_0 is the activation energy, k Boltzmann's constant and T the absolute temperature.

Values of E_0 for the individual temperature intervals derived from the above relation differ considerably between specimens of groups I and II (Table 2). For specimens of group I, with the exception of the temperature interval of 400°–600° C, activation energies between specimens are practically constant, and probably one type of electric conductivity predominates. For specimens of the second group, values of E_0 tend to be considerably higher above 600° C than those of the first group; we may, therefore, assume the existence of other types of electric conductivity.

Discussion

According to Verwey et al. (1947) the generation of electric current in ferrites can be envisaged as a process of an electron transition between ions according to the following general pattern:

$$M_{(1)}^{q+} + M_{(2)}^{r+} \rightleftarrows M_{(1)}^{(q+1)+} + M_{(2)}^{(r-1)+}. \qquad (5)$$

According to this model of valence transfer, if ions of the same element exist in a crystal at equivalent locations and differ in valency by unit, transition of electrons will be easy and the electrical conductivity will be high. Values of the activation energy E_0, on the other hand, will be low.

In titanomagnetites Fe^{2+} and Fe^{3+} ions may exist in octahedral (B), as well as tetrahedral (A) positions. Valence transfer will be posible only between ions in either (A–A) or (B–B) positions. These measurements (Tables 1 and 2) have shown that the value of $\log\sigma$(200° C) is small dependent on parameter x and more dependent for temperature $t = 600°$ C. Therefore, it is probable that part of the Fe^{2+} ions will be located in the tetrahedral (A) positions and the cations will be distributed as

$$Fe_k^{3+} Fe_l^{2+} [Fe_m^{3+} Fe_n^{2+} Ti_d^{4+} \square_f] O_4^{2-},$$

where $k+m=(2-2x+p)/L$, $l+n=(1+x)(1-z)/L$, $d=x/L$, $f=3p/8L$ and $p=z(1+x)$, $L=1+p/8$.

As indicated by Eqs. (2) and (3), $\log\sigma$ reflects the effect of the degree of oxidation z expressively, i.e., with increasing non-stoichiometry the electrical conductivity decreases markedly. This can be explained by a decrease of Fe^{2+} ions with increasing value of the number of vacancies f (or of z), and consequently a decrease of $Fe^{2+} - Fe^{3+}$ pairs, because the created vacancies do not contribute to the conductivity. The decrease of $\log\sigma$ with decreasing content of Fe^{2+} ions is determined by Eqs. (1) and (4). For specimens of the second group, the tendencies observed are very similar; however, values of the correlation coefficients are lower than the critical values for a 95 % probability.

It is clear that the content of Ti^{4+} ions which distinctly prefer octahedral positions will also play a role in the mechanism of electrical conductivity because one may assume rearrangements of the type

$$\beta Fe^{3+} + \gamma Fe^{3+} + \delta Ti^{3+} \rightleftarrows \delta Ti^{4+} + \gamma Fe^{2+} + \beta Fe^{3+}, \qquad (6)$$

$\beta \gg \gamma, \delta$, i.e., the mechanism of controlled valence with the aid of admixture ions (Krupička 1969). In this case current carriers will be electrons localized on Fe^{2+} cations and generated by ionization of the $Ti^{4+} - Fe^{2+}$ pairs, which represent donor centres. The number of carriers will be proportional to the number of Ti^{4+} cations, i.e., to the value x. The energy required to ionize the $Ti^{4+} - Fe^{2+}$ pair will be higher than that required to ionize the $Fe^{2+} - Fe^{3+}$ pair, but it will be markedly lower than for pure maghemite ($\gamma - Fe_2O_3$), where transitions between levels $Fe^{3+} + Fe^{3+} \rightleftarrows Fe^{2+} + Fe^{4+}$ mainly occur to preserve the condition of electron neutrality. With the titanomagnetites the value of $\log\sigma$ should be proportional to the concentrations of Ti^{4+} cations.

Also the values for selected specimens are: No. 220 ($x=0.14$, $f=0.33$; $\log\sigma(200° C) = -8.283\,[\Omega m]^{-1}$) and No. 298 ($x=0.50$, $f=0.34$; $\log\sigma(200° C) = -6.230\,[\Omega m]^{-1}$) which differ significantly only in the x-value and not in concentration of vacancies.

The observed smaller mean values of the activation energy E_0 for the first group as compared to the second group (Table 2) in all temperature intervals also can be explained on the basis of these mechanism of electrical conductivity.

Conclusion

1. In the process of high-temperature oxidation of titanomagnetites at temperatures of up to 900° C and partial oxygen pressure of 0.1–1.0 Pa two types of changes occur:

   (a) Increase of the degree of stoichiometry without the generation of other phases;

   (b) oxidation of titanomagnetites with the generation of new phases of hemoilmenite, pseudobrookite, rutile and anatase.

2. The electrical conductivity of titanomagnetites dependends on the degree of oxidation z. For $z<0.4$ and $x<0.7$ the mechanism of valence transfer ($Fe^{2+} + Fe^{3+} \rightleftarrows Fe^{3+} + Fe^{2+}$) with low values of the activation energy predominates. For $z>0.4$–0.5 and $x>0.7$ the mechanism of controlled valency [Eq. (6)]. The activation energy is higher.

References

Azároff, L.V., Buerger, M.J.: The powder method in X-ray crystalography. New York: McGraw-Hill 1958

Bleil, V.: Cation distribution in titanomagnetites. Z. Geophys. **37**, 305–319 (1971)

Bleil, V.: An experimental study of the titanomagnetite solid solution series. Pure Appl. Geophys. **114**, 165–175 (1975)

Gendler, T.S., Kropáček, V.: Change of structure, phase composition and magnetic properties with the oxidation of titanomagnetites. Studia Geophys. Geod. **23**, 42–54 (1979)

Kropáček, V.: Kinetics of oxidation processes in titanomagnetites, J. Geophys, **40**, 329–340 (1974)
Kropáček, V.: Magnetické vlastnosti vulkanických hornin. Čas. Mineral. Geol. **21**, 113–135 (1977)
Krupička, S.: Fyzika ferritů. Praha: Academia 1969
Laštovičková, M., Kropáček, V.: Electrical conductivity of the Fe–Ti minerals. Studia Geophs. Geod. **23**, 199–203 (1979)
O'Donovan, J.B., O'Reilly, W.: Cation distribution in synthetic titanomagnetites. Phys. Earth. Planet. Inter. **16**, 200–208 (1978)
O'Reilly, W., Banerjee, S.K.: Oxidation of titanomagnetites and self-reversal. Nature **211**, No. 5044, 26–28 (1966)
Ozima, M., Joshima, M., Kinoshita, H.: Magnetic properties of submarine basalts and the implication on the structure of the oceanic crust. J. Geomagn. Geoelectr. **26**, 335–347 (1974)
Readman, P.W., O'Reilly, W.: The synthesis and inversions of non-stoichiometric titanomagnetites. Phys. Earth Planet. Inter. **4**, 121–128 (1970)
Readman, P.W., O'Reilly, W.: Magnetic properties of oxidized (cation-deficient) titanomagnetites (Fe, Ti, $\square)_3O_4$. J. Geomagn. Geoelectr. **24**, 64–90 (1972)
Verwey, E.J., Haayman, P.W., Romeijn, F.C.: Physical properties and cation arrangement of oxides with spinel structures. J. Chem. Phys. **15**, 181–187 (1947)

Received October 23, 1978; Revised Version November 23, 1979

J Geophys. 48, 47–53, 1980

Journal of Geophysics

Maximum Velocity Amplitudes of *P* and *S* Waves From Filtered Broadband Records

M.V.D. Sitaram[1], A. Plešinger[2], and J. Vaněk[2]

[1] National Geophysical Research Institute, Hyderabad, 500007, India
[2] Geophysical Institute, Czechoslovak Academy of Sciences, 14131 Praha 4, Czechoslovakia

Abstract. The maximum velocity amplitudes (A_{vmax}) of *P* and *S* waves in different period bands are studied using filtered broadband records from the KHC Seismic Station. Two sequences of shallow earthquakes from the Tadzhik-Sinkiang border and from North-eastern China with magnitudes m_b between 5.0 and 6.3 are investigated. Amplitude-periodband (APB) diagrams are constructed for each event. The largest values of A_{vmax} are found in the period band from 2.2 to 10 s for *P* waves and from 3.4 to 23 s for *S* waves. A shift of the maximum in the APB diagrams towards longer periods for earthquakes with $m_b > 6.0$ is observed. For events with $m_b < 6.0$ the position of the maximum virtually does not depend on the magnitude. This phenomenon seems to be related with the process of wave generation in the focal region.

Key words: Seismology – Earthquake magnitude – Maximum velocity amplitudes – Broadband records.

Introduction

A more profound knowledge of the physical principles of the earthquake magnitude seems to be one of the substantial conditions for an improvement of routine magnitude determinations. Especially, a comprehensive investigation into the properties of the observational basis of the magnitude, the maximum ratio of amplitude and period $(A/T)_{max}$ for the individual wave types, is very desirable. In this context the behaviour of the quantity $(A/T)_{max}$ in different period ranges appears to be essential. Several papers reporting on results of investigations in this direction have been published (Korchagina and Moskvina 1972; Zapolskij et al. 1974; Aranovich et al. 1977; Roglinov et al. 1978).

The aim of the present study is to analyze the maximum velocity amplitudes of *P* and *S* waves in the period range from 1 to 170 s on the basis of a sequence of filtered broadband records for two series of shallow earthquakes.

1 Observational Material and Method of Data Processing

For the present study two sequences of shallow earthquakes from the Tadzhik-Sinkiang border (TSB) and from North-eastern China (NEC) were used. The parameters of the individual earthquakes are listed in Table 1. The respective observational material was obtained from the magnetic tape library of the FBV Broadband Seismograph System operating at the KHC Seismic Station (Kašperské Hory in South Bohemia, $\varphi = 49°07.8'$N, $\lambda = 13°34.8'$E). The variation of the epicentral co-ordinates is so small that the wave propagation paths to KHC are nearly the same for all events in each series. The epicentral distances for the TSB and NEC events are, respectively, about 43° and 69.5°

The FBV system has a flat-velocity response in the period range $0.3 < T < 300$ s. Each component (Z, NS, EW) is recorded bi-level in an overall dynamic range of 80 dbs on 12-track FM magnetic tape. For the processing of the tapes the DPS Data Processing System containing proper analogue and hybrid devices is used in the Geophysical Institute at Prague. Both systems are described in a previous paper by Plešinger and Horálek (1976).

To obtain the data in the form needed for the present study the FBV records were processed in the following way. The vertical and the horizontal components of each event were passed through a set of twelve one-octave 4th-degree Butterworth filters. The passbands of the individual filters and their centre periods are given in Table 2. The centre periods cover the range from 1.5 to 128 s in half-octave steps. Simultaneously, a procedure of particle motion analysis that discriminates *P*, *S*, and *LR* waves was performed. The outputs of the individual filters and the output of the particle motion analyzer were recorded by a linear multichannel high-speed heat-pen recorder (see Fig. 1). On each of the filtered records three quantities of the *P* and *S* wave groups were measured for every component: the maximum amplitude from trough to peak or vice versa, the period corresponding to the maximum amplitude, and the time difference between the first onset of the wave group and the occurrence of the maximum amplitude. In the majority of cases the signal-to-noise ratio was high enough to enable an undistorted evaluation. It became necessary, however, to employ also few wave groups with amplitudes comparable with those of the noise in the respective period band in order to find out the general trends and the maxima of the curves investigated in Sect. 2.

Since the FBV system has a broadband flat-velocity response the maximum amplitudes determined in the way outlined above, converted into ground velocity and divided by 2π, directly give the value of $(A/T)_{max}$ which is the basic parameter for magnitude estimations.

2. Amplitude-Periodband Diagrams

For the graphical representation of the dependence of the maximum velocity amplitudes on the filter passbands we use the term 'amplitude-periodband diagrams' in this paper and denote them briefly APB diagrams. These diagrams are analogous to the so-called 'velocity spectra' introduced by Bune et al. (1973) and Za-

0340-062X/80/0048/0047/$01.40

Table 1. Parameters of earthquakes

Event	Date		Origin time	Latitude	Longitude	Depth (km)	m_b	Source
TSB 1	1974 August	11	01:13:55	39.34N	73.76E	7	6.2	ISC
2		11	05:12:35.1	39.33N	73.75E	48	5.3	
3		11	05:23:57.3	39.32N	73.75E	72	5.4	
4		11	07:02:07	39.34N	73.80E	18	5.1	
5		11	20:05:30.9	39.44N	73.67E	41	5.7	
6		11	21:21:37.1	39.46N	73.62E	26	5.8	
7		27	12:56:01.0	39.52N	73.82E	19	5.7	
8		27	17:33:58.6	39.34N	73.86E	37	5.1	
9	1974 September	03	19:41:21.1	39.42N	73.74E	43	5.3	
NEC 1	1976 July	27	19:42:54.6	39.57N	117.98E	22.8	6.3	NEIS
2		27	23:17:31.4	39.36N	117.82E	31.4	5.4	
3		28	00:58:46.9	39.41N	117.78E	33.0	5.0	
4		28	10:45:35.2	39.66N	118.40E	26.4	6.3	
5		28	15:35:55.3	39.85N	118.66E	13.0	5.3	
6		29	01:01:03.2	39.92N	118.88E	35.4	5.1	
7		30	21:23:15.0	39.82N	118.33E	33.0	5.4	
8	1976 August	31	03:25:27.8	39.80N	118.86E	31.8	5.3	
9	1976 November	15	13:53:00.6	39.44N	117.69E	15.2	6.0	

Table 2. System of filters

Number of filter	Passband (seconds)	Centre period (seconds)
I	1.0– 2.0	1.5
II	1.5– 3.0	2.2
III	2.2– 4.5	3.4
IV	3.4– 6.8	5.1
V	5.0– 10.0	7.5
VI	7.5– 11.0	11.0
VII	11.5– 23.0	17.0
VIII	17.0– 34.0	25.0
IX	25.0– 50.0	35.0
X	37.0– 75.0	56.0
XI	56.0–112.0	84.0
XII	85.0–170.0	128.0

polskij et al. (1974) on the basis of observations with Frequency Selecting Seismic Stations (ČISS). It must be noticed, however, that APB diagrams cannot be considered to be identical with velocity amplitude spectra because they express the dependence of two quite different quantities (maximum amplitude instead of the root of spectral density, filter passband instead of frequency). In general, no prominent similarity must exist between APB diagrams and amplitude spectra of *P* and *S* waves for teleseismic events. A special study of the relation between APB diagrams and amplitude spectra is underway and the results will be published in a separate paper.

In the present study, the APB diagrams were constructed for the vertical and horizontal components (*Z*, *NS*, and *EW*) of *P* and *S* waves for both the TSB and NEC events. To illustrate the reliability of the results two typical APB diagrams of the seismic noise level at KHC are shown in Fig. 2.

2.1. APB Diagrams for P Waves

The resulting APB diagrams for *P* waves are shown in Fig. 3, maximum velocity amplitudes in microns per second being plotted in the logarithmic scale. Number of event and the corresponding magnitude m_b are indicated for every curve. The solid circles indicate values of A_{vmax} which were safely above the noise level. Values of A_{vmax} evaluated from readings with comparable signal and noise amplitudes are denoted by open circles. Some of these cases were included in order to find out the trends of the APB curves, especially in their long-period flanks.

The general shape of the APB diagrams for all three components of *P* waves obviously is similar for both the TSB and NEC regions. The diagrams usually have a single pronounced maximum in the range of periods represented by filters III, IV, and V. An exception is the NS component for event No. 1 from the TSB region. For events with larger magnitudes the maxima tend to shift towards longer periods. The levels of the curves correspond to the order of magnitude m_b with the exception of NEC event No. 7 for which $m_b = 5.4$ given by NEIS seems to be too large in comparison with our observations. The anomalous level of the curve for the NS component of NEC event No. 5 is probably due to the disturbing effect of noise in the period band of filters III and IV.

The decrease of maximum velocity amplitudes for passbands in the short-period range is more expresssive for the *Z* and *EW* components of events from the NEC region than for those from the TSB region.

2.2. APB Diagrams for S Waves

The APB diagrams for *S* waves are shown in Fig. 4. In general, they have a similar character as those for *P* waves but the maxima of curves occur at longer periods, i.e., in the range of filters IV to VII for the TSB events and V to IX for the NEC events.

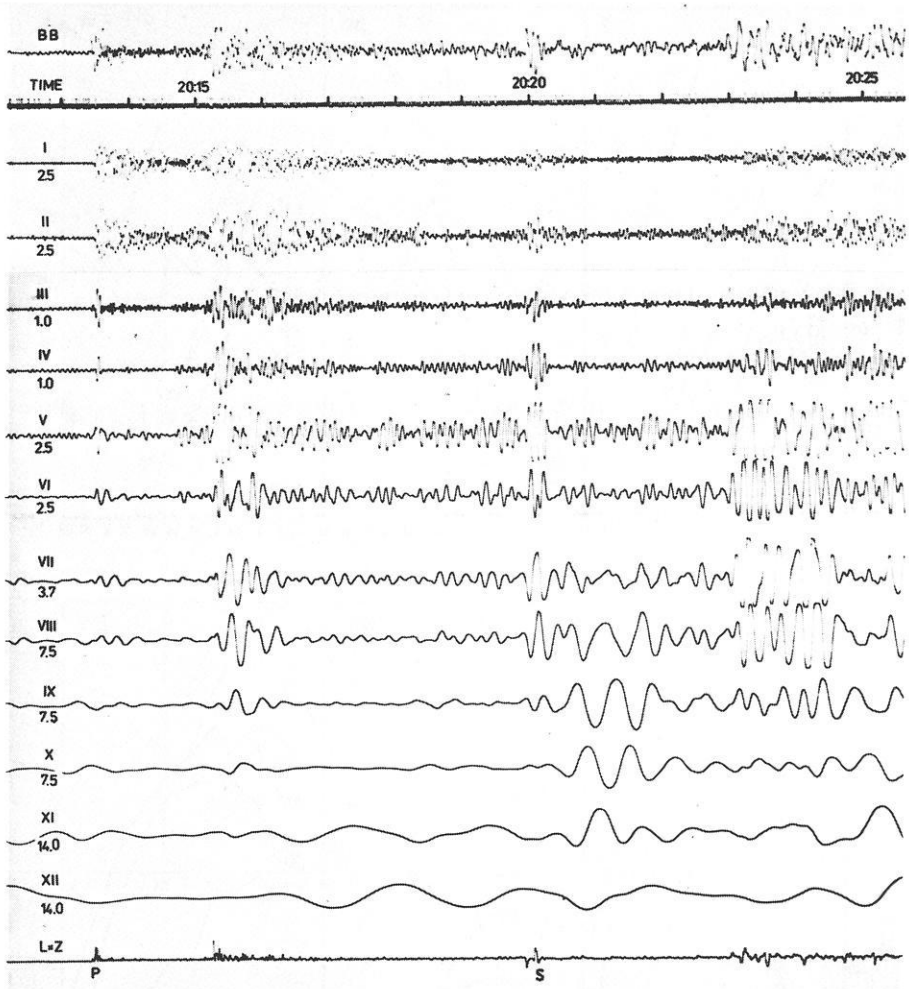


Fig. 1. Example of filtered broadband records (event TSB 5, component EW). Trace 1: broadband signal (0.3 to 300 s), trace 2: coded time, traces 3 to 14: outputs of filters I to XII (the arabic numbers are the relative amplification coefficients of the individual channels), trace 15: result of polarization analysis (product of longitudinal and vertical broadband components)

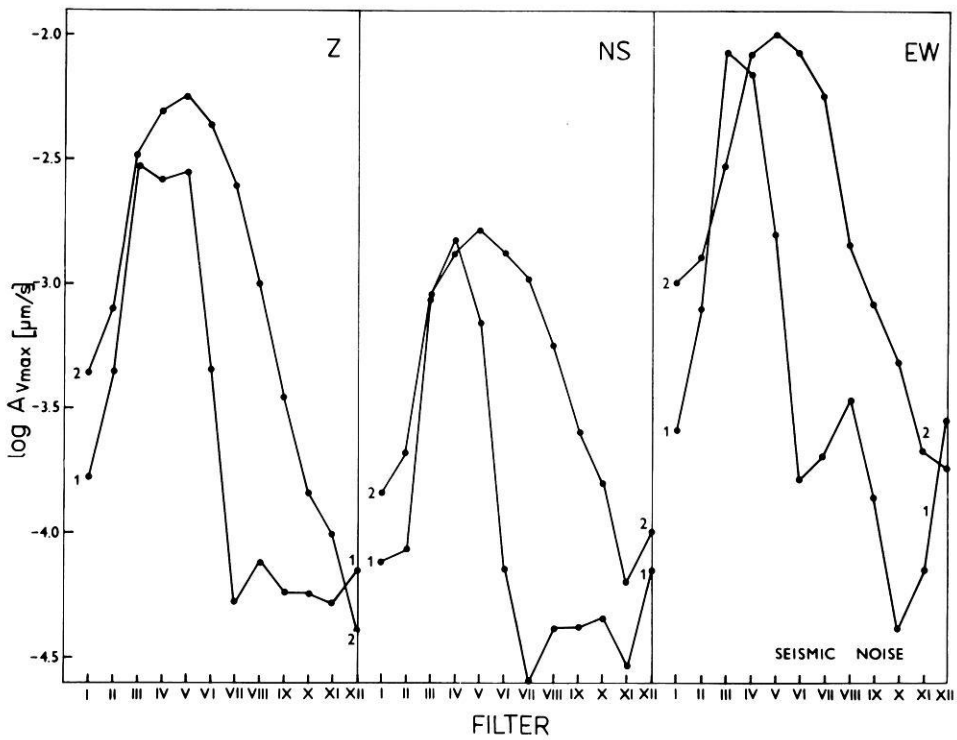


Fig. 2. APB diagrams of the seismic noise preceding events TSB 4 (1) and NEC 3 (2). A_{vmax} are the maximum values observed in a time interval of 3 min before the onset of the *P* wave

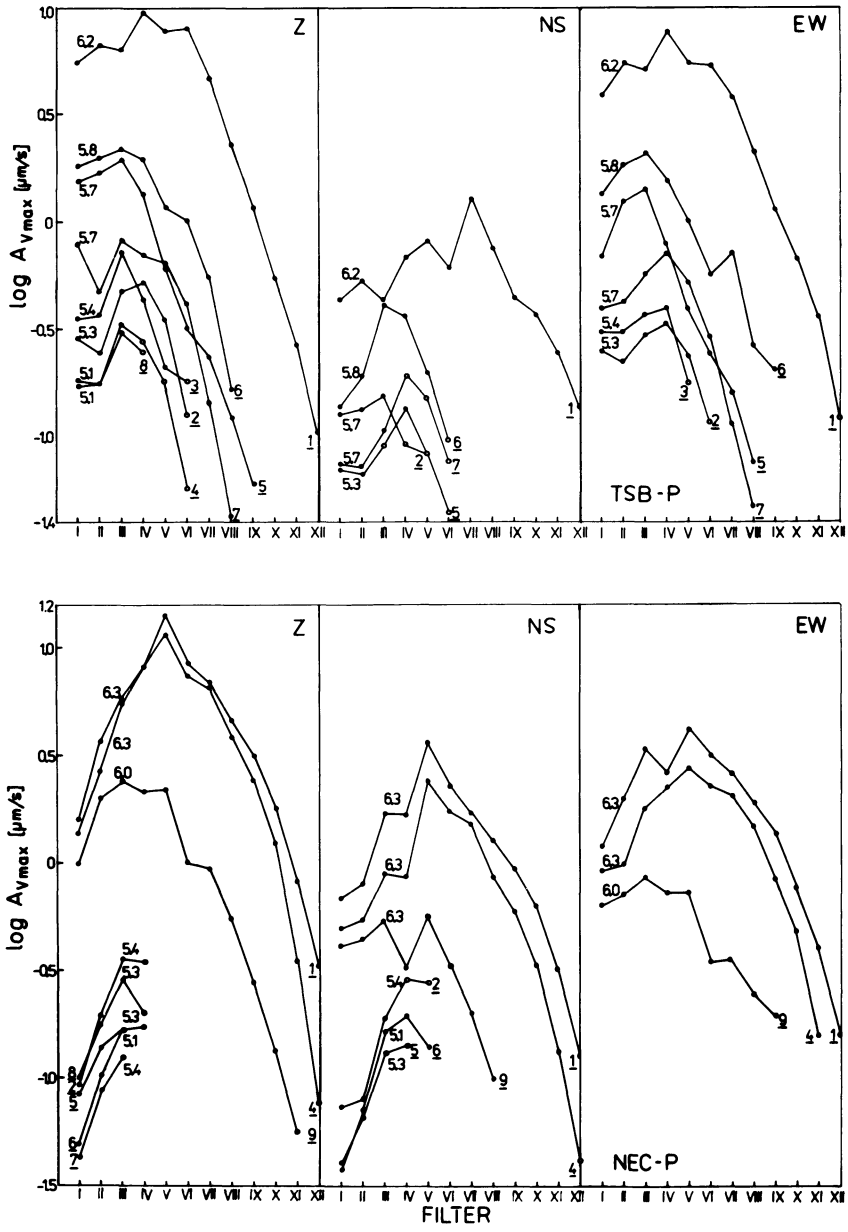


Fig. 3. Amplitude-periodband diagrams of *P* waves from the TSB (Tadzhik-Sinkiang border) and NEC (North-eastern China) events. A_{vmax} values above the noise level are denoted by *solid points*, those comparable with the noise level by *circles*. The underlined numbers correspond to the event numbers in Table 1. The decimal numbers are the m_b magnitudes of the individual events (the correct m_b for diagram 9 in figure NEC-P, component NS, is 6.0)

In comparison with *P* waves the maxima of the diagrams are not so well pronounced for most of the events. The A_{vmax} values are comparable in more passbands so that the maximum cannot be unambiguously attached to an individual filter (see, e.g., TSB event No. 6 and NEC events Nos. 1 and 4 in the NS component or TSB event No. 1 in the *EW* component).

The shape of the APB diagrams is also influenced by the internal structure of the wave group investigated. For instance, a detailed polarization analysis of the *S* wave group of TSB event No. 1 has revealed that this group consisted of a pure SH pulse with a half-period of approximately 12 s and a slightly later wave group of rather SV type with a predominant period of about 36 s. This means that A_{vmax} in different passbands can correspond to different wave types.

A tendency of shifting of the maxima of the APB curves towards longer periods for events with larger magnitudes is again observed also for *S* waves.

2.3. Time of Occurrence of Maximum Velocity Amplitude

In our study the problem of the time of occurrence τ of maximum velocity amplitude in the *P* and *S* wave groups for different filters has been also investigated. Knowledge of τ values to be expected in different passbands is important for the routine determination of earthquake magnitudes on the basis of maximum amplitudes. The quantity τ was measured from the onsets of the *P* and *S* wave group. The values for *P* waves were determined from the vertical components, those for *S* waves from the *EW* components of the filtered broadband records. The behaviour of τ for a set of selected events is shown in Fig. 5.

It is evident that τ increases with increasing centre periods of the filters as well as with increasing magnitude of the events. Its variation appears to be negligible up to the period range of filter VII. In the passband from 1 s up to approximately 20 s the time of occurrence of the maximum velocity amplitude usually

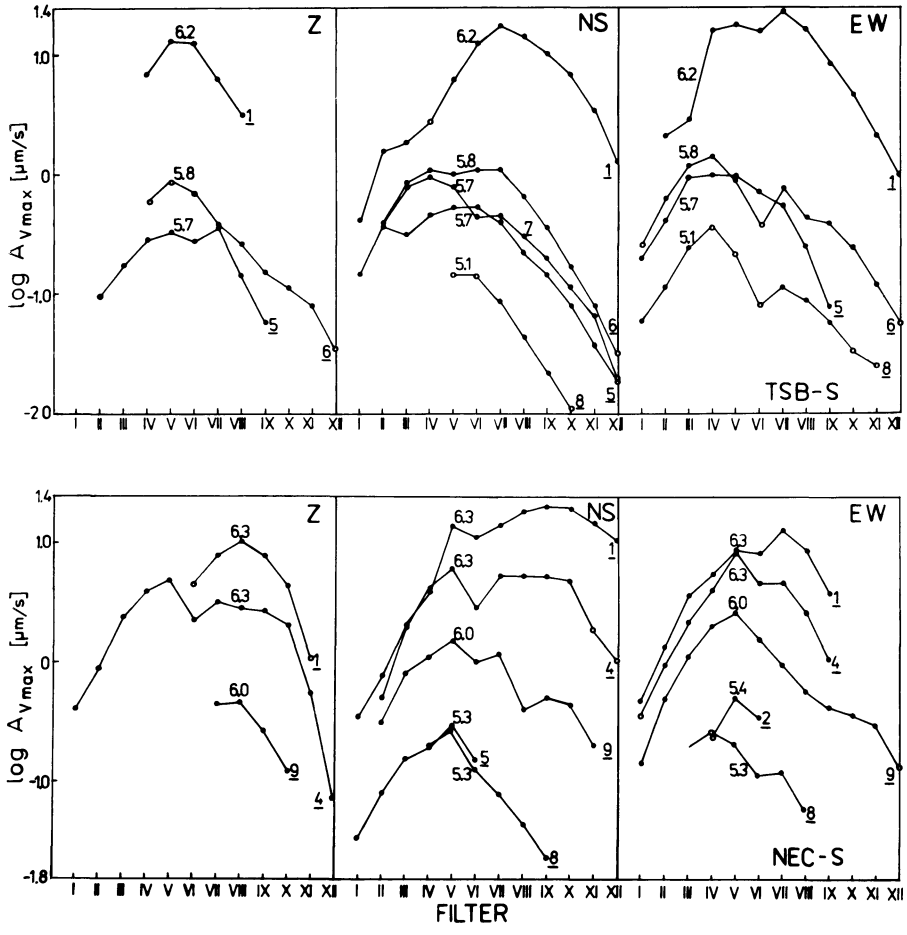


Fig. 4. Amplitude-periodband diagrams of S waves from the TSB and NEC events. The individual values and curves are denoted in the same manner as in Fig. 1

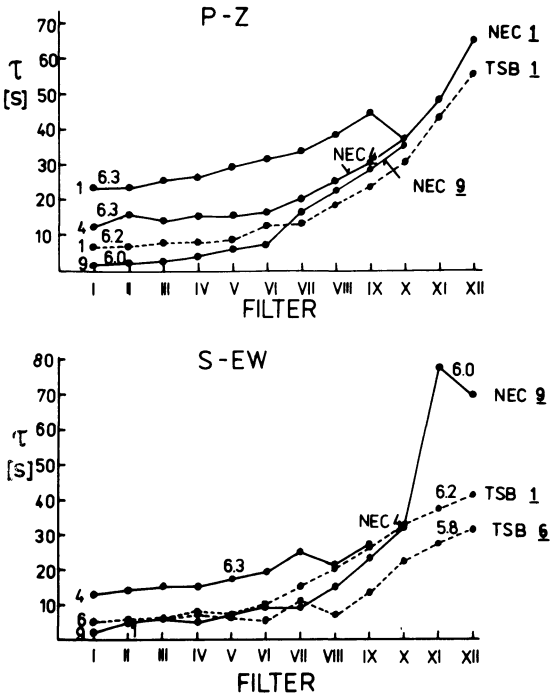


Fig. 5. Time of occurrence τ of maximum velocity amplitude vs. period-passband (filter number) for a set of selected events

does not exceed 30 s for both P and S waves. The variation of τ in the long-period range is considerable, reaching up to 80 s for S as well as P waves.

3. Discussion

A comparison of the APB diagrams in Figs. 3 and 4 indicates that there exists a stable relation between the three components of P and S waves for any of the events investigated. It should, therefore, be sufficient to select only one component for investigations of this kind. For P waves the vertical component and for S waves one of the horizontal components appear to be the most appropriate.

The shapes of the APB diagrams for P and S waves have a very similar character for all the events. However, for S waves the maxima of the curves are shifted towards longer periods. It was experienced that for events having magnitudes $m_b < 5.0$ the APB diagrams for both P and S waves can be constructed only in exceptional cases, the noise background, especially long-period microseisms, being the limiting factor.

For both the Tadzhik-Sinkiang and North-eastern China sequences the maxima of the APB curves vary within the period range of filters III to IX, i.e., the largest value of A_{vmax} occurs in the intermediate period band. Furthermore, the general shape of the APB diagrams is practically the same for both regions without any exception.

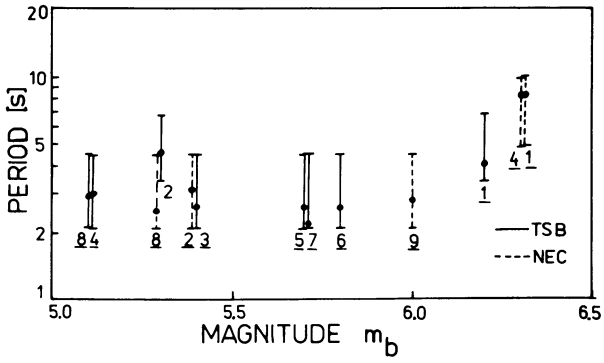


Fig. 6. Dependence of the passband corresponding to the largest A_{vmax} value on m_b magnitude for the vertical component of P waves. The periods T_v corresponding to A_{vmax} (*solid circles*) and the numbers of the TSB and NEC events are indicated

Table 3. Short and medium period HMS magnitudes for the vertical component of P waves

Event	M_{PVs}	M_{PV}	Event	M_{PVs}	M_{PV}
TSB 1	6.85	6.9	NEC 1	6.5	7.25
2	5.55	5.65	2	5.25	5.65
3	5.65	5.8	4	6.4	7.15
4	5.35	5.45	5	5.2	–
5	6.3	6.25	6	4.95	5.35
6	6.35	6.3	7	4.9	–
7	6.0	5.85	8	5.3	5.55
8	5.35	5.45	9	6.25	6.45
9	5.7	–			

The results pertaining to the vertical component of P waves, mentioned in the last paragraph, are in general agreement with those obtained by Bune et al. (1973) and Zapolskij et al. (1974) on the basis of observational material from Frequency Selecting Seismic Stations. The APB diagrams for the vertical component of P waves are also similar to the majority of diagrams constructed in the same way for a sequence of aftershocks from the Hokkaido region (Roglinov et al. 1978). In spite of the fact that the shapes of the APB diagrams for some of the Hokkaido events were found to be exceptionally quite different, it seems that the type of diagrams observed in the present study represents the process of seismic wave generation that prevails for shallow earthquakes.

An interesting phenomenon is the shift of the maxima of the APB diagrams towards longer periods with increasing magnitude of the events (see Figs. 3 and 4). Figure 6 shows the dependence of the passband, in which the largest A_{vmax} of the vertical component of P waves is found, on the magnitude m_b. It appears that the maxima of A_{vmax} shift towards longer periods for magnitudes $m_b > 6.0$ whereas for $m_b < 6.0$ the occurrence of the largest A_{vmax} in the passband from 2.2 to 4.5 s is independent of magnitude and region. The shift to filter IV for TSB event No. 2 is an exception the probable cause of which was constructive signal-noise interference. The same tendency of shifting of maxima of the APB diagrams to longer periods for events with larger magnitudes can be observed also for S waves but the phenomenon is not so expressive as for P waves due to the flatness of the maxima.

Using calibrating functions of the Eurasian homogeneous magnitude system HMS (Vaněk et al. 1978; Christoskov et al. 1978) the magnitudes M_{PVs} and M_{PV} for all the events were re-estimated on the basis of short-period and medium-period observations of the vertical component of P waves. The short-period observations are approximately simulated by A_{vmax} in the output of filter I and the medium-period observations by A_{vmax} from those filters for which the maxima of the APB diagrams are observed. The HMS magnitudes, rounded to 0.05, are given in Table 3 without applying the station correction of KHC. It can be seen that for TSB events the agreement between M_{PVs} and M_{PV} is fairly good, the maximum differences being smaller than 0.15. On the contrary, for NEC events the differences between M_{PVs} and M_{PV} are up to 0.75 of the magnitude unit. These discrepancies can be explained by the more intense decrease of A_{vmax} values in the short-period range from the maxima of the APB curves for NEC events (see Fig. 3). The differences between magnitudes m_b and the HMS magnitudes M_{PVs}, M_{PV} are caused by the different magnitude level for the Eurasian HMS in comparison with the ISC system of magnitudes (for details see Vaněk et al. 1978; Christoskov et al. 1978).

The study of times of occurrence τ of maximum velocity amplitudes reveals that in the short-period and intermediate-period range up to the band from 11.5 to 23 s the quantitiy τ usually is below 30 s for both P and S waves. This observation supports the recommendation accepted by the IASPEI Commission on Seismological Practice (1973) at the XVth General Assembly of the I.U.G.G. in Moscow. According to this recommendation measurements of $(A/T)_{max}$ for the determination of body wave magnitudes should be carried out in the range of 25 s from the wave onset. However, for the long-period range an increase of τ up to 70–80 s may be considered.

Conclusions

From the study of amplitude-periodband diagrams of P and S waves obtained on the basis of filtered broadband records for two sequences of shallow earthquakes from the Tadzhik-Sinkiang border and from North-eastern China the following conclusions can be inferred.

The general shape of the APB diagrams for the Z, NS, and EW components of P and S waves is similar for both regions. The diagrams are characterized by a pronounced maximum in the range of intermediate periods. The largest maximum velocity amplitudes for the regions investigated are observed in the period bands from 2.2 to 10 s for P waves and from 3.4 to 23 s for S waves. Because earthquake magnitude estimations are based on the principle of the largest amplitude of the wave group, observations in these period-passbands should be used for magnitude determinations from P and S waves. Exceptions constitute earthquakes with an anomalous behaviour of APB diagrams (e.g., Roglinov et al. 1978). This conclusion applies for shallow events in a limited magnitude range ($5.0 < m_b < 6.5$) and its general validity must be verified by further investigations.

A shift of the maximum of APB diagrams towards longer periods, observed for earthquakes with $m_b > 6.0$, and its independence on magnitude for $m_b < 6.0$, seem to be related with the process of wave generation in the focal region.

It is observed that the time of occurrence τ of maximum velocity amplitudes usually is shorter than 30 s for the short-period and intermediate-period band, and up to 80 s for the long-period band.

Acknowledgements. The authors wish to express their thanks to J. Horálek for processing the observational material for this study on the FBV Data Processing System. One of the authors (M.V.D.S.) thanks the Czechoslovak Academy of Sciences for providing the necessary facilities in the Geophysical Institute at Prague in order to carry out the present study under the scientific exchange programme between the Czechoslovak Academy of Sciences and the Council of Scientific and Industrial Research of India.

References

Aranovich, Z.I., Korchagina, O.A., Horálek, J., Plešinger, A.: On the influence of seismograph responses on the estimation of earthquake magnitudes (in Russian). Travaux Inst. Géophys. Acad. Tchécosl. Sci. No. 486. Geofys. Sb. 1977 (in press)

Bune, V.I., Vvedenskaja, N.A., Gorbunova, I.V., Zapolskij, K.K., Kondorskaja, N.V., Fedorova, I.V.: To the problem of earthquake magnitude determination. Pure Appl. Geophys. **103**, 350–361, 1973

Christoskov, L., Kondorskaja, N.V., Vaněk, J.: Homogeneous magnitude system of the Eurasian continent. Tectonophysics, **49**, 131–138, 1978

Korchagina, O.A., Moskvina, A.G.: Comparative study of input signal distortions in standard class seismographs (in Russian). Izv. AN SSSR, Fiz. Zem. No. 8, 75–85, 1972

Plešinger, A., Horálek, J.: The seismic broadband recording and data processing system FBV/DPS and its seismological applications. J. Geophys. **42**, 201–217, 1976

Recommendations on Magnitude Determination in Practice of Seismic Observations by the IASPEI Commission on Seismological Practice, XV General Assembly of IUGG. Pure Appl. Geophys. **103**, 433–434, 1973

Roglinov, A., Horálek, J., Plešinger, A., Vaněk, J.: Magnitude scale structure from filtered broadband *P* wave records. Pure Appl. Geophys. **117**, 816–826, 1978

Vaněk, J., Kondorskaja, N.V., Christoskov, L.: Earthquake Magnitude in Seismological Practice, vol. 1, PV and PVs Waves (in Russian with extensive English summary). Sofia: 1978 Publ. House Bulgarian Acad. Sci. (in press)

Zapolskij, K.K., Nersesov, I.L., Rautian, T.G., Khalturin, V.I.: Physical principles of earthquake magnitude classification (in Russian). In: Magnituda i energetičeskaja klassifikacija zemletrjasenij, 79–131, Moskva 1974

Received December 12, 1977; Revised Version June 5, 1979

J. Geophys. 48, 54–55, 1980

Journal of Geophysics

Letter to the Editor

Comment on: On the Sources of the 12-Month Wave in the *an* and *as* Geomagnetic Activity Indices

by P.N. Mayaud
J. Geophys. **46**, 261–271, 1979

D. Damaske[1] and J. Meyer[2]

[1] Institut für Geophysikalische Wissenschaften, Freie Universität Berlin, Rheinbabenallee 49, D-1000 Berlin 33
[2] Institut für Geophysik, Universität Göttingen, Herzberger Landstraße 180, D-3400 Göttingen, Federal Republic of Germany

Key words: Geomagnetic activity – Annual wave – Hemispheric Activity Indices.

In an interesting paper Mayaud (1979) discusses possible sources of the annual wave in hemispheric activity indices (*Kn*, *Ks* or *an*, *as*) and comments the results of Damaske (1978b) obtained with the harmonic analysis method. Mayaud claims that his own analysis is the more comprehensive one and that the source model he suggests provides a better fit to the observed activity variations. It seems to us that his arguing against the concept of a modified modulation is based on a number of misunderstandings and misinterpretations and that also his merely qualitative judging deserves some critical remarks, although we can mention only a few details here.

1. Mayaud (1979) objects against the use of 27-day means and the restriction to the annual wave in the paper by Damaske (1978b), emphasizing that his own analysis is more comprehensive because his Fig. 1 displays also the daily variation, within groups of only 6 days. He should remember that it was the declared purpose of Damaske's paper to investigate merely the annual wave in both hemispheres and its behaviour within the eight hemispheric longitude sectors from which *Kn*, *Ks* (or *an*, *as*) are derived. Computing an annual wave from consecutive daily or even hourly values would not yield the slightest increase of accuracy besides much more computation labour. The advantage of 27-day means has been described elsewhere (Meyer 1973). The relation between diurnal and semidiurnal UT waves (including their annual amplitude modulation) and the annual and semiannual waves of geomagnetic activity, as well as their unified deduction from the modified modulation function $\sin^2(\beta+\beta_0)$, has been set forth extensively in previous papers (Damaske 1976, 1977).

2. Mayaud (1979) obviously uses the term 'modulation' in a way different to its common meaning in physics. It is a clear misinterpretation when he states that the modified modulation function $\sin^2(\beta+\beta_0)$ was introduced by Damaske (1977) 'in order to interpret the systematic amplitude modulation (with opposite sign in both hemispheres) of the 24-hour UT wave'. There does not exist such an amplitude modulation with opposite sign or phase. Instead, there exists a systematic *shift* (with opposite sign or direction) in the amplitude modulation curves for the diurnal UT wave in both hemispheres (see Damaske 1976, Figs. 25 and 26; or Damaske 1977, Figs. 13 and 14). Moreover, it is definitely wrong when Mayaud states that it is the 'asymmetry in the amplitude of the daily variation from one solstice to the other in each hemisphere .. what Damaske calls the amplitude modulation of (!) the diurnal UT wave' With this understanding a wave *without* such amplitude asymmetry between the two solstices as, e.g., the top curve in Mayaud's Fig. 2, would not undergo any amplitude modulation at all. Suffice it to say that we are using the term 'modulation' in its generally accepted sense.

3. We cannot follow Mayaud when he assumes that any change of the diurnal wave amplitude modulation would necessarily affect the annual variation of activity, e.g., bringing about a 12-month wave if such a change is confined to a certain season. A primary change of the daily variation would affect the annual variation only if also the average daily activity has changed. One can easily figure a diurnal wave with quite arbitrary amplitude modulation but no annual variation (of daily values) at all. It seems that Mayaud mixes the annual amplitude modulation of the diurnal UT wave with the annual variation of geomagnetic activity itself. Both have to be clearly distinguished. It is just the modified modulation function $\sin^2(\beta+\beta_0)$ that relates the diurnal with the annual activity variation in a definite and confirmed form.

4. Mayaud (1979) judges on the validity of the modified modulation function suggested by Damaske (1976, 1977) by a merely qualitative comparison of the figures he presents, Fig. 1 showing observational data and Fig. 2 illustrating different functions $\sin^2(\beta+\beta_0)$ by a polygonal approximation, although not for the empirical value of β_0. Nevertheless, he admits that one of the two specific features of hemispheric activity modulation, i.e., his feature (b), is similar in both figures. In fact also feature (a) can clearly be recognized in the lower curve of Fig. 2, as far as the existence of a significant daily variation at the equinoxes is concerned.

Mayaud's final objection mainly bears upon the envelopes of the curves in Figs. 1 and 2, which is certainly insufficient if not misleading. The modified modulation function has been derived from the quasi-logarithmic *Kn*, *Ks* and thus should not be compared with the linear *an*, *as* in all aspects. This does not simply concern the angle β_0. It applies to the whole formulation of activity modulation which is more complex for *an*, *as*, e.g., involving a dependency of all modulation amplitudes on the general level of activity (Damaske 1978a). Above all one has to make allowance

0340-062X/80/0048/0054/$01.00

for the fact that the scale for the square-sine functions in Fig. 2 is limited at its upper end, whereas the scale for *an* and *as* in Fig. 1 is still continuous and unlimited (contrary to that for *Kn*, *Ks* which would also be finite). Hence, any comparison of the upper envelope in Fig. 2, especially near the upper end of the scale, with the one in Fig. 1 is irrelevant.

Likewise Mayaud's statement that the solar longitude distance between the days of vanishing diurnal UT wave in the theoretical curve would not fit the observation, cannot be accepted. A quantitative evaluation of the correct curves derived from the exact formula and the results for the diurnal constituent in *Kn* and *Ks*, as well as in *an* and *as*, clearly shows that both are in good accordance with only minor deviations (cf. Damaske 1976, Figs. 25, 26, and 30, 31; Damaske 1977, Figs. 13, 14 and 17, 18; Damaske 1978a, Fig. 15).

5. Estimating the additional hemispheric effects by directly comparing the observational results in Fig. 1 with the modulation curve $\sin^2\beta$ for planetary activity (upper curve in Fig. 2) is certainly insufficient because the arbitrary scale in Fig. 2 involves the possibility of either exaggerating or suppressing the effect for a certain season, depending on the fitting of the relative amplitude scale. An unbiassed estimate of the specific hemispheric effects requires primarily a comparison with the *average* of both curves in Fig. 1, i.e., with the corresponding curve for *am*. If Mayaud had done this, he might have recognized that an additional diurnal UT wave, with nearly constant amplitude and phase, exists not only during local summer but indeed throughout the year. Since the phase of the *planetary* wave in summer and winter is opposite, its amplitude appears to have increased (in local winter) or decreased (in local summer) as an effect of this additional hemispheric wave, in full accordance with the results of harmonic analysis. Thus it seems that Mayaud is giving separate explanations of only certain features which he sees (and of *how* he sees them), thereby disregarding other features which he does not recognize by his merely qualitative judging.

6. Finally, it is unsatisfactory if Mayaud judges the relative validity of the modified modulation function by a qualitative comparison between the observational results in Fig. 1 and the modified square-sine functions in Fig. 2 (yet, strictly speaking, in an irrelevant way as has been shown above), without presenting and discussing a corresponding synthetic modulation curve for his own source model. In fact, his main source constituent, i.e., the one he relates with the *DP2* fluctuations (the other one does not significantly affect the annual wave in *hemispheric* data) would lead to an alteration congruent to the two envelopes of the upper curve in Fig. 2 'without any change in the range of the UT daily variation due to $\sin^2\beta$'. For a 12-month wave of this type culminating in local summer the upper envelope then would display a 'belly' during that season, instead of the 'trough' actually noticed in Fig. 1. In addition, this conception would match neither of the two specifically hemispheric features (a) and (b) in Mayaud's paper. It is hard to believe that all these systematic discrepancies respecting the daily variation should be explained just as 'irregularities' due to a longitude dependence of the source effect in connection with the non-homogeneous net of stations. If this were true, it would indeed severely question the utility of the hemispheric activity indices which Mayaud himself has introduced.

Summing up we may infer that Mayaud's conclusions are all together untenable. It will be shown in a forthcoming paper that the source model he suggests, though being tentatively acceptable as a working hypothesis, does not contribute significantly to the actually observed hemispheric annual wave.

References

Damaske, D.: Der Weltzeitgang der erdmagnetischen Aktivität als magnetosphärischer Modulationseffekt. Geophys. Abh., Inst. Geophys., Freie Universität Berlin, Heft 5, Berlin: Reimer 1976

Damaske, D.: Magnetospheric modulation of geomagnetic activity, I. Harmonic analysis of quasi-logarithmic indices *Km*, *Kn*, and *Ks*. Ann. Géophys. **33**, 461–478, 1977; – II. Harmonic analysis of linear indices *am*, *an*, and *as*. Ann. Géophys. **34**, 231–242, 1978a

Damaske, D.: On the annual wave of hemispheric geomagnetic activity. J. Geophys. **45**, 81–90, 1978b

Mayaud, P.N.: On the sources of the 12-month wave in the *an* and *as* geomagnetic activity indices. J. Geophys. **46**, 261–271, 1979

Meyer, J.: Zur Modulation der erdmagnetischen Aktivität. Geophys. Abh., Inst. Geophys., Freie Universität Berlin, Heft 3, Berlin: Reimer 1973

Received October 23, 1979; Revised Version January 9, 1980

J. Geophys. 48, 56, 1980

Journal of Geophysics

Book Reviews

Inkohlung und Geothermik. Beziehungen zwischen Inkohlung, Illit-Diagenese, Kohlenwasserstoff-Führung und Geothermik. Mit 96 Abbildungen, 51 Tabellen, 7 Tafeln. Krefeld, geologisches Landesamt Nordrhein-Westfalen 1979, I–XI, 1–372 S., 8°. Fortschritte in der Geologie von Rheinland und Westfalen, Band 27. Kunstleder DM 77,–.

Im Zuge der Suche nach neuen Energieformen und -quellen sowie noch unbekannten Lagerstätten fossiler Energiestoffe hat ein in der Vergangenheit etwas vernachlässigtes Teilgebiet der Geophysik, die Geothermik, einen beträchtlichen Aufschwung erlebt. Doch noch weit mehr als in anderen Gebieten der Angewandten Geophysik ist in der Angewandten Geothermik die interdisziplinäre Verflechtung mit anderen Zweigen der Geowissenschaften wie Geologie, Sedimentologie, Petrographie, Petrologie, Mineralogie und Geochemie deutlich, und eine isolierte Betrachtung physikalischer geothermischer Prozesse ohne Berücksichtigung der sonstigen geowissenschaftlichen Randbedingungen, Ursachen und Folgen ist nicht möglich. Für die Energieforschung ist die Geothermik interessant, einmal im Hinblick auf die Auffindung möglicher Lagerstätten von Erdwärme in Gebieten geothermischer Anomalien, zum anderen hinsichtlich der Genese von Erdöl, Erdgas, Kohle, der Diagenese von Speichergesteinen und letzten Endes auch der physikalisch-chemischen Zustandsbedingungen in Erdöl-Erdgas-Lagerstätten.

In all diesen Anwendungsfällen sind aber nicht nur die rezenten geothermischen Verhältnisse, sondern vor allem deren gesamte historische Entwicklung über geologische Zeiten wichtig. Wir benötigen „Paleogeothermometer", geochemische Reaktionen, die uns in Form von dauerhaften Reaktionsprodukten Auskunft über P,T-Bedingungen in der geologischen Vergangenheit geben können. Mit Untersuchungsmethoden auf dieser Basis und deren Anwendung befaßt sich der vorliegende Band, an dem eine Reihe namhafter Autoren aus dem gesamten Spektrum der Geowissenschaften mitgewirkt haben. Das Werk bietet sich damit als ein wertvolles Rüstzeug für den angewandten wie allgemeinen Geothermiker an.

Ausführlich werden die Methodik und Anwendung der Kohlepetrologie behandelt, die es ermöglicht, durch mikroskopische Reflektivitäts- oder Fluoreszenz-Messungen den thermischen Umwandlungsgrad (Inkohlungsgrad) von fein verteilter organischer Substanz in Sedimentgesteinen zu bestimmen. Daneben gehen verschiedene Autoren auch auf die Tonmineraldiagenese als Maß für die thermische Einwirkung in der Vergangenheit ein. Auswertung und Interpretation entsprechender Laboratoriumsuntersuchungen ist breiter Raum gewidmet, wobei eine Synthese von theoretischen Überlegungen – unter Zugrundelegung des Zeitintegrals einer Temperaturfunktion als bestimmendem Parameter – und empirischen Studien – mittels Korrelation von Inkohlung, Tonmineraldiagenese, Kohlenwasserstoffgenese, geothermischen Gradienten, geothermischer Geschichte und Absenkungsgeschichte – angestrebt ist. Auch der unterschiedliche Einfluß des Hangenddruckes auf Inkohlung und Tonmineraldiagenese wird in die Betrachtungen einbezogen.

Die empirischen Untersuchungen sind vorwiegend am Oberrhein-Graben durchgeführt. Damit gibt das Werk auch gleichzeitig wichtige Auskünfte zur geologischen und geothermischen Geschichte dieses Gebietes. Weitere Artikel befassen sich mit Nordwestdeutschland, Urach und anderen Gebieten.

In einer Serie von 14 Artikeln unter Beteiligung von 12 Autoren, mit vielen Graphiken, Karten und Tabellen ist das anstehende Thema in vielfältiger Form ausführlich behandelt.

Der Band schließt an einen früheren Band 24 „Inkohlung und Erdöl" des gleichen Herausgebers an. Er liefert eine zusammenfassende Übersicht des gesamten Komplexes und ist somit auch dem Geophysiker als nützliches Werkzeug sehr zu empfehlen.

Jürgen R. Schopper, Clausthal-Zellerfeld

Peter Krzonkalla: Wie finde ich Literatur zu den Geowissenschaften? 223 Seiten, Berlin Verlag, 1979, ISBN 3-87061-178-2.

In sehr klarer, übersichtlicher und allgemeinverständlicher Form zeigt der Autor die verschiedensten Wege zur Literatur der Geowissenschaften. Es ist ein sehr nützliches und empfehlenswertes Buch für alle interessierten Laien, Angestellte in entsprechenden Bibliotheks- und Dokumentationsdiensten sowie für Studenten. Um für Wissenschaftler und Doktoranden ähnlich nützlich zu sein, bedürfte es noch einiger kleinerer Ergänzungen.

Ein solches Buch muß, wie der Autor richtig feststellt, Auswahlcharakter haben. Nur sollte die Auswahl mehr von Fachleuten getroffen werden, als aufgrund von Planungen des IuD-Programmes der Bundesregierung und aufgrund von Beständen der Universitätsbibliothek der Technischen Universität Berlin. So kommt es z.B., daß es eine Geophysik ohne Aeronomie bzw. Atmosphärenphysik gibt. Insbesondere bei der Fächeraufteilung bzw. Fächerzuordnung sind Ergänzungen notwendig, insbesondere wenn an fachübergreifende Nutzung gedacht wird.

Mit der Auflistung von Informationszentren, Informationsdiensten und Informationsleistungen hat der Autor unmerklich zusätzlich zur reinen und klassischen Literaturdokumentation (bibliographische Daten) auch noch den heute immer wichtiger werdenden Bereich der entsprechenden Datendokumentation (numerische Daten, Meßdaten) ins Spiel gebracht. Es wird unversehens auch mit der entsprechenden Terminologieproblematik konfrontiert. Die ganz zum Schluß des Buches benutzten Begriffe Datenbasen und Datenbanken bedürfen deshalb zur Vermeidung von Mißverständnissen unbedingt entsprechender Zusatzbemerkungen.

Da diese Vermischung von Literaturdokumentation und Datendokumentation – im obigen Sinne – immer unvermeidbarer wird, häufiger auftritt und die Wissenschaftler teilweise mehr nach geeigneten Daten, als nach entsprechender Literatur suchen, sollte z.B. die Liste der Informationsdienste ergänzt werden, in der Geophysik z.B. um die zwei folgenden:

1. Guide to International Data Exchange through the World Data Centers. ICSU Panel on World Data Centers, Dec. 1973
2. CODATA Bulletin. ICSU Committee on Data for Science and Technology

G. Hartmann, Katlenburg-Lindau

J. Geophys. 48, 57–59, 1980

Journal of Geophysics

A Late Cretaceous $^{40}Ar-^{39}Ar$ Age for the Lappajärvi Impact Crater, Finland

E.K. Jessberger[1] and W.U. Reimold[2]

[1] Max-Planck-Institut für Kernphysik, Postfach 103980, D-6900 Heidelberg, Federal Republic of Germany
[2] Institut für Mineralogie, Gievenbecker Weg 61, D-4400 Münster, Federal Republic of Germany

Abstract. We report on a $^{40}Ar-^{39}Ar$ study of kärnäite from the ~17 km Lappajärvi impact crater, Finland. Four samples from a 3,000 m profile across the crater center give rather well defined age plateaux and indicate complete degassing at the time of the impact event. The mean age is 77 m.y., much younger than geologically derived age estimates.

Key words: Impact cratering – $^{40}Ar-^{39}Ar$ dating – Meteorites – Kärnäite – Rb–Sr dating – Impact melt – Rate of impacts – Chronology – Melt sheet.

Meteorite impact cratering is a fundamental process in surface feature formation of terrestrial planets as has been shown by the lunar, martian, and mercurian missions of the last decade (Neukum and Wise 1976). In a most recent attempt to define a crater production rate for the evolution time of Earth and Moon (Grieve and Robertson 1979) absolute dating of terrestrial impact events has been demonstrated to be essential. In contributing to that tedious but important task we dated the Lappajärvi impact structure, Finland, by the $^{40}Ar-^{39}Ar$ technique and obtained an age of 77 m.y., much younger than the geologically derived upper age limit of 600 m.y. (Grieve and Robertson 1979).

Fredriksson and Wickman (1963) first suggested that the Lappajärvi structure might have been produced by meteorite impact, which was supported later by the identification of shock metamorphism (Svensson 1968; Lehtinen 1976). Recently, the projectile of the Lappajärvi impact has been identified as a chondritic meteorite (Reimold 1979; Göbel et al. 1980).

Lake Lappajärvi (Fig. 1) is located in the Svecofennian gneiss mass of Central Finland (63°09′N/23°42′E) about 300 km northwest of Helsinki. From gravity measurements the original crater diameter of the glacially deeply eroded crater has been suggested to be about 17 km (Elo 1976). The basement rocks of that area consist mainly of mica schists, granite-pegmatite, and smaller bodies of granodiorite and amphibolite. According to Rb–Sr dating the ages of these rocks range from 1,559–1,745 m.y. (Reimold 1979; Maerz et al. 1979).

Autochthonous impact breccias and melt ('kärnäite') are exposed on several islands in the center of the crater area, whereas allochthonous boulders of kärnäite and suevite are found in glacial deposits some kilometres southeast of the lake.

Kärnäite appears as a dense aphanitic rock that contains abundant mineral and lithic clasts of all grades of shock metamorphism

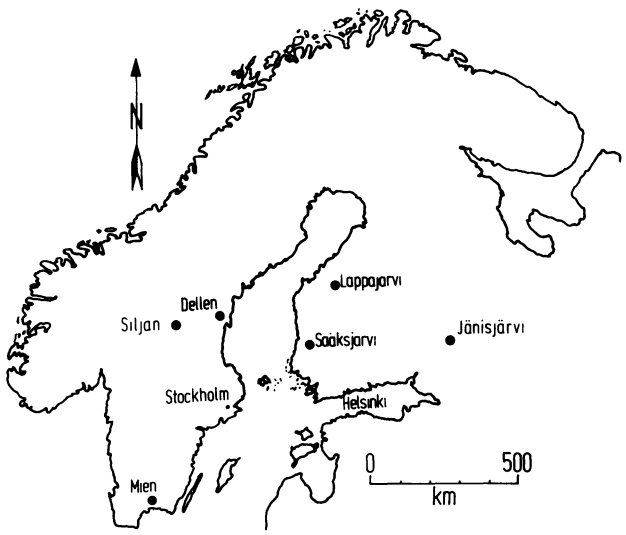


Fig. 1. Locations of the Scandinavian impact structures

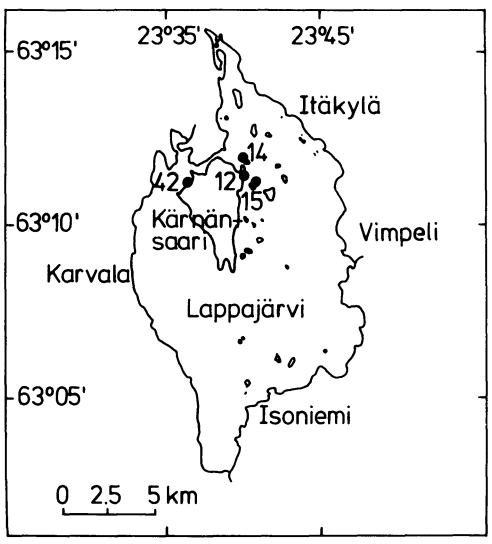


Fig. 2. Locations of the four samples La 42, La 12, La 14, and La 15 from the melt sheet exposed on Kärnänsaari studied with the $^{40}Ar-^{39}Ar$ dating technique

0340-062X/80/0048/0057/$01.00

in a fine-grained, mesostasis-rich matrix. Following the petrographic descriptions of Lehtinen (1976) and Maerz (1979) the inclusions consist mainly of quartz that often exhibit 'ballenstructure'. Frequently these fragments are surrounded by rims of orthopyroxene. Feldspar clasts consist of plagioclase and alkali-feldspar. The former often shows 'checkerboard'-texture also known from other impact sites. Lithic clasts are medium to coarse grained granite, and rarely granodiorite or amphibolite. The matrix consists of plagioclase, often exhibiting H-type crystal growth, also zoned and rimmed by sanidine, of orthopyroxene, in part altered into clinopyroxene or biotite, of magnetite, ilmenite, and twinned cordierite (Maerz 1979). The mesostasis is very K-rich and most likely consists of alkali-feldspar and quartz.

We have chosen 4 samples of authochthonous kärnäite for age determinations. These samples are part of a 3,000 m profile across the coherent melt sheet of the crater center (Fig. 2). Rb–Sr dating of kärnäite failed because of the extremely small range of the $^{87}Rb/^{86}Sr$ ratio of kärnäite whole-rock samples (Maerz et al. 1979). Mineral separation of the matrix was not possible because of its small grain size (< 35 µm). For the $^{40}Ar-^{39}Ar$ study we selected inclusion-poor kärnäite and analyzed whole-rock samples routinely as has been described previously (Jessberger et al. 1976)[1]. The results are summarized in Table 1 and the age patterns shown in Fig. 3.

All samples exhibit rather similar age patterns with no evidence for major ^{40}Ar diffusion losses after the time of complete degassing. The integrated K–Ar ages (Table 1) are only marginally different from the $^{40}Ar-^{39}Ar$ ages. There are also no elevated ages at higher extraction temperatures indicating the absence of inherited ^{40}Ar such as that encountered in dating lunar impact breccias (Jessberger et al. 1974; 1976). Obviously, the mineral or

[1] Irradiation at the FR2 reactor, Karlsruhe, together with muscovite monitor of known age, Bern 4M (Jäger et al. 1963). Ages are calculated using the ^{40}K decay constants and K isotopic composition recommended by Steiger and Jäger (1977)

Table 1. Summary of $^{40}Ar-^{39}Ar$ results for Lappajärvi impact melt samples. The uncertainty of the K- and Ca-contents are 5%. K–Ar ages are obtained from the integrated ^{40}Ar and ^{39}Ar amounts released in all extraction fractions. $^{40}Ar-^{39}Ar$ ages are calculated from argon released in extraction fractions of the given temperature range. The quoted uncertainty corresponds to the 1σ standard deviation from the weighed average $^{40}Ar/^{39}Ar$ ratio. The quality of the $^{40}Ar-^{39}Ar$ ages as plateau ages may be visualized from Fig. 3

Sample no.	Sample Weight (g)	K (%)	Ca (%)	$^{40}Ar_{atm}$ (%)	K–Ar Age (m.y.)	$^{40}Ar-^{39}Ar$ Age and Range	
						(m.y.)	(° C)
La 12	0.570	3.2	1.8	19.5	78.5	78.6±0.8	660–1,360
La 14	0.553	3.4	1.9	38.2	74.8	76.6±0.3	820–1,550
La 15	0.543	3.3	1.8	23.2	76.8	77.2±0.3	770–1,550
La 42	0.447	3.6	2.0	26.0	76.2	77.5±0.2	700–1,520

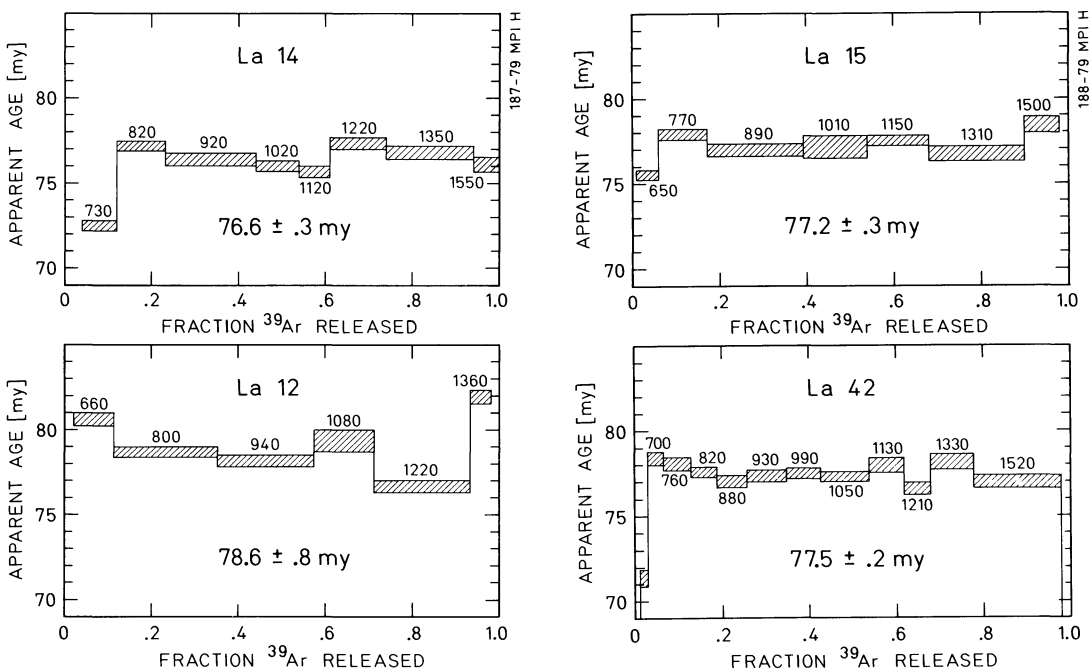


Fig. 3. $^{40}Ar-^{39}Ar$ age patterns for four samples from the Lappajärvi impact structure. Each bar represents the apparent age calculated from the ($^{40}Ar/^{39}Ar$) ratio of argon released within about 45 min at the extraction temperature given in °C. The heating procedure was incremental. The width of the bar is the fractional amount of ^{39}Ar released, the height is the statistical uncertainty of the K-derived ($^{40}Ar/^{39}Ar$) ratio of that fraction

lithic clasts – relicts of the Precambrian target rocks – did not retain ^{40}Ar or they do not contain much K. Many of the inclusions consist of quartz, and Rb–Sr analyses of kärnäite inclusions have also shown that most of the feldspathic inclusions have been equilibrated with the surrounding melt (Reimold and Stöffler 1979).

Sample La 12 shows a slightly different age pattern (Fig. 3) with more scatter of the apparent ages than the other samples. This may be due to the higher portion of clastic material mixed to the melt. This sample appears as a melt agglomorate which is transitional to a suevite breccia rather than a coherent melt as the other samples. There are also differences in the chemical composition especially in the REE patterns (Reimold and Stöffler 1979). The mean age, however, is not affected by the difference, and is well defined to be 77.3 ± 0.4 m.y.

The present study which gives the first lateral age profile across an impact melt body indicates that the whole melt body has been totally degassed in the course of the impact. This extends even to the autochthonous suevitic sample La 12. Other evidence for complete homogenization of the Lappajärvi melt body is presented in (Reimold 1979; Reimold and Stöffler 1979).

On the Scandinavian shield there are six structures classified as 'probable' impact structures (Grieve and Robertson 1979): Mien, Dellen, and Siljan in Sweden, Sääksjärvi and Lappajärvi in Finland, and Jänisjärvi in the USSR (Fig. 1). $^{40}Ar-^{39}Ar$ ages of Mien and Siljan are 119 m.y. and 362 m.y., respectively (Bottomley et al. 1978). Lappajärvi is 77 m.y. old. An approximate age of 700 m.y. for the Jänisjärvi event is mentioned by Grieve and Robertson. Also for Dellen and Sääksjärvi no accurate ages are available (Bottomley et al. 1978). We therefore propose to date the latter three structures also to further substantiate their impact origin and to obtain the full impact rate for the Scandinavian shield.

Acknowledgement. We are greatly indebted to Dr. D. Stöffler for fruitful discussions and to him and Dr. G. Turner for reviews of the manuscript. This work was supported by the Gesellschaft für Kernforschung mbH, Karlsruhe.

References

Bottomley, R.J., York, D., Grieve, R.A.F.: $^{40}Ar-^{39}Ar$ dating of Scandinavian impact craters. Contrib. Mineral. Petrol. **68**, 79–84, 1978

Elo, S.: A study of the gravity anomaly associated with Lake Lappajärvi, Finland. Geol. Surv. Finland, 1976

Fredriksson, K., Wickman, F.E.: Meteoriter, Sven. Naturvetensk. **16**, 121–157, 1963

Göbel, E., Reimold, W.U., Baddenhausen, H., Palme, H.: The projectile of the Lappajärvi impact crater, Naturwissenschaften in press, 1980

Grieve, R.A.F., Robertson, P.B.: The terrestrial cratering record I. Current status of observation. Icarus **38**, 212–229, 1979

Jäger, E., Niggli, E., Baethge, H.: Two standard minerals, biotite and muscovite for Rb–Sr and K–Ar age determinations, samples Bern 4B and Bern 4M from a gneiss from Brione, Valle Verzasca (Switzerland). Schweiz. Mineral. Petrogr. Mitt. **43**, 465–470, 1963

Jessberger, E.K., Huneke, J.C., Wasserburg, G.J.: Evidence for a 4.5 AE age of plagioclase clasts in a lunar highland breccia. Nature **248**, 199–202, 1974

Jessberger, E.K., Kirsten, T., Staudacher, Th.· Argon-argon ages of consortium breccia 73215. Proc. 7th Lunar Sci. Conf. 2201–2215, 1976

Lehtinen, M.: Lake Lappajärvi, a meteorite impact site in western Finland. Geol. Surv. Finl. Bull. **282**, Dissertation, 1976

Maerz, U.: Petrographisch-chemische Untersuchungen von Impaktschmelzen und Brekzien skandinavischer Meteoritenkrater. Diploma Thesis, Institut für Mineralogie, Universität Münster, 1979

Maerz, U., Reimold, W.U., Stöffler, D.: Preliminary results on the petrography and Rb–Sr isotope geochemistry of the Lappajärvi impact melt. Lunar Planet. Sci., 769–771, The Lunar Planet. Inst., Houston, 1979

Neukum, G., Wise, D.U.. Mars – a standard crater curve and possible new time scale. Science **194**, 1381–1384, 1976

Reimold, W.U.: Rb–Sr- und Multielement-Analysen von Gesteinen des Lappajärvi-Meteoritenkraters. Fortschr. Mineral. **57** (Beiheft 1), 123–124, 1979

Reimold, W.U., Stöffler, D.: Isotope, major and trace element chemistry of the Lappajärvi impact melt. Meteoritics in press, 1979

Steiger, R.H., Jäger, E.: Subcommission on Geochronology: Convention on the use of decay constants in geo- and cosmochronology. Earth Planet. Sci. Lett. **36**, 359–362, 1977

Svensson, N.B.: Lake Lappajärvi, Central Finland: A possible meteorite impact structure. Nature **217**, 438, 1968

Received January 15, 1980; Revised Version February 15, 1980

J. Geophys. 48, 60–66, 1980

Journal of Geophysics

Magnetic Properties of Swiss Flysch*

J.E.T. Channell, R. Freeman, and F. Heller
Institut für Geophysik, ETH Hönggerberg, CH-8093 Zürich, Switzerland

Abstract. The principal magnetic mineral present in the flysch units of central Switzerland is magnetite which exhibits susceptibility anisotropy corresponding to a sedimentary fabric. Remanent magnetization directions are consistent for the finer grained (upper) parts of the turbidity flows. Laboratory controlled build-up of viscous remanent magnetization indicates that the natural remanence could have been viscously acquired during the Brunhes epoch.

Key words: Rock magnetism – Viscous remanent magnetization – Anisotropy of magnetic susceptibility – Flysch.

1. Introduction

The flysch units of Central Switzerland are made up of several 1,200–1,600-m-thick series of alternating fine-grained sandstones and shales, which were deposited during the Maastrichtian to Middle Eocene in a South Penninic deep marine ('oceanic') environment (Schaub 1951; Hsü 1960; Caron 1976; Morel 1978; Van Stuijvenberg 1979). The palaeogeographic position of these units is not altogether clear. In the Late Eocene to Oligocene, during the Mesoalpine orogeny (Trümpy 1973), they were torn from their roots and transported northwards over considerable distances and now lie in thrust sheets at the base of the Prealps (Fig. 1).

Outcrop is poor throughout the external Prealps and the sampling was restricted to four quarries in the Gurnigel, and three river beds in the Schlieren and Wägital flysch. The dip of the strata is similar throughout each of the three units, and therefore the palaeomagnetic results from the outcrops were deemed representative for the whole unit.

Originally we had hoped to concentrate on the finest-grained layers by sampling at the top of each turbidite cycle in addition to the calcitic hemipelagic interbeds. Most of these rocks proved too friable to drill successfully or too clay-rich to survive desication and cutting in the laboratory. As a result, most of the samples are from the upper parts of each turbidite cycle and are in general fine grained ($^1/_4$–$^1/_8$ mm) arkosic sandstones with a significant (up to 30%) contribution of volcanic and plutonic detrital fragments (Hubert 1967).

2. Magnetic Properties

The flysch beds are generally turbiditic and therefore show grain size decrease from bottom to top. Prior to the main sampling, two flows from the Gurnigel flysch (Zollhaus quarry) were sampled in detail, and consistent NRM directions were found towards the top of the flows. However, the magnetization intensity decreases to the noise level of the magnetometer at the base of the flows where the magnetization directions cannot be precisely measured. This may reflect either a decrease in ferromagnetic mineral content or increasing turbulent deposition towards the bottom of the turbidite flows.

* Contribution No. 253, Institut für Geophysik, ETH Zürich

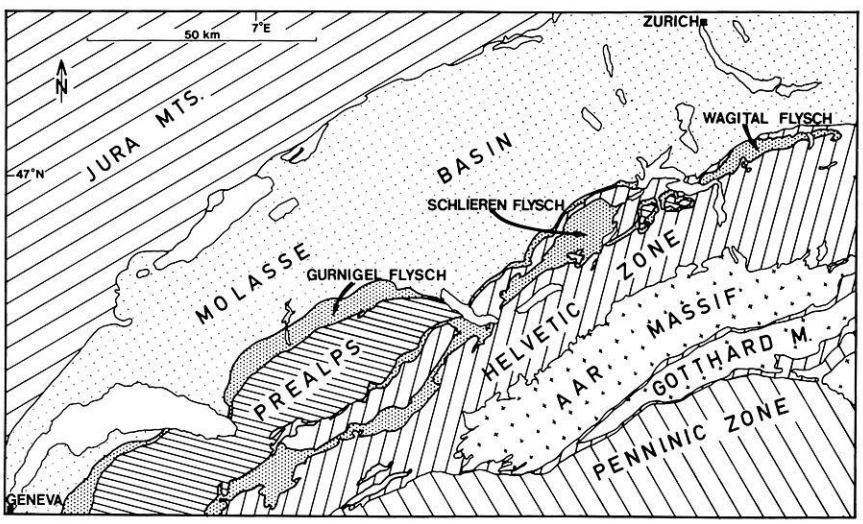


Fig. 1. Location Map

0340-062X/80/0048/0060/$01.40

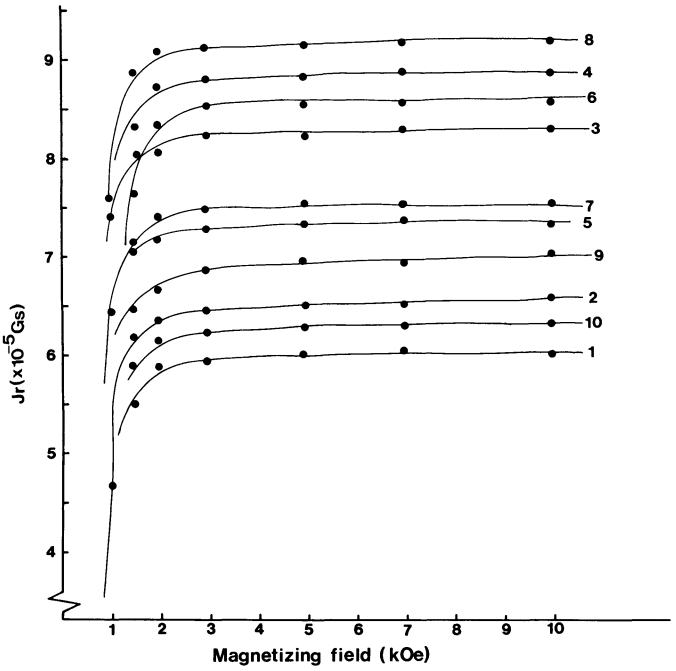


Fig. 2. Acquisition curves of Isothermal Remanent Magnetization (IRM) for a single turbidite flow from the Gurnigel flysch (Zollhaus Quarry), the labels refer to the order of the samples from the top (1) to the base (10) of the flow

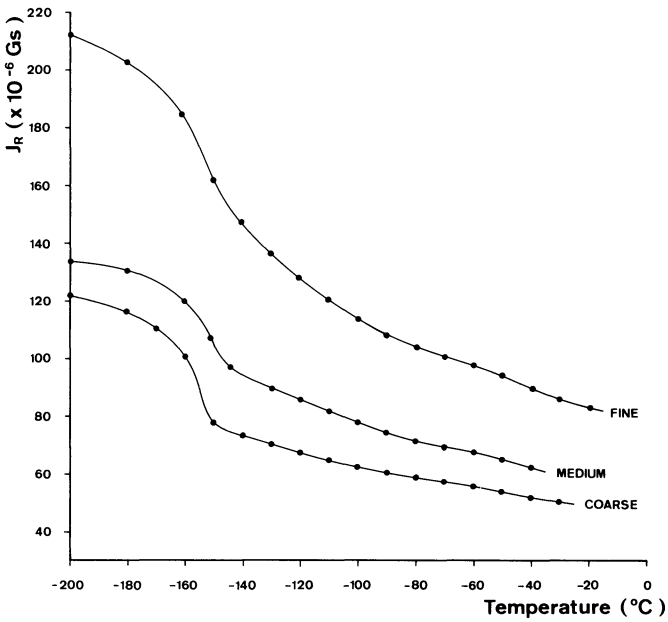


Fig. 3. Decay of an IRM (imposed at liquid nitrogen temperature) during warming to room temperature for three samples from the same flow. Approximate mean grain size. Very fine to fine ($^1/_4$–$^1/_{10}$ mm), Medium ($^1/_2$–$^1/_4$ mm), Coarse (1–$^1/_2$ mm)

(a) Isothermal Remanent Magnetization (IRM)

The acquisition of IRM provides a very useful guide to the coercivity spectrum of magnetic minerals present in a rock sample. The shape of the acquisition curves from a single turbidite flow (Fig. 2), and also from throughout the sampling area, indicates that low coercivity minerals are present in the flysch. All samples are saturated at 2–3 kOe. The value of saturation remanence at room temperature shows no consistent relationship with position of the sample within the graded bed (Fig. 2), suggesting that the weight percent of magnetic mineral does not vary consistently, and that the decrease in NRM intensity is either a function of grain size and/or the alignment efficiency of detrital grains.

The variation in magnetic grain size within the beds is reflected during low temperature treatment (Fig. 3). The samples were given a saturation IRM (in a 10 kOe field), at liquid nitrogen temperature (−196° C) and the remanence was measured during warming to room temperature. In the coarser grained sediments, a sudden decrease of remanence is observed corresponding to the change of sign of the principal magnetic anisotropy constant K_1 of magnetite at about −150° C (Nagata et al. 1964). The transition temperature depends on the concentration of impurity cations in magnetite (Syono 1965) and is depressed by small titanium contents. It is suppressed entirely if the magnetite is single domain, the magnetic anisotropy of magnetite being mainly shape controlled. The form and appearance of the transition depends on the grain size of the sample (Fig. 3). The transition is most distinctly observed in the coarse grained samples where about 30% of the low temperature IRM has been lost. This percentage corresponds to the contribution of multidomain magnetite, therefore, a major portion of IRM must be carried by single domain (perhaps pseudo-single domain) grains. The IRM intensity decreases further when warming from −150° C to room temperature indicating a small contribution from superparamagnetic magnetite. The medium grain-sized samples (Fig. 3) show a slightly broadened, but still pronounced, transition. The portion of the IRM carried by multidomain magnetite is reduced (< 25%), and the gradient of the curve above the transition is steeper. Therefore the contribution of single domain and superparamagnetic (at room temperature) grains is enhanced in the medium grained as opposed to the coarse grained samples. In the fine grained samples (Fig. 3), the transition is difficult to recognise, indicating that the multidomain fraction is largely absent. The low temperature IRM intensity is about twice the usual room temperature value (Fig. 2). The shape of the curve indicates a large contribution from very fine grained magnetite which is superparamagnetic at room temperature.

Thus the magnetic mineralogy in the flysch samples is dominated by superparamagnetic and single domain magnetite. Multidomain magnetite is an important contributor to the magnetization of the coarser grained material, but the vast majority of the samples were collected from the fine-grained parts of the turbidite flows.

A maximum blocking temperature of about 540° (Fig. 4) was obtained by giving a sample a saturation IRM at room temperature and monitoring the decay of this IRM progressively during heating, using apparatus designed by Heiniger and Heller (1976). This maximum blocking temperature together with the IRM acquisition curves (Fig. 2) and the low temperature characteristics (Fig. 3) indicate that magnetite is the only magnetic mineral present in the flysch.

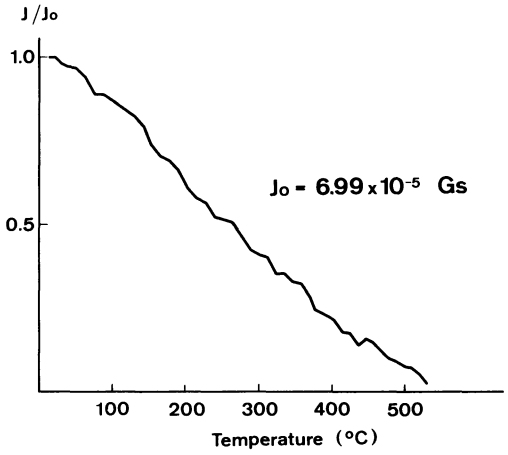


Fig. 4. Continuous measurement of the decay of an IRM (imposed at room temperature) during heating for a fine-grained sample

(b) Natural Remanent Magnetization (NRM)

The NRM directions are consistent for finer grain sizes within a single flow, and for fine grained samples from all exposures (Fig. 5). On AF demagnetization the scatter of directions increases, as the magnetization intensities approach the noise level of the ScT magnetometer. The ScT cryogenic magnetometer measures the magnetization along three orthogonal directions at the same time and almost instantaneously, regardless of the intensity of the magnetization. The samples were measured in the upright and inverted position and the six magnetization component readings were combined to produce eight estimates of the remanent magnetization vector. The parameter ψ is a measure of the deviation of the eight directions about their mean and is a useful reliability parameter for individual sample measurements (Lowrie et al. 1980)

where $\psi = 81\ [(N-1)/(N-R)]^{-1/2}$
and N = number of estimates (8) of magnetic moment vector,
R = vector sum of eight estimates of magnetization vector.

The uncertainty in individual sample measurements, quantified by ψ, increases with increasing peak demagnetization field as magnetization intensity decreases. This is due to unstable magnetziation of these samples as well as to the fact that the noise level of the magnetometer is approached. The weak magnetization intensity of the NRM (mean value 10^{-7}G) and the low median destructive field (100 Oe) often causes the magnetization intensity to fall beneath the noise level of the magnetometer after demagnetization at 200 Oe or 300° C (see Fig. 6). However, in many cases the samples picked up a secondary magnetization after treatment at higher fields and temperatures, indicating a tendency for viscous remagnetization.

The significance of the apparently stable NRM directions was tested in two ways, the first being a field test and the second a laboratory test of VRM acquisition.

For the field test, we sampled a drag fold at the base of the Schlieren flysch. From AF and thermal cleaning results we isolated the most stable directions using a criterion based on the minimum change of magnetization direction between two demagnetization steps. The results are plotted in Fig. 7. The directions before fold correction are scattered to some extent and their mean is located near the present Earth's field direction. After correcting for sample attitude in the fold, the dispersion is greatly increased, and the magnetizations from the overturned limb now lie in the upper hemisphere (Fig. 7). This failed fold test indicates that the NRM was acquired after Alpine emplacement.

(c) Viscous Remanent Magnetization (VRM)

The acquisition of VRM for six samples from the main (fine grained) sample collection was monitored by allowing the samples

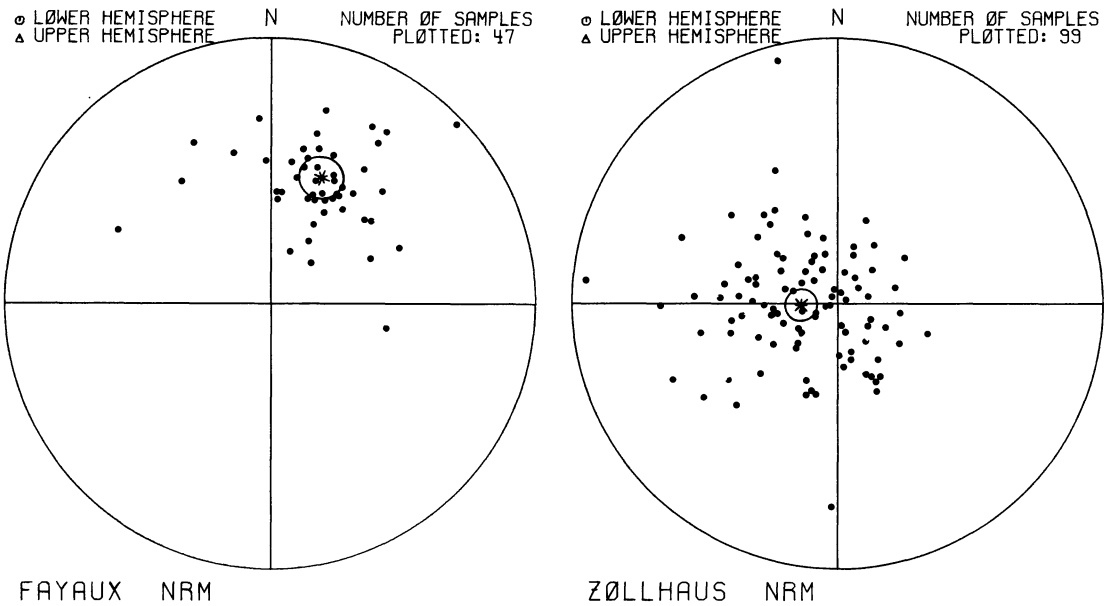


Fig. 5. Natural remanent magnetization directions of fine-grained samples from two Gurnigel flysch quarries before tectonic correction

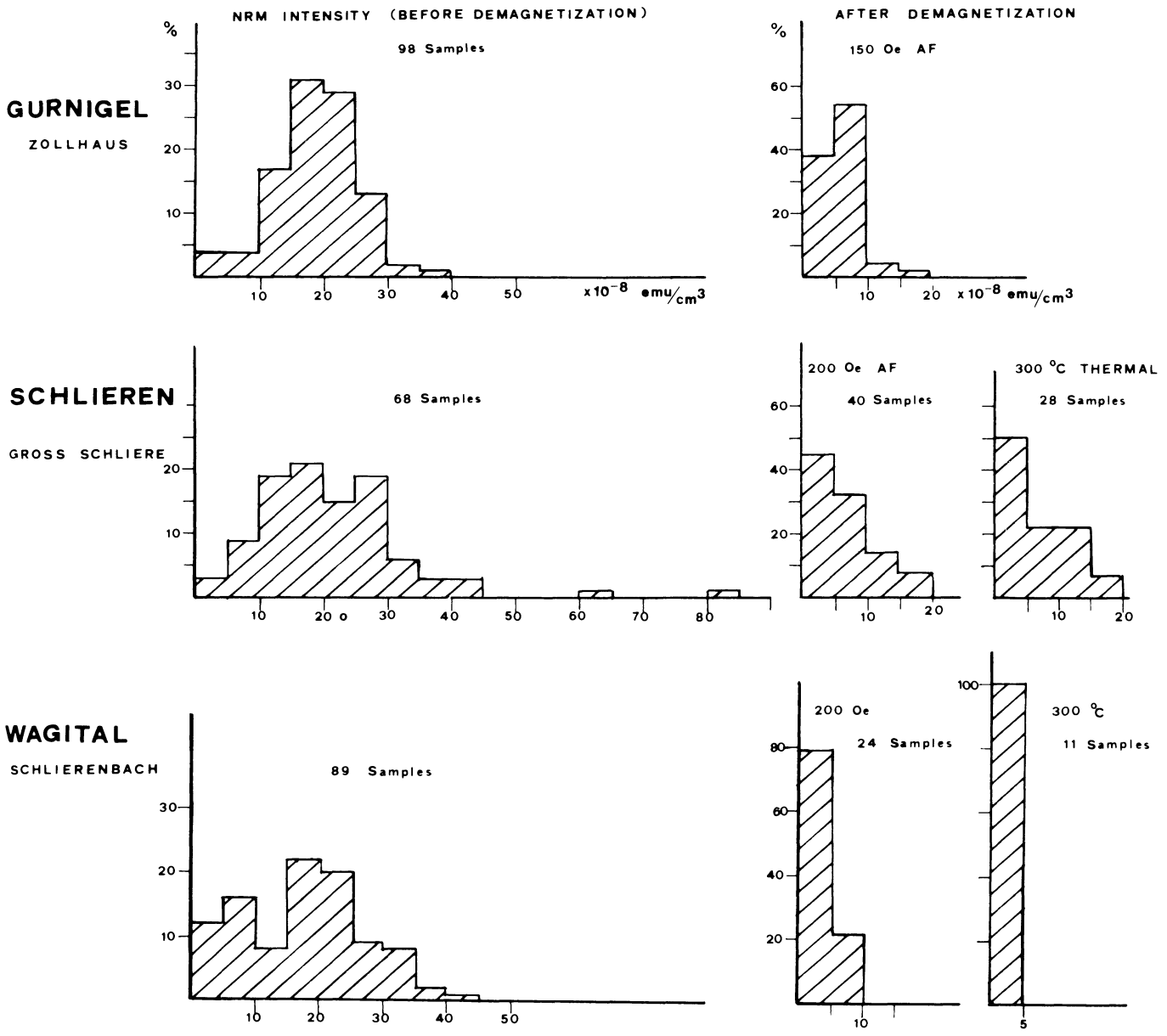


Fig. 6. Histograms of the magnetization intensity before and after AF or thermal demagnetization

to remain in the Earth's field and observing the build-up of magnetic remanence over a 6-month period. These samples were randomly selected and represent the general VRM properties of the flysch. Starting from the momentary NRM state (both undemagnetized and AF cleaned with peak fields of 200 Oe), the acquisition of VRM appears to depend on the logarithm of time (Fig. 8). If we assume no long term changes in the magnetic viscosity behaviour, we can extrapolate the straight-lines to values of log t corresponding to the onset of the Brunhes epoch. It is indicated that considerable build-up remanence may have occurred since the last reversal of the earth's magnetic field. The extrapolated intensity of VRM exceeds the magnitude of the NRM.

The AF demagnetization of the VRM (Fig. 9) shows that the VRM-components acquired during a 6-month period in the laboratory can be totally obliterated in peak fields of 40–80 Oe. The remanence vectors removed (Fig. 9) during AF cleaning are perfectly aligned along the present day field direction (Dec=358°, Inc=63°) after the first demagnetization step ($H=20$ Oe). As demagnetization proceeds, the removed vectors become more and more dispersed reaching a nearly random distribution after demagnetization in peak fields of 100 Oe with values of $\alpha_{95}=40°$ ($N=17$). The relationship

$$\frac{H_1}{H_2}=\frac{\log t_1}{\log t_2} \quad \text{(Néel 1955)}$$

where H_1 and H_2 are the peak alternating fields necessary to obliterate the viscous build-up in time t_1 and t_2 respectively, can give an estimate of the field H_2 necessary to remove a VRM acquired during the Brunhes epoch. If $H_1=60\pm20$ Oe, $t_1=6$ months, $t_2=700{,}000$ years then $H_2=160\pm50$ Oe. The values of

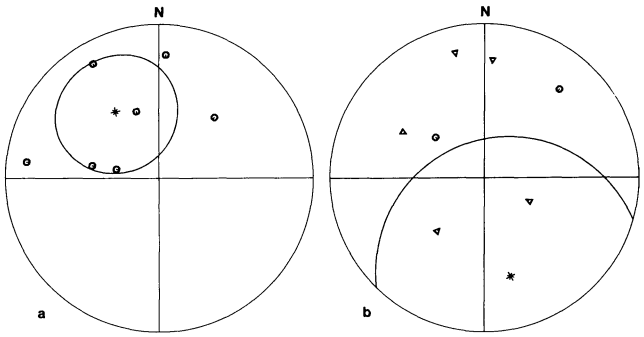


Fig. 7a and b. Fold test in the Schlierenflysch. The small circles represent sample directions in the lower hemisphere while the small triangles are directions in the upper hemisphere. The star locates the mean direction and the ellipse is the projection of the 95% confidence circle. **a** Before fold correction. **b** After fold correction

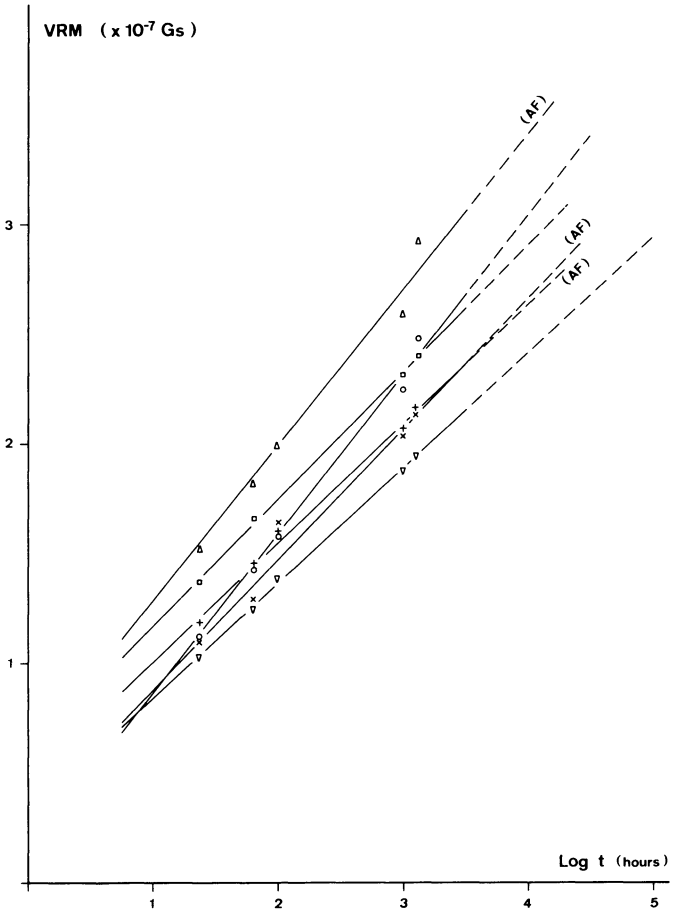


Fig. 8. The intensity of VRM acquired in time (t) plotted against log t. Three of the samples were demagnetized (AF) up to peak fields of 200 Oe at time $t=0$

H_2 coincide with the range of peak demagnetizing fields at which the NRM reached instrumental noise level. Therefore both build-up of VRM during storage and its subsequent destruction by alternating fields indicate that the NRM of these flysch samples

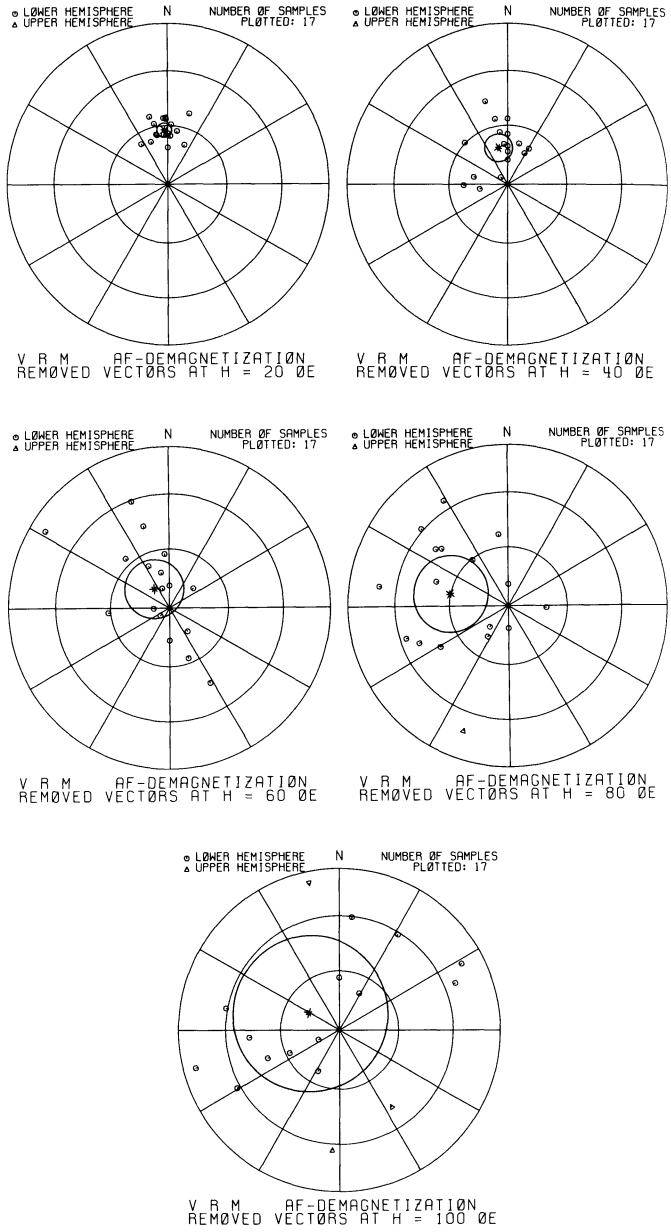


Fig. 9. Removed vectors during various stages of alternating field demagnetization of the VRM acquired in 6 months

is entirely of viscous origin and acquired during the Brunhes normal epoch.

In order to investigate the viscous behaviour in more detail, we studied several samples which were much more viscous than the norm. In this experiment we continuously recorded the decay of viscous magnetization acquired during 5–400 sec emplacement in the Earth's field (Fig. 10).

Dunlop and West (1969) reviewed the theoretical and experimental evidence showing that multidomain grains have a decay rate of VRM half as big as the growth rate. In these flysch samples, the decay rate was about 1.5 times greater than the growth rate. This evidence further supports the hypothesis that single domain magnetite grains with short relaxation times at room temperature are responsible for the VRM in the flysch.

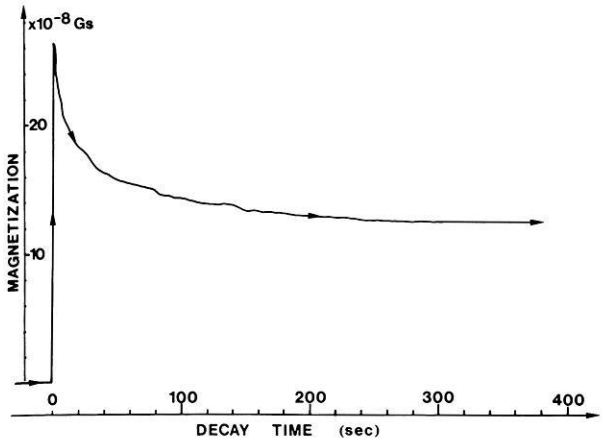


Fig. 10. Decay of the z-component of magnetization of a flysch sample held 360 sec in the Earth's field and then immediately lowered into the measuring region of the cryogenic magnetometer

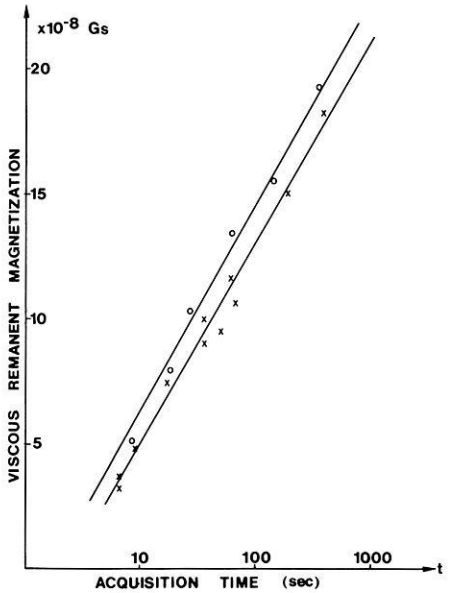


Fig. 11. Short time acquisition of VRM. The slopes are identical for both the undemagnetized (*crosses*) and the demagnetized (*circles*) sample

A plot of the acquired VRM against log t (sec) for an undemagnetized and demagnetized pair of these very viscous flysch samples (Fig. 11) demonstrates the very fast acquisition of VRM. The acquisition behaviour is similar for both magnetization states, and these samples acquired 50% of their NRM values in about 10 min.

The viscosity of the flysch samples may be related to their relatively high content of volcanic and plutonic detritus. X-ray diffraction analysis of a magnetically separated fraction, optical examination of polished sections and microprobe investigations failed to identify any magnetic mineral. Thus the magnetite grains are probably smaller than the resolving power of these methods (ca. 1 μm). However, a large albite component was present in the magnetically separated fraction. Since albite itself is not magnetic, this observation indicates that magnetite might be present as exsolved phases within the twinning lamellae of the plagioclase. Indeed, in thin sections several large grains of albite did exhibit tiny opaque phases – possibly magnetite – aligned along the twinning interfaces. These phases were in the micron to submicron size and may be responsible for the particular viscosity of the rocks investigated.

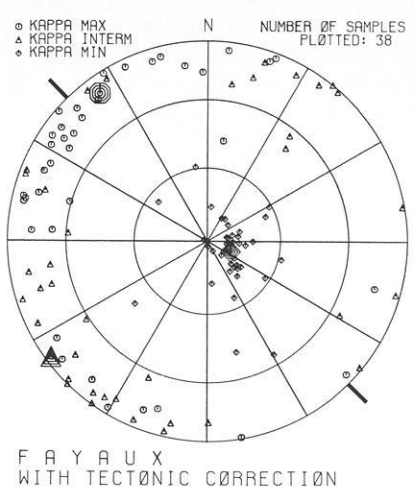


Fig. 12. Axes of the magnetic susceptibility ellipsoid for samples from the Fayaux quarry of the Gurnigel flysch. The *thick lines* at the edge of the circle give the mean elongation axis of flute casts

The viscous nature of the NRM explains the lack of reversed NRM directions (Fig. 5). The time span represented by the deposition of the samples is such that reversed directions would be expected if the magnetizations were primary. The NRM directions before tectonic correction (Fig. 5) are close to, but do not coincide with, the present-earth's field direction. This is probably due to the influence of VRM acquired in the laboratory which would have some consistency in direction for cores drilled with a preferred orientation in a quarry face, and subsequently stored upright in the laboratory.

(d) Magnetic Anisotropy

In magnetite bearing rocks such as the Gurnigel flysch, shape anisotropy in magnetite usually controls the form of the susceptibility ellipsoid. After alternating field demagnetization, the susceptibility of the samples was measured using the ScT cryogenic magnetometer in susceptibility mode as described by Scriba and Heller (1978). The susceptibility in the flysch is anisotropic. The predominant feature of the plots of the major axes of the susceptibility ellipsoids (Fig. 12) is that the minimum susceptibility axes are oriented perpendicular to bedding. The susceptibility ellipsoids are oblate (disc shaped), in other words, the ratio $(k_{int}-k_{min})/(k_{max}-k_{int})$ is greater than unity. k_{max} and k_{int} are not always distinguished (Channell et al. 1979) but in the case of Fayaux (Fig. 12) the maximum susceptibility axes coincide with the axes of flute casts. This oblate sedimentary fabric is typical for magnetite bearing rocks but the mode of occurrence of magnetite in these flysch leads us to believe that the fabric may not be due to the shape alignment of the magnetite grains themselves. The magnetite grains are smaller than a few microns and occur along the twin planes (010) of albite grains. The predominant cleavage planes of albite

are (001) and (010), therefore the elongated clusters of magnetite grains will tend to be oriented parallel to the maximum elongation axes of the host albite. The sedimentary fabric in the flysch is due to the orientation of albite and other detrital grains by gravitational and hydrodynamic forces. It is the coincidence of elongation axes of albites and the elongation of magnetite clusters that gives rise to a susceptibility anisotropy which reflects sedimentary processes.

3. Conclusion

Although not positively identified by X-ray and optical examination, the presence of magnetite in the flysch has been established by coercivity and blocking temperature spectrum analysis. VRM tests and IRM low temperature characteristics suggest that the magnetite grain size is predominantly in the single domain to superparamagnetic range. VRM has totally obliterated any primary magnetization which may have been present, and therefore the NRM of the flysch of Central Switzerland no longer records the direction of the ambient magnetic field at time of deposition. The susceptibility anisotropy records a sedimentary fabric and records preferential alignment of grains and hence palaeocurrent direction.

Acknowledgements. Jan van Stuijvenberg and Bill Lowrie provided help and useful comments.

References

Caron, C.: La nappe du Gurnigel dans les Préalpes. Eclogae Geol. Helv. **69**, 297–308, 1976

Channell, J.E.T., Heller, F., Van Stuijvenberg, J.: Magnetic susceptibility anisotropy as an indicator of sedimentary fabric in the Gurnigel Flysch. Ecologae Geol. Helv. **72**, 781–787, 1979

Dunlop, D.J., West, G.F.: An experimental evaluation of single domain theories. Rev. Geophys. **7**, 709–757, 1969

Heiniger, Ch., Heller, F.: A high temperature vector magnetometer. Geophys. J.R. Astron. Soc. **44**, 281–288

Hsü, K.J.: Paleocurrent structures and palaeogeography of the Ultrahelvetic Flysch Basins, Switzerland. Geol. Soc. Am. Bull. **71**, 577–610, 1960

Hubert, J.F.: Sedimentology of Prealpine Flysch sequences, Switzerland. J. Sediment. Petrol. **73**, 885–905, 1967

Lowrie, W., Channell, J.E.T., Heller, F.: On the credibility of Remanent Magnetization Measurements. Geophys. J.R. Astron. Soc. **60**, 493–496, 1980

Morel, R.: Geologie du Massif du Nirement (Préalpes romandes) et ses abords. Thesis Inst. Géol. Univ. Fribourg, 1978

Nagata, T., Kobayashi, K., Fuller, M.D.: Identification of magnetite and hematite in rocks by magnetic observations at low temperature. J. Geophys. Res. **69**, 2111–2120, 1964

Néel, L.: Some theoretical aspects of rock magnetism. Philos. Mag. [Suppl. Adv. Phys.] **4**, 191–243, 1955

Schaub, H.: Stratigraphie und Mikropaläontologie des Schlierenflysch, mit besonderer Berücksichtigung der paleocänen und untereocänen Nummiliten und Assilien. Schweiz. Palaeontol. Abh. **68**, 222p., 1951

Scriba, H., Heller, F.: Measurements of anisotropy of magnetic susceptibility using inductive magnetometers. J. Geophys. **44**, 341–352, 1978

Stuijvenberg, J. Van: Geology of the Gurnigel area (Préalpes, Switzerland). Matér. Cate géol. Suisse. Nouvelle série, 151^{e} livraison. p. 111, 1979

Syono, Y.: Magnetocrystalline anistoropy and magnetostriction of $Fe_3O_4-Fe_2TiO_4$ series, with special application to rock magnetism. Jpn. J. Geophys. **4**, 71–143, 1965

Trümpy, R.: The timing of orogenic events in the Central Alps. In: Gravity and tectonics, K.A. DeJong and R. Scholten, eds.: pp. 229–251. New York: John Wiley 1973

Received July 17, 1979; Revised Version March 18, 1980

J. Geophys. 48, 67–79, 1980

Journal of Geophysics

Source Parameters of Earthquakes From the Himalayan Region*

S.K. Upadhyay[1] and S.J. Duda[2]

[1] Department of Earth Sciences, University of Roorkee, Roorkee, U.P., India

[2] Institut für Geophysik, Universität Hamburg, Bundesstraße 55, D-2000 Hamburg 13, Federal Republic of Germany

Abstract. Source parameters of eight recent earthquakes from the Himalayan region are investigated, using single station digital data from the Central Seismological Observatory in Erlangen (FRG). Spectral displacement densities are obtained for both *P*- and *S*-waves and compensated for the effects of geometrical spreading and the instrument response in the applicable frequency range. Consideration is given to the effects of anelasticity and the radiation pattern. The compensated spectral densities are analysed in terms of theoretical faulting models. The seismic moment, the fault length and area, the stress drop and related quantities are estimated. Some evidence is presented as to the source time function and the anelasticity of the area.

Only one half of the earthquakes analysed are found to comply with the one-corner frequency model, the remaining earthquakes revealing two corner frequencies. The source time function as retrieved from the far field displacement shows approximately a cubic increase with time ($t^3 \exp(-vt)$). The earthquakes are found to be of the low stress drop/low moment type.

Key words: Earthquake source models – Seismic moment – Stress drop – Source time function – Himalaya region.

Introduction

Following Reid's (1910) 'Elastic Rebound Theory', various models have been proposed for the description of the process in the earthquake focus. The approaches can be classified as static and as kinematic. In the first case earthquakes are treated as the solution of a problem of elastostatics. Here no consideration is given to the action of inertial forces, and attention is stressed on the permanent deformation incurred subsequently to the earthquake process in the medium surrounding the focus. The empirical verification of such models resorts primarily to geodetic measurements of deformation around faults. The components of the strain tensor can then be determined and related to the seismic moment, the fault dimension, the stress drop and the seismic energy of earthquakes (e.g., Tsuboi 1939; Byerly and De Noyer 1958; Chinnery 1961, 1963). In practice, this approach is restricted to observations made at short epicentral distances and to near-surface events.

On the other hand, in kinematic models of the earthquake process, the inertial forces are taken into account and the earthquake is treated as the solution of a boundary value problem. In addition to the source parameters previously mentioned, new parameters enter the problem, as the polarisation angle of *S*-waves, the sense of first *P*-wave motion, the rupture velocity, the friction on the fault planes and the source time function. Thereby the source parameters are, at least in principle, retrievable from the seismograms of earthquakes.

Kinematic models have received considerable attention in recent years, due to the greater flexibility of the approach, and due to the relative ease of making observations at both small and large epicentral distances. In line with this approach the present paper is based on the analysis of seismograms for the purpose of obtaining source parameters of a suit of earthquakes from a particular region – the Himalayas.

Mathematically treatable source models assume either a displacement discontinuity across the fault plane (e.g., Haskell 1964, 1966, 1969; Savage 1966; Aki 1967), or a stress discontinuity due to a distribution of elementary forces acting at each point of the fault plane (e.g., Ben-Menahem 1962; Hirasawa and Stauder 1965). The source time function, i.e., the displacement or stress at each point of the fault plane as function of time, has been originally postulated to have the form of either a step (Aki 1968), or a ramp (Haskell 1964). Brune (1970) related – more realistically – the source time function to the effective stress available to accelerate the two sides of the fault.

On this basis, the aims of the present paper are the following:

(i) verification of known theoretical earthquake models with the help of eight well-recorded earthquakes from the Himalaya-region,

(ii) retrieval of the source parameters, and

(iii) analysis of the source time function characteristic of the earthquakes.

Data and Method of Analysis

Digital magnetic tape recordings of the earthquakes listed in Table 1 were obtained at the Central Seismological Observatory (Gräfenberg) at Erlangen (FRG). The position of the earthquakes with respect to the observatory is shown in Fig. 1. *P*- and *S*-waves were analysed for time intervals of 40 and 80 s and of 60 and 120 s respectively, beginning at the expected arrival time of each of the waves. The time intervals determine the lowest usable frequency. On the other hand, the Nyquist frequency (10 Hz) is determined by the digitisation interval, amounting to 1/20 s. Displacement amplitude spectral densities are obtained by way of the FFT-method. Corrections for the instrument response, the geometric spreading of the wave front and the effect

* The paper was prepared while Dr. Upadhyay stayed at the Institute of Geophysics of Hamburg University, as awardee of a fellowship of the A.v.Humboldt-Foundation

0340-062X/80/0048/0067/$02.60

Table 1. List of earthquakes analyzed

Earthquake No.	Date	Origin time	Magnitude[a] M_B	Magnitude[a] M_S	Depth (km)	Location of epincentre	Epicentral distance from Central Seismological Observatory	Source azimuth from Central Seismological Observatory
1	29.5.76	$12^h 23^m 18.7^S$	6.1	6.9	8	24.57° N 98.953° E	70.11°	75.103°
2	29.5.76	$14^h 00^m 18.5^S$	6.0	7.0	10	24.531° N 98.71° E	69.98°	75.303°
3	12.8.76	$23^h 26^m 46.2^S$	6.4	5.8[b]	27	26.68° N 97.07° E	67.41°	74.847°
4	6.11.76	$18^h 04^m 08.9^S$	5.8	6.5	33	27.605° N 101.05° E	69.21°	71.415°
5	21.7.76	$15^h 10^m 45.6^S$	5.8	6.3	9	24.782° N 98.698° E	69.79°	75.123°
6	9.6.76	$00^h 20^m 39.5^S$	5.7	5.9	33	24.894° N 98.752° E	69.75°	75.007°
7	1.1.77	$21^h 39^m 41.3^S$	5.9	6.3	27	38.146° N 91.007° E	55.86°	69.240°
8	19.1.77	$0^h 46^m 18.3^S$	5.9	5.8	33	37.022° N 95.697° E	59.42°	67.383°

Coordinates of recording station of the Central Seismological Observatory: 49° 41′ 31″ N, 11° 13′ 18″ E;

[a] Magnitudes published by USGS, unless otherwise mentioned

[b] Magnitude calculated from Erlangen records

of the free surface are introduced into the spectra by way of well-known methods (see Båth 1974). The velocity response of the instruments of the Central Seismological Observatory is shown in Fig. 2.

It may be mentioned that the earthquakes investigated are the strongest in the region under consideration, in recent time. Furthermore, it has been made sure that no microseismic storm took place during any of the earthquakes. In consequence of that, and in view of the quality of the recording station, the signal-to-noise ratio is as favourable as it can be for respective earthquakes observed in Central Europe.

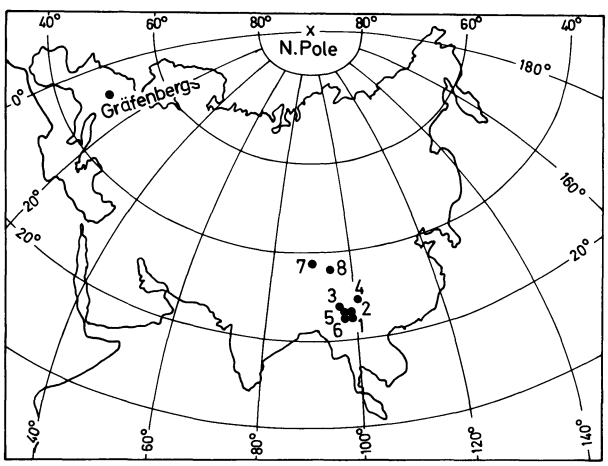


Fig. 1. Position of the epicentres in the Himalayan region with respect to the recording station. *Numbers* adjacent to the epicentre locations correspond to those used in the remaining figures and tables

Theory

The kinematic source models developed by Haskell (1964), Aki (1967), Berckhemer and Jacob (1968), Brune (1970) and Savage (1972) predict a far field radiation with the following common features:

(i) the low-frequency trend in the displacement density spectrum forms a line parallel to the frequency axis;
(ii) the displacement density spectrum decays at high frequencies;
(iii) the high-frequency and low-frequency trends in the spectrum intersect at one or at two frequencies, forming corners in the spectrum and defining corner frequencies;
(iv) corner frequencies are related to the dimensions of the source;
(v) the low-frequency value of the spectrum is determined by the seismic moment of the source;
(vi) the high-frequency decay of the spectrum is controlled by stress drop or by the effective stress (Brune 1970).

The following relations exist between the source parameters and the parameters characterising the displacement density spectrum of the radiated signal.

Seismic Moment M_0. According to Keilis-Borok (1959):

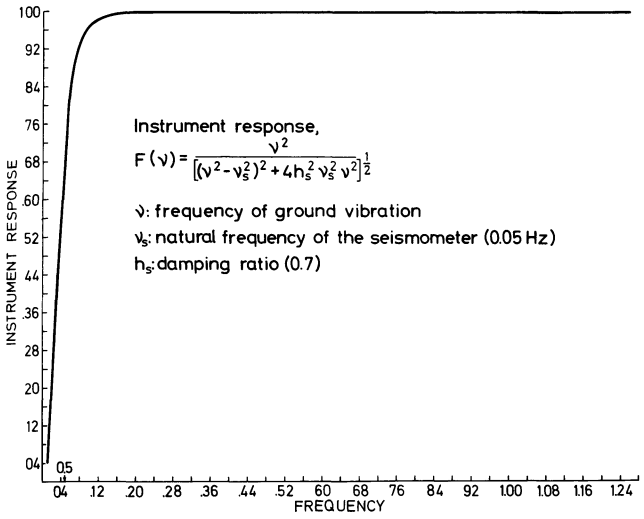


Fig. 2. Instrument response of seismometers used

$$M_0=\frac{4\pi\rho v^3\Omega_0(\omega)}{2RR_{\theta\phi}} \tag{1}$$

where

ρ = density near the seismic source, in gm/cm^3
R = factor accounting for the geometrical spreading of the wave front in a spherically stratified earth, in cm^{-1}
v = velocity of P- (α) and S-waves (β) in the source region, in cm/s
$\Omega_0(\omega)$ = low-frequency level of the displacement density spectrum, in cm/s
$R_{\theta\phi}$ = factor accounting for the radiation pattern of P- and of S-waves; $0\leqq|R_{\theta\phi}|\leqq 1$.

The amplification of the amplitude at the free surface is taken care of by the factor 2 in the denominator.

In Eq. (1) R is given by

$$R=\frac{1}{r_0}\left|\frac{\sin i_h}{\sin\Delta\cos i_0}\frac{di_h}{d\Delta}\right|^{\frac{1}{2}} \tag{2}$$

with

r_0 = radius of the Earth, in cm
i_h = angle of emergence at the focus at depth h of the ray emerging at epicentral distance Δ
i_0 = angle of incidence at the Earth's surface, at epicentral distance Δ (Shimshoni and Ben Menahem 1970).

Formula (1) is used for the determination of the seismic moment. For this purpose, in the present paper, the following numerical values are assumed for the respective quantities:

$\rho=2.7$ gm/cm^3
$\alpha=6.0$ km/s
$\beta=3.5$ km/s
$\mu=3.3\times10^{11}$ dynes/cm^2.

Fault Dimension. Formulae involving the fault dimensions have been derived for various source models. Depending on the model, the fault dimensions are functions of the corner frequency or frequencies f_i ($i=1,2,3$), the seismic velocities α, β at the focus, the rupture velocity v_R and the fractional stress drop ε. Table 2 shows the formulae for the models of Haskell (1964),

Table 2. Relation between corner frequencies f_i and fault dimensions for various fault models (Savage 1972)

Model	S-wave	P-wave
Haskell $\left(\frac{v_R}{\beta}=0.9\right)$		
One corner frequency:	$f_3=3.8\frac{\beta}{\sqrt{A}}$	$f_3=1.7\frac{\alpha}{\sqrt{A}}$
Two corner frequencies	$f_1=3.6\frac{\beta}{L}$	$f_1=1.2\frac{\alpha}{L}$
	$f_2=4.1\frac{\beta}{w}$	$f_2=2.4\frac{\alpha}{w}$
Brune ($L=2a$)		
One corner frequency	$f_3=4.1\beta\sqrt{\frac{1.6-0.6\varepsilon}{\varepsilon A}}$	$f_3=4.1\alpha\sqrt{\frac{1.6-0.6\varepsilon}{\varepsilon A}}$
Two corner frequencies	$f_1=4.7\frac{\beta}{L}$	$f_1=4.7\frac{\alpha}{L}$
	$f_2=4.7\frac{\beta}{L}\left(\frac{1.6-0.6\varepsilon}{\varepsilon}\right)$	$f_2=4.7\frac{\alpha}{L}\left(\frac{1.6-0.6\varepsilon}{\varepsilon}\right)$
Berckhemer and Jacob $\left(L=2a,\ \frac{v_{\max}}{\beta}=0.9\right)$		
One corner frequency	$f_3=2.5\frac{\beta}{L}$	$f_3=1.5\frac{\alpha}{L}$

v_R = average rupture velocity in cm/s. L, w, A = length (cm), width (cm) and area (cm^2) of fault, ε = (dimensionless) fractional stress drop. While Haskell assumes a rectangular model, both Brune, and Berckhemer and Jacob assume equivalent circular models (radius = a) of the fault. $f_3=\sqrt{f_1 f_2}$ applies

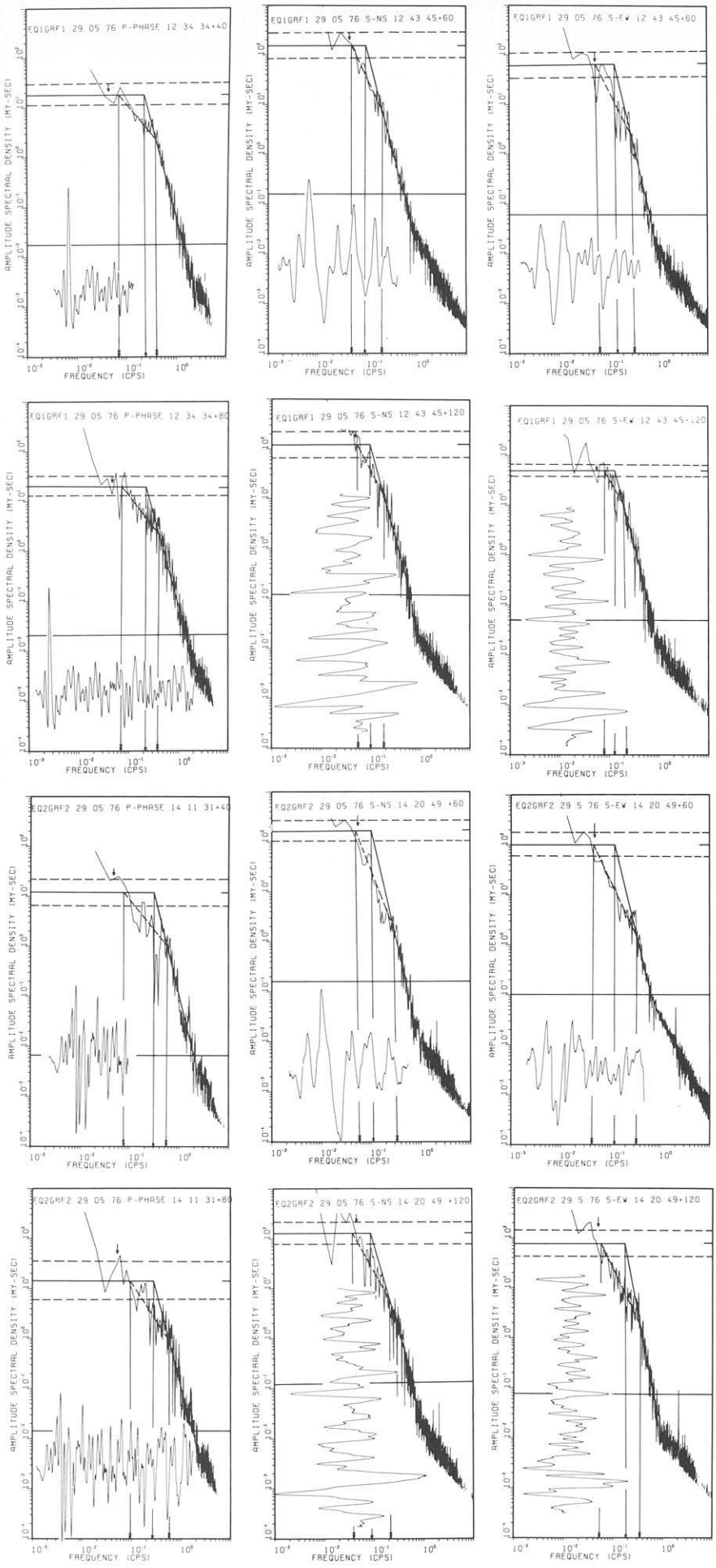


Fig. 3.
Seismograms and spectra corresponding to earthquakes No. 1 (upper portion) and No. 2 (lower portion).
Note: *P*-waves (vertical component) and N – S and E – W component of *S*-waves appear in the first, second and third column, respectively. For further explanation, refer to the text

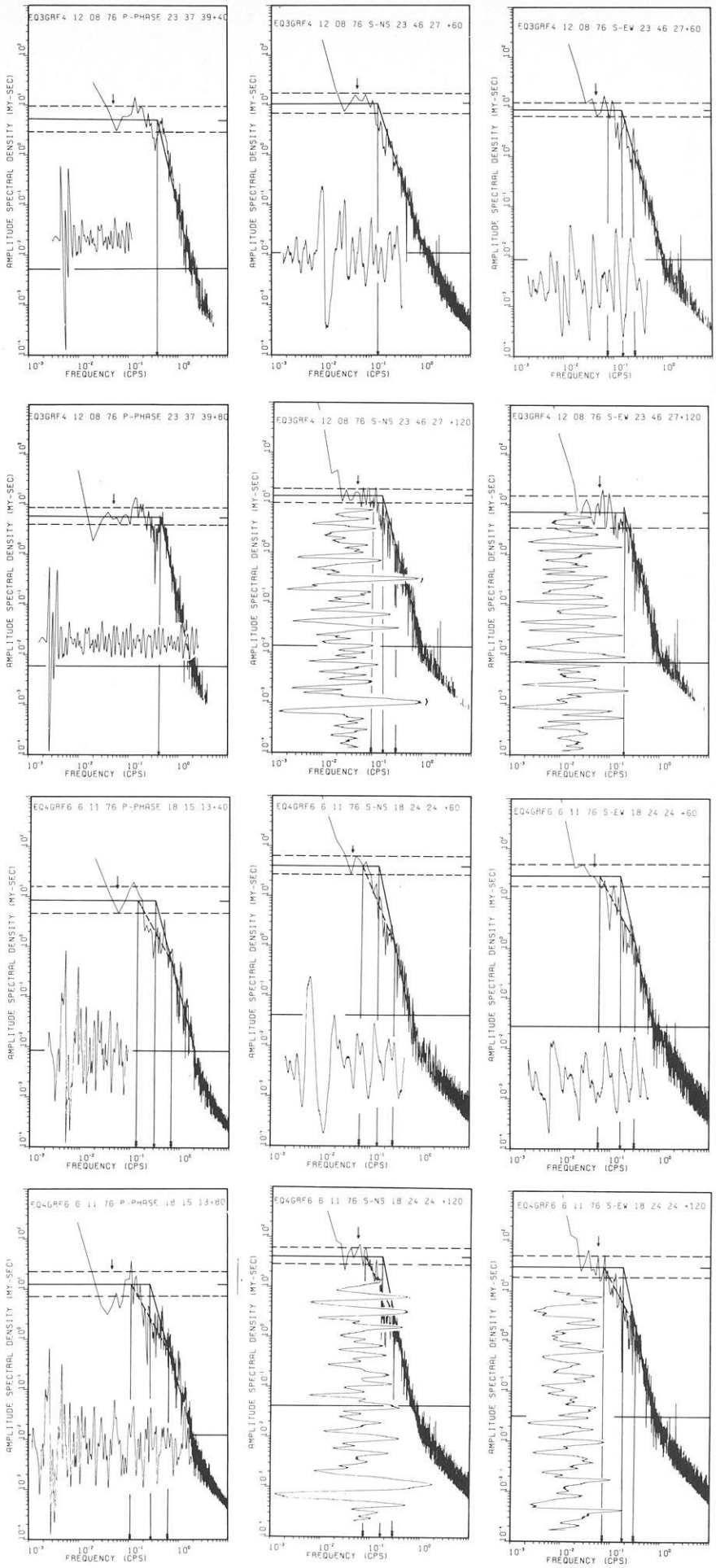


Fig. 4.
Seismograms and spectra corresponding to earthquakes No. 3 (upper portion) and No. 4 (lower portion). See Note of Fig. 3

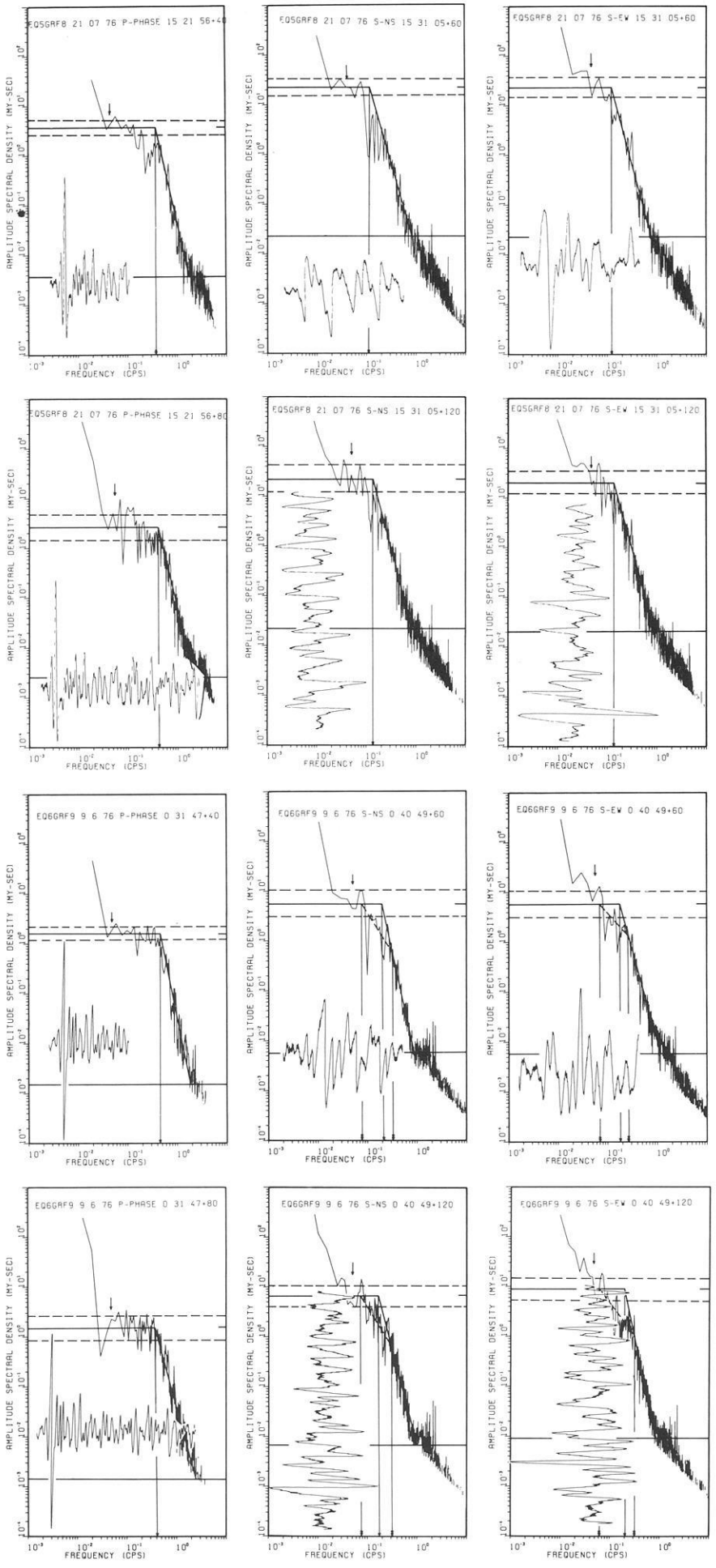


Fig. 5.
Seismograms and spectra corresponding to earthquakes No. 5 (upper portion) and No. 6 (lower portion). See Note of Fig. 3

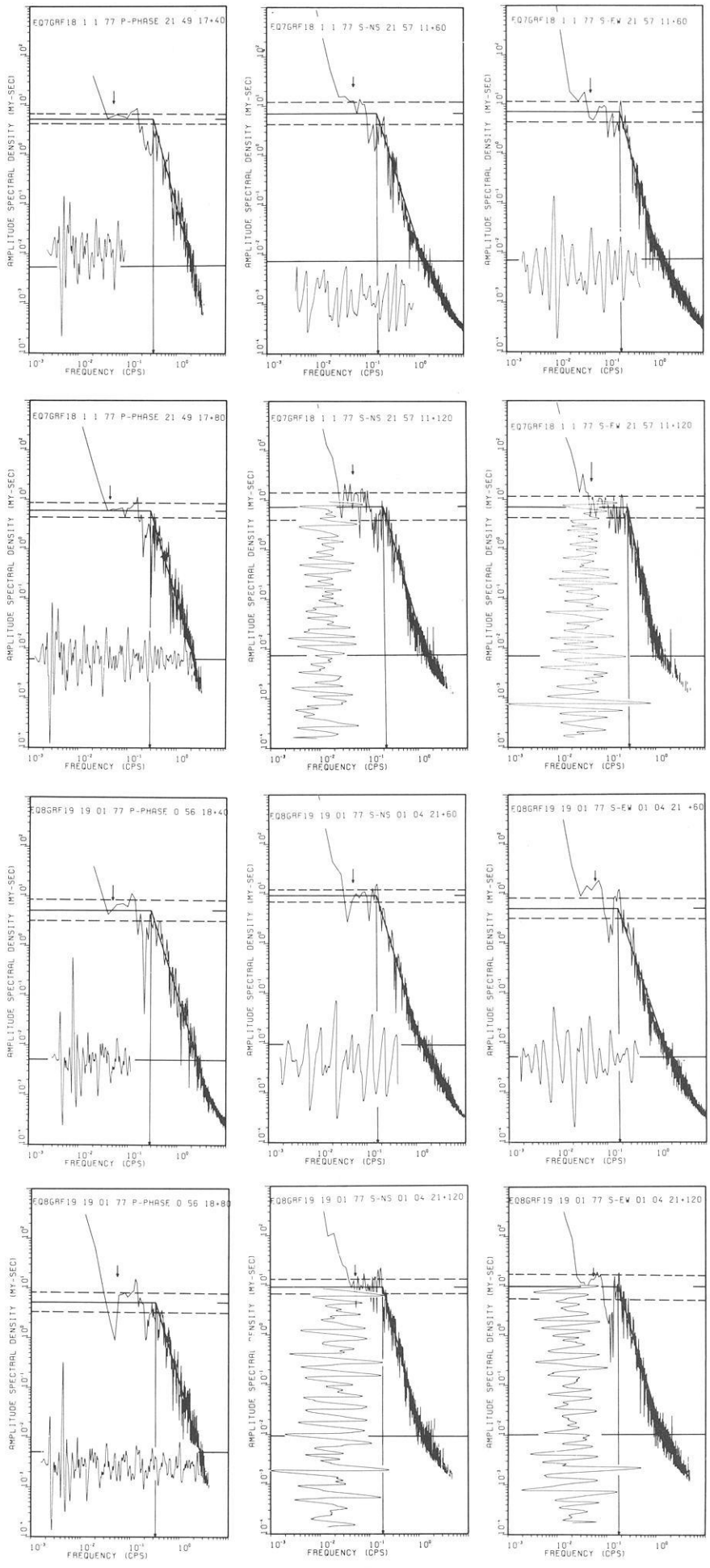


Fig. 6.
Seismograms and spectra corresponding to earthquakes No. 7 (upper portion) and No. 8 (lower portion). See Note of Fig. 3

Berckhemer and Jacob (1968), Brune (1970) and Savage (1972). In Haskell's model, the fault plane is assumed to be of rectangular shape, the width and length being ω and L respectively. Also, unilateral fracture propagation with a constant rupture velocity is here assumed. In Brune's model, the rupture propagation is neglected. Instead, the parameter ε is introduced which accounts for the fractional stress drop along the fault plane during the earthquake. In Berckhemer and Jacob's model, the effect of maximum reputure velocity, as opposed to average rupture velocities in the previous cases, is considered. In a generalisation of Haskell's (1964) model, a bilateral fracture propagation in opposite directions is discussed by Savage (1972). In this case, as also in Haskell's model, two corner frequencies are expected, provided that the ratio of fault length to width is large.

Average Dislocation and Stress Drop. Aki (1966) has related the seismic moment M_0 to the average dislocation $\bar{u}$, and the fault area A $(=L\cdot\omega)$ through the formula

$$\bar{u}=\frac{M_0}{\mu A} \tag{3}$$

where μ is the rigidity of the material in the earthquake focus. The seismic moment so found can be used for the estimation of the stress drop, $\Delta\sigma=(\sigma_1-\sigma_2)$, during an earthquake at the focus, whereby σ_1 and σ_2 are the average initial and final shear stresses on the fault surface respectively. According to Brune (1970) the corresponding relation for a circular fault plane with radius r is

$$\Delta\sigma=\frac{7}{16}\cdot\frac{M_0}{r^3}, \tag{4}$$

which readily can be estimated.

Structure of the Spectra and Numerical Results

Figures 3–6 show body wave displacement density spectra, corrected for the instrument response. Each of the graphs identifies in the top line the earthquake as to its number in Table 1, and the occurrence date, as well as the wave-type, the seismometer-component under consideration, the arrival time of the respective phase and the record length (40, 80 or 60, 120 s). The wave trains analysed are included in the lower left portion of each diagram. The arrow in the upper part of the diagrams indicates the frequency at which the response of the seismometer falls below 0.707 of the maximum value. Below this frequency the spectra become progressively less reliable. The steep rise of the spectra at the lowest frequencies is attributed to the truncation of the signal with a box car or Hamming window, in accordance with results of Linde and Sacks (1971). Nonetheless, a plateau is identified in all cases, though of variable quality. An approximation of the constant level and of the decaying part of the spectrum is attempted, as shown by the solid lines. The intersection of the two lines defines the corner frequency. Wherever applicable, the spectra are approximated by an additional, intermediate line (dashed). Two corner frequencies (f_1, f_2) are defined by the intersections of the dashed and the two solid (level and decaying) lines. The portion of the spectrum with values three decades below the level portion is neglected as a rule in the approximations. An error estimate of the level portion of the spectrum is indicated by horizontal, parallel lines (dashed) above and below the optimum value. The corner frequencies are shown by arrows on the frequency axis.

The spectra are found to have a definable constant level at low frequencies. The high frequency decay is controlled by one segment in case of the earthquakes Nos. 5, 7 and 8 yielding one corner frequency (type A earthquakes), and by two segments in case of earthquakes Nos. 1, 2, and 4 (type B earthquakes) yielding two corner frequencies. For earthquake No. 6, only the spectra of *S*-waves show two corner frequencies. Earthquake No. 3 yields spectra with one corner frequency though with two exceptions (*S*-wave on EW-(60 s) and NS-(120 s) components). Both earthquake Nos. 3 and 6 are labeled as type C (see Table 3).

The high-frequency decay of the spectra is approximated by the function const. $f^{-\gamma}$. Table 3 summarizes all pertinent values concerning the spectra. The values given in the table are averages of determinations made independently for the *P*-wave signals 40 and 80 s in length, and for the *S*-wave signals 60 and 120 s in length. For *S*-waves, the values from the two horizontal component records were averaged too. The respective corner frequencies as well as Ω_0 and $+\gamma$ revealed a good agreement, even prior to averaging (see Figs. 3–6), proving that these values are independent of the record length for a particular phase within the range investigated.

Table 4 gives the numerical values of the seismic moments, the stress drop and the average dislocation, as found with the formulae (1), (4), and (3), and the Table 2. Thereby, the moment from *S*-waves is obtained by way of a vector addition of the values found from the NS- and EW-component, as usually done. Also, for comparison, the seismic moments are given in accordance with the formula,

$$\log M_0=M_S+19.2 \tag{5}$$

valid in the magnitude range $6\leqq M_S\leqq 7$, with M_0 in dyne cm, and correlating the moment with the surface wave magnitude (Aki 1973). It can be seen that the moments of the 8 Himalaya-earthquakes determined from *S*- and from *P*-waves are in fair agreement, for each of the earthquakes, while at the same time they are lower on the average, by one order of magnitude, if compared with the ones found from Aki's formula.

The question arises as to the achievable accuracy of seismic moment determinations. In view of the difficulties in separating *P*- and *S*-phases from others, as *pP*, *sS*, *sP*, *pS*, core phases, etc., it is quite possible that the signals Fourier-analysed are the result of interference between phases. This can lead to an overestimate of the moment by a factor of 2–3. Thus, individual moment determinations can be accurate only to within about one order of magnitude. Even so, though, the deviation of moments of the Himalaya-earthquakes here investigated, from Aki's average relation appears to remain valid.

No clear dependence of either the stress drop or the average dislocation on magnitude is seen. All eight earthquakes feature however relatively small values of the stress drop, which fact labels the earthquakes as being of low stress drop type.

Table 4 further shows the fault length and area, as found from the spectra when using the three source models indicated. As a rule, the fault plane area found from *P*-waves is smaller than that from *S*-waves. In addition, differences among the estimates of upto one order of magnitude are seen, depending on the model employed. A comparison is made in Fig. 7 of all fault plane areas found here from *S*-waves, with fault plane areas given by Båth and Duda (1964) and by Chinnery (1969).

Table 3. Characteristics of displacement amplitude density spectra for *P*- and *S*-waves

Earthquake		Spectral characteristic of *P*-wave					Spectral characteristic of *S*-wave					
No.	Type	Ω_0 ($\times 10^{-4}$ cm-s)	$+\gamma$	f_1 (Hz)	f_2 (Hz)	f_3 (Hz)	Ω_0(NS) ($\times 10^{-4}$ cm-s)	Ω_0(EW) ($\times 10^{-4}$ cm-s)	$+\gamma$	f_1 (Hz)	f_2 (Hz)	f_3 (Hz)
1	B	19	3.47	0.07	0.39	(0.24)	150	55	4.57	0.061	0.23	(0.12)
2	B	14	3.92	0.086	0.57	(0.29)	150	89	4.52	0.047	0.30	(0.13)
3	C	5.65	4.49	–	–	0.41	12	8	3.75	0.075(?)	0.28(?)	(0.17)
4	B	10.25	3.61	0.13	0.65	(0.30)	40	30	3.82	0.065	0.29	(0.17)
5	A	3.15	4.33	–	–	0.40	22	22	3.66	–	–	0.13
6	C	1.6	4.33	–	–	0.41	6	7	4.27	0.071	0.30	(0.18)
7	A	5.65	3.47	–	–	0.30	7	7	4.38	–	–	0.22
8	A	5.1	3.14	–	–	0.30	9	7	3.86	–	–	0.18

γ = slope (negative) of high-frequency portion of spectrum
A = one corner frequency, *B* = two corner frequencies, *C* = mixed spectra – one or two corner frequencies

Table 4. Calculated values of seismic source parameters

Earthquake			Phase used	Moments		Stress drop (bars)	Average dislocation (cm)	Brune's model		Haskell's model			Berckhemer and Jacob's model	
No.	M[a]	Type[b]		Determined from spectra	from Aki (1973)			Length (km)	Area (km^2)	Length (km)	Width (km)	Area (km^2)	Length (km)	Area (km^2)
				$\times 10^{25}$	(dyne·cm)									
1	6.9	B	P	1.42	12.6	7.2	15.1	19	284	16	6	96	6	29
			S	2.05		3.6	12.8	24	451	33	10	324	12	106
2	7.0	B	P	0.80	15.8	8.5	14.5	15	168	13	4	54	5	19
			S	2.23		6.0	16.8	22	380	43	8	325	11	96
3	5.8	C	P	0.39	(1.0)	10.3	12.4	11	96	4		16	4	10
			S	0.18		3.3	2.9	16	189	27	8	218	8	53
4	6.5	B	P	0.69	5.0	7.1	11.8	15	177	9	4	31	5	18
			S	0.57		1.0	3.6	24	467	31	8	245	8	55
5	6.3	A	P	0.23	3.2	5.6	7.0	11	101	4		17	4	10
			S	0.39		1.3	3.9	20	305	16		272	11	87
6	5.9	C	P	0.10	(1.3)	2.9	3.3	11	96	4		16	3	10
			S	0.09		0.7	1.6	15	165	29	8	218	8	47
7	6.3	A	P	0.33	3.2	3.6	5.7	15	177	5		29	5	18
			S	0.10		1.7	2.6	12	113	10		95	6	31
8	5.8	A	P	0.32	(1.0)	3.0	5.3	15	184	5		29	5	19
			S	0.30		2.4	5.3	15	170	12		142	8	49

For one-corner frequency earthquake (Type A) Haskell's model predicts equal length and width of fault plane.
Moments outside of range of applicability of Aki's (5) formula are given in parantheses

[a] Compare Table 1
[b] Compare Table 3

The fault plane areas from the three models show a clear though gentle dependence on the surface wave magnitude. For the larger earthquakes, the areas are exceeded by those corresponding to Chinnery's, and to Båth and Duda's empirical formulae, while for the smaller magnitudes a better agreement of all estimates is found. This fact may bear some evidence concerning the strain release. The determination of the fault plane area by Båth and Duda was based on Benioff's hypothesis, according to which the aftershocks delineate the fault plane activated during the main shock. While this hypothesis is sup-

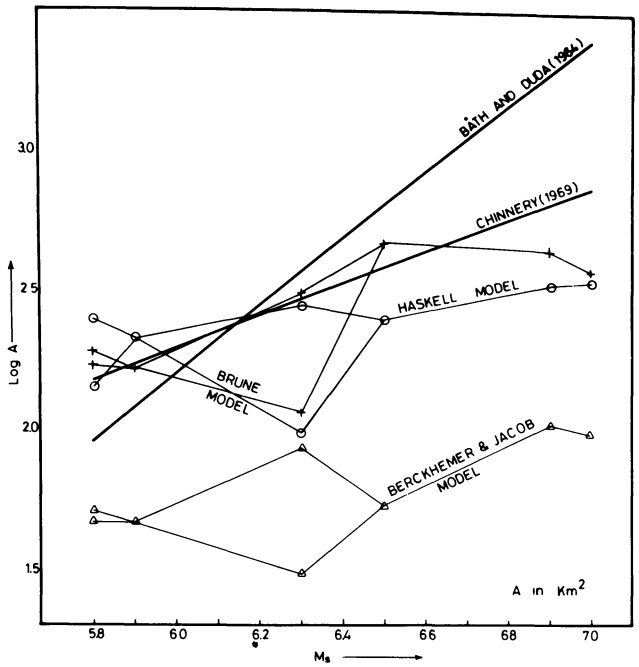


Fig. 7. Fault plane area as determined in the present study according to three different models of earthquake process. Included are results obtained earlier by different methods shown in *thick solid lines*

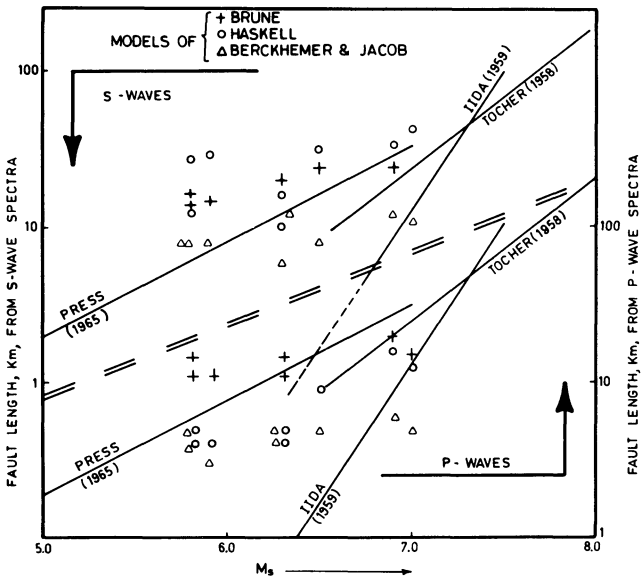


Fig. 8. Fault length from *S*-waves (upper portion) and from *P*-waves (lower portion) as determined in the present study according to three different models of the earthquake process. Included are the results obtained earlier by different methods shown by *solid lines*

ported by the present measurements only in the lower magnitude range, say from 5.8 to 6.4, it is contradicted for large magnitudes. In the latter range the main shock occurs on a smaller fault plane than is the plane delineated by subsequent aftershocks. The aftershocks let the originally activated fault plane grow, its area yielding an overestimate of the fault plane area of the main shock. This discrepancy of fault plane areas, as seen in Fig. 7 for magnitudes above 6.4, is viewed on the background of the relative abundance of aftershocks following the larger earthquakes. It indicates that the stress drop in a large magnitude main shock is less complete than in case of a main shock in the magnitude range, say, 5.8–6.4.

The disagreement of the Båth and Duda formula on one side, and the estimates of the fault plane area in the upper magnitude range on the other (Fig. 7), can also be a consequence of the fact that the coherence length K_L^{-1} and the coherence time K_T^{-1}, as defined by Aki (1967), are systematically exceeded by the fault length L and the characteristic time T. The possibility of an underestimate of the fault dimension under this condition was pointed out by Savage (1972). If this is the case, Benioff's hypothesis mentioned above continues to be valid in the entire magnitude range, along with Båth and Duda's fault plane formula.

Figure 8 compares the fault lengths found here from the three models, with empirical formulae which have been published by Tocher (1958), Iida (1959) and Press (1965). The estimates found in the present study are given separately for *P*- and for *S*-waves.

The formula by Iida shows the poorest agreement with the present results, while those by Tochter and Press appear to feature a too strong dependence on magnitude. The overall agreement of the results with the two empirical formulae is, however, satisfactory.

Discussion

Theoretical considerations lead to opposing views on the behaviour of seismic spectra at low frequencies. A constant level down to zero frequency is predicted in models of Haskell (1964), Aki (1967), Berckhemer and Jacob (1968), Brune (1970) and Savage (1972). On the other hand, a dominant period should prevail according to Kasahara (1957), Tsujiura (1967), Archambeau (1968) and Linde and Sacks (1971, 1972). In the first case, the constant level in the spectrum is supposed to reflect the moment of the equivalent double couple point source, while in the second case, the dominant period is thought to be a function of the source dimension. The empirical verification of either of the models is overdue, mainly because of the difficulties of obtaining reliable instrument responses at long periods. This limitation is equally faced in the present investigation. Though the spectra do reveal a constant level in part of the frequency range investigated, a subsequent drop at still lower frequencies cannot be excluded. The interpretation of spectra given in the present paper favours models predicting a constant level down to lowest frequencies. The empirical spectra obtained in course of the present study do not preclude however the possibility of the models predicting a dominant period to apply at least to some class of seismic events.

The spectra presented in Figs. 3–6 are corrected for the instrument response and for the geometric spreading of the wave front. They are thus more reliable in the low-frequency part, in the sense that they reflect the spectrum radiated from the source.

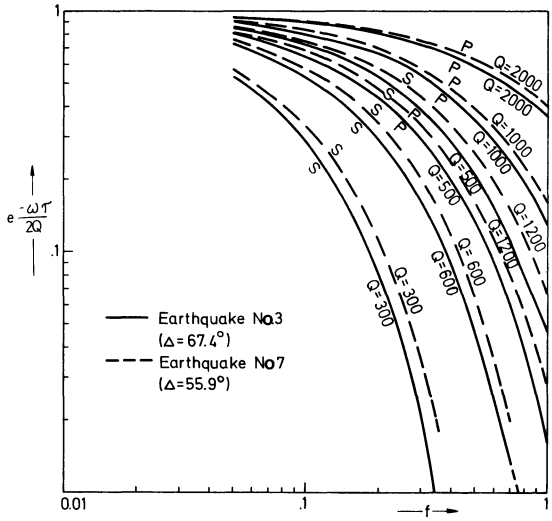


Fig. 9. Absorption effect as a function of frequency for two of the investigated earthquakes. A range of Q-values is assumed for P- and S-waves, as shown

A further correction for the attenuation would have to be applied, particularly in the high-frequency part if a progressive deterioration of reliability with frequency is to be prevented. The correction would require the knowledge of the distribution of the Q-factor along the ray path from the Himalayas to the recording station.

In Fig. 9 the factor is shown, by which the spectral densities of P- and S-waves would have to be divided in the applicable frequency range, for average Q-values along the ray path ranging from 300 to 2000. In Fig. 10a, b, as an example, the spectrum of the earthquake No. 3, as given in Fig. 4, is compensated for the absorption. Thereby three different Q-values are assumed in each case, as applicable for the respective wave. It can be readily seen that high average Q-values lead to only a slight change in the observed spectrum, while a decrease in the average Q-value yields a spectrum which not only progressively deviates from the observed one, but appears to become less reliable. This is due to the fact that Q is not likely to be constant along the ray path. To avoid the related uncertainties, observation of spectra at near epicentral distances are in need. In any case, the level portion of the spectra suffers least from absorption, which fact leaves moment estimates, and estimates of the corner frequency, virtually unaffected.

The problem of a global Q-profile, and of regional perturbations of such a profile, is under investigation since some time. Recently, Anderson and Hart (1978) have published four Q-profiles for S- and for P-waves, on the basis of abundant, world-wide estimates. Their values tend to be generally lower than hitherto believed. The average $\bar{Q}_\beta$ and $\bar{Q}_\alpha$ above the mantle-core boundary for the models SL2, SL3, SL7, and SL8 amounts to 240 and 600 respectively. Both values are in fair agreement with estimates of Nortmann (1977), who finds average Q-values of 300 and 500 resp. underneath the Eurasian continent.

For earthquakes at epicentral distances of about 65°, as in the present investigation, the ray paths have their lowest point at 1,600 km. The Q-profiles of Anderson and Hart predict that in this case the ray paths have longest segments in layers with Q_β close to 400, and Q_α close to 1,000. Thus, it is likely that the waves analysed here were mainly transmitted through layers with Q-values higher than the respective average values. We conclude that estimates of seismic moments and of corner frequencies, as given here, are only mildly distorted by anelasticity.

Moment (as well as magnitude and energy) determinations made at one station suffer from uncertainties due to the radiation pattern. If, however, a double couple source model is assumed, the radiation maxima of both kinds of body waves are rotated by 45° with respect to each other and the maximum of one should coincide with the minimum of the other. Though some information on the tectonic trend in the investigated area is known (Tandon and Srivastava 1975), a transposition of it into the probable orientation of the fault planes of the investigated earthquakes, and a subsequent determination of $R_{\theta\phi}$ [see (1)] is futile in view of other uncertainties. We observe systematically lower values of moments from P-waves, if compared with those from S-waves (see Table 4). Both are, however, always of the same order of magnitude for a given earthquake,

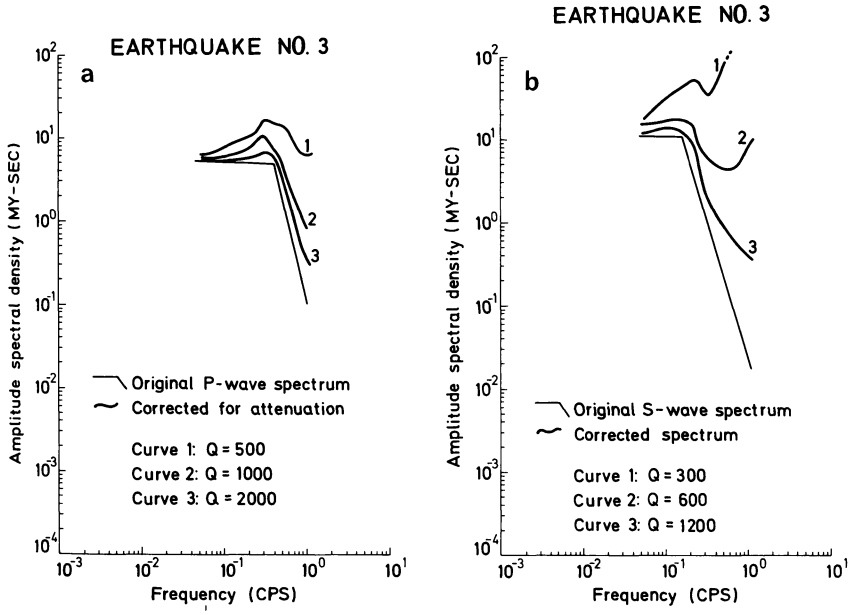


Fig. 10a and b.
Distortion of P-wave spectrum (**a**) and S-wave spectrum (**b**) for earthquake No. 3 (compare Fig. 4). Three different Q-values are assumed, as shown

indicating that the observation is made far away from a nodal line of either the *P*- or the *S*-wave. We consider this fact as justifying the neglection of the radiation pattern in the present case. For numerical calculations $R_{\theta\phi}$ is set equal to one.

The asymptotic behaviour of the spectra at high frequencies is reflected by the γ-coefficient. The numerical values for the eight earthquakes considered are given in Table 3. The averages of γ are 3.8 for *P*-waves and 4.1 for *S*-waves. If high *Q*-values are assumed along the ray path (e.g., 1,200 and 2,000 for *S*- and *P*-waves respectively; see Fig. 10), the observed γ-values can be interpreted in terms of build-up of slip at any point of the fault plane. As pointed out by Savage (1972), γ is determined then by the order of discontinuities in the pulse shape. Discontinuities in the displacement, velocity and acceleration of the initial motion result in $\gamma = 1, 2, 3$ respectively. A $\gamma \cong 4$, as found here, would imply an initial increase with time of the approximate form $t^3 \exp(-vt)$.

It is more likely however, that the γ-values determined from the spectra are influenced not only by the radiated spectrum, but by *Q* along the ray path as well. Whether the radiated spectrum in fact had a $\gamma = 3$, as corresponding to an initial displacement of the form $t^2 \exp(-vt)$ (Savage model), or $\gamma = 2$, corresponding to $t \exp(-vt)$ (Brune and Berckhemer and Jacob model), or even $\gamma = 1$ corresponding to a Heaviside unit step increase, could only be decided if detailed *Q*-values along the ray path were available.

We finally remark that, at variance with Haskell's model, but in agreement with the findings of Wyss and Hanks (1972) and Hanks and Wyss (1972), the corner frequencies for *P*-waves are about twice those for *S*-waves in case of the investigated Himalaya-earthquakes.

Conclusions

1. The spectra of 8 Himalaya-earthquakes are of both one-corner-frequency and two-corner-frequency type. While the fault length exceeds the width in the latter case, both, length and width are comparable in the former case.
2. The earthquakes are found to be of low-moment, low-stress drop type.
3. A discrepancy is seen between the fault plane areas as determined from the spectra and from the aftershock distribution for earthquakes with magnitudes above 6.4, but in agreement for lower magnitudes. From this, an only partial strain release is concluded in the large-magnitude earthquakes, which usually are followed by long aftershock sequences.
4. The stress drop and the fault length show only an unconspicuous dependence on magnitude.
5. Assuming no absorption losses along the ray path, the asymptotic behaviour of the spectra suggests an initial displacement proportional to $t^3 \exp(-vt)$.

Acknowledgement. The research was carried out while one of the authors (S.K.U) was an awardee of a fellowship of Alexander von Humboldt Foundation, Bonn-Bad Godesberg (FRG). The author expresses his thanks to the Director of the Institute of Geophysics, Hamburg University, Professor Dr. H. Menzel, for providing facilities to carry out the work.

Dipl. Geoph. R. Nortmann from the Institute of Geophysics, Hamburg University, made available some of his computer programmes. The basic data for the research were provided by the Central Seismological Observatory, Erlangen. Part of the research was performed under a grant of Deutsche Forschungsgemeinschaft, Bonn-Bad Godesberg.

All support is gratefully acknowledged.

References

Aki, K.: Generation and propagation of *G* waves from the Niigata earthquake of June 16, 1964. Part 2: Estimation of earthquake moment, released energy and stress-strain drop from the *G* wave spectrum. Bull. Earth. Res. Inst. **44**, 73–88, 1966

Aki, K.: Scaling law of seismic spectrum. J. Geophys. Res. **72**, 1217–1231, 1967

Aki, K.: Seismic displacements near a fault. J. Geophys. Res. **73**, 5359–5376, 1968

Aki, K.: Scaling law of earthquake source time function. Geophys. J. R. Astron. Soc. **31**, 3–25, 1973

Anderson, D.L., Hart, R.S.: *Q* of the Earth. J. Geophys. Res. **83**, 5869–5882, 1978

Archambeau, C.B.: General theory of elastodynamic source fields. Rev. Geophys. Space Phys. **6**, 241–288, 1968

Båth, M.: Spectral Analysis in Geophysics. Amsterdam, Oxford, New York: Elsevier Scientific Publishing Company 1974

Båth, M., Duda, S.J.: Earthquake volume, fault plane area, seismic energy, strain, deformation and related quantities. Ann. Geofis. **17**, 353–368, 1964

Ben-Menahem, A.: Radiation of seismic body waves from a finite moving source in the earth. J. Geophys. Res. **67**, 345–350, 1962

Berckhemer, H., Jacob, K.H.: Investigation of the dynamical process in earthquake foci by analyzing the pulse shape of body waves. Bericht des Institutes für Meteorologie und Geophysik der Universität Frankfurt/Main, Nr. 13, 1968

Brune, J.N.: Tectonic stress and the spectra of seismic shear waves from earthquakes. J. Geophys. Res. **75**, 4997–5009, 1970

Byerly, P., DeNoyer, J.: Energy in earthquakes as computed from geodetic observations. Contribution to Geophysics in honour of B. Gutenberg; H. Benioff et al. (eds.): pp. 17–35. New York: Pergamon Press 1958

Chinnery, M.A.: The deformation of the ground around surface faults. Bull. Seismol. Soc. Am. **51**, 355–372, 1961

Chinnery, M.A.: The stress changes that accompany strike slip faulting. Bull. Seismol. Soc. Am. **53**, 921–932, 1963

Chinnery, M.A.: Earthquake magnitude and source parameters. Bull. Seismol. Soc. Am. **59**, 1909–1982, 1969

Hanks, T.C., Wyss, M.: The use of the body wave spectra in the determination of seismic source parameters. Bull. Seismol. Soc. Am. **62**, 561–589, 1972

Haskell, N.A.: Total energy and energy spectral density of elastic wave radiation from propagating faults. Part I: Bull. Seismol. Soc. Am. **54**, 1811–1841, 1964. Part II: Bull. Seismol. Soc. Am. **56**, 125–140, 1966

Haskell, N.A.: Elastic displacements in the near field of a propagating fault. Bull. Seismol. Soc. Am. **59**, 865–908, 1969

Hirasawa, T., Stauder, W.: On seismic body waves from a finite moving source. Bull. Seismol. Soc. Am. **55**, 237–262, 1965

Iida, K.: Earthquake energy and earthquake faults. J. Earth. Sci. Nagoya Univ. **7**, 98–107, 1959

Kasahara, K.: The nature of seismic origins as inferred from seismological and geodetic observations. Bull. Earth. Res. Inst. Tokyo Uni. **35**, 473–532, 1957

Keilis-Borok, V.I.: An estimation of the displacement in an earthquake source and of the source dimensions. Ann. Geofis. **12**, 205–214, 1959

Linde, A.T., Sacks, I.S.: Errors in the spectral analysis of long period seismic body waves. J. Geophys. Res., **76**, 3326–3336, 1971

Linde, A.T., Sacks, I.S.: Dimension, Energy and Stress release for South American deep earthquakes. J. Geophys. Res. **77**, 1439–1451, 1972

Nortmann, R.: Die Anelastizität des Erdmantels im Europäischen Raum, Diplomarbeit, Institut für Geophysik der Universität Hamburg, Dezember 1977

Press, F.: Dimension of the source region for small shallow earthquakes, Proceedings of the Vesiac Conference on the current status, and future progress for understanding the source mechanism of shallow seismic events in the 3 to 5 magnitude range, 1965

Reid, H.: The mechanism of earthquakes; California earthquake of April 8, 1906, Report of State Investigation Commission (Carnegie Institution of Washington), Vol. 2, 1910

Savage, J.C.: Radiation from a realistic model of faulting. Bull. Seismol. Soc. Am. **56**, 577–592, 1966

Savage, J.C.: Relation of corner frequency to fault dimensions. J. Geophys. Res. **77**, 3788–3795, 1972

Shimshoni, M., Ben-Menaham, A.: Computation of the divergence coefficient for seismic phases. Geophys. J. R. Astron. Soc. **21**, 285–294, 1970

Tandon, A.N., Srivastava, H.N.: Fault plane solutions as related to known geological faults in and near India. Ann. Geofis. **28**, 13–27, 1975

Tocher, D.: Earthquake energy and ground breakage. Bull. Seismol. Soc. Am. **48**, 147–173, 1958

Tsuboi, C.: Deformation of the earth crust as disclosed by geodetic measurements. Beitr. Geophys., Ergeb. Kosm. Phys. **4**, 106–168, 1939

Tsujiura, M.: Frequency analysis of seismic waves, 2. Bull. Earth. Res. Inst. Tokyo Univ. **45**, 973–995, 1967

Wyss, M., Hanks, T.C.: The source parameters of the San Fernando earthquake inferred from teleseismic body waves. Bull. Seismol. Soc. Am. **62**, 591–602, 1972

Received May 22, 1979; Revised Version February 7, 1980

J. Geophys. 48, 80–83, 1980

Journal of Geophysics

Teleseismic Evidence for Velocity Heterogeneity Beneath the Rhenish Massif *

S. Raikes
Geophysikalisches Institut, Universität Karlsruhe, Kaiserstr. 12, D-7500 Karlsruhe 1, Federal Republic of Germany

Abstract. Observations of teleseismic *P* wave residuals for 56 stations in the vicinity of the Rhenish Massif show that arrivals within the Massif may be up to 0.6 s later than those immediately outside. Stations within the Massif also tend to have delays which are strongly azimuthally dependent (up to 1 s variation) in marked contrast to those outside (maximum 0.3 s variation). The strongest variation and delays are associated with the area of the Massif west of the Rhine, and preliminary modelling suggests they are caused by a low velocity region in the uppermost mantle (ca. 50–150 km depth) centred beneath the West Eifel volcanic field. Delays of up to 0.8 s, but with little azimuthal variation, are also found within the Vogelsberg volcanics, and are attributed to a shallower ($\lesssim$60 km) low velocity region.

Key words: Teleseismic *P* delays – Velocity heterogeneity – Upper mantle structure – Rhenish Massif.

Introduction

The Rhenish Massif is a relatively small (200×150 km^2) area of plateau uplift that has risen some 150 m during the Quaternary. The nature of this uplift and its causes are currently the subject of an intensive multidisciplinary research programme in Germany. As part of this, a detailed analysis of the variation of teleseismic *P* wave delays for stations in the area (Fig. 1) with azimuth and distance to the event is now in progress in order to obtain information about velocity variations within the lower crust and upper mantle beneath the Rhenish Massif. The technique has been described in detail elsewhere (e.g., Raikes 1976; Engdahl et al. 1977), and this paper only presents a summary of the observations, and their implications for the velocity structure.

Observations

Approximately 300 well-recorded events during the period 1976–1979 were chosen for this study, including 54 at distances over 120°; the azimuthal distribution is fairly good, except to the SE and NNW, although most events occur in the NE quadrant (Table 1). First arrivals were read with a precision of 0.1 s at most of the stations, and this data set was supplemented by bulletin times for a few extra stations (Fig. 1). Delays were calculated with respect to the Jeffreys-Bullen arrival times, corrected for the earth's ellipticity and station elevation, using the U.S.G.S. hypocentral parameters. For the core phases Bolt's (1968) times were used since these gave a better fit to the observed $dT/d\Delta$. In order to minimise effects other than those arising from structure beneath the stations, the residuals were then normalised by subtracting the delays at BUH. For a given source region the scatter in relative residuals at a single station was ± 0.05 to ± 0.15 s, except for stations where bulletin data were used, where it was generally ± 0.15 to ± 0.4 s.

Table 1. Distribution of events studied

Distance, degrees	Azimuth, degrees			
	0–90	90–180	180–270	270–360
0– 40	8	20	3	4
40– 80	59	12	21	12
80–105	64	–	12	25

Total 240+54 events at greater than 125°

A contour map of the average relative residuals for PKP phases from events in the South Pacific is shown in Fig. 2. The values of these residuals reflect structure directly beneath each station, because the rays are essentially vertically incident, and they are thus of great value in fixing the horizontal location of any anomalies. The main features of this map are the late arrivals associated with the Rhenish Massif and the station GLS in the Vogelsberg. Outside the Massif the residuals lie largely between 0 and 0.3 s except for a small area in the southern Rhinegraben and stations such as WTS where sediment corrections are required. Observations for other source regions are similar to Fig. 2, although the late arrivals associated with the western Rhenish Massif appear to be shifted directly away from the source, giving rise to large azimuthal variation of residuals (e.g. BNS -0.1 to $+0.9$ s, STB 0.02 to 0.92 s, BIR 0.1 to 0.9 s, ELG -0.09 to $+0.81$ s). There is a marked contrast between the western Massif and the eastern half which has earlier arrivals with less azimuthal variation (e.g. TNS 0.16 to 0.56 s). Outside the Rhenish Massif the residuals vary little with azimuth (e.g. CLZ -0.01 to $+0.23$ s, WTS 0.37 to 0.63 s, BFO -0.15 to $+0.07$ s, WLS -0.2 to $+0.1$ s) implying that the large variations in the western Massif are not simply the result of changes in the level of residuals at the reference station BUH. One exception is the station GRF which, as already

* Contribution No. 224, Geophysikalisches Institut, Universität Karlsruhe

0340-062X/80/0048/0080/$01.00

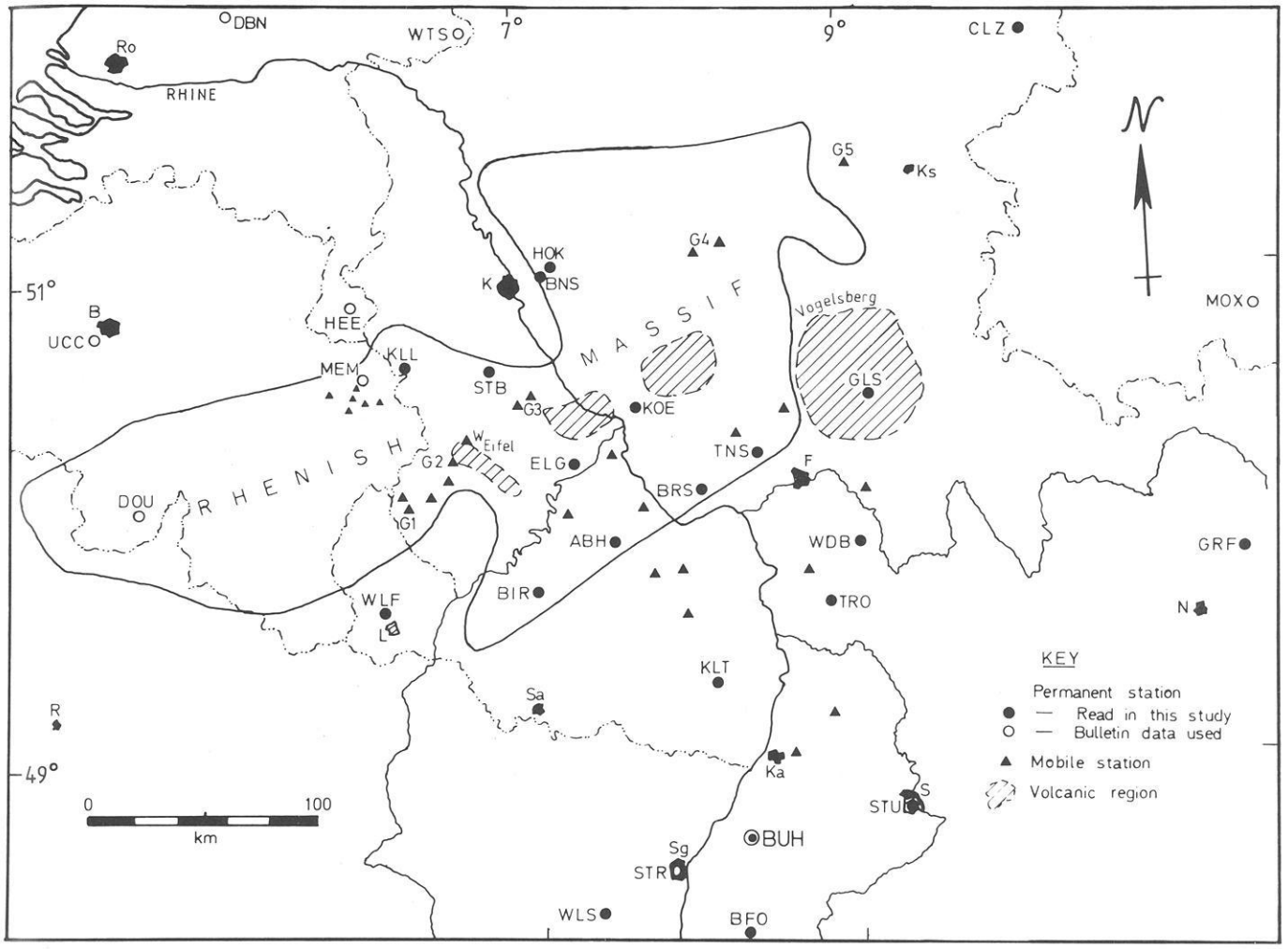


Fig. 1. Seismic stations used in this study

noted by Wenderoth (1978), has large positive residuals for events in the NE quadrant (0.6 to 1.2 s for azimuths 20° to 90° compared with 0.25 s for PKP and 0.3 to 0.5 s for azimuths 250° to 360°).

Implications

The observed pattern of teleseismic delays suggests that there are regions of low velocity beneath the Vogelsberg and the western half of the Rhenish Massif. The azimuthal variation (over 0.5 s) of residuals for stations in the latter area implies that they are caused by a velocity decrease in the upper mantle. Figure 3 shows a cross section through the Massif, with ray paths for events in the North Atlantic, Hindu Kush, and Sumatra regions, and the contours of residuals for the vertically incident PKP phases. The location of these ray paths and of those for other source regions suggests that the main low velocity regions lie at ca. 50 to 150 km depth beneath the western Massif, with the shallowest point beneath the West Eifel, and a shallower ($\lesssim 60$ km) region beneath the Vogelsberg. The velocity decrease depends on the depth extent of the low velocity region, but beneath the west Eifel it must be at least 3%. The residual variation is also compatible with a ca. 50 km thinning of the lithosphere beneath the western Rhenish Massif: a study of Rayleigh wave dispersion, which is more sensitive to S velocity variations, by Panza et al. (1979) indicates considerable variation in lithospheric thickness throughout Europe with thinning in the Rhinegraben – Rhenish Massif area.

More detailed modelling using automatic inversion procedures, and different normalisation schemes, plus increased mobile station coverage in the Belgian part of the Rhenish Massif, will allow better constraints to be placed on the anomalous regions. However, it is interesting that the should be great differences between the areas east and west of the Rhine. Low velocities are often associated with increased temperatures so it is perhaps reasonable that the major anomaly should be associated with the west Eifel, which is the area of most recent vulcanism. Currently available gravity data do not show a correspondingly strong anomaly in this area, so the velocity decrease is unlikely to be caused by a phase change (e.g. eclogite to garnet granulite, $\Delta p \sim 0.15$ gcm^{-3}). A more likely explanation appears to be the presence of partial melt ($\lesssim 1\%$, $\Delta p \sim 0.03$ gcm^{-3}) which would also be detected by deep electrical sounding techniques as it causes a marked increase in electrical conductivity.

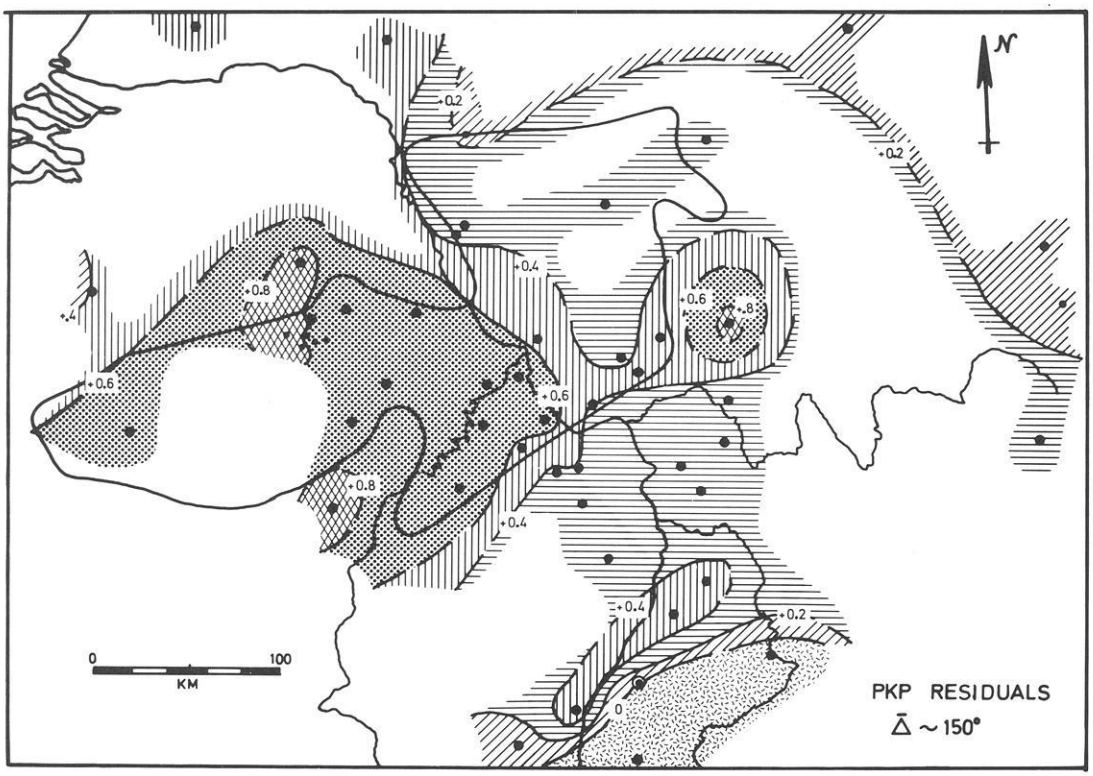


Fig. 2. Contour map of average delays, relative to BUH, for PKP phases from events in the South Pacific. The contour interval is 0.2 s; stations are indicated by *closed circles*

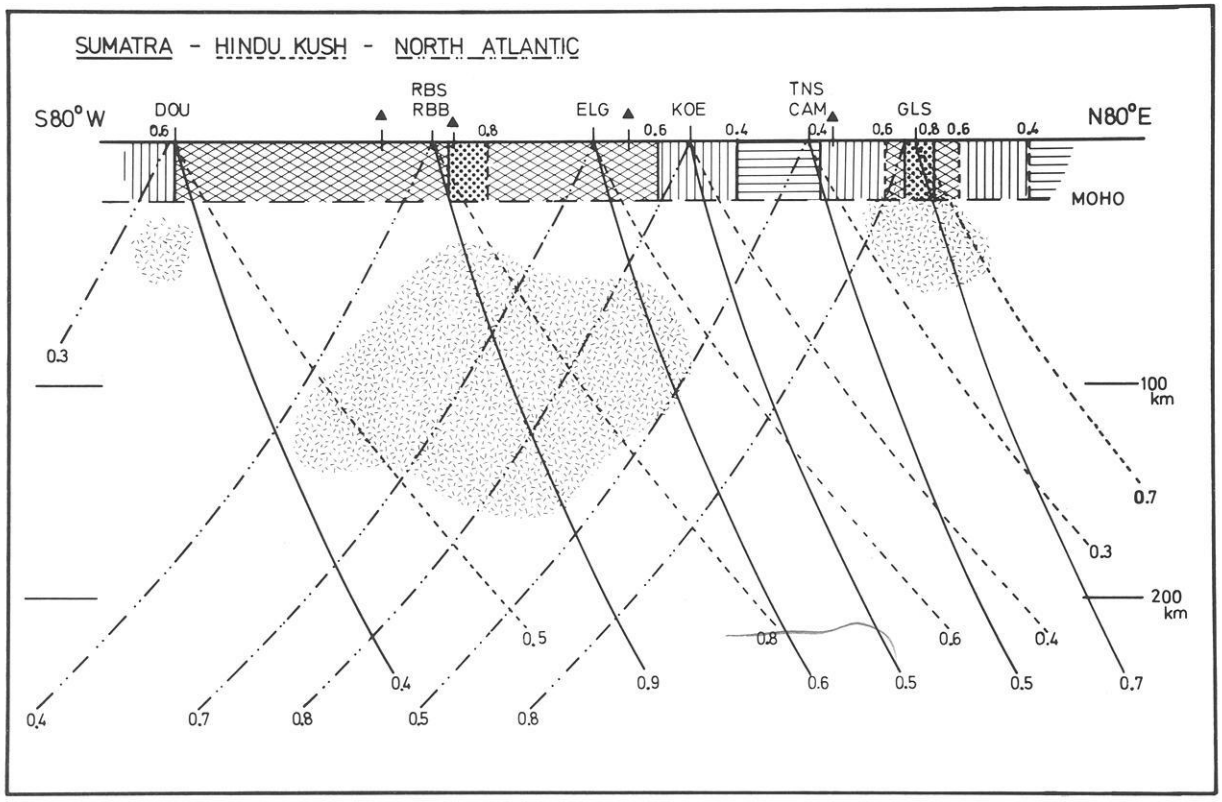


Fig. 3. Cross section through the Rhenish Massif showing ray paths to the named stations from events in the N. Atlantic (-·-·-·-), Hindu Kush (-----) and Sumatra (——) regions. Numbers at the end of each ray indicate the mean relative residual for that region, and the contours for PKP delays are projected on the crust. Further stations along the section, whose rays were omitted for clarity, are shown by *triangles*, and the inferred regions of low velocity are *hatched*

Acknowledgments. This work is supported by the Deutsche Forschungsgemeinschaft within the priority programme investigating the Rhenish Massif. I am indebted to Professor Dr. K. Fuchs for making this work possible, and to him and Dr. D. Bamford and Dr. K.-P. Bonjer for many interesting discussions. The following people kindly made data available: Drs. L. Ahorner, B. Baier, J. Flick, F. Keller, and J. van Gils. The work of many people, especially Dr. K.-P. Bonjer, Y. Godard, Dr. R. King, M. Koch, J. Mechie, and T. Moldoveanu, in operating the mobile stations is gratefully acknowledged. The Geostores were provided by the N.E.R.C. and operated in cooperation with Birmingham University. Mrs. Eichler typed the manuscript.

References

Bolt, B.A.: Estimation of PKP travel times. Bull. Seismol. Soc. Am. **58**, 1305–1324, 1968

Engdahl, E.R., Sinndorf, J.G., Eppley, R.A.: Interpretation of relative teleseismic *P* wave residuals. J. Geophys. Res. **82**, 5671–5682, 1977

Panza, G.F., Mueller, St., Calcagnile, G.: The gross features of the lithosphere-asthenosphere system in the European Mediterranean area. Phys. Earth Planet. Inter. in press, 1979

Raikes, S.A.: The azimuthal variation of teleseismic P-wave residuals in Southern California. Earth Planet. Sci. Lett. **29**, 367–372, 1976

Wenderoth, R.: Laufzeitanomalien teleseismischer *P* Wellen im Gebiet des Oberrheingrabens. Diplomarbeit, Geophys. Institut Universität Karlsruhe, p. 170, 1978

Received January 18, 1980; Revised Version March 20, 1980; Accepted March 26, 1980

J. Geophys. 48, 84–93, 1980

Journal of Geophysics

The Simulation Problem for Broad-Band Seismograms

D. Seidl

Seismologisches Zentralobservatorium Graefenberg, Krankenhausstr. 1–3, D-8520 Erlangen, Federal Republic of Germany

Abstract. A fundamental problem in the numerical data pre-processing of digital broad-band seismograms is the simulation of arbitrary analog seismograph systems, especially seismometer-galvanometer combinations. A special case of this simulation problem is the deconvolution or restitution problem as the realization of a wide-band seismograph system with a transfer function proportional to ground displacement, velocity or acceleration. The simulation problem can be solved by a digital cascade recursive filter using the bilinear z-transformation. Applications of the simulation filter are: a combined interpretation of digital broad-band and analog narrow-band recordings, a routine analysis of broad-band seismograms consistent with ordinary analog stations, the determination of the local magnitude from simulated Wood-Anderson seismograms and the restitution of broad-band recordings. The relationship between bandwidth, fine structure and information content of seismograms can be demonstrated in an obvious way by comparing broad-band recordings of the Graefenberg-array with simulated seismograms for different standardized seismometer-galvanometer systems.

Key words: Graefenberg-array – Broad-band seismograms – Simulation and restitution problem – Bandwidth and fine structure of seismograms.

Introduction

A fundamental problem in the pre-processing of digital broad-band seismograms is the simulation of arbitrary band-limited seismograph systems. A special case of this simulation problem is the restitution problem as a realization of a seismograph system whose transfer function is directly proportional to ground displacement, velocity or acceleration in the broadest possible frequency range.

While the application of simple band-pass filters is sufficient for routine analysis of broad-band seismograms (determination of arrival time, body wave magnitude, surface wave magnitude, etc.), for simultaneous processing of digital broad-band recordings and analog narrow-band seismograms it often proves necessary or desirable to have the most exact possible simulation of the corresponding analog seismograph system. The analysis of the Bukarest earthquake on March 4, 1977, is an instructive practical example of this simulation problem (Müller et al. 1978).

Data for the realization of the simulation problem are digital broad-band seismograms from the 3-component station A1 (A1-Z, A1-N, A1-E) of the Graefenberg subarray A. The recording is proportional to ground velocity between the 3-db frequencies 0.05 Hz (seismometer) and 5 Hz (anti-aliasing filter). The dynamic range of the final version of the data acquisition system is 132 db (not 138 db as it was intended; Harjes and Seidl 1978), the resolution is 66 db. Due to the large dynamic range and high resolution, as well as the small low-frequency noise level of the feedback leaf-spring seismometer STS-1V and STS-1H (Wielandt 1973, 1975; Streckeisen 1978) it is possible to vary the shape and bandwidth of the transfer function by digital methods to a large degree.

Simulation of Seismometer-Galvanometer Systems

Seismometer-galvanometer combinations are the most important narrow-band seismograph systems. A summary and classification of the standardized response functions is included in the Manual of Seismological Observatory Practice (Willmore and Karnik 1970). Figure 1 shows the amplitude response to ground displacement for the WWSSN-LP, WWSSN-SP, GALITZIN and KIRNOS system. The appropriate instrument parameters T_1, h_1, T_2, h_2 (natural period and damping constant of seismometer and galvanometer), as well as the period at maximum (T_{max}) and the periods at the 3-db points of the amplitude response (T_l, T_u) are summarized in Table 1.

An exact simulation of the WWSSN-SP system would require the incorporation of the inductive coil reactances of the seismometer (BENIOFF-seismometer of the variable reluctance type) and the galvanometer (Savill et al. 1962; Chakrabarty et al. 1964). For the practical processing of seismograms (determination of arrival time and body wave magnitude, improvement of signal-to-noise ratio) an approximation using a suitable induction free system is sufficient. As an example, Fig. 1 shows the amplitude response of a seismometer-galvanometer combination ($T_1=1.05$ s, $h_1=0.67$, $T_2=0.75$ s, $h_2=0.55$) which is a very good approximation of the magnification curve between 0.1 and 5.0 Hz for the WWSSN-station Stuttgart.

Neglecting the galvanometer coupling factor and omitting constant amplitude factors, the transfer function of the simulation filter $H_{SG}(s)$ which transforms the broad-band seismogram u into the simulated seismogram w is given by

$$H_{SG}(s)=H_0^{-1}(s)\cdot H_1(s)\cdot H_2(s). \tag{1}$$

$$H_0^{-1}(s)=\frac{s^2+2h_0\omega_0 s+\omega_0^2}{s^3} \tag{2}$$

0340-062X/80/0048/0084/$02.00

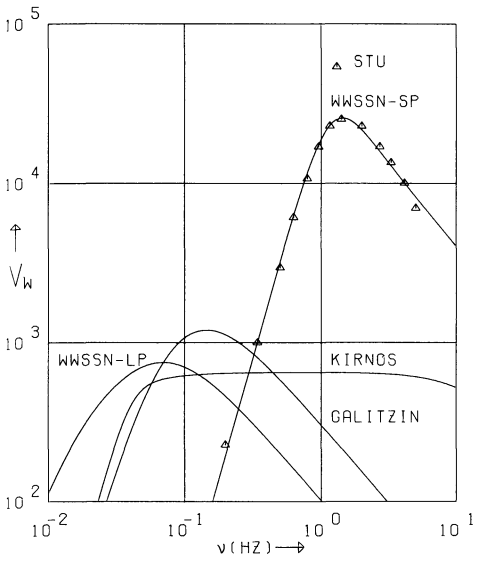


Fig. 1. Magnification curves of the seismometer-galvanometer systems WWSSN-SP, WWSSN-LP, GALITZIN and KIRNOS. Instrument parameters in Table 1. Δ measured magnifications of WWSSN-SP station Stuttgart (STU)

is the inverse transfer function of the broad-band seismometer (velocity transducer with the natural period $T_0=2\pi/\omega_0$ and damping constant h_0),

$$H_1(s)=\frac{s^3}{s^2+2h_1\omega_1 s+\omega_1^2} \tag{3}$$

is the transfer function of the simulated seismometer (velocity transducer with the natural period $T_1=2\pi/\omega_1$ and damping constant h_1), and

$$H_2(s)=\frac{\omega_2^2}{s^2+2h_2\omega_2 s+\omega_2^2} \tag{4}$$

is the transfer function of the simulated galvanometer (natural period $T_2=2\pi/\omega_2$, damping constant h_2).

Combining Eqs. (1) to (4) gives

$$H_{SG}(s)=H_E(s)\cdot H_G(s)=\frac{s^2+2h_0\omega_0 s+\omega_0^2}{s^2+2h_1\omega_1 s+\omega_1^2}\cdot\frac{\omega_2^2}{s^2+2h_2\omega_2 s+\omega_2^2}. \tag{5}$$

The first factor $H_E(s)$ in Eq. (5) is the transfer function of an inverse filtering circuit as described by Wielandt (1970). The second factor $H_G(s)$ corresponds to the low-pass characteristic of the galvanometer or a corresponding electronic filter (Wielandt and Mitronovas 1976).

Since $H_{SG}(s)$ in Eq. (5) takes the form of the product of two polynomial fractions in s, it is obvious that simulation can be approximated by a cascade of two recursion filters in the time domain (Beauchamp 1973). The transition from the s-plane of the continuous system to the z-plane of a sampled-data filter with the digitization interval Δt can be made with the standard z-transform or the bilinear z-transformation (Kaiser 1963). The second, so-called z-form method has the advantage that no aliasing problems occur. On the other hand, it has the disadvantage that the entire imaginary frequency axis of the s-plane is mapped onto the unit circle of the z-plane, producing a non-linear distortion of the frequency scale. If the transfer function $H(s)$ can be pieced together from pass- and rejection-bands with approximately constant slopes, then the frequency distortion can be compensated for by a prewarping of the corner frequencies ω_i $(i=0, 1, 2)$ of the continuous system to $\bar{\omega}_i$ according to the conversion formula (Kaiser 1963; Beauchamp 1973)

$$\bar{\omega}_i=\frac{2}{\Delta t}\cdot\tan\frac{\Delta t\cdot\omega_i}{2}=\frac{2}{\Delta t}\omega_i'. \tag{6}$$

The bilinear z-transformation

$$s=\frac{2}{\Delta t}\cdot\frac{z-1}{z+1} \tag{7}$$

then shifts the prewarped values $\bar{\omega}_i$ to the desired values ω_i.

Substituting ω_i for $\bar{\omega}_i$ after Eq. (6) and inserting Eq. (7) into Eq. (5) the z-transform $H_E^*(z)$ of the inverse filter factor becomes

$$H_E^*(z)=\frac{V^*(z)}{U^*(z)}=\frac{\left(\frac{z-1}{z+1}\right)^2+k_0'\frac{z-1}{z+1}+\omega_0'^2}{\left(\frac{z-1}{z+1}\right)^2+k_1'\frac{z-1}{z+1}+\omega_1'^2}=\frac{a_0+a_1z^{-1}+a_2z^{-2}}{1+b_1z^{-1}+b_2z^{-2}} \tag{8}$$

where $\omega_i'=\tan(\Delta t\cdot\omega_i/2)$ after Eq. (6) and $k_i'=2h_i\omega_i'$ $(i=0,1)$.

$V^*(z)$ and $U^*(z)$ are the z-transforms of the inverse filter output v_t and of the broad-band seismogram u_t respectively.

The filter coefficients in Eq. (8) are given by

Table 1. Parameters of standardized seismometer-galvanometer systems

	T_1 (s)	h_1	T_2 (s)	h_2	T_l (s)	T_{max} (s)	T_u (s)
WWSSN-LP (V_w)	15.0	1.0	100.0	1.0	6.06	14.40	32.29
WWSSN-SP (V_w)	1.05	0.67	0.75	0.55	0.42	0.69	1.02
GALITZIN (V_w)	12.0	1.0	12.0	1.0	3.15	6.93	12.98
KIRNOS (V_w)	25.0	0.5	1.2	8.0	0.08	1.45	22.47
BENIOFF-LP (V_G)	1.0	0.5	100.0	0.5	0.87	70.75	116.9

T_1, T_2: Natural period of seismometer resp. galvanometer
h_1, h_2: Damping constant of seismometer resp. galvanometer
T_{max}, T_l, T_u: Maximum period, lower and upper 3-db period of magnification curve for ground displacement (V_w) resp. ground velocity (V_G) in Fig. 1

$$a_0=\frac{1}{b_0}(\omega_0'^2+k_0'+1)$$

$$a_1=\frac{2}{b_0}(\omega_0'^2-1)$$

$$a_2=\frac{1}{b_0}(\omega_0'^2-k_0'+1)$$

$$b_1=\frac{2}{b_0}(\omega_1'^2-1)$$

$$b_2=\frac{1}{b_0}(\omega_1'^2-k_1'+1) \tag{9}$$

with

$$b_0=\omega_1'^2+k_1'+1.$$

Correspondingly, the z-transform of the galvanometer factor $H_G^*(z)$ in Eq. (5) becomes

$$H_G^*(z)=\frac{W^*(z)}{V^*(z)}=\frac{\omega_2'^2}{\left(\frac{z-1}{z+1}\right)^2+k_2'\frac{z-1}{z+1}+\omega_2'^2}$$

$$=\frac{c_0(1+2z^{-1}+z^{-2})}{1+d_1z^{-1}+d_2z^{-2}} \tag{10}$$

where $\omega_2'=\tan(\Delta t\cdot\omega_2/2)$ after Eq. (6) and $k_2'=2h_2\omega_2'$.

$W^*(z)$ is the z-transform of the simulated seismogram w_t. The filter coefficients in Eq. (10) are

$$c_0=\frac{1}{d_0}\omega_2'^2$$

$$d_1=\frac{2}{d_0}(\omega_2'^2-1)$$

$$d_2=\frac{1}{d_0}(\omega_2'^2-k_2'+1) \tag{11}$$

with

$$d_0=\omega_2'^2+k_2'+1.$$

Rearranging Eqs. (8) and (10) gives

$$V^*(z)\cdot(1+b_1z^{-1}+b_2z^{-2})=U^*(z)\cdot(a_0+a_1z^{-1}+a_2z^{-2}) \tag{12}$$

$$W^*(z)\cdot(1+d_1z^{-1}+d_2z^{-2})=V^*(z)\cdot c_0(1+2z^{-1}+z^{-2}). \tag{13}$$

Application of the shifting theorem

$$X^*(z)\cdot z^{-k}\equiv x_{t-k} \tag{14}$$

converts Eqs. (12) and (13) into difference equations:

$$v_t=a_0u_t+a_1u_{t-1}+a_2u_{t-2}-b_1v_{t-1}-b_2v_{t-2} \tag{15}$$

$$w_t=c_0(v_t+2v_{t-1}+v_{t-2})-d_1w_{t-1}-d_2w_{t-2}. \tag{16}$$

Eqs. (15) and (16) describe a two-step recursion filter which transforms the broad-band seismogram u_t into the simulated seismogram w_t. The output of the first step corresponds to the output of a seismometer with natural period $T_1=2\pi/\omega_1$ (digital inverse filtering). The upper cut-off frequency of w_t is determined by the anti-aliasing filter of the data acquisition system (corner-frequency 5.0 Hz, slope 42 db/octave). Simulation of long period systems is best accomplished by data reduction with a low-pass filter followed by decimation. The z-form method can also be applied to the simulation of the SRO-system resulting in a seven-step recursion filter for the long-period channel. A detailed description will be given in a subsequent paper.

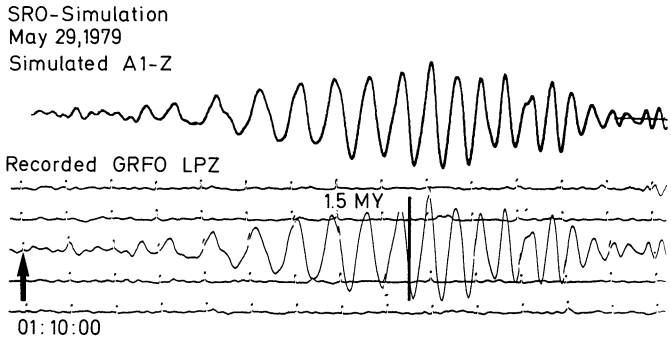
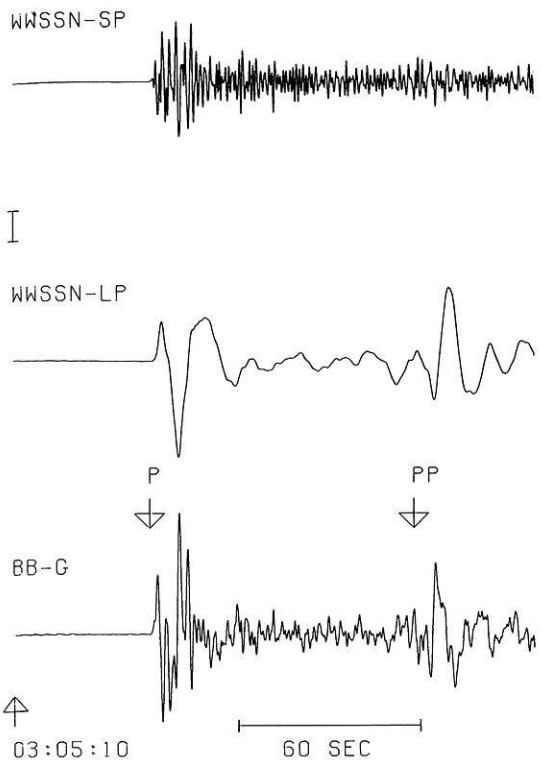


Fig. 2. Analog seismogram recorded by the LPZ-channel of the SRO-station GRFO and corresponding seismogram simulated from digital broad-band data of array station A1-Z. The horizontal distance between the stations is about 30 m, the vertical distance about 110 m

The simulation filter approximates the amplitude characteristic of the simulated system. The accuracy of the filter can be demonstrated most clearly by comparing simulated with recorded seismograms. Figure 2 shows an example for the long-period SRO-system. An example for a seismometer-galvanometer combination was given by Harjes and Seidl (1978). The most important applications of simulation filters are the simultaneous analysis of narrow-band and broad-band seismograms, routine processing compatible with analog stations, the approximate solution of the restitution problem, determination of local magnitude with simulated Wood-Anderson seismograms, as well as investigation of the dependence of the fine structure of seismograms as a function of the band-width and dynamic range of the seismograph system.

Broad-band seismograms contain a pronounced fine structure in the spectral range between 2 s and 20 s that is suppressed to a great extent by both the narrow-band analog stations and the new SRO-stations (Peterson et al. 1976). The main purpose of the Graefenberg-array is to record this fine structure in the frequency-wave number space with large dynamic range and high resolution. The broad-band seismograms in Figs. 3–11 are recordings of the 3-component station A1 of subarray A. Comparison between these broad-band recordings and the simulated seismograms of narrow-band systems makes it possible to demonstrate the relationship between fine structure or information content, band-width and dynamic range. The first example is given in Fig. 3. It shows the seismograms of the *P*-wave group from an Uzbekistan earthquake on May 7, 1976, for a broad-band recording BB-G proportional to ground velocity as well as the simulated seismograms for the WWSSN-LP and WWSSN-SP systems. The seismograms, which were calculated with a recursion filter, can also be visualized as the convolution of the broad-band seismogram with the velocity impulse response function of the seismometer-galvanometer system. The pulse-width of this function for the WWSSN-LP system is about 7 s. Fine structures that have a smaller time constant are smeared by convolution. The *P*-wave group in Fig. 3 consists of two signals which are 6 s apart. They appear in the WWSSN-LP seismogram only as a broadening of the impuls and are lost to direct interpretation (pP-phase of double shock). Only the arrival time



and the spectral amplitude in the frequency band around 1 Hz can be read from the amplitude modulated, quasi-monochromatic WWSSN-SP seismogram.

Figure 4 shows the total broad-band recording BB-G as well as the simulated seismograms for the KIRNOS, GALITZIN and WWSSN-LP systems and demonstrates the superiority of the WWSSN-LP system for recording long-period surface waves.

Figures 5 and 6 show seismograms corresponding to those shown in Figs. 3 and 4, taken from the Bukarest earthquake on March 4, 1977. The *P*-wave group of the broad-band seismogram BB-G (Fig. 5) is made up of at least three impulses, whose separation in time is less than the width of the impulse response of the WWSSN-LP system, so that the individual impulses only appear as shoulders in the simulated WWSSN-LP seismogram. By comparing the BB-G and WWSSN-LP seismograms in Fig. 5 Müller et al. (1978) identified and correlated corresponding shoulders in the recordings of many WWSSN-stations and were

Fig. 3. Broad-band recording BB-G (proportional to ground velocity, upper 3-db period $T_u = 30$ s; $V_G = 10.4$ μ/s) and simulated seismograms of WWSSN-LP ($V_{max} = 750$) and WWSSN-SP ($V_{max} = 25{,}000$) system of an Uzbekistan earthquake (May 17, 1976, 02 h 58 m 40.6 s, 40.4 °N, 63.5 °E, $m_b = 6.3$, $\Delta = 37.4°$). Magnification V_G is related to marked amplitude scale, maximum magnification V_{max} to 1 cm

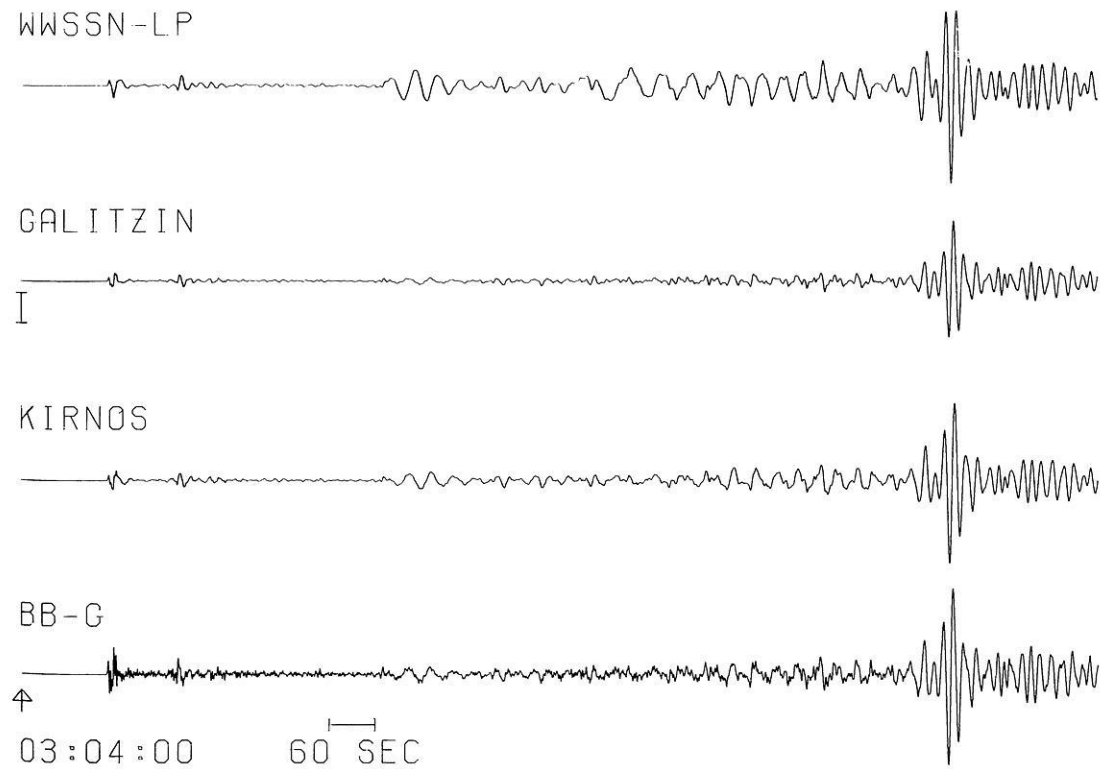


Fig. 4. Broad-band recording BB-G (proportional to ground velocity, $T_u = 30$ s; $V_G = 42$ μ/s) and simulated seismograms of KIRNOS, GALITZIN, and WWSSN-LP system ($V_{max} = 100$) for the same earthquake as in Fig. 3. Instrument parameters in Table 1

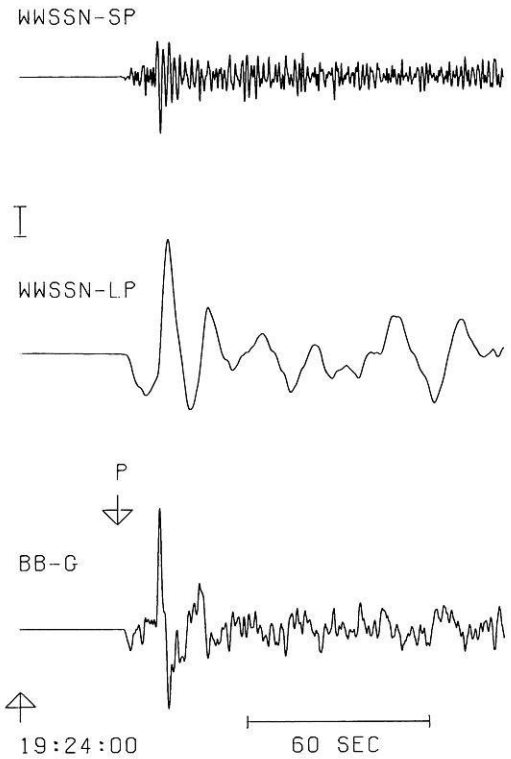


Fig. 5. Broad-band recording BB-G (proportional to ground velocity, $T_u = 30$ s, $V_G = 138$ μ/s) and simulated seismograms of WWSSN-LP ($V_{max} = 60$) and WWSSN-SP ($V_{max} = 2{,}500$) system of the Bukarest earthquake (March 4, 1977, 19 h 21 m 54.1 s, 45.8 °N, 26.8 °E, $m_b = 6.4$, $\Delta = 11.2°$)

thus able to show that the Bukarest earthquake was a multiple-shock.

Figure 7 shows a summary of simulated broad-band and narrow-band seismograms for a deep focus earthquake in the North Korea region on March 9, 1977. The *P*-wave group in the broad-band recording BB-G is made up of a low-frequency signal and a high-frequency phase arriving shortly after the first onset. The WWSSN-seismograms separate these two wave groups and contain in this example unlike to Fig. 3 and Fig. 5 the same information as the broad-band recording BB-G.

The seismograms in Figs. 3–7 demonstrate clearly the relationship between bandwidth, signalshape and information content. The narrow-band seismograms are more or less smooth whereas the broad-band recordings show a pronounced fine structure (for example multi-pulse signals, variation of amplitude and frequency with time) which cannot be completely reconstructed from digitized analog seismograms by deconvolution. The interpretation and inversion of this fine structure is the main object of broad-band seismology. Some examples should be mentioned: Phase identification and determination of source depth with theoretical seismograms (Kind 1979), determination of source parameters by spectral analysis (Stoll 1979) and by using the broad-band data as reference seismograms for the interpretation of analog recordings (Müller et al. 1978).

Simulation of a Wood-Anderson Seismograph

The local magnitude M_L is related to the maximum amplitude of a seismogram recorded on a Wood-Anderson seismograph. The usual practice of deriving M_L from some other kinds of

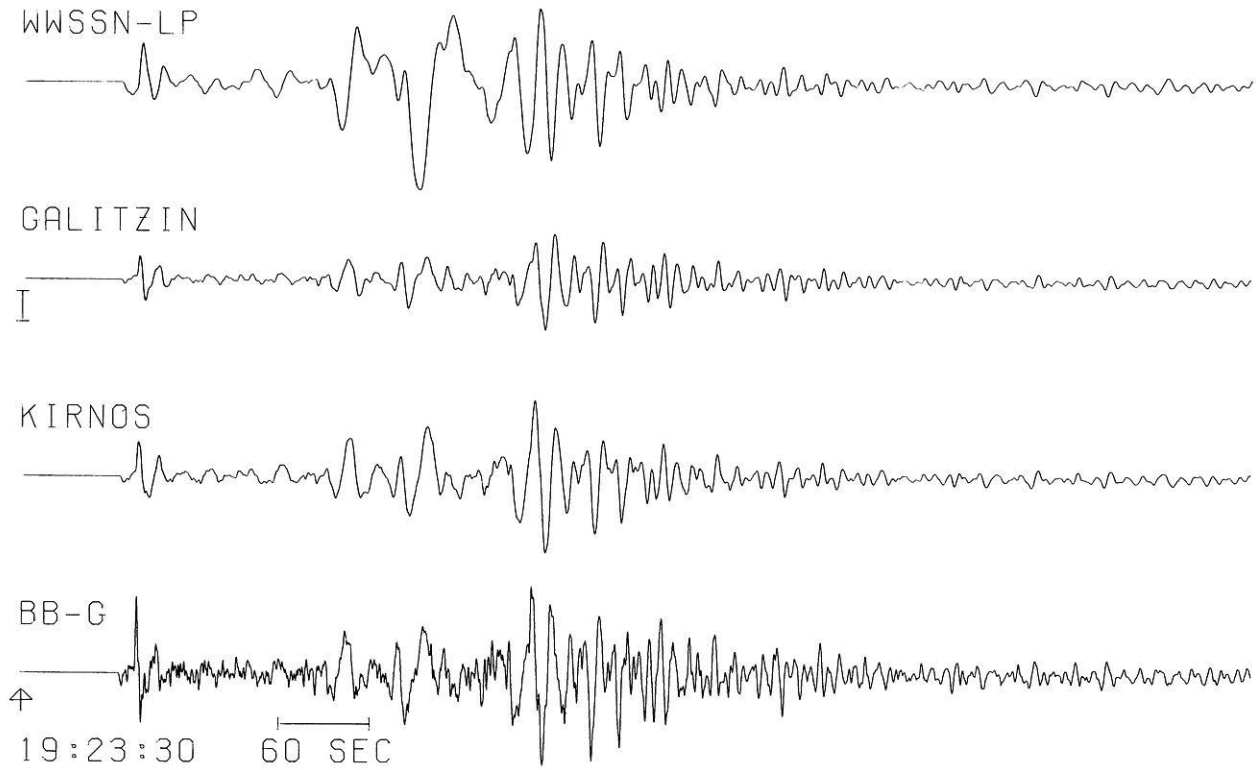


Fig. 6. Broad-band recording BB-G (proportional to ground velocity, $T_u = 30$ s, $V_G = 219$ μ/s) and simulated seismograms of KIRNOS, GALITZIN, and WWSSN-LP system ($V_{max} = 20$) for the same earthquake as in Fig. 5. Instrument parameters in Table 1

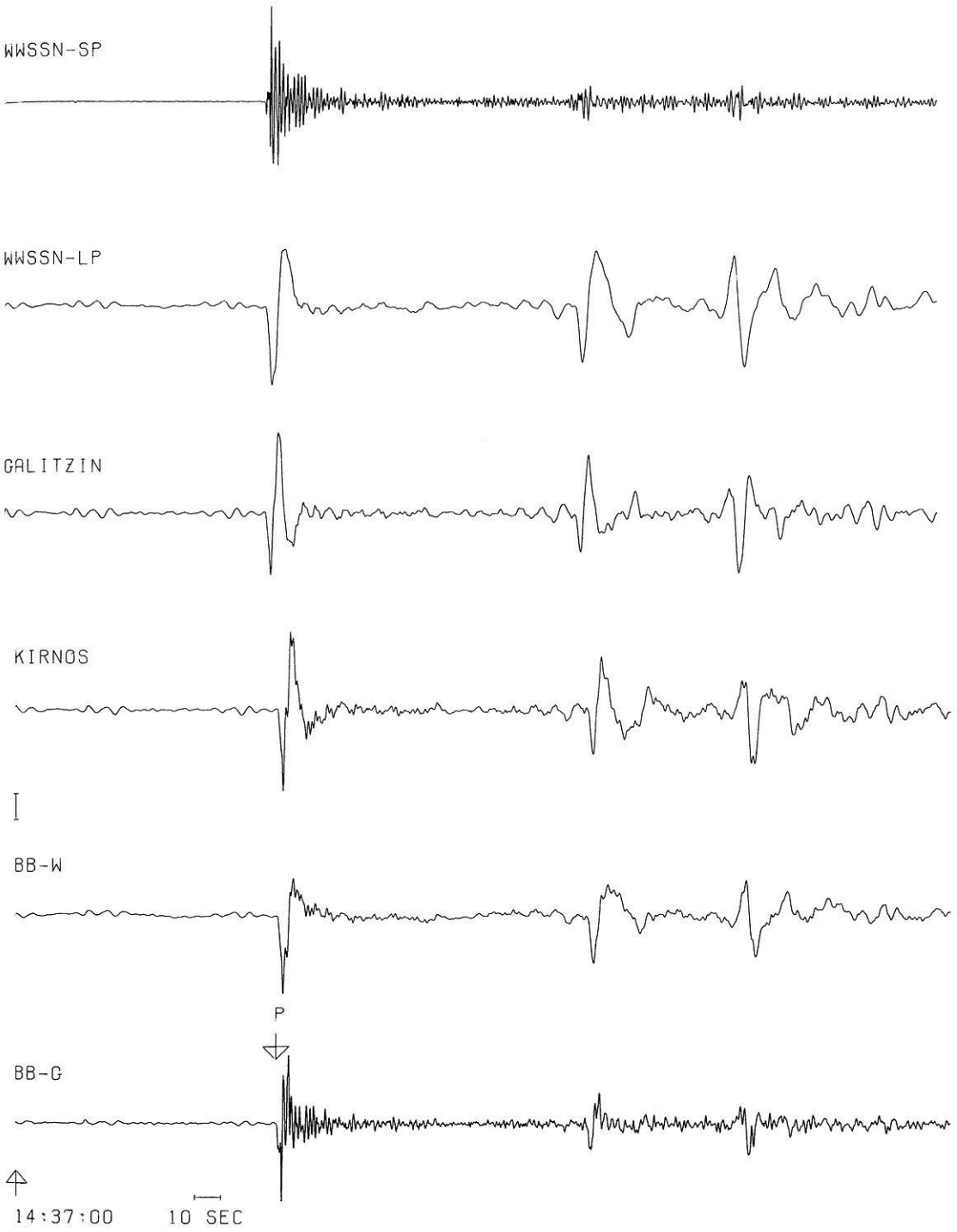


Fig. 7. Broad-band recording BB-G (proportional to ground velocity, $T_u = 100$ s, $V_G = 5.7$ μ/s) and BB-W (proportional to ground displacement, $T_u = 60$ s, $V_{max} = 1{,}750$) and simulated seismograms of the systems KIRNOS ($V_{max} = 2{,}500$), GALITZIN ($V_{max} = 3{,}000$), WWSSN-LP ($V_{max} = 3{,}000$) and WWSSN-SP ($V_{max} = 25{,}000$) of a deep focus earthquake in the North Korea region on March 9, 1977 (14 h 27 m 53.6 s, 41.6 °N, 130.9 °E, $h = 528$ km, $m_b = 5.9$, $\Delta = 74.8°$). Instrument parameters in Table 1

instruments is the calculation of equivalent Wood-Anderson amplitudes by point-by-point amplitude conversion. The transformation of short period records into synthetic Wood-Anderson seismograms by Fourier transform was described by Bakun et al. (1978). In the routine seismogram analysis at the Seismological Observatory Graefenberg M_L is calculated by simulating a Wood-Anderson seismograph in the time domain with a two-step recursion filter corresponding to Eqs. (15) and (16).

The transfer function $H_{WA}(\dot{s})$ of a continuous filter which transforms the broad-band seismogram u into the simulated seismogram x of a Wood-Anderson seismograph is given, in analogy to Eq. (1) by

$$H_{WA}(s) = H_0^{-1}(s) \cdot H_3(s) \tag{17}$$

$H_0^{-1}(s)$ is the inverse transfer function of the broad-band seismograph in Eq. (2) and

$$H_3(s) = V_3 \frac{s^2}{s^2 + 2h_3\omega_3 s + \omega_3^2} \tag{18}$$

is the transfer function of the Wood-Anderson seismograph (displacement transducer with the natural period $T_3 = 0.8$ s, damping constant $h_3 = 0.8$ and magnification $V_3 = 2{,}800$). From Eqs. (2), (17), and (18), and omitting the gain factor V_3 it follows

$$H_{WA}(s) = \frac{s^2 + 2h_0\omega_0 s + \omega_0^2}{s^2 + 2h_3\omega_3 s + \omega_3^2} \cdot \frac{1}{s}. \tag{19}$$

Application of the z-form method to Eq. (19) enables the simulation of a Wood-Anderson seismograph as a two-step recursion filter in the time domain in analogy to Eqs. (6) to (16):

$$v_t = f_0 u_t + f_1 u_{t-1} + f_2 u_{t-2} - g_1 v_{t-1} - g_2 v_{t-2} \tag{20}$$

$$x_t = h_0(v_t + v_{t-1}) + x_{t-1}. \tag{21}$$

The filter coefficients f_i and g_i are given by the coefficients a_i and b_i in Eq. (9) when ω_3' takes the place of ω_1' from Eq. (6) and k_1' is replaces by $k_3' = 2h_3\omega_3'$.

The recursion Eq. (21) which is derived from the factor 1/s in Eq. (19) corresponds to the trapezoidal rule for numerical integration with $h_0 = \Delta t/2$.

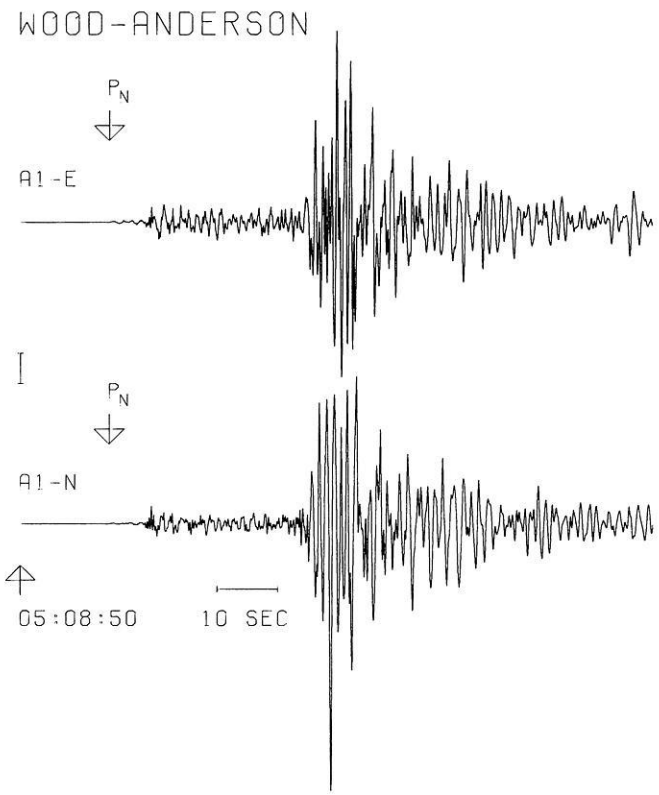


Fig. 8. Simulated Wood-Anderson seismograms of the Albstadt earthquake on September 03, 1978 (05 h 08 m 31.8 s, 48.28 °N, 9.03 °E, $h = 6.6$ km, $M_L = 5.9$, $\Delta = 224$ km) $V_{max} = 1{,}000$ related to marked amplitude scale of 1 cm

The Graefenberg-array, which was designed to record teleseismic signals has a frequency cut-off at 5.0 Hz determined by the anti-aliasing filter of the data acquisition system. Thus systematic errors can occur in the determination of M_L for near earthquakes with small epicenter distances. The seismic active area nearest to the Graefenberg-array is the Swabian Jura (mean distance about 200 km). The M_L values calculated with the Wood-Anderson simulation filter from Graefenberg broad-band recordings are generally in very good agreement with the local magnitudes derived from analog seismograms of the Swabian Jura station network (recording proportional to ground displacement between the 3-db periods 1.5 s and 50 s).

As an example Fig. 8 shows the simulated Wood-Anderson seismograms for the Albstadt earthquake on September 3, 1978. M_L was determined to be $M_L = 5.9$.

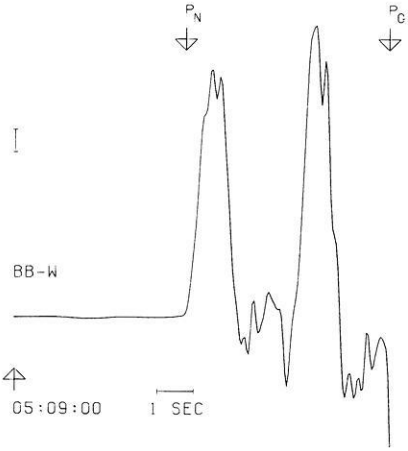


Fig. 9. Broad-band recording BB-W (proportional to ground displacement, $T_u = 22.5$ s, $V_{max} = 15{,}000$) of P_n-wave for the same earthquake as in Fig. 8

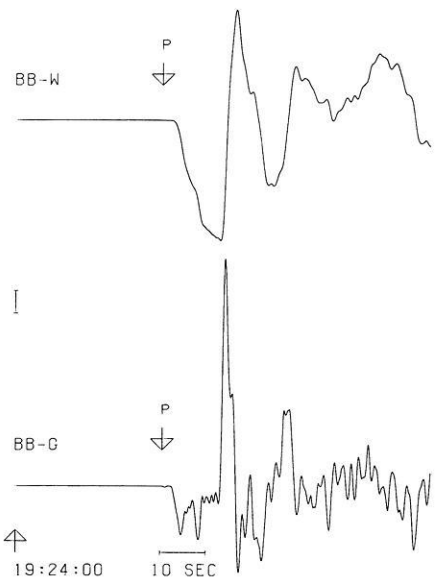


Fig. 10. Broad-band recording BB-W (proportional to ground displacement, $T_u = 172$ s, $V_{max} = 100$) and BB-G (proportional to ground velocity, $T_u = 200$ s, $V_G = 59$ μ/s) of P-wave for the same earthquake as in Fig. 5

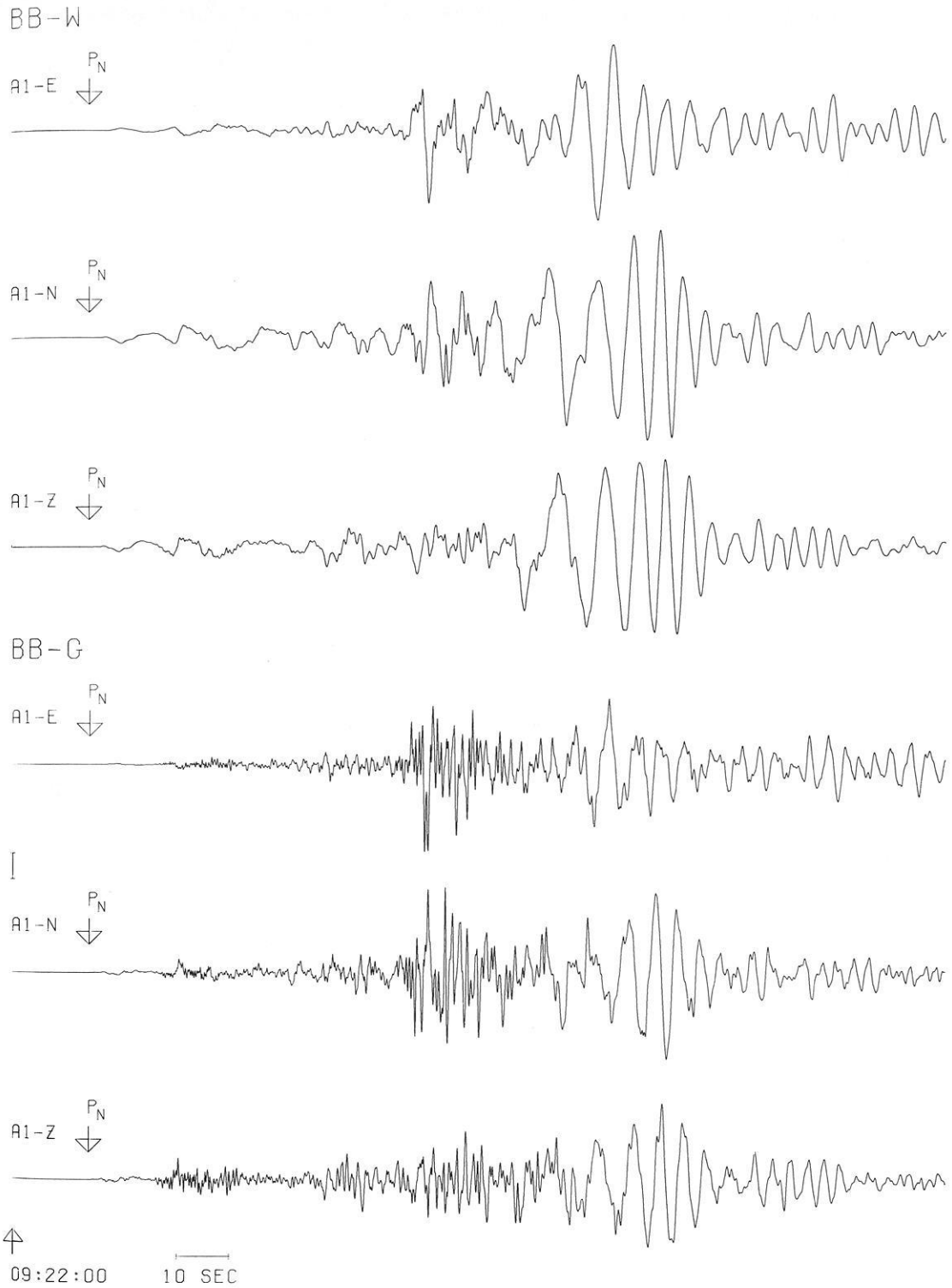


Fig. 11. 3-component recording BB-W (proportional to ground displacement, $T_u = 22.5$ s, $V_{max} = 50$) and BB-G (proportional to ground velocity, $T_u = 30$ s, $V_G = 367$ μ/s) of a Friuli earthquake on September 15, 1976 (09 h 21 m 19.1 s, 46.3 °N, 13.1 °E, $m_b = 5.4$, $\Delta = 3.6°$)

Restitution by Simulation

The restitution or restoration problem is the problem of reconstructing the input signal from the output signal of a band-limited system with a known transfer function.

The formal solution of this problem results from multiplication of the spectrum of the output signal with the inverse system transfer function and transformation to the time domain. Since the inverse transfer functions of seismograph systems with displacement or velocity transducer have a pole of second

or third order at $s=0$, the restitution problem in seismology, that is, the determination of the true ground motion in terms of displacement, velocity or acceleration, can only be solved approximately by shifting the pole to the left s-half plane. The accuracy of the approximation is determined less by the mathematical method than by the quality of the data (spectrum of the input signal, bandwidth, linearity, dynamic range, resolution and low-frequency instrument noise of the entire seismograph and recording system). A comparison of various methods (J. Geophys. 39, 501–626, 1973) has shown that a numerical restitution of digitized narrow-band analog seismograms usually does not produce useful results. In principal, the restitution of digitized broad-band seismograms has the same limitations, but better data quality allows sufficient accuracy for practical purposes (for example determination of seismic moment). A summary of various restitution methods in the time and frequency domain can be found in the above mentioned special issue 'Seminar on deconvolution of seismograms and high-fidelity seismometry' (J. Geophys. 39, 501–626, 1973).

Since the parameters for a seismometer-galvanometer system can always be chosen so that the transfer function is proportional to ground displacement, velocity or acceleration in a specified frequency range the restitution problem can also be solved by the simulation of an appropriately chosen seismometer-galvanometer combination.

The parameters (T_1,h_1, and T_2, h_2, natural period and damping constant of seismometer and galvanometer, respectively) of a seismometer-galvanometer system whose transfer function is proportional to ground displacement, velocity or acceleration in the period interval $T_A < T < T_B$ can be deduced from the following approximation formulas (Savarenski and Kirnos 1960):

Displacement

$$T_1 \gg T_B \quad h_1 = 0.4 \quad T_2 = \sqrt{T_A T_B} \quad h_2 \gg 1$$

Velocity

$$T_1 \gg T_B \quad h_1 = 0.4 \quad T_2 \ll T_A \quad h_2 \leqq 1.$$

Acceleration

$$T_1 \ll T_A \quad h_1 < 1 \quad T_2 = \sqrt{T_A T_B} \quad h_2 \gg 1.$$

In the above equations, the seismometer and galvanometer parameters can be interchanged.

An example of a displacement proportional system is given by the KIRNOS characteristic in Table 1. For the simulation filter, the upper 3-db period T_u can be raised to a maximum value determined only by the low-frequency signal to noise ratio. The maximum value for a specific seismogram is best determined by systematically raising T_u until the limit of the stable region is reached.

Using as an example the parameters

$$T_1 = 100\,\text{s}, h_1 = 0.4, \quad T_2 = 2\,\text{s}, h_2 = 10.0$$

one gets $T_u = 58$ s as an upper 3 db-period, whereas for

$$T_1 = 180\,\text{s}, h_1 = 0.5, \quad T_2 = 10\,\text{s}, h_2 = 8.0$$

this upper bound is $T_u = 172$ s.

A seismometer-galvanometer system with broad-band velocity recording is described in Table 1 by the long-period BENIOFF-LP characteristic. The velocity proportional seismograms BB-G in Figs. 3–11 were calculated from broad-band Graefenberg recordings through digital inverse filtering using Eq. (15). A seismometer-galvanometer system recording proportional to ground acceleration between the 3 db-periods $T_l = 0.08$ s and $T_u = 40$ s can be realised for example with the parameters $T_1 = 0.05$ s, $h_1 = 0.5$, $T_2 = 2.0$ s, $h_2 = 10.0$.

An important application of the restitution problem is the calculation of the displacement – time integral of the *P*-wave to determine the seismic moment. Figure 9 shows the displacement proportional seismogram of the P_n-wave group for the Albstadt earthquake on September 3, 1978. The first half-oscillation gives a value of $6.3\ 10^{-6}$ m s. Figures 10 and 11 show the restituted displacement and velocity proportional seismograms of the *P*-wave for the Bukarest earthquake on March 4, 1977, and a 3-component recording of the Friuli earthquake on September 15, 1976.

Acknowledgements. I thank H.-P. Harjes for many valuable and stimulating discussions and Mrs. M. Hellweg for translating the manuscript. I am greatly indebted to C. Kisslinger for critically reading the manuscript and for several helpful suggestions. The GRF-array project is supported by the Deutsche Forschungsgemeinschaft (German Research Council) and the Bundesanstalt für Geowissenschaften und Rohstoffe (Federal Institute for Geosciences and Natural Resources).

References

Bakun, W.H., Houck, S.T., Lee, W.H.K.: A direct comparison of synthetic and actual Wood-Anderson seismograms. Bull. Seismol. Soc. Am. **68**, 1199–1202, 1978

Beauchamp, K.G.: Signal processing. London: George Allen & Unwin LTD 1973

Chakrabarty, S.K., Choudhury, G.C., Roy Choudhery, S.N.: Magnification curves of electromagnetic seismographs. Bull. Seismol. Soc. Am. **54**, 1459–1471, 1964

Harjes, H.-P., Seidl, D.: Digital recording and analysis of broad-band seismic data at the Graefenberg (GRF)-array. J. Geophys. **44**, 511–523, 1978

Kaiser, J.F.: Design methods for sampled data filters. Proc. 1st Annu. Allerton Conf. Circuit System Theory pp. 221–236, 1963

Kind, R.: Observations of sPn from Swabian Alb earthquakes at the GRF array. J. Geophys. **45**, 337–340, 1979

Müller, G., Bonjer, K.-P., Stöckl, H., Enescu, D.: The Romanian earthquake of March 4, 1977. I. Rupture process inferred from fault-plane solution and multiple-event analysis. J. Geophys. **44**, 203–218, 1978

Peterson, J., Butler, H.M., Holcomb, L.G., Hutt, C.R.: The Seismic Research Observatory. Bull. Seismol. Soc. Am. **66**, 2049–2068, 1976

Savarenski, E.F., Kirnos, D.P.: Elemente der Seismologie und Seismometrie. Berlin: Akademie-Verlag 1960

Savill, R.A., Carpenter, E.W., Wright, J.K.: The derivation and solution of indicator equations for seismometer-galvanometer combinations including the effect of seismometer inductance. Geophys. J. R. Atron. Soc. **6**, 409–425, 1962

Stoll, D.: Spektralanalyse seismischer Wellen zur Beschreibung des Scherbruches bei den Erdbeben in Friaul (1976), 48 pp. Diplomarbeit, Universität Stuttgart 1979

Streckeisen, G.: Wide-band feedback seismograph system STS-1V/STS-1H. Prospekt der Firma G. Streckeisen

Meßgerätebau, Buchsweg 17, CH-8400 Winthertur, Schweiz 1978
Wielandt, E.: Ein einfacher elektronischer Entzerrer für Seismometer. J. Geophys. **36**, 763–769, 1970
Wielandt, E.: Noise in electronic seismograph systems. J. Geophys. **39**, 597–602, 1973
Wielandt, E.: Ein astasiertes Vertikalpendel mit tragender Blattfeder. J. Geophys. **41**, 545–547, 1975
Wielandt, E., Mitronovas, W.: An electronic long-period seismograph for surface-wave dispersion studies. Bull. Seismol. Soc. Am. **66**, 987–996, 1976
Willmore, P.L., Karnik, V.: Manual of seismological observatory practice. International Seismological Centre Edinburgh, Scotland 1970
Proceedings. Seminar on deconvolution of seismograms and high-fidelity seismometry. J. Geophys. **39**, 501–626, 1973

Received June 22, 1979; Revised Version February 15, 1980

J. Geophys. 48, 94–100, 1980

Journal of Geophysics

Load-Induced Stresses and Their Relation to the Initial Stress Field

G. Bock*
Geophysikalisches Institut, Universität Karlsruhe, Hertzstr. 16, D-7500 Karlsruhe 21, Federal Republic of Germany

Abstract. The additional stresses under a reservoir are computed by the analytical integration of the Boussinesq equations. Expressions are given for (1) a constant and (2) a linearly varying vertical load over a rectangular area. Given the direction of the primary stress field a simple calculation can be performed to check whether the additional stresses drive the total stress field towards the limiting failure threshold or not. This is demonstrated by a model lake of three-dimensional geometry. It is shown that not only in the case of normal faulting, but also in the case of thrust and transcurrent mechanisms, stresses are changed in the direction favourable for the occurrence of earthquakes by the water load alone. An application of the method to the seismic observations at the Schlegeis-Reservoir, Austria, indicates that a thrust-type mechanism of the seismic events as postulated in an earlier model is not compatible with the results of the theoretical calculations.

Key words: Initial stress field – Reservoir induced seismicity – Failure mechanism.

Introduction

Up to now more than 30 cases are known where changes of the natural seismicity are probably caused by the filling of large reservoirs (Simpson 1976; Gupta and Rastogi 1976). Earthquakes associated with dams cover a wide magnitude range, with an upper limit, as far observed, of about 6.25. Destructive ones like in Koyna, India, (Gupta et al. 1969) and Kremasta, Greece, (Comninakis et al. 1968) as well as minor earthquakes and changes in micro-earthquake activity have been related to the presence of artificial lakes. For a review see Simpson (1976). As indicated by observations at several sites, filling of a reservoir may even bring about a decrease of the natural seismicity (Simpson 1976; Bufe 1976). In spite of this diversity of the characteristic features of reservoir induced seismicity only two mechanisms are mainly discussed (Kisslinger 1976): Earthquakes near reservoirs can be triggered (1) by the incremental stresses induced by the water load (e.g., Gough and Gough 1970b) and (2) as a result of reservoir water migrating into the rock and increasing the pore pressure, which reduces the effective stresses (Hubert and Rubey 1959) thus facilitating failure. The last mechanism has also been proposed to explain the occurrence of seismic events associated with fluid injections into deep wells (e.g. Raleigh et al. 1976). According to Kisslinger (1976) a consensus is emerging that the increasing pore pressure is the most important factor controlling induced seismicity associated with dams. But no mechanism can be rejected up to now. In all cases the initial stress field must be already close enough to the limiting failure stress so that additional forces due to the reservoir, which are relatively low compared with the primary ones, are sufficient to trigger an earthquake. This idea is confirmed by fault-plane solutions of induced earthquakes which are consistent with the regional stress field.

In this paper the modification of the primary stress field by the water load is discussed. Several authors have described methods for the computation of the incremental stresses under artificial lakes (Gough 1969; Gough and Gough 1970a; Lee 1972; Nyland and Withers 1976; Withers and Nyland 1976). The modification of the initial tectonic stress field by the water load has been discussed by Snow (1972). Following Snow, earthquakes may be triggered by the load only in tectonic environments of normal faulting whereas transcurrent and thrust faults are stabilized. In his discussion Snow assumes a lake extended infinitely in the horizontal directions. Taking into account a lake with finite dimensions it may be imagined, however, that also earthquakes with strike slip and overthrust mechanism can be triggered by the water load. Recently, this has been demonstrated by Bell and Nur (1978) with two-dimensional lake models. It is the goal of this paper to illustrate the modification of the primary stress field by water loads with three-dimensional geometry.

Method of Computation

One method for the computation of the incremental elastic stresses under a lake of arbitrary shape has been presented by Gough and Gough (1970a). The water load of the lake is approximated by an array of vertical point forces which act at the surface of an elastic, homogeneous and isotropic halfspace. At a point within the halfspace the stresses due to one point force are given by the Boussinesq-equations. The total stress field is obtained by summation over all point forces. In this computation scheme singularities arise at the origin of the point forces. Therefore, the calculation of stresses is inaccurate at points situated close to the bottom of the lake. One method to

Contribution No. 221 of the Geophysical Institute, University of Karlsruhe. Sonderforschungsbereich Felsmechanik of Karlsruhe University

* *Present address:* Research School of Earth Sciences, Australian National University, P.O. Box 4, Canberra ACT 2600, Australia

0340-062X/80/0048/0094/$01.40

avoid these singularities makes use of spatially harmonic surface loads and the Fourier transform of the lake depth (Nyland and Withers 1976). Another method is the analytical integration of the Boussinesq-equations over areas of known load distributions. This concept has been applied in the present study. It proved to be very efficient in the computation of simple geometrical lake models as presented in a later section. The results of the analytical integration are given in the appendix for the following two load distributions over a rectangular area: (a) constant load and (b) linearly varying load. In the cartesian coordinate system used, x points to the East, y to the North and z in downward direction (left-handed system). To allow for other directions a local coordinate system x', y', z' is also introduced. In this x' is tangent to the lake bottom in the direction of increasing load, if case (b) of the above mentioned load distributions is investigated. The z' direction may be different from the z direction by an angle ψ which describes the dip of the lake bottom. The y'-axis is horizontal. If φ denotes the angle between the y- and y'-direction a point with coordinates (x, y, z) in the first system will have coordinates $(x' y', z')$ in the local system, given by

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = A \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

The 3×3 transformation matrix A is given by

$$A = \begin{pmatrix} -\sin\varphi\cos\psi & \cos\varphi\cos\psi & -\sin\psi \\ \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi\sin\psi & \cos\varphi\sin\psi & \cos\psi \end{pmatrix}.$$

The components of the stress tensor are computed in the local (x', y', z')-system. The stress tensor S in the (x, y, z)-system is obtained from the stress tensor S' in the local system by the transformation equation

$$S = A^T S' A$$

where T means transposition. The total stress field is obtained by summation over all elements by which the water load of the lake is approximated.

It is desirable now to compare this stress field with the initial one. In most cases nothing is known about the primary stress field. From geological and seismological observations, however, sometimes its direction can be deduced. Then a simple numerical calculation can be carried out to check whether the filling of a reservoir drives the stresses towards the failure envelope in a Mohr diagram or not. If σ_{10} and σ_{30} denote the greatest and least compressional stress and a_1, b_1, c_1 and a_3, b_3, c_3 their direction cosines in the (x, y, z)-system, the additional stresses in σ_{10} and σ_{30} direction are given by:

$\Delta\sigma_i = a_i^2\sigma_x + b_i^2\sigma_y + c_i^2\sigma_z + 2b_i c_i \tau_{yz} + 2c_i a_i \tau_{xz} + 2a_i b_i \tau_{xy}$ $(i=1$ and $3)$

σ_x, σ_y, σ_z, τ_{xy}, τ_{xz}, τ_{xy} are the stress components at a field point in the halfspace due to the water load. The Mohr circle of stress moves away from the failure envelope (stabilization) if

$$\frac{\Delta\sigma_3}{\Delta\sigma_1} > \frac{1-\sin\phi}{1+\sin\phi} \qquad \text{(A)}$$

and moves towards the failure envelope (destabilization) if

$$\frac{\Delta\sigma_3}{\Delta\sigma_1} < \frac{1-\sin\phi}{1+\sin\phi} \qquad \text{(B)}$$

(Snow 1972). For deriving these relations the Mohr envelope of failure is approximated by a straight line. Its slope is given by the angle of internal friction ϕ, which for rocks is commonly less than 45° (e.g. Handin 1966). Condition B is equivalent to the condition that $\Delta\tau/\Delta\sigma$ rises above $\tan\phi$. $\Delta\tau/\Delta\sigma$ is the ratio of the shear to the normal stress across a plane passing through the direction of the intermediate principal stress and making an angle $\pi/4-\phi/2$ with the direction of σ_{10}. Knowing the orientation of the fault plane and the slip vector in an earthquake, this ratio can be calculated without assuming the direction of the initial stress field. In the examples of the next section the change $\Delta\tau_0$ of the initial maximum shear stress $\tau_0 = 0.5(\sigma_{10}-\sigma_{30})$ has been chosen as parameter describing the measure of stabilization or destabilization:

$$\Delta\tau_0 = 0.5(\Delta\sigma_1 - \Delta\sigma_3)$$

$\Delta\tau_0$ is a meaningful quantity, if it is small compared with τ_0, i.e., if the additional stress field is much weaker than the initial stress field. $\Delta\tau_0$ is between $-\tau_{max}$ and $+\tau_{max}$ where τ_{max} denotes the maximum shear stress induced by the water load. The initial maximum shear stress τ_0 is always increased, if condition B is fulfilled (compression reckoned positive). Condition A can be connected with both positive and negative values of $\Delta\tau_0$ (if $\phi > 0$).

A Simple Geometrical Model

A simple three-dimensional model has been studied to test the method. This model is a lake of rectangular cross section (Figs. 1-3). The length of the lake is 0.8 km, the width is 0.2 km and the maximum water depth is 100 m. The stresses are symmetrical about the two vertical planes which cross the lake at the horizontal centre lines. A value of 30° has been chosen for the angle of internal friction of the crustal rock. The distribution of the additional stress $\Delta\tau_0$ in the direction of the initial maximum shear stress τ_0 is shown in Figs. 1–3 for different initial stress directions. They represent tectonic environments of normal, thrust and transcurrent faulting. The regions in which the stresses move towards the failure envelope (i.e., condition B is fulfilled) are emphasized by the stippled pattern.

In a region of normal faulting the zones destabilized by the water load are situated below the lake bottom (Fig. 1). It is remarkable to note that the likelihood for failure is less directly below the lake compared with greater depths. At depths greater than about one third of the lake width the risk for failure increases. There the initial maximum shear stress τ_0 is increased by more than 3 bar. In regions situated farther away from the lake τ_0 is decreased by more than 0.5 bar. It is unlikely that there an earthquake of normal faulting mechanism is triggered by the water load alone. Nevertheless, it could be triggered by the increasing pore pressure (Simpson 1976).

For thrust faulting, the likelihood of failure is increased in regions situated very close to the surface and extending landwards (Fig. 2). Epicenters of overthrusting earthquakes induced by the water load alone would be located outside the reservoir region. The focal depths would be very shallow. The extension of the destabilized volume increases with the lake width. Thus focal depths could be of the order of 1 km if the lake width is of the order of several kilometers. The increase of τ_0 is much less

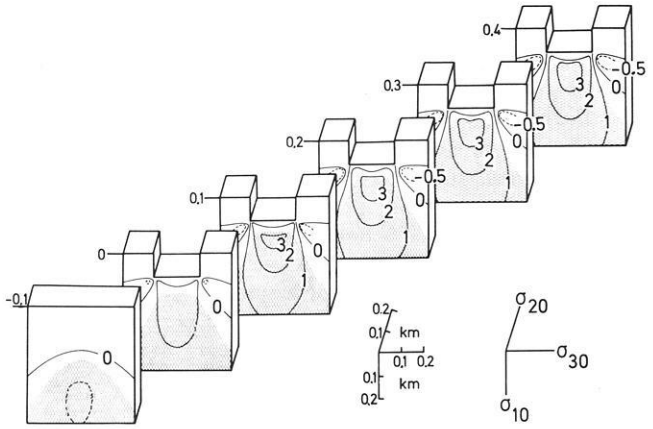


Fig. 1. Change of initial maximum shear stress in bar, in a tectonic environment of normal faulting under a lake of rectangular cross section, 0.2 km wide, 0.8 km long and 100 m deep. Regions destabilized by the water load are *stippled*. Angle of internal friction = 30°

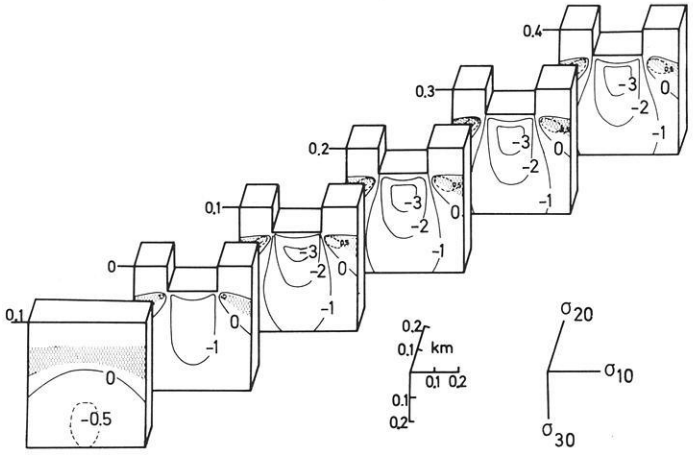


Fig. 2. Same as Fig. 1, for a tectonic environment of thrust faulting

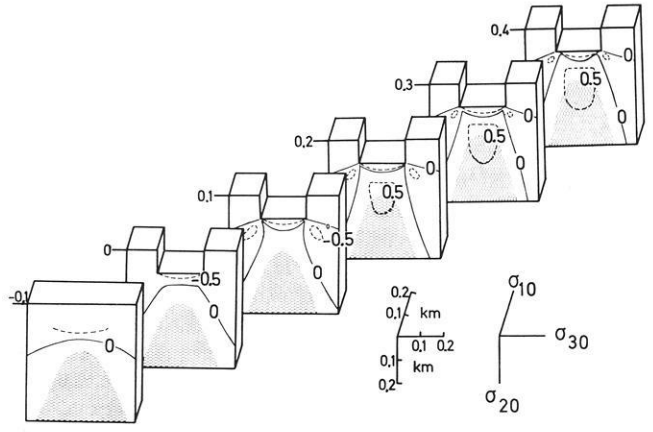


Fig. 3. Same as Fig. 1, for a tectonic environment of transcurrent faulting

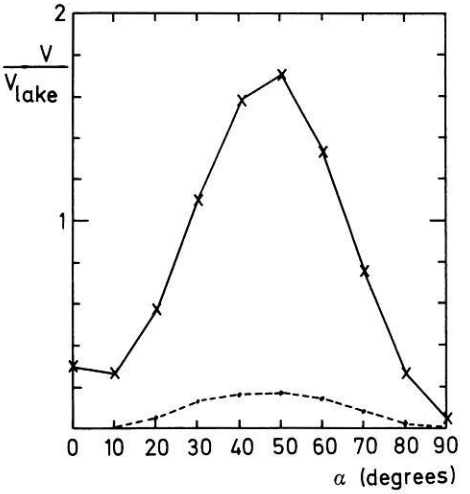


Fig. 4. Volume destabilized in a tectonic environment of transcurrent faulting under a lake of rectangular cross section and 100 m deep. Length and width of the lake are in the ratio 4:1. The volume is normalized to the volume of the lake. *Full line:* volume with $\Delta\tau_0 \geqq 0.5$ bar; *broken line:* volume with $\Delta\tau_0 \geqq 1$ bar. α is the angle between the direction of the greatest principal compressional stress and the longer axis of the lake

than in a normal faulting region; its maximum value is about 0.5 bar. The crustal parts situated below the reservoir region are stabilized in a thrust faulting stress regime.

In a third example an initial stress field has been assumed which favours transcurrent faulting (Fig. 3). In this case the likelihood of failure increases in regions which are situated below the reservoir. However, as in the case of normal faulting the very shallow parts of the crustal rocks are less subjected to failure than the deeper parts. The transition to the destabilized region occurs approximately at a depth of two thirds of the lake width. In the example presented the lake width is 0.2 km and the stresses approach the failure envelope at depths greater than 0.13 km below the bottom of the lake. If the lake width is 3 km this depth is 2 km. For transcurrent faulting, the volume destabilized by the water load depends strongly on the direction of the greatest principal compressional stress σ_{10} with respect to the longer axis of the reservoir. In the example of Fig. 3 both directions are identical. The variation of the destabilized volume with the direction of σ_{10} is shown in Fig. 4. It reaches a maximum value if the angle α between the longer axis of the lake and σ_{10} is between 40 and 60°. Already for $\alpha > 10°$ τ_0 is increased by more than 1 bar (broken line). The example of Fig. 3 corresponds to $\alpha = 0°$ in Fig. 4.

An Example (Schlegeis-Reservoir, Austria)

The Schlegeis arch buttress dam on the Zemm river was completed in 1971. The height of the dam is 131 m. The storage volume of the reservoir is 127.7×10^6 m^3 (Widmann et al. 1972). Continuous seismic monitoring was started in 1971 by a short period vertical station (Blum and Fuchs 1974). Localization of seismic events was made possible by additional stations installed temporarily in the vicinity of the dam (Blum et al. 1977). The local seismic activity observed is characterized by numerous weak tremors ($M_L < 0$) occurring during periods of lowering water level. During periods of rising and maximum water level

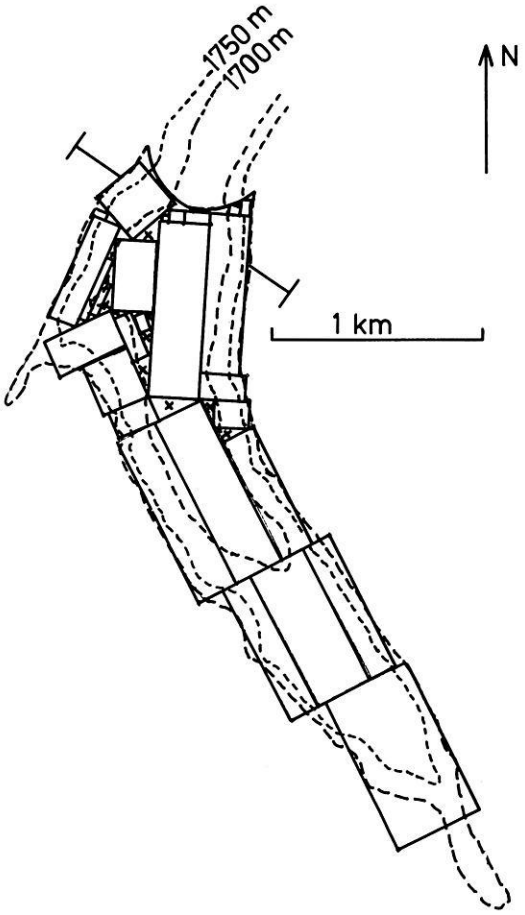


Fig. 5. Schlegeis-Reservoir, approximation of the water load and *contour lines* given in intervals of 50 m. Maximum lake level is reached at 1782 m. *Crosses* mark single point forces

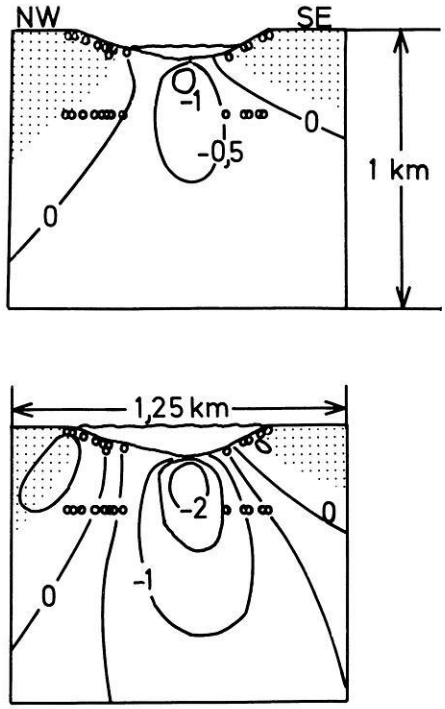


Fig. 6. Change of the initial maximum shear stress along a vertical cross section under the Schlegeis-Reservoir. The position of the section is given in Fig. 5. Stresses are in bar. *Upper part:* water level at 1722 m, *lower part:* at 1782 m. *Stippled:* regions destabilized by the filling of the reservoir. Angle of internal friction = 30°. *Circles* denote the position of hypocenters in the vicinity of the cross section. Both, the maximum (lower row of *circles*) and minimum (upper row of circles close to the surface) possible focal depth of each tremor is shown as calculated by Blum (1975)

local seismic activity is relatively low. A detailed description of the seismic phenomena observed until 1975 has been given by Blum (1975). Since then the periodic occurrence of the local seismic activity did not change. Hypocenters of local tremors could be determined for the period March–April 1974 (Blum 1975; Blum et al. 1977). Focal depths are very shallow (< 300 m) and the epicenters were located clearly outside the region of greatest water depth. Focal mechanism data are not available. Blum et al. (1977) proposed a model of dynamic stress equilibrium to explain the tremor mechanism. They assumed a tectonic environment of thrust faulting with the greatest compressional stress steadily increasing. The upper bound of stress accumulation is given by the failure stress. Without water load excess stresses were suggested to be released by slow deformations. With the additional water load the underlying rock is stabilized because the least principal stress (vertical) is increased more than the greatest principal stress (horizontal). Now additional horizontal stresses can be stored by the rock, moving the stresses again towards the failure threshold. During rapid lowering of the lake level the least principal stress is reduced much more than the greatest principal stress so that failure can occur. According to this model the foci of the seismic events should be located in regions which are stabilized by the water load and destabilized during periods of falling water level. In the following this essential assumption is checked by calculating the load induced stresses and comparing them with the initial stress field assumed by Blum et al. (1977). The effect of elastic anisotropy as well as anisotropy caused by the jointed body of rock is not considered in the calculations.

The initial stress field in the vicinity of the Schlegeis-Reservoir is not known. In their model Blum et al. (1977) proposed maximum compression in the direction SSE-NNW which agrees with the regional stress field in Central Europe and in the Alps as derived from fault plane solutions of earthquakes (Ahorner et al. 1972; Müller 1977). It is also confirmed by geological observations of postcrystalline deformation along shear planes dipping approximately to the North and to the South (Mignon 1972). Therefore this direction of the maximum compression was also chosen for the model calculations. The least principal stress was assumed to be directed vertically.

The water load of the reservoir was approximated by 29 rectangular elements with constant and linearly varying load distribution and by 8 point forces. The arrangement of the elements is depicted in Fig. 5. Additionally, the bathymetry of the lake is shown by contour-lines at intervals of 50 m. The maximum water level is reached at 1782 m above sea level.

Results of the calculations are shown along a vertical and a horizontal cross section (Figs. 6–7). The vertical plane crosses the reservoir about 0.2 km south of the dam (see Fig. 5). In the upper part of Fig. 6 the distribution of $\Delta\tau_0$ is shown for a water level of 1772 m which was reached in April 1974 at the beginning of the tremor series from which epicenter determi-

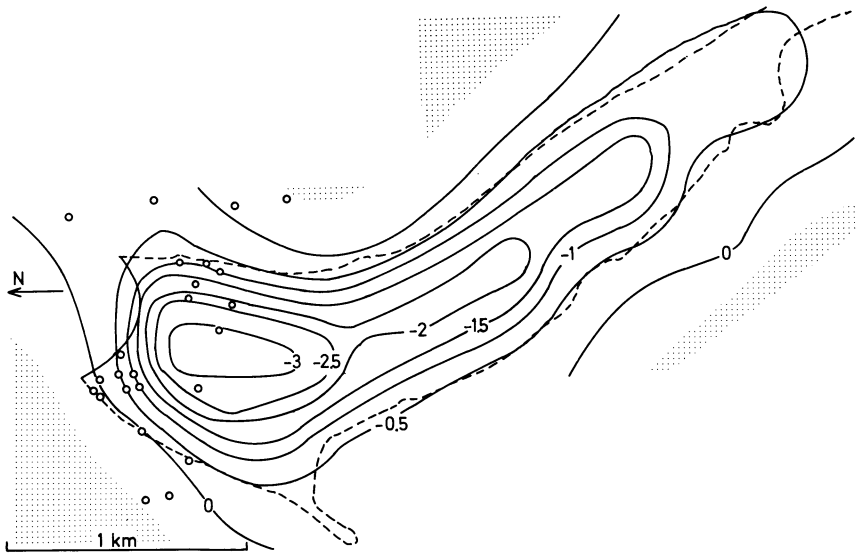


Fig. 7. Change of initial maximum shear stress, given in intervals of 0.5 bar, along a horizontal cross section at a depth of 1482 m. The lake level is at 1782 m. *Stippled:* regions destabilized by the filling of the reservoir. Circles denote epicenters of local tremors as located by Blum (1975)

nations are available. In the lower part of the figure the distribution of $\Delta\tau_0$ is shown for the maximum water level of 1782 m. Circles mark the epicenters of the tremors; they were projected on the cross section. Focal depths could not be determined as accurately as the epicenters. Therefore, upper and lower bounds of focal depths are given. With the initial stress direction assumed the rock within the reservoir region is stabilized by the water load ($\Delta\tau_0$ negative). The initial maximum shear stress τ_0 is diminished by more than 2 bar (Fig. 6), south of the profile even by more than 3 bar (see Fig. 7). The transition to the zones destabilized by the water load begins near the shore and extends landward to greater depths. At the maximum lake level τ_0 is increased by more than 0.5 bar. The tremor foci are situated in the transition zone between the stabilized and the destabilized region. During periods of falling water level the zones within the reservoir region are destabilized with respect to the period of maximum water level. If the model of Blum et al. (1977) were correct, an accumulation of hypocenters would be expected in these zones and not where they are actually located. Similar results have been found for two different initial stress fields: (1) for a stress field leading to EW strike slip motions along the nearly vertically dipping schist plane at Schlegeis and (2) for the stress field due to the weight of the surrounding mountains (Bock 1978). Therefore the model of dynamic stress equilibrium is not confirmed by the stress calculations carried out.

The increase of pore pressure at depth produced by the rise of the lake level is also not suited to explain the Schlegeis tremors for the following reason. This model would require rock permeabilities of the order of 0.01 millidarcy to explain the time lag of 4–6 months between the maximum water level and the occurrence of seismic events (Bock 1978). Such a low permeability is unlikely. Permeabilities for compact granite are of the order of 1 millidarcy and are increased by several orders of magnitude if the rock is fractured (Louis 1967). Seepage flow measurements in the vicinity of the Schlegeis dam also indicate higher permeabilities of the basement rock (Widmann and Heigerth 1976).

The characteristic features of the seismic activity observed at Schlegeis can be explained better by a model which was suggested after seismic events quite similar to those at Schlegeis have been also observed at other reservoirs in the Alps and in the Rumanian Carpathians (Merkler et al. 1978; Bock 1978). In the following this model is briefly described. A broader discussion will be given in a forthcoming paper (Merkler et al. in preparation). Cracks very close to the surface are water-filled by the reservoir. In the period of falling water level – i.e., in winter and springtime – the thermal insulation provided by the lake is removed and the water in the cracks can freeze at lower temperatures. This process leads to an increase in water volume. The stresses near the end of the cracks can be high enough to exceed the rock strength. This model does not only explain the temporal but also the spatial distribution of epicenters situated clearly outside the region of the actual shoreline (see Fig. 6 top).

Conclusions

The significance of the primary stress field for load induced seismicity has been demonstrated for a lake model of simple geometry. If the primary stress directions are known regions can be marked where the stresses are moved closer to the failure threshold. Taking into account the finite dimensions of a reservoir, earthquakes can in principle also be triggered in a tectonic environment of transcurrent and thrust faulting by the water load alone. Under the assumption of an initial stress direction favouring thrust faulting a model was checked which was proposed by Blum et al. (1977) to explain the mechanism of seismic events observed at the Schlegeis-Reservoir, Austria. It could be shown that the tremor foci are situated predominantly outside the regions which are stabilized by the increasing water load, where the foci should be located according to the model. Therefore, this model has been rejected as a viable hypothesis to explain the Schlegeis tremors.

Acknowledgements. These studies have been supported by the Deutsche Forschungsgemeinschaft (German Research Society) within the rock mechanics research program SFB 77 at Karlsruhe University. Computing facilities were made available by the computer center of Karlsruhe University. The author is indebted to Professor K. Fuchs, Dipl. Geol. G. Merkler, and Dr. G. Müller for many helpful discussions, and Professor K. Fuchs, Professor C. Kisslinger, and Dr. G. Müller for reading the manuscript.

Appendix

The stresses due to a distributed normal loading $p(x, y)$ on the surface $z=0$ of the semi-infinite solid $z>0$ are obtained by integration of the Boussinesq equations over a region F at the surface. In cartesian coordinates this leads to (Jaeger and Cook 1971, p. 282):

$$\sigma_x=\frac{1}{2\pi}\iint_F\left\{\frac{3(x-\xi)^2 z}{\rho^5}+(1-2\nu)\left[\frac{(y-\eta)^2+z^2}{\rho^3(z+\rho)}-\frac{z}{\rho^3}-\frac{(x-\xi)^2}{\rho^2(z+\rho)^2}\right]\right\}\cdot p(\xi,\eta)\,d\xi\,d\eta$$

$$\sigma_y=\frac{1}{2\pi}\iint_F\left\{\frac{3(y-\eta)^2 z}{\rho^5}+(1-2\nu)\cdot\left[\frac{(x-\xi)^2+z^2}{\rho^3(z+\rho)}-\frac{z}{\rho^3}-\frac{(y-\eta)^2}{\rho^2(z+\rho)^2}\right]\right\}\cdot p(\xi,\eta)\,d\xi\,d\eta$$

$$\sigma_z=\frac{3z^3}{2\pi}\iint_F\frac{p(\xi,\eta)}{\rho^5}\,d\xi\,d\eta$$

$$\tau_{yz}=\frac{3z^2}{2\pi}\iint_F\frac{(y-\eta)}{\rho^5}p(\xi,\eta)\,d\xi\,d\eta$$

$$\tau_{xz}=\frac{3z^2}{2\pi}\iint_F\frac{(x-\xi)}{\rho^5}p(\xi,\eta)\,d\xi\,d\eta$$

$$\tau_{xy}=\frac{1}{2\pi}\iint_F\left\{\frac{3z(x-\xi)\,(y-\eta)}{\rho^5}-(1-2\nu)\cdot\frac{(x-\xi)(y-\eta)(z+2\rho)}{\rho^3(z+\rho)^2}\right\}\cdot p(\xi,\eta)\,d\xi\,d\eta.$$

The normal force $p(\xi,\eta)\,d\xi\,d\eta$ acts at the point $(\xi,\eta,0)$. The distance to a field point (x, y, z) in the halfspace is given by $\rho=[(x-\xi)^2+(y-\eta)^2+z^2]^{1/2}$. For a constant vertical load $p(\xi,\eta)=p_0$ over a rectangular area $(0\leqq\xi\leqq l, 0\leqq\eta\leqq b)$ the following expressions for the stresses at the point $(0, 0, z)$ are obtained:

$$\sigma_z=\frac{p_0}{2\pi}\left[\arctan\frac{lb}{zR_3}+\frac{lbz}{R_3}\left(\frac{1}{R_1^2}+\frac{1}{R_2^2}\right)\right]$$

$$\sigma_x=\frac{p_0}{2\pi}\left[\arctan\frac{bl}{zR_3}-\frac{blz}{R_1^2R_3}-(1-2\nu)K_x\right]$$

$$\sigma_y=\frac{p_0}{2\pi}\left[\arctan\frac{lb}{zR_3}-\frac{lbz}{R_2^2R_3}-(1-2\nu)L_y\right]$$

$$\tau_{yz}=\frac{p_0}{2\pi}\left[\frac{l}{R_1}-\frac{z^2l}{R_2^2R_3}\right]$$

$$\tau_{xz}=\frac{p_0}{2\pi}\left[\frac{b}{R_2}-\frac{z^2b}{R_1^2R_3}\right]$$

$$\tau_{xy}=\frac{p_0}{2\pi}\left[1+\frac{z}{R_3}-z\left(\frac{1}{R_2}+\frac{1}{R_1}\right)-(1-2\nu)\ln\frac{(z+R_1)(z+R_2)}{2z(z+R_3)}\right]$$

with

$$R_1=\sqrt{l^2+z^2}$$
$$R_2=\sqrt{b^2+z^2}$$
$$R_3=\sqrt{l^2+b^2+z^2}$$
$$K_x=\arctan\frac{zb}{lR_3}+\arctan\frac{bl}{zR_3}-\arctan\frac{b}{l}$$
$$L_y=\arctan\frac{zl}{bR_3}+\arctan\frac{lb}{zR_3}-\arctan\frac{l}{b}.$$

For points other than $(0, 0, z)$ stresses are obtained by employing the principle of superposition (Poulos and Davis 1974, p. 12).

For a linearly varying vertical load over a rectangular area it is assumed that the load increases from zero at $(0, 0\leqq y\leqq b)$ to p_0 at $(l, 0\leqq y\leqq b)$. This leads to $p(\xi,\eta)=p_0(\xi/l)$. The stresses at (x, y, z) are given by

$$\sigma_x=\frac{p_0}{2\pi l}\left\{\frac{(x-l)zl}{r_1^2}\left(\frac{y}{R_1}-\frac{y-b}{R_3}\right)+xA+2zB+(1-2\nu)\left[lA_3-xA-zB+(y-b)C-yD+\frac{x-l}{2}\ln\frac{(x-l)^2+(y-b)^2}{(x-l)^2+y^2}-\frac{x}{2}\ln\frac{x^2+(y-b)^2}{x^2+y^2}+\frac{y}{2}E-\frac{y-b}{2}F+\frac{z}{2}G\right]\right\}$$

$$\sigma_y=\frac{p_0}{2\pi l}\left\{xA+\frac{(y-b)z}{r_2^2}\left(R_2-\frac{R_2^2}{R_3}+\frac{xl}{R_3}\right)-\frac{yz}{r_4^2}\left(R_0-\frac{R_0^2}{R_1}+\frac{xl}{R_1}\right)+zB-(1-2\nu)\left[x(A+A_1-A_2-C+D)-\frac{y-b}{2}F_1-(y-b)F_2+yE_2+\frac{y}{2}E_1+\frac{z}{2}G\right]\right\}$$

$$\sigma_z=\frac{p_0}{2\pi l}\left\{xA+\frac{x(x-l)(y-b)z}{R_3}\left(\frac{1}{r_2^2}+\frac{1}{r_1^2}\right)-\frac{x(x-l)yz}{R_1}\left(\frac{1}{r_4^2}+\frac{1}{r_1^2}\right)-\frac{x^2(y-b)z}{R_2}\left(\frac{1}{r_2^2}+\frac{1}{r_3^2}\right)+\frac{x^2yz}{R_0}\left(\frac{1}{r_4^2}+\frac{1}{r_3^2}\right)+\frac{z^3}{r_1^2}\left(\frac{y-b}{R_3}-\frac{y}{R_1}\right)-\frac{z^3}{r_3^2}\left(\frac{y-b}{R_2}-\frac{y}{R_0}\right)\right\}$$

$$\tau_{yz}=\frac{p_0}{2\pi l}\left\{\frac{xz^2}{r_4^2}\left(\frac{x-l}{R_1}-\frac{x}{R_0}\right)+\frac{xz^2}{r_2^2}\left(\frac{x}{R_2}-\frac{x-l}{R_3}\right)+z^2\left(\frac{1}{R_1}-\frac{1}{R_0}-\frac{1}{R_3}+\frac{1}{R_2}\right)\right\}$$

$$\tau_{xz}=\frac{p_0}{2\pi l}\left\{\frac{z^2l}{r_1^2}\left(\frac{y}{R_1}-\frac{y-b}{R_3}\right)-zA\right\}$$

$$\tau_{xy}=\frac{p_0}{2\pi l}\left\{zl\left(\frac{1}{R_3}-\frac{1}{R_1}\right)+zH-(1-2\nu)\left[\frac{x}{2}(F_1-E_1)+x(F_2-E_2)-y(A_2-D)+(y-b)(A_1-C)-zH\right]\right\}$$

with

$$R_0^2=x^2+y^2+z^2$$
$$R_1^2=(x-l)^2+y^2+z^2$$
$$R_2^2=x^2+(y-b)^2+z^2$$
$$R_3^2=(x-l)^2+(y-b)^2+z^2$$
$$r_0^2=x^2+y^2$$
$$r_1^2=(x-l)^2+z^2$$
$$r_2^2=(y-b)^2+z^2$$
$$r_3^2=x^2+z^2$$
$$r_4^2=y^2+z^2$$

$$A = \arctan\frac{(x-l)(y-b)}{zR_3} - \arctan\frac{(x-l)y}{zR_1} - \arctan\frac{x(y-b)}{zR_2} + \arctan\frac{xy}{zR_0}$$

$$A_1 = \arctan\frac{z(x-l)}{(y-b)R_3} - \arctan\frac{zx}{(y-b)R_2}$$

$$A_2 = \arctan\frac{z(x-l)}{yR_1} - \arctan\frac{zx}{yR_0}$$

$$A_3 = \arctan\frac{y-b}{x-l} - \arctan\frac{y}{x-l} - \arctan\frac{z(y-b)}{(x-l)R_3} + \arctan\frac{zy}{(x-l)R_1}$$

$$B = \ln\frac{y-b+R_3}{y-b+R_2} + \ln\frac{y+R_0}{y+R_1}$$

$$C = \arctan\frac{x-l}{y-b} - \arctan\frac{x}{y-b}$$

$$D = \arctan\frac{x-l}{y} - \arctan\frac{x}{y}$$

$$E = E_1 + E_2$$

$$E_1 = \ln\frac{(r_4^2 - zR_1)^2 + y^2(x-l)^2}{(r_4^2 - zR_0)^2 + y^2x^2} \qquad E_2 = \ln\frac{x^2+y^2}{(x-l)^2+y^2}$$

$$F_1 = \ln\frac{(r_2^2 - zR_3)^2 + (y-b)^2(x-l)^2}{(r_2^2 - zR_2)^2 + (y-b)^2x^2} \qquad F_2 = \ln\frac{x^2+(y-b)^2}{(x-l)^2+(y-b)^2}$$

$$G = \ln\frac{(r_4^2 - yR_1)^2 + z^2(x-l)^2}{(r_4^2 - yR_0)^2 + z^2x^2} - \ln\frac{(r_2^2 - [y-b]R_3)^2 + z^2(x-l)^2}{(r_2^2 - [y-b]R_2)^2 + z^2x^2}$$

$$H = \ln\frac{x-l+R_3}{x-l+R_1} + \ln\frac{x+R_0}{x+R_2}.$$

References

Ahorner, L., Murawski, H., Schneider, G.: Seismotektonische Traverse von der Nordsee bis zum Apennin. Geol. Rundsch. **61**, 915–942, 1972

Bell, M.L., Nur, A.: Strength changes due to reservoir-induced pore pressure and stresses and application to Lake Oroville. J. Geophys. Res. **83**(B9), 4469–4483, 1978

Blum, R.: Seismische Überwachung der Schlegeis-Talsperre und die Ursache induzierter Seismizität. Diss., Universität Karlsruhe, 1975

Blum, R., Bock, G., Fuchs, K., Merkler, G., Widmann, R.: Correlation between micro-activity and variation of water level at the Schlegeis-Reservoir. J. Geophys. **43**, 561–567, 1977

Blum, R., Fuchs, K.: Observation of low-magnitude seismicity at a reservoir in the Eastern Alps. Eng. Geol. 8, 99–106, 1974

Bock, G.: Induzierte Seismizität: Modelle und die Beobachtungen am Schlegeis- und Emosson-Stausee. Diss., Universität Karlsruhe, 1978

Bufe, C.G.: The Anderson Reservoir seismic gap – induced aseismicity? Eng. Geol. **10**, 255–262, 1976

Comninakis, P., Drakopoulos, J., Moumoulidis, G., Papazachos, B.: Foreshock and aftershock sequences of the Kremasta earthquake and their relation to the water-loading of the Kremasta artificial lake. Ann. Geofis. **21**, 39–71, 1968

Gough, D.I.: Incremental stress under a two-dimensional artificial lake. Can. J. Earth Sci. **6**, 1067–1075, 1969

Gough, D.I., Gough, W.I.: Stress and deflection in the Lithosphere near Lake Kariba – I. Geophys. J.R. Astron Soc. **21**, 65–78, 1970a

Gough, D.I., Gough, W.I.: Load-induced earthquakes at Lake Kariba – II. Geophys. J. R. Astron. Soc. **21**, 79–101, 1970b

Gupta, H.K., Narain, H., Rastogi, B.K., Mohan, I.: A study of the Koyna earthquake of December 10, 1967. Bull. Seismol. Soc. Am. **59**, 1149–1168, 1969

Gupta, H.K., Rastogi, B.K.: Dams and earthquakes. Amsterdam: Elsevier 1976

Handin, J.: Strength and ductility. In: Handbook of physical constants, S.P. Clark Jr., ed.: pp. 223–289. Geol. Soc. Am. Mem. **97**, 1966

Hubbert, M.K., Rubey, W.W.: Mechanic of fluid-filled porous solids and its application to overthrust faulting. Bull. Geol. Soc. Am. **70**, 115–166, 1959

Jaeger, J.C., Cook, N.G.W.: Fundamentals of rock mechanics. London: Chapman and Hall Ltd., Science Paperback 1971

Kisslinger, C.: A review of theories of mechanisms of induced seismicity. Eng. Geol. **10**, 85–98, 1976

Lee, T.: A method for computing the deformation of the crust caused by the filling of large lakes. Bull. Seismol. Soc. Am. **62**, 1597–1610, 1972

Louis, C.: Strömungsvorgänge in klüftigen Medien und ihre Wirkung auf die Standsicherheit von Bauwerken und Böschungen im Fels. Veröffentlichungen Inst. f. Bodenmechanik und Felsmechanik der Universität Karlsruhe **30**, 1967

Merkler, G., Bock, G., Fuchs, K.: Ice-induced seismic microactivity at reservoirs (Abstract). ESC-Meeting, Strasbourg, September 1978

Mignon, K.: Überblick über die geologischen Verhältnisse. Österr. Z. Elektrizitätswirtschaft **25**, 432–436, 1972

Müller, G.: Fault-plane solution of the earthquake in Northern Italy, 6 May 1976, ans implications for the tectonics of the Eastern Alps. J. Geophys. **42**, 343–349, 1977

Poulos, H.G., Davis, E.H.: Elastic solutions for soil and rock mechanics. New York: J. Wiley & Sons 1974

Nyland, E., Withers, R.J.: A fast method for computing load induced stress in the Earth. Geophys. J.R. Astron. Soc. **44**, 689–698, 1976

Raleigh, C.B., Healy, J.H., Bredehoeft, J.D.: An experiment in earthquake control at Rangely, Colorado. Science **191**, 1230–1237, 1976

Simpson, D.W.: Seismicity changes associated with reservoir loading. Eng. Geol. **10**, 123–150, 1976

Snow, D.T.: Geodynamics of seismic reservoirs. Proc. Symp. on percolation through fissured rocks, Int. Soc. Rock Mechanics, Stuttgart, T2-J, 1–19, 1972

Widmann, R., Heigerth, G.: Rock deformation and seepage flow in the foundation of the Schlegeis arch dam. Comm. Inter. des Grands Barranges, 12ième Congrès, Mexico, Q.45, R.39, 1976

Widmann, R., Schlosser, J., Stäuble, H.: Die Bogengewichtsmauer Schlegeis. Österr. Z. Elektrizitätswirtschaft **25**, 395–404, 1972

Withers, R.J., Nyland, E.: Theory for the rapid solution of ground subsidence near reservoirs on layered and porous media. Eng. Geol. **10**, 169–185, 1976

Received July 25, 1979; Revised Version December 20, 1979

J. Geophys. 48, 101–115, 1980

Journal of Geophysics

Ground-Based Observations of an Onset of Localized Field-Aligned Currents During Auroral Breakup Around Magnetic Midnight

H.J. Opgenoorth[1,2], R.J. Pellinen[1], H. Maurer[3], F. Küppers[2], W.J. Heikkila[4], K.U. Kaila[1], and P. Tanskanen[5]

[1] Finnish Meteorological Institute, Division of Geomagnetism, Vuorikatu 24, Box 503, SF-00100 Helsinki 10, Finland
[2] Institut für Geophysik der Universität Münster, Gievenbecker Weg 61, D-4400 Münster, Federal Republic of Germany
[3] Institut für Geophysik und Meteorologie der Technischen Universität Braunschweig, Mendelsohnstrasse 1 A, D-3300 Braunschweig, Federal Republic of Germany
[4] University of Texas at Dallas, Center for Space Sciences, Richardson, Texas 75080, USA
[5] Air Force Geophysics Laboratory, Hanscom Air Force Base, Massachusetts 01731, USA
(on Leave from Department of Physics, University of Oulu, Linnanmaa, SF-90570 Oulu 57, Finland)

Abstract. The substorm on 2 March 1978 was selected for study as a relatively weak substorm, starting at about local magnetic midnight, that could be observed with instruments in Northern Scandinavia. The analysis is based on a comparative study of data from the IMS magnetometer network, all-sky cameras, pulsation magnetometers, and riometers in the Scandinavian area. In addition other data are used to support the results, e.g., a photograph from the DMSP-F2 satellite, showing the auroral situation over Scandinavia, and further west, immediately after the substorm onset.

The substorm was preceded by a weak activation of aurora and magnetic disturbance about 3 min before the onset. After a fading that lasted for 20 s and could be observed only in optical aurora, the substorm onset led to a strong brightening of the aurora, an enhancement of the westward electrojet, a sudden rise in the ionospheric D-layer absorption, and Pi B type pulsations. Immediately after the onset, the ground magnetic data suggest the appearence of a pair of oppositely directed, localized, field-aligned currents (FACs). The main development of the signatures of the downward FAC was clearly delayed by about 3 min. There were significant correlations between the magnetic signatures of the two FACs and different features and spectra of the optical aurora, both in time and location. The observed Pi B type pulsations lasted as long as a growth in the local onset-connected FACs could be inferred.

Within the first three minutes the localized three dimensional current system developed into a more sheet-like configuration. An expansion to the west, possibly accompanied by a westward travelling surge, was traced with riometers and magnetometers on Iceland and Greenland.

Key words: Auroral substorm – Magnetic substorm – Substorm onset – Field-aligned currents – Auroral electrojet.

1. Introduction

The global features of magnetospheric substorms have been described in many papers (e.g. Akasofu 1964; McPherron et al. 1973; Akasofu 1974; Fukunishi 1975; Kamide and Akasofu 1975; Kisabeth and Rostoker 1974; McPherron 1979). The numerous case studies reported in these papers have led to the knowledge of substorm phenomena that we have today. However, more detailed event analyses are needed to increase our understanding of the physics involved in localized and transient effects appearing during substorms. Careful consideration of parameters such as local magnetic time, latitude, amplitude of the observed disturbance, and general magnetic activity around the event may permit a more general interpretation that can test the global picture of substorm phenomena. In this paper we shall give an example of such a study.

The coupling mechanism between the magnetospheric processes and the ionospheric phenomena observed from the ground is nowadays understood to be represented by currents driven along the magnetic field lines. Boström (1964) introduced two field-aligned current (FAC) configurations, which could possibly be responsible for the polar electrojets. In the first case (later referred to as the Boström I model), the auroral electrojet is the ionospheric closure-current between a pair of field-aligned line-currents at the eastern and western end of the electrojet, while in the second case (later referred to as the Boström II model) the auroral electrojet is the ionospheric Hall-current that is driven between a poleward and an equatorward field-aligned sheet-current. In recent literature it seems to be generally accepted that a FAC-configuration of the Boström II type exists permanently and is responsible for the quiet time eastward and westward electrojets. On the other hand, the substorm-related three-dimensional current system seems to be represented to a greater extent by the Boström I type, involving an upward FAC on the evening- and a downward FAC on the morning side of the location of substorm onset (Clauer and McPherron 1974; Iijima and Potemra 1976; Rostoker and Boström 1976; Kisabeth and Rostoker 1977; McPherron 1979).

In several magnetospheric substorm models (Atkinson 1971; McPherron et al. 1973; Akasofu 1977; Boström 1977; Heikkila and Pellinen 1977) the substorm related FAC-system is explained as being caused by deviation of the magnetospheric cross-tail current, partially closing via field lines and the ionosphere. The models differ mainly in their attempts to explain

0340-062X/80/0048/0101/$03.00

- possible trigger mechanisms for the substorm onset,
- energization of particles to the observed energies up to MeV (Pellinen and Heikkila 1978a and references therein),
- anisotropy in the precipitating particle energy spectra (Lui et al. 1977),
- well defined distribution patterns of precipitating electrons and protons (Fukunishi 1975),
- mechanisms to maintain the deviation of the cross-tail current,
- expansion of the disturbance in the neutral-sheet plane.

The ionospheric phenomena on the evening side of the expanding substorm region have already been studied to great extent. Very early scientists were attracted by such striking phenomena as the auroral westward travelling surge (WTS) (Akasofu et al. 1965a) and sharp transitions from positive to negative variation in the H-component of magnetic recordings (Heppner 1954; Harang 1946). It has been found that the area covered by the WTS can be regarded as the local area of substorm-connected electron precipitation (Akasofu et al. 1969), carrying the western upward FAC (Kamide and Akasofu 1975; Wescott et al. 1975). It forms the leading edge of the substorm-enhanced westward electrojet (Pytte et al. 1976; see also Swift 1979; McPherron 1979, and references therein).

Less is known about comparable features produced by the downward FAC on the eastern, expanding side of the substorm region. Certainly the much wider scatter of precipitating protons and positive ions (Vallance Jones 1976, p. 54) will be a reason for the lack of sharp auroral and magnetic features. Nevertheless photometer observations show a substorm related enhancement of H_β-emissions (Fukunishi 1975; Oguti et al. 1974) where the downward FAC is expected to be located (Pellinen and Heikkila 1978a).

Due to spatial limitations of instrument networks, ground-based observers have recorded the onset of a substorm mainly close to its central meridian. Since the three-dimensional current system can be expected to be quite small in extent at the very start of a substorm (Pellinen and Heikkila 1978a), it should be possible, as this study will show, to observe the features of nearly the entire onset current-system with suitably located instruments.

The studies of the local behaviour of the auroral electrojets with a meridian chain of magnetometers, in connection with other observations (Kisabeth and Rostoker 1971, 1973; Kisabeth 1972; Akasofu 1974), resulted in a better understanding of the substorm-connected current systems as well as in an improvement of analysis methods. The next step in ground-based substorm studies was the installation of a two-dimensional magnetometer array in an area that was already covered by networks of other instruments, such as all-sky cameras, riometers, photometers, pulsation magnetometers and others. This type of network was first installed by Bannister and Gough (1977, 1978) in Canada. Later Küppers et al. (1979) started a more extensive network in northern Scandinavia. The area of the Scandinavian array is sufficiently large (ca. $1{,}000 \times 1{,}000\ km^2$, see Fig. 1) to cover a good portion of the entire substorm-onset phenomenon if one restricts analyses to weak or moderate substorms, recorded near to the initial onset location. As there seems to be little or no principal difference between weak and intense substorms (Kamide et al. 1975), this restriction still allows the generalization of the results.

Recently Untiedt et al. (1978) described the observation of a substorm onset near magnetic midnight with the first part of the Scandinavian IMS-network of stations. Since the auroral breakup was, in this particular case, initiated over the eastern part of their observation area, they recorded only the signatures

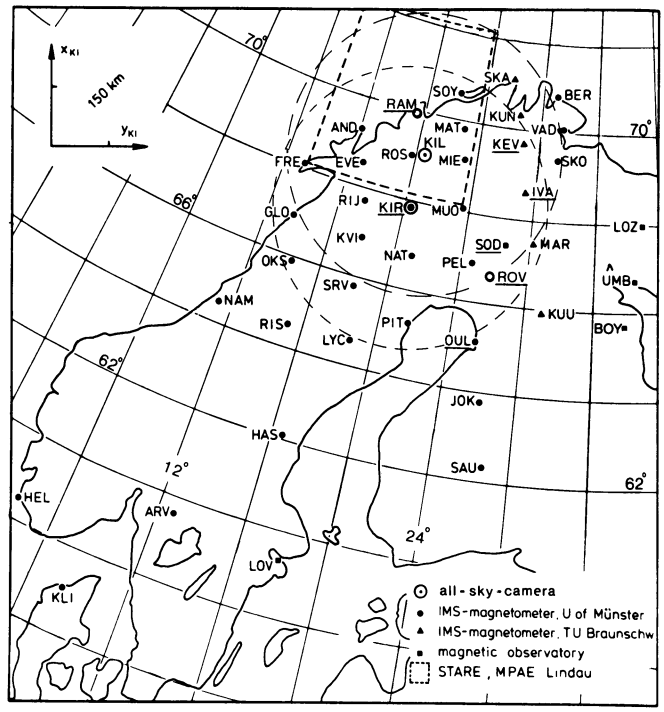


Fig. 1. Map of recording stations within the Scandinavian area from which data have been used in the present study. The different symbols that are used to indicate the various instruments are given in the right bottom corner. The stations with riometers are underlined. The observational coverage by all-sky cameras is indicated by broken circles (as defined by 15° elevation angle and 100 km height). The broken lines mark the area within which irregularity drift-velocities have been computed from the STARE-data. (for further details concerning the stations see Table 1)

of the western edge of the area of upward FAC. In the hope of recording events with the region of substorm onset being optimally located with respect to the ground-based network of stations, the so called Auroral Breakup Campaign (ABC-1) was carried out in February and March 1978 (IMS-newsletter No. 12, 1977). Of the events recorded within the three weeks of the campaign a moderate substorm that occured on 2 March 1978 was selected for the present comparative study.

In the next section we shall describe the instruments from which data has been used in our analysis. We also introduce the methods of data processing and presentation. In Sect. 3 a general overview of the geomagnetic situation during the substorm analysed will be given, followed by the description of observations during the growth phase, substorm onset, and expansion phase. In the final section we will discuss how far the observations presented are in correspondence with observations of other scientists, and whether they can be explained by the models for the magnetospheric substorm.

2. Instrumentation and Methods of Data Processing

Due to the enhanced recording activity coordinated amongst the European IMS-participants, many kinds of data-sets have been available for this study. In this section we will describe only the instruments used by the authors even though various other data have been studied and considered in the final interpretation. These

Table 1. Permanent or temporary stations from which data have been used in the present study

Station	Symbol	Geogr. coord.		Type of instrument[a] and institution[b]
Andenes	AND	69.3° N	16.0° E	M (UM)
Arvika	ARV	59.6	12.6	M (UM)
Bear Island	BJN	74.5	19.2	M (AOT)
Berlevag	BER	70.9	29.1	M (UM)
Boyarskaya	BOY	65.8	33.8	M (PGI)
Evenes	EVE	68.5	16.8	M (UM)
Fredvang	FRF	68.1	13.2	M (UM)
Glomfjord	GLO	66.9	13.6	M (UM)
Hassela	HAS	62.1	16.5	M (UM)
Hellvik	HEL	58.5	5.8	M (UM)
Ivalo	IVA	68.6	27.5	M (TUB), R (SGO)
Jokikylä	JOK	63.8	26.1	M (UM)
Kevo	KEV	69.8	27.0	M (TUB), R (SGO)
Kilpisjärvi	KIL	69.1	20.8	A (FMI)
Kiruna	KIR	67.8	20.4	M (UM), A, P, R (KGI)
Klim	KLI	57.1	9.2	M (UM)
Kunes	KUN	70.3	26.5	M (TUB)
Kuusamo	KUU	65.9	29.0	M (TUB)
Kvikkjokk	KVI	66.9	17.9	M (UM)
Lavangsdalen	LAV	69.7	19.9	P (UO)
Lovö	LOV	59.4	17.8	M (SGU)
Lovozero	LOZ	68.0	35.0	M (PGI)
Lycksele	LYC	64.6	18.7	M (UM)
Martti	MAR	67.5	28.3	M (TUB)
Mattisdalen	MAT	69.9	22.9	M (UM)
Mieron	MIE	69.1	23.3	M (UM)
Muonio	MUO	68.3	23.6	M (UM)
Namsos	NAM	64.5	11.1	M (UM)
Narssarssuaq	NAR	61.2	314.6	R (DMI)
Nattavaara	NAT	66.8	21.0	M (UM)
Okstindan	OKS	65.5	14.3	M (UM)
Oulu	OUL	65.1	25.5	M (UM), R (SGO)
Pello	PEL	66.9	24.7	M (UM)
Pitea	PIT	65.3	21.6	M (UM)
Ramfjord	RAM	69.8	19.6	R (AOT)
Risede	RIS	64.5	15.1	M (UM)
Ritsemjokk	RIJ	67.7	17.5	M (UM)
Rostadalen	ROS	69.0	19.7	M (UM)
Rovaniemi	ROV	66.6	25.8	R (SGO)
Sauvamäki	SAU	62.3	26.7	M (UM)
Skarsvag	SKA	71.1	25.8	M (TUB)
Skogfoss	SKO	69.4	29.4	M (UM)
Sodankylä	SOD	67.4	26.6	R (SGO)
Söröya	SOY	70.6	22.2	M (UM)
Storavann	SRV	65.8	18.2	M (UM)
Tjörnes	TJÖ	66.2	342.9	R (NTNF/UB)
Umba	UMB	66.7	34.5	M (PGI)
Vadsö	VAD	70.1	29.7	M (UM)

[a] M = magnetometer, A = all-sky camera, P = photometer, R = riometer

[b] If not explicitly explained in the text: AOT = Auroral Observatory at Tromsö, UM = University at Münster, SGU = Swedish Geological Survey, KGI = Kiruna Geophysical Institute, SGO = Geophysical Observatory at Sodankylä, TUB = Technical University at Braunschweig, PGI = Polar Geophysical Institute at Apatity, UO = University at Oslo, NTNF = Royal Norwegian Council for Scientific and Industrial Research, UB = University at Bergen, DMI = Danish Meteorological Institute, FMI = Finnish Meteorological Institute

instruments will be described later, when necessary. Results from magnetic pulsation data for this event are published elsewhere (e.g. Wedeken et al. 1979; T. Bösinger et al., paper in preparation).

The locations of the different kinds of instruments are shown in Fig. 1. More detailed information concerning the stations is listed in Table 1. During the night of 2 March 1978 two all-sky cameras (ASC) of the same design (Hyppönen et al. 1974) were operating at Kiruna and Kilpisjärvi, recording at rates of 1 frame per minute and 3 frames per minute respectively. These cameras record on 16 mm colour film, specially processed to a sensitivity of ASA 640, thus allowing an exposure time of about 2 s. The digital time display is accurate to the nearest second. Radioactive sources are used to activate calibration surfaces, visible on each photograph, for calibrations in both the red and the green part of the spectrum.

Distinct auroral structures in the ASC-pictures have been digitized along their lower borders, rectified and mapped under the assumption of a normal height of 100 km (Boyd et al. 1971).

Where possible, the final locations of aurorae have been averaged over the results of both cameras. The error in the location of the auroral forms so derived is less than 20 km, according to our experience with several stations (Opgenoorth 1978).

Data from 36 stations of the IMS-magnetometer network in Scandinavia have been used in the present study. A detailed description of this network and of the instrument types has recently been given by Küppers et al. (1979) and Maurer and Theile (1978). Additional instruments run by observatories or other institutions that were used in this study are given in Fig. 1 and Table 1.

The magnetic as well as the auroral data will be presented in a special cartesian coordinate-system, the so called KIRUNA-system. This was introduced by Küppers et al. (1979) for easier display and processing of data from the Scandinavian area. The system is centered at Kiruna in a plane tangent to the globe at the same location. The familiar horizontal components H and D are rotated to A and B respectively; B is parallel to the y_{Ki}-axis which is the tangent to the revised corrected geomagnetic latitude in Kiruna (Gustafsson 1970), A is orthogonal to B along the x_{Ki}-axis, and Z remains unchanged. The use of the KIRUNA-system in a number of studies (e.g. Untiedt et al. 1978; Baumjohann et al. 1978) showed good results in reducing the geometrical effects while the physical effects become clearer.

The magnetic disturbance vectors have been determined in comparison to a baseline which has been taken in the early morning hours of two quiet days before and after the considered night, in order to avoid Sq-effects. The data will be presented mostly in the form of so-called equivalent current vectors (TEC = *t*otal *e*quivalent *c*urrent), derived by turning the horizontal magnetic disturbance vector clockwise through 90° The vertical component will be indicated by different symbols for positive and negative variations or by isolines.

The need to differ between quiet-time current-systems and substorm-connected ones, as well as the need to recognize the character of changes in an existing current-system, made the use of so-called differential equivalent current vectors (DEC) necessary. The basic idea of this method is the assumption that a current system observed at a time t_1 stays constant until a time t_2, while an additional current system develops during the same time interval. According to the free superposition of magnetic fields, the difference between the magnetic disturbance field at the time t_2 and the field at the time t_1 can be regarded as the relatively independent effect of the new current system. The final result is naturally very much dependent on a reasonable choice of the times t_1 and t_2. However, even when the current system at the time t_1 does not stay constant until t_2 the patterns of the resulting DEC-vectors still reveal clearly the changes in the pre-existing system (e.g. Fig. 4). It should be pointed out that, due to the implied ambiguity, DEC-vectors have to be interpreted with care. In most cases other data (e.g. optical observations) allow restriction of the various possible causes of a certain pattern. The idea of splitting the total magnetic disturbance field was first introduced in the analysis of magnetic-latitude profiles by Kisabeth and Rostoker (1973) to study the small scale magnetic effects connected with auroral loops and surges. It was later applied by Untiedt et al. (1978) to data from a two-dimensional magnetometer array, and has shown its power in separating transient local phenomena from the more intense global magnetic disturbances.

Other ground based data considered in the analysis of the present event were available from the Danish Ionlab riometer network (Stauning 1978), the riometer operated in Kiruna, the Finnish chains of riometers (for technical details see Ranta 1978) and pulsation magnetometers, and the *S*candinavian-*T*win-*A*uroral-*B*ackscatter-*R*adar (STARE) at Hankasalmi, Finland and Trondheim, Norway (Greenwald et al. 1978). Unfortunately the observation area of STARE was slightly too far to the north for the special aims of this study. A meridian scanning photometer equipped with filters for measuring the wavelengths 486.1 nm (H_β), 490.7 nm (N_2^+), 630.0 nm (O_1), and 480.1 nm (background) was in operation in Kiruna, in addition to the all-sky camera. Because of the change of filters, the time resolution for this instrument was 4 min for a constant elevation. A faster, 4-channel, simultaneous-scanning photometer was operated in Lavangsdalen (Norway), but some operational changes during the time interval considered restricted the use of the data for our purposes.

Due to fortunate circumstances the ground-based data can be supported by a photograph from the DMSP-F2 satellite taken during a pass over Scandinavia (pass. no. 3837) with a camera scanning perpendicular to the satellite orbit (see Fig. 6). The picture shows the auroral situation over Scandinavia at about the time of the substorm onset, and the situations over Iceland and Greenland at later times. The movement of the sub-satellite point could be used to infer a time axis. The given coordinates have been corrected to a sphere at the altitude of the visible auroral forms, i.e. 100 km, with allowance for the varying aspect angles of the scanning camera. A more detailed description of the use of DMSP-satellite data has been given by Pike and Whalen (1974) and Eather (1979).

3. Observations

General Situation

The substorm analysed from the night of 2 March 1978, between 2200 and 2300 UT, is not isolated in the classical sense, but well separated from other activations in the course of the night. It's clear and distinct features also support the assumption that it is independent from the previous substorm, which started around 1830 UT.

This earlier substorm has a duration of more than two hours, reaching an intensity of about 500 nT. During this time considerable auroral activity is observed north of Scandinavia. Also, the main part of the substorm-connected westward electrojet is situated north of the area of observation. After these preceding substorm activities, the magnetic disturbances at stations all along the nightside of the auroral oval return to almost zero level until 2200 UT Correspondingly, auroral observations for that time show clear dark sky.

In Fig. 2 the three components of magnetograms along profile 4 (i.e., from SOY to SAU) including BJN are plotted in a stacked form to give an overview over the magnetic signatures of the substorm studied in this paper. It should be noted that they are relatively weak, reaching a maximum intensity at MUO of only 200 nT The onset of the substorm can be recognized by the sudden deflections of all components at about 2223 UT The disturbance is very localized and at that time restricted mainly to the Scandinavian area. Only later, in the course of the substorm expansion-phase, do magnetic stations further east (e.g., Dixon, USSR) and west (e.g., Leirvogur, Iceland) record the start of a negative H-bay (data not shown here). At stations in mid-latitudes and even sub-auroral latitudes, the effects of this substorm are too small to permit a reliable analysis.

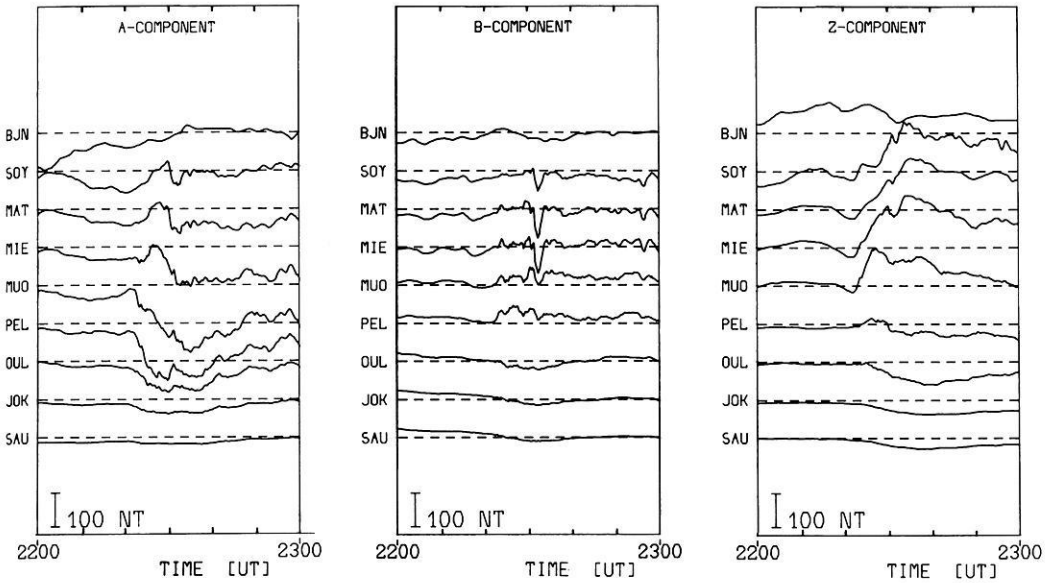


Fig. 2. Magnetic variations on 2 March 1978 from 2200 to 2300 UT at stations along a meridian. *A* and *B* denote horizontal components parallel to the x_{Ki} – and y_{Ki} – axis respectively. *Z* is positive downwards. *Horizontal lines* indicate undisturbed levels (see Sect. 2)

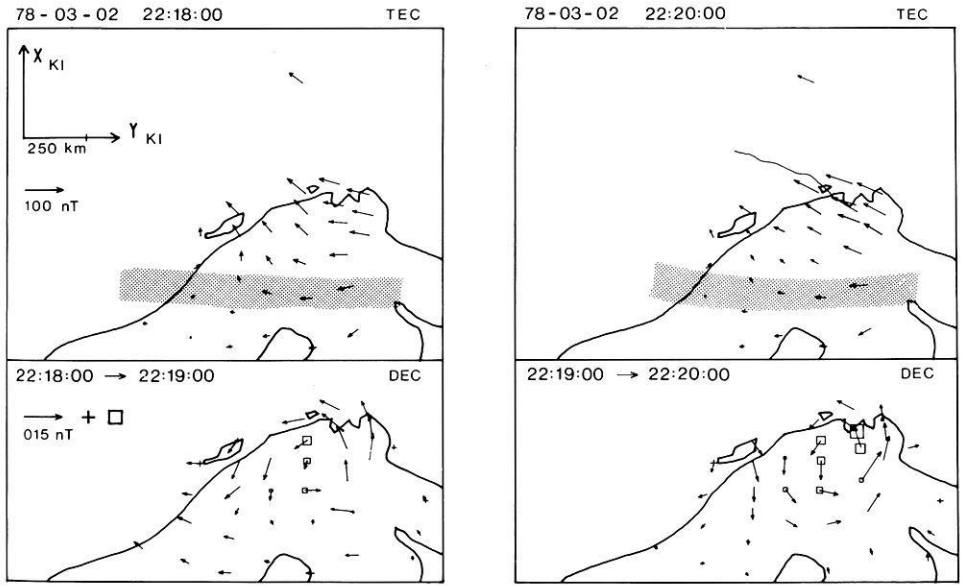


Fig. 3. Total horizontal equivalent current vectors at stations within Scandinavia before and after the appearance of a new discrete auroral form (aurora indicated by *solid lines* and *shaded area*) (*upper panels*).

Differential equivalent current vectors for the two minutes during which the new auroral form becomes appearent and grows in intensity. The vertical components are indicated by the size of the station symbol (□ = negative, + = positive) (*lower panels*)

Prephase

Before the actual substorm onset, very faint growth-phase phenomena can be observed. Soon after 2200 UT a very weak westward equivalent current starts to grow slowly in intensity. At the same time a broad, diffuse auroral band becomes visible in the ASC-pictures and the Lavangsdalen photometric recordings, extending from the western to the eastern border of the observation area. The location at 2218 UT, as shown in Fig. 3 (upper left panel), is about $x_{Ki} = -50$ to -150 km. After 2218 UT some discrete, but nevertheless still faint auroral forms suddenly appear over northern Norway. This change in auroral activity causes no clear response in the total equivalent current vectors (Fig. 3, upper right panel).

It is more interesting to look at the DEC-vectors for the time interval during which the discrete aurora appears (Fig. 3, lower panels). They show that a circular, fully closed, equivalent current with a counterclockwise sense of rotation is added to the pre-existing current system at approximately the position of the new aurora. From the ASC-pictures it cannot be clearly deduced that the development of the new auroral form takes about 2 min, as should be expected from the DEC-patterns. The Lavangsdalen

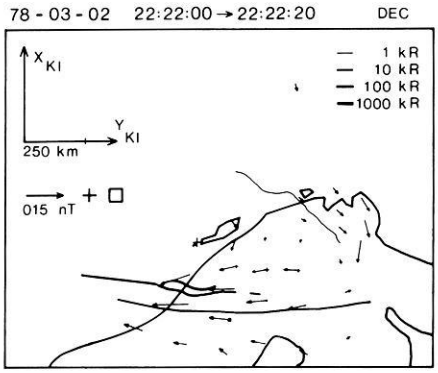


Fig. 4. Horizontal components of the differential equivalent current vectors for a 20 s interval of sudden auroral brightening. (The *brightened aurora* is represented by *solid curves* of different thicknesses, indicating different intensities.)

photometer, however, reveals that, in fact, the growth in the 557.7 nm intensity is 3 kR from 2228 to 2219 UT and still 2 kR from 2219 to 2220 UT. Referring to the papers by Fukushima (1976) and Untiedt et al. (1978) such a ground magnetic signature may be attributed to an onset of a localized upward FAC and the associated current flow in the ionosphere. Due to the ambiguity of DEC-vectors, a switch-off of a localized downward FAC could also be responsible for the same pattern. In this case the concurrent observation of the appearance of a new discrete aurora, with good correlation in time and location, makes it possible to decide that the feature observed in the DEC-vectors is most likely to be produced by a localized upward FAC, carried by precipitating electrons, which also excite the observed auroral emission.

At 2222:00 UT the diffuse band further to the south brightens and develops into discrete auroral arcs, as shown in Fig. 4; the DEC-vectors for this time interval are shown in the same figure. They reveal that, at this time, the magnetic ground signature at the position of the auroral brightening is due to a pure enhancement of the westward electrojet without a perturbation in the B-component. In the meantime the equivalent current in the northeast decreases.

Substorm Onset

Immediately before the aurora breaks up, the southernmost arc fades again for a very short time of about 20–30 s. This fading can be seen in the ASC-photograph at 2222:43 UT in Fig. 5. Perhaps due to the very short time and narrow latitudinal extent of the fading, no corresponding features in the magnetic recordings can be uniquely indentified.

The start of the breakup over Scandinavia can be fixed by ASC-data to shortly after 2222:43 UT (see Fig. 5). It starts clearly to the west of the observation area and expands very fast at first in a longitudinal direction. Even though we observe, in this case, only the eastward expansion, it seems reasonable to assume that the same effect happens on the western side as well. Since this longitudinal expansion along the brightening arc occurs within approximately 40 s it would probably appear in ASC-data with 1 min time resolution as an instantaneous brightening of the equatorward arc, as reported by Akasofu (1964). The northward expansion speed of the aurora immediately after the breakup is about 2 km/s, later decreasing to 1 km/s, according to the digitized ASC-data. The northward expansion over western Scandinavia can also

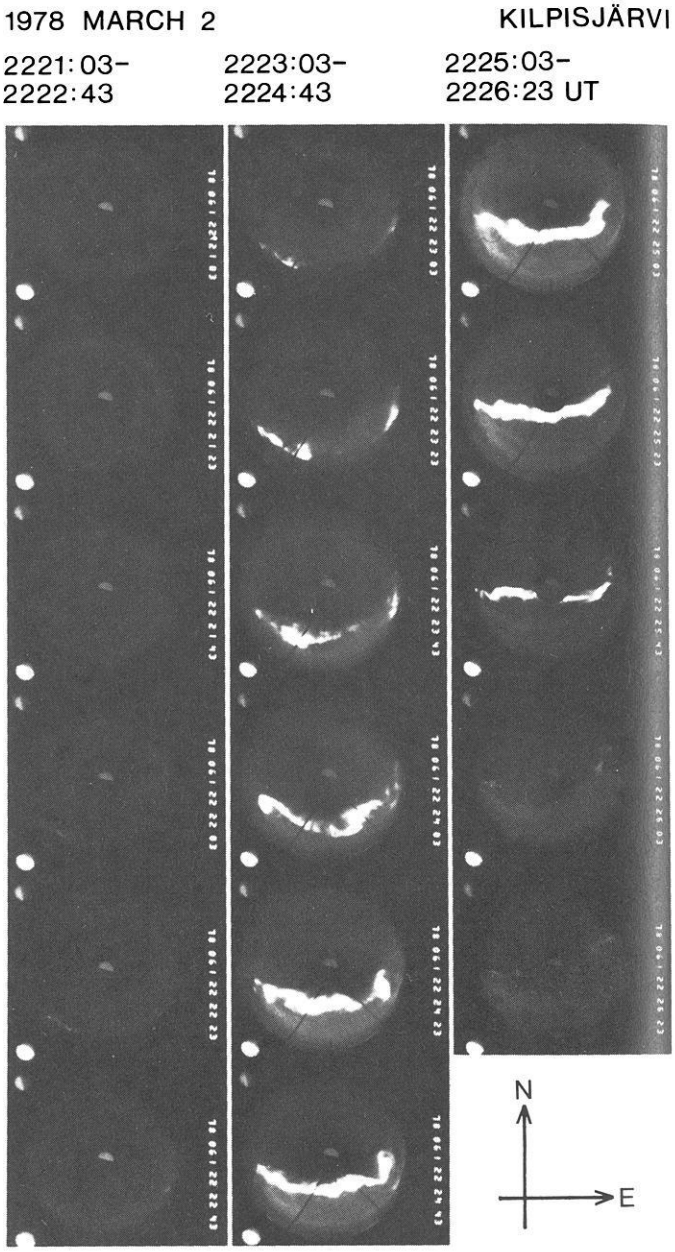


Fig. 5. All-sky camera photographs at Kilpisjärvi, Finland, for the time interval discussed on March 2, 1978

be recognized in the DMSP-picture in Fig. 6. From the picture it is possible to see that the main part of the intense aurora, and thus probably the initial location of the breakup, lies west of the Norwegian coast at about 0°–5° eastern geographical longitude, i.e., very close to local magnetic midnight (Whalen 1970).

The auroral breakup is accompanied by a strong burst of Pi-type pulsations (T. Bösinger, private communication) and a sudden increase in the ionospheric D-layer absorption. The riometer recordings in Fig. 7 show that the sharpest absorption onset occurs simultaneously with the auroral breakup and is only recorded by stations at the same latitude, i.e., Kiruna and Ivalo. A comparison of shape and amplitude of the onset at both stations supports the idea that the main area of particle precipitation lies nearer to Kiruna than Ivalo and may possibly correspond to the area of initial auroral breakup. It should be noted that the enhancement

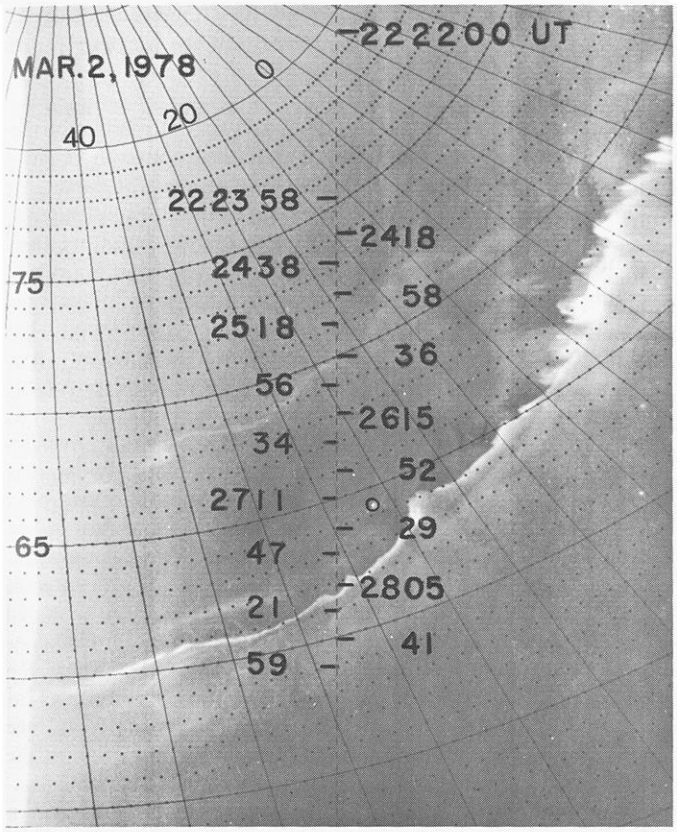


Fig. 6. Photograph from the DMSP-F2 satellite taken over the area between Scandinavia (*upper right corner*) and Greenland (*left side*) at about the time of the substorm onset. The city lights of Reykjavik are encircled, The time axis refers to the movement of the sub-satellite point. The coordinates are corrected for a height of 100 km above the earth's surface

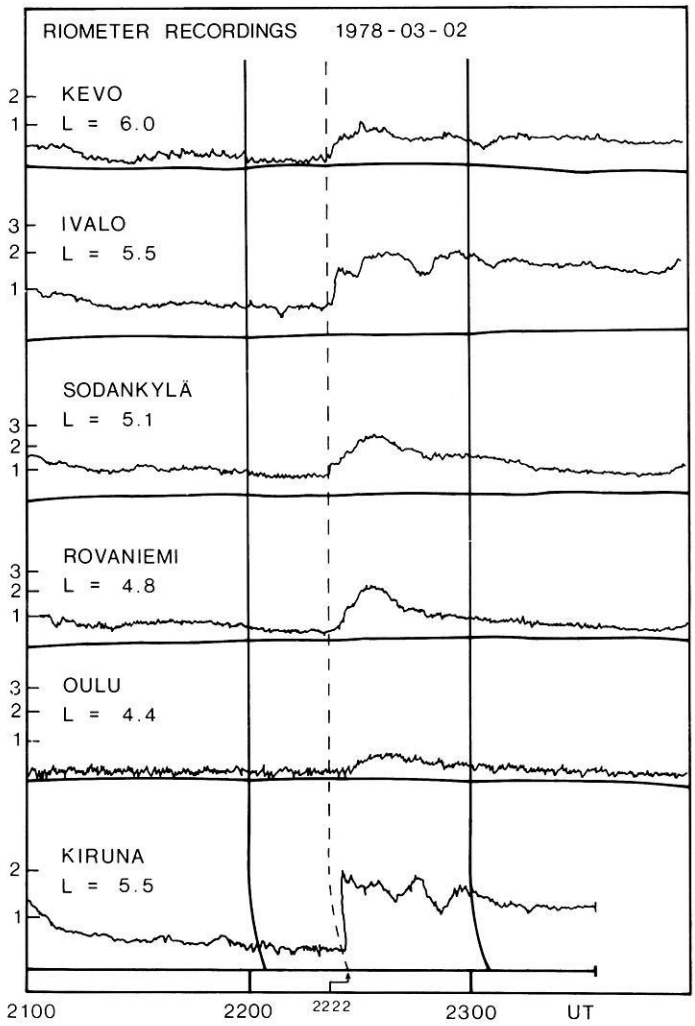


Fig. 7. Relative ionospheric absorption during the night of March 2, 1978 recorded by some riometers within the Scandinavian area. (The curve from Kiruna is taken from a different type of strip-chart recording as can be seen from the curved time marks.)

of the diffuse arc at 2222:00 UT (Fig. 4) may correspond to the smaller absorption onset observed at Sodankylä, slightly before the main absorption increase.

In Fig. 8 the total equivalent current systems before and after the substorm onset are shown. The TEC-vectors at 2222:00 UT (upper panel) are strongest in the vicinity of the visible aurora. The electrojet is still clearly separated into two major portions, but, as pointed out before, the southern portion grows towards the onset, while the northern one decreases. About two minutes after the onset at 2225:00 UT (lower panel) the westward equivalent current flow at the latitude of the auroral intensification is drastically enhanced.

To get an idea about the current system that is directly related to the breakup, and superposed on the pre-existing one, it is again reasonable to consider the DEC-system, based on the time of the moment when the aurora faded prior to the breakup, i.e., 2222:40 UT. The time by which a new current system has possibly developed seems at first glance to be 2225:20 UT, when the aurora reaches it's first maximum in northward expansion and increase of brightness. The resulting DEC-vectors in Fig. 9 (central panel) show, in the western part, a distinct U-shaped equivalent-current loop around the location of initial auroral intensification. The sense of rotation in this loop is counterclockwise with growing negative Z variations towards it's center. But in the northeastern part of the picture there is another equivalent current loop with a clockwise sense of rotation. Correspondingly, here the enclosed Z variation is positive. To understand the development of this system in space and time we analysed several series of DEC-vectors for varying time intervals as well as based on different starting times. Finally we were able to trace the following history of the substorm-connected equivalent current system during the first few minutes after the onset:

At the time of the auroral breakup a counterclockwise equivalent current loop develops around the location of the initial auroral intensification. The development of this loop lasts until the aurora reaches a first maximum of brightness at 2225:20 UT. The formation of the clockwise equivalent current loop starts simultaneously with the other one, but develops much more slowly. As can be seen in the right-hand panel of Fig. 9, a good portion of it still seems to be added during the interval from 2225:20 UT to 2226:20 UT, when the development of the counterclockwise loop has already finished. During the same time a decrease of auroral intensity is observed over the Scandinavian area (see Fig. 5). From the DMSP-photograph in Fig. 6 it can be seen that, at this time, there is no decrease in auroral activity further to the evening side of the auroral oval. Also, from the Kiruna ASC-data, there is some evidence for an intense aurora remaining south-west of

the Lofot islands in the vicinity of the counterclockwise current loop. The approximate position of the remaining bright aurora is indicated in Fig. 9 (right panel). Unfortunately the Kiruna city lights disturb the western part of the ASC-pictures and prevent a more exact location of the aurora in this part of the observation area. From Fig. 9 it can also be noted that the position of the clockwise equivalent current loop is north of the breakup aurora during the first minute after substorm onset and later within the area where the aurora ceases.

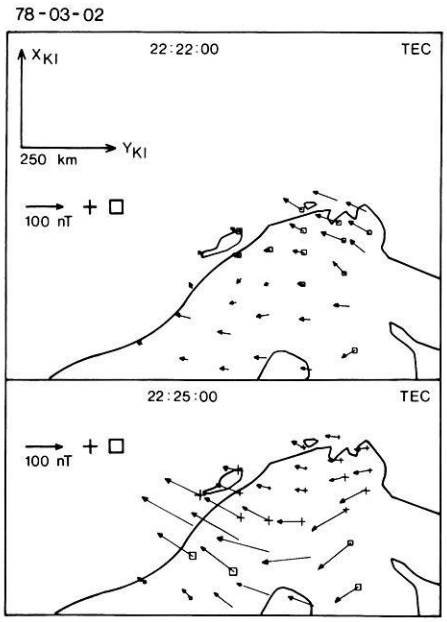


Fig. 8. Total equivalent current vectors for times before and after the substorm onset. (Vertical components again indicated by the same symbols as in Fig. 3.)

Considering the excellent temporal and spatial correlation between aurora and magnetic field as described above there must be a connection between the obviously local decrease of the intense discrete aurora and the main development of the clockwise equivalent current loop. Vice versa, there seems to exist a causal relationship between the counterclockwise equivalent current loop and the auroral breakup. But, even though a time shift in the final development of the two loops of about 2 min is obvious, it has not been possible to find a clear separation of the two loops in time. Such a separation would have indicated the possibility of a switch-on and switch-off of one and the same current system. However, in this case, the two oppositely directed equivalent-current loops clearly coexisted during the whole 3–4 min after the substorm onset. Thus it seems reasonable to consider both features as being connected with the ground signatures of the three-dimensional magnetospheric substorm onset current-system. The resulting pattern is shown by the DEC-vectors for the interval 2222:40 to 2226:20 UT in Fig. 9 (left panel).

Unfortunately the observation field of the STARE-radar does not reach far enough to the south to cover the area where the auroral breakup takes place, However, in the recordings at 2225:20 UT (shown in Fig. 10) there seems to be some evidence for a clockwise curvature in the irregularity drift-direction at the position where we observe the northern part of the counterclockwise equivalent current loop. Also, corresponding to the western part of the clockwise equivalent current loop, there is a counterclockwise irregularity drift at the same location.

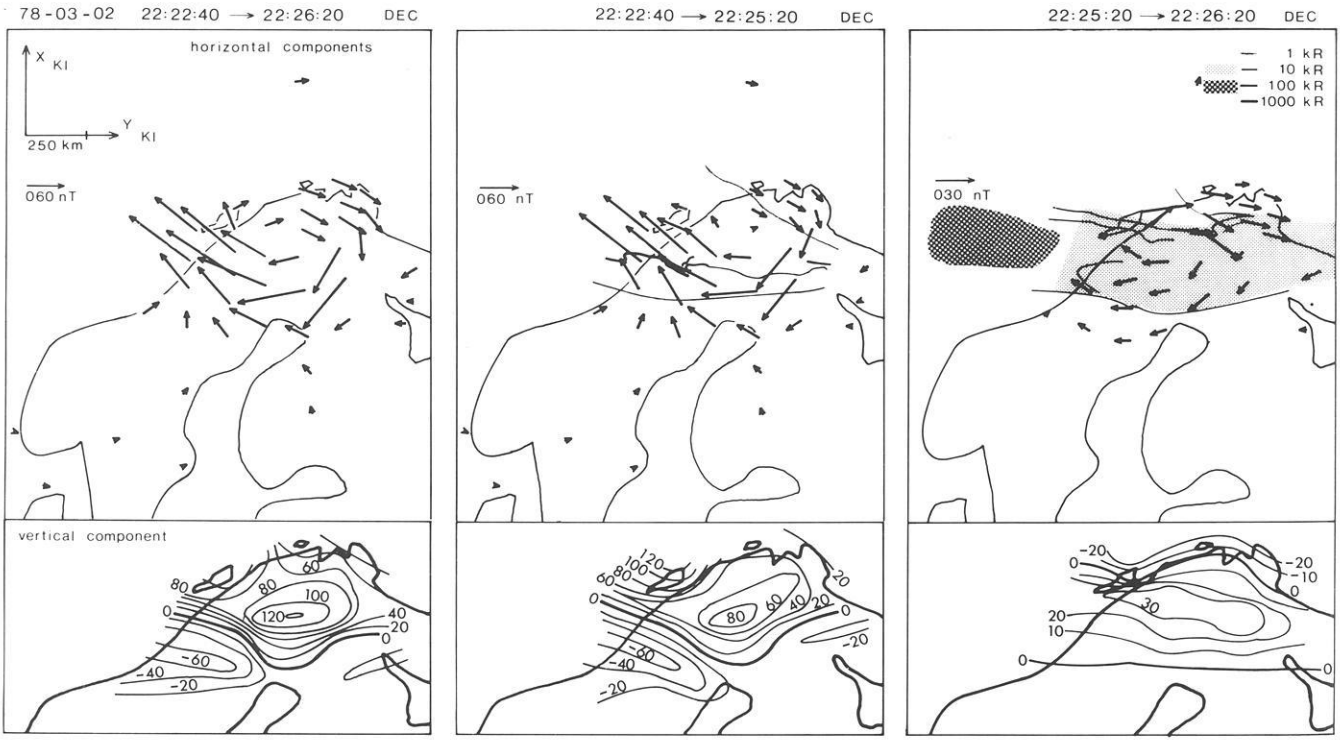


Fig. 9. Differential equivalent current vectors for selected intervals during the early time of the substorm development. The vertical components are shown separately in the lower panels by isolines computed from the values recorded at the stations. The aurorae at 2223:00 UT in the central and 2226:00 UT in the right panel are indicated by *solid curves* and *shaded areas* with intensity scaling

Expansion Phase

As pointed out in the previous chapter, the expansion speed of the aurora to the north is 2 km/s at the beginning of the breakup, and later, 1 km/s. This northward expansion of the poleward border of the visible aurora lasts until about 2235:00 UT, reaching a northernmost position of about $x_{Ki}=400$ km. The expansion speed decreases further during the later expansion phase. Until the end of this phase the aurora enhances and decreases several times and some possibly related phenomena can be found in the magnetic data. However, most of these very local features are embedded in the simultaneous growth of the westward electrojet. A further analysis will therefore be relatively complicated and we will devote this particular study to the substorm onset.

The limited area of ASC-observation does not allow observation of substorm expansion to the east and to the west, but in the magnetic data an expansion of the equivalent current system can be traced for at least the first three minutes after the substorm onset. The clockwise loop which is (immediately after the onset) relatively small and of nearly circular shape expands to the east. A series of DEC-vectors at one minute intervals, each of them based on the end time of the previous one, (shown in Fig. 11) reveals that the equivalent current added to this loop between 2224:00 and 2225:00 UT is already oval shaped, and that the DEC-vectors for the time from 2225:00 to 2226:00 UT no longer show a closure to the east, but are, by that time, already U-shaped. During the whole development of the eastern border, both the western border of this loop and, on the opposite side, the eastern border of the counterclockwise loop stay remarkably stable. It can be concluded, therefore, that the same kind of expansion to the west takes place on the western side of the counterclockwise loop, and is not observed due to a lack of stations in that area. Baumjohann (1979) reported an expansion of the same character as shown in Fig. 11 for a counterclockwise equivalent-current loop, which had been observed in relation to a substorm onset over the eastern part of his observation area. That a westward expansion of the substorm phenomena in fact occurs in the present case as well can be deduced from the following observations further to the west.

As mentioned before, the Finnish chain of riometers, as well as the riometer at Kiruna, recorded at the time of the substorm onset a steep increase in the ionospheric D-layer absorption. The same observation was made by Stauning (1978) with his riometer at Ramfjord, Norway. However, riometers further to the west show a significant delay in the absorption onset. Figure 12 shows how the area of increased absorption expands to the west, reaching

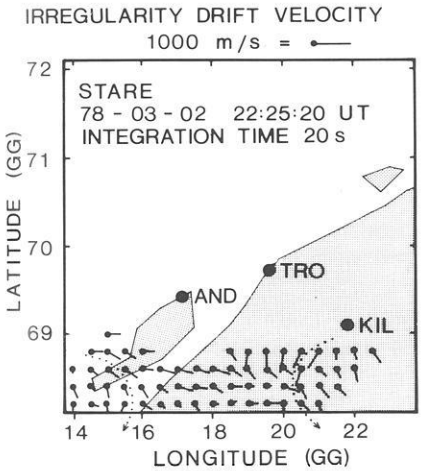


Fig. 10. Irregularity drift velocity vectors computed from the STARE-data at 2225:20 UT

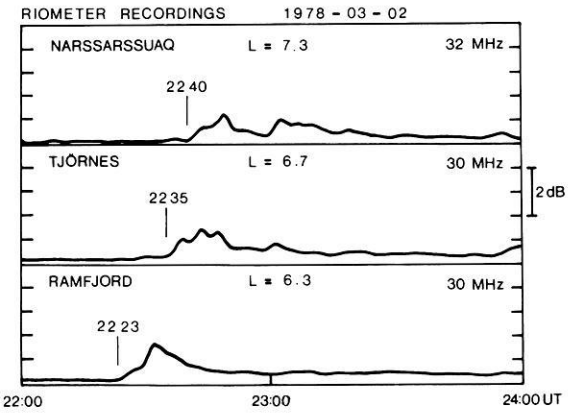


Fig. 12. Riometer recordings at stations in Scandinavia, Iceland, and Greenland on March 2, 1978 (from Stauning 1978).

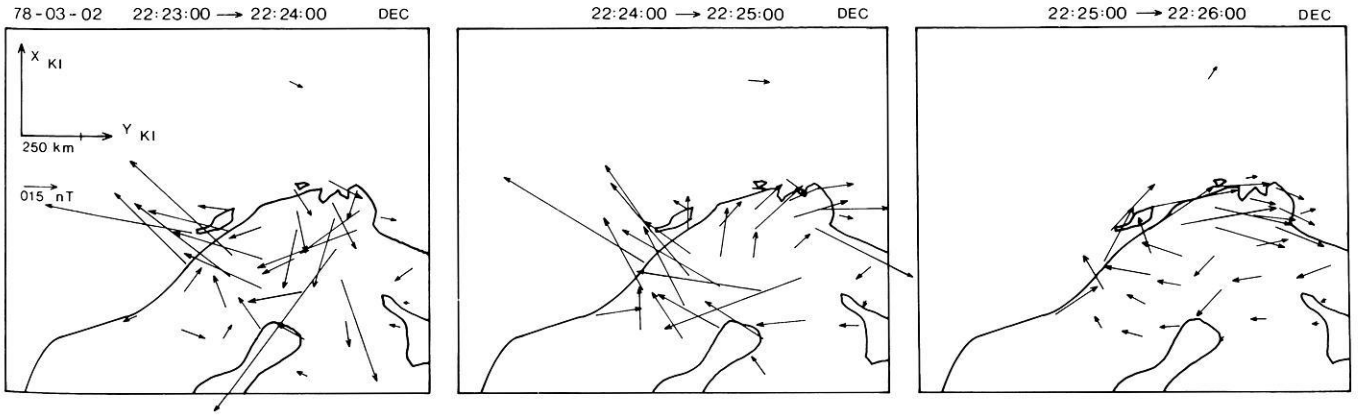


Fig. 11. Differential horizontal equivalent current vectors for three subsequent one-minute intervals during the early time after the substorm onset

Tjörnes on Iceland at about 2235:00 UT. The expansion possibly continues as far as Greenland, where the absorption over Narssarssuaq rises suddenly at 2240:00 UT. All three stations are on approximately the same geomagnetic latitude and, during this night, as can be seen from the satellite picture in Fig. 6, well situated within the auroral oval. Magnetic stations at the same location (Narssarssuaq) or very near to the riometer (Leirvogur) record the onset of a negative H bay at the same moment as the area of increased absorption arrives (data not shown here). In both cases the negative H variation is accompanied by a negative Z deviation and a sudden distortion in the D component which is clearly positive before the arrival. If one assumes that the western border of the substorm onset current system began at about 5° E geographic longitude, the resulting average expansion speed to the west would be about 1.3 km/s between Scandinavia and Iceland and about 4.5 km/s between Iceland and Greenland.

To the east, the only station from which data was available to the authors is Dixon, USSR, about 2,000 km east of Scandinavia. Here a small negative H onset was recorded at 2232:00 UT, which could possibly be related to the substorm onset over Scandinavia under discussion. Assuming the initial eastern border of the substorm onset current system at 30° E geographic longitude, this would correspond to an eastward expansion speed of about 3.4 km/s.

4. Discussion

Even though the substorm analyzed in this paper produces only very weak magnetic disturbances, it shows all the features that are known to occur during the course of a substorm of larger intensity. The clear separation from the previous intense substorm, as well as the faint growth-phase phenomena, prove that the event is independent and not an enhancement of an earlier system, as in a multiple-onset substorm reported by Pytte et al. (1976).

Development Before the Substorm Onset

The slow increase of westward equivalent current that starts after 2200:00 UT may be the result of a weak growth-phase preceding the substorm onset. Enhanced particle precipitation is indicated by the faint diffuse band of auroral luminosity that becomes visible in the ASC-pictures after about 2215:00 UT The diffuse and faint character of the visible aurora and the related westward equivalent current are not enough information to decide whether a southward movement of the aurora and the related current system takes place during the growth phase. Nevertheless the observed features clearly occur south of the latitude of the previous substorm activity.

The sudden appearance of discrete aurora at 2119:00 UT may be connected to a magnetospheric growth-phase. Several authors (Akasofu 1964; McPherron 1970; Pytte et al. 1976; Pellinen and Heikkila 1978b) have reported the observation of transient brightening or the appearance of discrete aurorae connected to small negative perturbations in the magnetic H component, taking place particularly during the growth-phase of a magnetospheric substorm. This phenomenon has been given different names, e.g., trigger bay (Rostoker 1968), pseudo-breakup (Akasofu 1964), etc., but in all cases it was believed to be of magnetospheric origin, and to be related to the same kind of instability which, in a more effective situation, leads to the onset of a real substorm. Our observations (Fig. 3) strongly suggest that these precursors are connected to a localized onset or enhancement of upward field-aligned current, which would support the opinion of the above cited authors (see also the observations of Untiedt et al. 1978).

At the second auroral intensification observed before substorm onset, taking place just after 2222:00 UT (see Fig. 4), the aurora enhances along the whole portion of the auroral oval that can be observed by our instruments, causing a pure strengthening of the westward equivalent-current flow at the narrow latitudinal region of brighter aurora. Immediately after this final auroral enhancement, before the actual onset, the auroral arc fades for a time shorter than 40 s, as can be seen from the ASC-picture at 2222:40 UT, in Fig. 5.

According to the theoretical model for substorm onset in the magnetotail by Pellinen and Heikkila (Heikkila and Pellinen 1977; Pellinen and Heikkila 1978a, b; Pellinen 1979) this fading occurs due to a localized decrease in the magnetospheric cross-tail electric field, which results in a reduction of the most effective acceleration mechanism for quiet time auroral particles, the Fermi (or curvature drift) acceleration (Sharber and Heikkila 1972). Pellinen (1979) has given a detailed description of the sequence of events that leads to auroral fading: the increase of the cross-tail current and the related thinning of the plasma sheet will be approximately linear during early times of the substorm growth phase, but, after a certain moment, the increase of the cross-tail current becomes locally non-linear. During the linear growth and the early phase of the non-linear growth, enhanced particle precipitation into the ionosphere can be observed, which results in a momentary increase in the D-layer absorption, ionospheric conductivity, and auroral brightness in a narrow latitudinal region. In fact, all these observations have been made around 2,222:00 UT

The non linearly growing cross-tail current will induce a growing electric field which is oppositely directed to the magnetospheric dawn-to-dusk field and will, after certain time, become locally stronger than the latter (Pellinen 1979). The short-lived situation of balance between the two fields will cause a momentary fading of the ionospheric phenomena associated with reduced particle precipitation (Pellinen and Heikkila 1978b). As soon as the induced electric field dominates the cross-tail electric field in a localized region of the neutral sheet, the mechanism described by Heikkila and Pellinen (1977) and Pellinen and Heikkila (1978a) may result in the onset of the actual substorm phenomena.

Field-Aligned Currents Associated With Substorm Onset

In the introduction, we have pointed out that, with a network of the type and spacing as used for this study, it should be possible to infer a good portion of the onset-connected three-dimensional current system. Based on the data presented in the previous section, we believe strongly that, in this case, our ground-based stations were situated directly under (or at least very near) the footprints of a localized pair of oppositely directed field-aligned currents, which were related to the substorm onset. Let us try to consider the magnetic and electric effects of such a pair of field-aligned currents on the ionosphere and magnetometers on the ground.

Under the assumption of homogeneous or axi-symmetric conductivity in the ionosphere, one should expect circular counter-clockwise (clockwise) ionospheric Hall-currents to be driven around the upward (downward) FAC, decreasing with growing distance from the center. In such a case only these Hall-currents

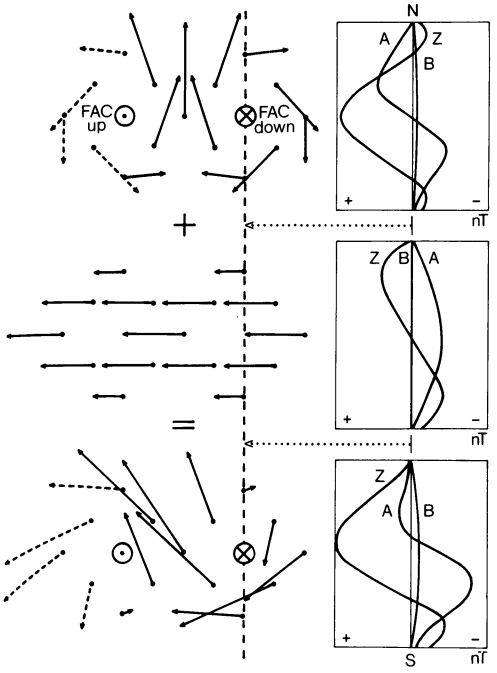


Fig. 13. Theoretical pattern of the magnetic ground effect of two vertically-incident field-aligned line-currents into a homogeneously conducting ionosphere in the forms of equivalent current vectors and a representative latitude profile (*upper panel*).

Ground perturbation of an ionospheric westward current (*central panel*). Combined ground effects of the above two panels (*lower panel*).

The chosen latitude profile is indicated by a *broken line, broken vectors* indicate the points for which no corresponding measurements could be presented in the experimentally observed pattern in Fig. 9

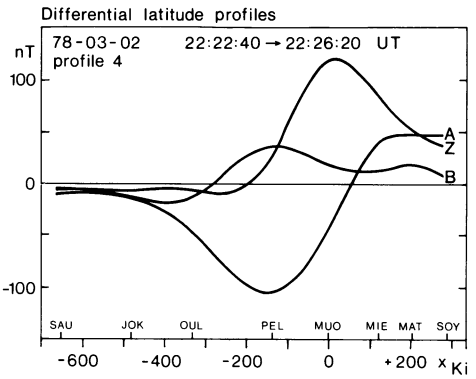


Fig. 14. Differential equivalent latitude profiles for the time interval during which the onset-connected current system develops. The meridian is located near the center of the clockwise equivalent-current loop

will produce a magnetic disturbance on the ground, while the effects of the radial Pedersen-currents, combined with the FACs, will vanish (Fukushima 1976). In Fig. 9 (left panel) two equivalent current loops can be seen, having the correct sense of rotation for the expected field-aligned current configuration at a substorm onset. The observation that, at the positions of the loops, the irregularity flow directions recorded by the STARE-radar (Fig. 10) are antiparallel to the equivalent currents supports the idea that the observed ground effect is produced by real ionospheric Hall-currents (Baumjohann et al. 1978).

Besides the two oppositely directed loops in Fig. 9, a certain asymmetry becomes apparent, which indicates the additional effect of a westward directed current. In Fig. 13 we demonstrate that, in fact, a superposition of two Hall-current loops and an additional westward current can account for the equivalent current pattern observed in connection with the substorm onset. The westward current will partly cancel the oppositely directed parts of the loop currents, while the parallel directed parts will be enhanced. The maximum and minimum of the vertical Z component will apparently be shifted to the north and south, respectively. In comparison to the theoretical latitude profiles for a meridian exactly through the center of the clockwise equivalent current loop, the observed latitude profiles for all three components along profile 4 (i.e., from SOY to SAU, see Table 1 and Fig. 1) are shown in Fig. 14. It can be seen that the observed profile is in fact very similar to the one that should be expected for the modelled superposition of a clockwise Hall-current loop and an additional westward current. The positive B-component in Fig. 14 indicates that profile 4 must be situated to the east of the clockwise loop center.

But what, in this case, is the reason for the observed westward equivalent current? The assumption of a homogeneous conducting ionosphere will not be valid for auroral latitudes, but (due to continuous particle precipitation along the auroral oval) there will always be a region of enhanced conductivity. Nevertheless, locally homogeneous conductivity will, to a first approximation, still be a reasonable assumption for the explanation of ground signatures of very localized FACs, as observed in connection with pseudo-breakups and auroral spirals, when the effective width of the FAC-connected disturbance is smaller than the latitudinal width of the auroral oval. In the case of the more intense, and possibly wider substorm onset connected FACs, however, a belt of higher conductivity has to be considered.

In a series of papers Fukushima (1974, 1975a, b) has discussed the ground magnetic effect of a single downward directed FAC which flows into a belt with Hall- and Pedersen-conductivity 3 and 2 times larger, respectively, than in the surrounding ionosphere. In order to maintain current-continuity he had to assume certain boundary conditions at the borders of the belt. The most probable conditions discussed are open field-lines at the poleward boundary, allowing secondary FAC to flow, and closed field-lines at the equatorward boundary, resulting in charge accumulation and dissipation via secondary ionospheric currents.

The basic results of his calculations are that, for such a configuration, the circular Hall-currents remain the same in principle as in the case of homogeneous conductivity, but are stronger inside the belt than outside. The Pedersen-currents, however, are no longer completely matched by the ground effect of the FAC, but have a dominating component along the belt in both directions away from the point of primary incoming FAC. If one adds to Fukushima's results the effects of an upward FAC to the west of the downward one, the resulting ground-effective Pedersen-current will be westward directed. The combined ground perturbations of the Hall- and Pedersen-currents in such a case will be somewhat similar to the pattern drafted in Fig. 13.

However, from Fukushima's numerical results it becomes clear that in all cases discussed, where the auroral-zone Hall-conductivity is larger than the Pedersen-conductivity, which according to Brekke et al. (1974) is a correct assumption, the magnetic ground perturbation caused by the ionospheric Pedersen currants will be several times smaller than the one caused by Hall-currents. Also,

other model calculations (W. Baumjohann, private communication) for a similar type of observed equivalent current pattern, assuming a conductivity channel with the ratio of Hall- to Pedersen-conductivity being 2/1, show that the westward Pedersen-current produced by a pair of field-aligned line currents alone can not account for the observed perturbation of the two Hall-current loops. Thus the only explanation for this discrepancy must be that, after the onset of the substorm, the quiet-time westward Hall-current, which is driven by the permanently existing global FAC-configuration, also increases. In fact, observations by Iijima and Potemra (1976) and Mozer (1971) have shown that the global FACs and associated electric fields do increase immediately after substorm onsets.

It should be mentioned here that all the models discussed above still neglect the inclination and curvature of the field-aligned currents, which should produce magnetic effects mainly on the equatorward side of the onset region. For a more sophisticated modelling of the real three-dimensional current configuration, the effects of currents induced in the ground have also to be considered more exactly than with only the assumption of a perfect conductor at a certain depth. In Fig. 9 (lower left panels) it can easily be seen that, e.g., the inductive coast effect is relatively strong at the Norwegian coast (station AND), producing distortions mainly in the Z component, during rapid changes in particular. Nevertheless erroneous contributions from the ground will not give rise to a wrong interpretation for the current pattern considered, the error being mostly in magnitudes.

Supporting Observations

So far we have considered only magnetic data in discussing whether the observed phenomena could have been produced by a localized pair of FACs. This is, in fact, in agreement with several other observations, as we will show below.

A sudden Pi-type pulsation burst is, according to other reports (e.g., Saito 1969; Kisabeth and Rostoker 1971, Pytte et al. 1976), very closely connected to the onset or enhancement of FAC and related ionospheric current systems. At this substorm onset a Pi-burst was observed to start simultaneously with it between 2222:00 and 2223:00 UT, to reach maximum amplitude at about 2225:00 UT, and to cease shortly after 2226:00 UT (T. Bösinger, University of Oulu, private communication). This is exactly the time interval during which we observe the development of the entire double loop system in the DEC vectors.

The very good spatial and temporal correlation between the optical auroral observations and the development of the onset-connected equivalent current system is in agreement with our interpretation. The initial auroral breakup, and later the remaining intense discrete aurora is very well situated within the counterclockwise current loop, i.e., under the upward directed FAC which should be carried mainly by energized precipitating electrons (Akasofu et al. 1969). Vice versa, the downward FAC should be carried at least to some extent by energized precipitating protons and positive ions (Potemra 1979; Klumpar et al. 1976; Pellinen and Heikkila 1978a) which are not able to produce bright aurora (Vallance Jones 1976, p. 54). This is in agreement with the observation that the aurora ceases over the Scandinavian area after 2225:20 UT, when there is still a large portion of clockwise equivalent current added to the onset-connected current system. The DMSP-photograph and the ASC-data show that this decrease in the aurora seems to be restricted to the area of probable downward FAC only.

The clear delay of about 2 min between the arrival of substorm-energized electrons (as indicated by the auroral breakup and the

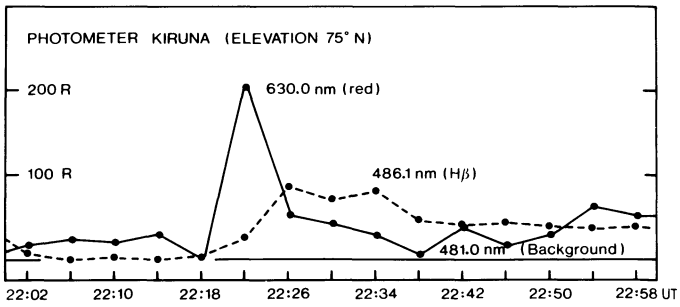


Fig. 15. Photometric recordings of the photometer in Kiruna for the wave lengths 486.1 nm (H_β) and 630.0 nm (red O_1), with background subtracted, for a representative elevation angle of 75° N. The original curve for the 630.0 nm line has been shifted 110 s back in time to correspond to the excitation of the aurora rather than its emission

counterclockwise loop representing upward FAC) and protons (as indicated by the localized cessation of the aurora and the final development and widening of the clockwise loop representing downward FAC) is an important observation. The delay had been reported earlier by Fukunishi (1975). If one accepts that the substorm onset connected energization of particles of both types must occur at the same time, by a symmetric mechanism capable of producing the observed distribution pattern, the relatively large delay in the arrival of the slower protons must mean that the energization mechanism is located far away from the earth, most probably in the magnetotail. Differences in arrival times of electron and proton beams produced in the magnetotail at substorm onset were also obtained in the simulation calculations by Pellinen and Heikkila (1978a).

Some photometric data from the instruments in Kiruna (Sweden) and Lavangsdalen (Ramfjord, Norway) show the delayed arrival of protons. The photometer at Lavangsdalen was not scanning during the time interval considered, but was viewing constantly along the magnetic field lines, i.e., towards an ionospheric region approximately over Rostadalen (ROS). The H_β-emissions from this region, which is very near to the possible location of downward FAC, start to rise after 2225:00 UT, when the visible aurora decreases in that area. Also, the photometer at Kiruna records H_β-emissions from nearly all elevation angles only some minutes after the substorm onset.

In Fig. 15 only the development of the H_β- and 630.0 nm emissions from a representative elevation angle of 75° over the northern horizon at Kiruna is shown. As we are mainly interested in comparing the moments of excitation of the two emissions, the curve for the 630.0 nm emission was shifted 110 s back in time corresponding to the average delay in the emission of this particular wavelength. Thus the two curves became approximately directly comparable, although not originally recorded synchronously. It can be seen that the rise in the H_β-emissions is preceded by a high short-lived peak of red auroral emissions, which is certainly related to the substorm onset. The entire photometric data set (not shown here) reveals that, in contrast to the H_β-emissions, this peak is observed at Kiruna only from about 40° to 90° N elevation. If one compares this observation with Fig. 9 and Fig. 14, it seems as if the red emissions are restricted to the area of initial downward FAC, while the proton-emissions observed later correspond in time and location to the final area of downward FAC, which has clearly widened and spread out further to the south. Since, according to Klumpar et al. (1976),

Kamide and Rostoker (1977), and Klumpar (1979), the downward equivalent current is carried to a large extent by upward going ionospheric electrons of low energy, the photometer observations, together with the ground signatures of localized downward FAC, give rise to some questions. It may be possible that the red auroral emissions that usually come from heights greater than 200 km (Vallance Jones 1976) have in this case been excited by upward going electrons. If this happens to be the case, it would suggest that the necessary upward directed field-aligned acceleration mechanism for ionospheric thermal electrons is related to the positive charge of the cloud of energized protons arriving from the magnetotail, as the red auroral peak is observed to decrease when the protons arrive at the ionosphere. However, the answer to this question has to be postponed until better observations are available.

The expansion of the onset-connected three-dimensional current system is in good agreement with the observations of other authors. To the west, the travelling disturbance resembles the features observed in connection with westward travelling surges (Kamide and Akasofu 1975; Kisabeth and Rostoker 1973), the speed, in the beginning at least, being of the same magnitude as given by Akasofu et al. (1965a). Figure 12 shows that the expansion of this relatively weak substorm reaches at least far as Narssarssuaq in Greenland. To the east, the expansion speed is considerably faster than to the west. The northward expansion stops after about 400 km, as reported by Akasofu et al. (1965b) for a moderate substorm.

No statement can be made on whether the central meridian of the three-dimensional substorm current system stayed stable during the whole development of expansion. According to Pytte et al. (1976) the position of the central meridian and the development of the substorm expansion can be deduced from a longitudinal profile of stations in mid-latitudes. An attempt was made to apply this method to the data from this substorm also, but due to the low intensity, the amplitudes at mid-latitudes were only of the order of nT, thus being too small to permit reliable analysis.

5. Summary and Conclusions

In this paper we reported observations of a moderate substorm with an extended network of ground-based stations. Even though the majority of the disturbances were relatively weak, the most typical features of a substorm have been observed. The results of this study can be summarized as follows:

- The extent of the IMS-network of stations in Scandinavia has proved to be close to the scale-size of substorm onset phenomena.
- The three-dimensional current system, which is connected to a substorm onset, is represented most likely by a localized pair of oppositely directed field-aligned line-currents of the configuration proposed by Boström (1964) (Boström I model).
- During the later course of the substorm, the very localized three-dimensional onset current system expands to east and west and develops into a more sheet-like configuration.
- The main magnetic ground perturbations observed at the substorm onset seem to be produced by ionospheric Hall-currents, driven by the electric fields of the local, as well as the global, field-aligned current-configurations.
- The auroral breakup is obviously related to the onset of a localized, upward directed, field-aligned current, carried by precipitating electrons, which represents the western part of the three-dimensional onset current system.
- The eastern part of the onset current system can reasonably be attributed to the effect of a downward field-aligned current. Ground based observations indicate that this downward current is carried by upward going ionospheric electrons during the first 2 min of the substorm development. Later, at least some part of this current is carried by precipitating magnetospheric protons, arriving delayed.
- For the first two minutes after the substorm onset the downward field-aligned current is well correlated with red auroral emissions, which may be excited by upward accelerated ionospheric electrons. Later the intensity of H_β-emission exceeds that of the red aurora.
- A burst of Pi-type pulsations is associated with the onset of the field-aligned currents, lasting for as long as the growth in the ground signatures of these currents can be observed.

The event reported in this paper has made us believe that we are dealing with a phenomenon characteristic of auroral substorm onset, which has never, up to now, been observed in it's entirety, even though it has been included in substorm models for a long time. We are sure that, with our experience in this study, we will be able to recognize signatures similar to our data again in the future and study them in more detail. The much better optical equipment that is planned for operation in the Scandinavian area in the future, will help us especially in clarifying some of the questions which have arisen in this study.

Acknowledgements. We wish to thank all past and present members of the magnetometer groups at the Universities of Münster and Braunschweig who were involved in collecting the magnetic data. The magnetic observations were performed in cooperation with the Aarhus University, the Department of Plasma Physics of the Royal Institute of Technology at Stockholm, the Finnish Meteorological Institute at Helsinki, the University of Bergen, the Geophysical Observatory of the Finnish Academy of Sciences and Letters at Sodankylä, the Kiruna Geophysical Institute, the University of Oulu, and the University at Tromsö. To these institutions our thanks are due for permanent support.

We also wish to thank the coordinating committee of the ABC – 1 campaign for making possible the unique data coverage, which we were able to use for the present study. We are indebted to all participants of the Second Regional IMS workshop in Bad Lauterberg, Federal Republic of Germany who took part in the first discussions concerning this particular substorm, and especially to those who supported our study with additional data, such as T. Bösinger and J. Kangas, University of Oulu, who supplied us with information about pulsations, R. Greenwald, Max Planck Institut at Lindau, who provided copies of the STARE-data, G. Gustafsson and Å. Steen, Kiruna Geophysical Institute, who sent us all-sky camera, riometer, and photometer data, K. Måseide, University at Oslo, from whom we received photometric data from Lavangsdalen, H. Ranta, Geophysical Observatory at Sodankylä, who sent us the data from the Finnish riometer chain, and P. Stauning, Danish Meteorological Institute, who made riometer data from several stations in Scandinavia, Iceland, and Greenland available to us.

The DMSP photograph was kindly provided by the Air Force Geophysical Laboratory, Hanscom Air Force Base, Massachu-

setts. The Polar Geophysical Institute at Apatity put magnetograms of stations on the Kola peninsula at our disposal.

Standard magnetograms that have been available to the authors were received from the World Data Center in Copenhagen. One of the authors (WJH) is indebted to the Atmospheric Research Section of the National Science Foundation for research support, while another (HJO) wishes to thank the Deutscher Akademischer Austauschdienst for a research grant allowing him to work in Finland.

The magnetometer array observations were supported financially by the Deutsche Forschungsgemeinschaft.

References

Akasofu, S.-I.: The development of the auroral substorm. Planet. Space Sci. **12**, 273–282, 1964

Akasofu, S.-I.: A study of auroral displays photographed from the DMSP-2 satellite and from the Alaska meridian chain of stations. Space Sci. Rev. **16**, 617–725, 1974

Akasofu, S.-I.: Physics of magnetospheric substorms. Dorndrecht, Boston: Reidel Publ. Comp. 1977

Akasofu, S.-I., Eather, R.H., Bradbury, J.N.: The absence of the Hydrogen emission H_β in the westward traveling surge. Planet. Space Sci. **17**, 1409–1412, 1969

Akasofu, S.-I., Kimball, D.S., Meng, Ch.-I.: The dynamics of the aurora II: westward traveling surges. J. Atmos. Terr. Phys. **27**, 173–187, 1965a

Akasofu, S.-I., Kimball, D.S., Meng, Ch.-I.: The dynamics of the aurora V: Poleward motions. J. Atmos. Terr. Phys. **27**, 497–503, 1965b

Atkinson, G.: Magnetosphere flows and substorms, In: Proceedings of Advanced Study Inst. on Magnetosphere Ionosphere Interactions, Dalseter, Norway 1971

Bannister, J.R., Gough, D.I.: Development of a polar magnetic substorm: A two-dimensional magnetometer array study. Geophys. J.R. Astron. Soc. **51**, 75–90, 1977

Bannister, J.R., Gough, D.I.: A study of two polar magnetic substorms with a two-dimensional magnetometer array. Geophys. J.R. Astron. Soc. **53**, 1–26, 1978

Baumjohann, W.: Spatially inhamogeneous current configurations as seen by the Scandinavian magnetometer array. In: Proceedings of the International Workshop on Selected Topics of Magnetospheric Physics. Tokyo, March 1979

Baumjohann, W., Greenwald, R.A., Küppers, F.: Joint magnetometer array and radar backscatter observations of auroral currents in northern Scandinavia. J. Geophys. **44**, 373–383, 1978

Boström, R.: A Model of the auroral electrojets. J. Geophys. Res. **69**, 4983–4999, 1964

Boström, R.: Current systems in the magnetosphere and ionosphere. In: Radar probing of the auroral plasma. A. Brekke, ed.: Proc. EISCAT Summer School, Tromsö, Norway, June 5–13 1975, Universitätsfortlaget, Tromsö-Oslo-Bergen, 1977

Boyd, J.S., Belon, A.E., Romick, G.J.: Latitude and time variations in precipitated electron energy inferred from measurements of auroral heights. J. Geophys. Res. **76**, 7694–7700, 1971

Brekke, A., Doupnik, J.R., Banks, P.M.: Incoherent scatter measurements of E region conductivities and currents in the auroral zone. J. Geophys. Res. **79**, 3773–3789, 1974

Clauer, C.R., McPherron, R.L.: Mapping the local time – universal time development of magnetospheric substorms using mid-latitude observations. J. Geophys. Res. **79**, 2811–2820, 1974

Eather, R.H.: DMSP calibration. J. Geophys. Res. **84**, 4134–4144, 1979

Fukunishi, H.: Dynamic relationship between proton and electron auroral substorms. J. Geophys. Res. **80**, 553–574, 1975

Fukushima, N.: Equivalent current pattern for a field-aligned current into the auroral belt of enhanced electric conductivity, Part I: Case of no charge accumulation at the auroral-zone boundaries. Rep. Ionos. Space Res. Jpn. **28**, 207–213, 1974

Fukushima, N.: Equivalent current pattern for a field-aligned current into the auroral belt of enhanced electric conductivity, Part II: Case of charge accumulation at the auroral-zone boundaries. Rep. Ionos. Space Res. Jpn. **29**, 31–38, 1975a

Fukushima, N.: Equivalent current pattern for a field-aligned current into the auroral belt of enhanced electric conductivity, Part III: Case of open field lines on the poleward side and closed field lines on the equatorward side of the auroral belt. Rep. Ionos. Space Res. Jpn. **29**, 39–46, 1975b

Fukushima, N.: Generalized theorem for no ground magnetic effect of vertical currents connected with Pedersen currents in the uniform-conductivity ionosphere. Rep. Ionos. Space Res. Jpn. **30**, 35–40, 1976

Greenwald, R.A., Weiss, W., Nielsen, E., Thomson, N.R.: Stare, a new radar auroral backscatter experiment in northern Scandinavia. Radio Sci. **13**, in press, 1978

Gustafsson, G.: A revised corrected geomagnetic coordinate system. Ark. Geotys. **5**, 595–617, 1970

Harang, L.: The mean field of disturbance of polar geomagnetic storms. Terr. Magn. Atmos. Electr. **51**, 353–380, 1946

Heikkila, W.J., Pellinen, R.J.: Localized induced electric field within the magnetotail. J. Geophys. Res. **82**, 1610–1614, 1977

Heppner, J.P.: Time sequences and spatial relations in auroral activity during magnetic bays at College, Alaska. J. Geophys. Res. **59**, 329–338, 1954

Hyppönen, M., Pellinen, R.J., Sucksdorff, C., Torniainen, R.: Digital all-sky camera. Technical report No. 9, Finnish Metereol. Inst. 1974

Iijima, T., Potemra, T.A.: The amplitude distribution of field-aligned currents at northern high latitudes observed by Triad. J. Geophys. Res. **81**, 2165–2174, 1976

Kamide, Y., Akasofu, S.-I.: The auroral electrojet and global auroral features. J. Geophys. Res. **80**, 3585–3602, 1975

Kamide, Y., Rostoker, G.: The spatial relationship of field-aligned currents and auroral electrojets to the distribution of nightside auroras. J. Geophys. Res. **76**, 5589–5608, 1977

Kamide, Y., Akasofu, S.-I., Deforest, S.E., Kisabeth, J.L.: Weak and intense substorms. Planet. Space Sci. **23**, 579–587, 1975

Kisabeth, J.L.: The dynamical development of the polar electrojets. Edmonton: PhD Thesis, University of Alberta 1972

Kisabeth, J.L., Rostoker, G.: Development of the polar electrojet during polar magnetic substorms. J. Geophys. Res. **76**, 6815–6828, 1971

Kisabeth, J.L., Rostoker, G.: Current flow in auroral loops and surges inferred from ground-based magnetic observations. J. Geophys. Res. **78**, 5573–5584, 1973

Kisabeth, J.L., Rostoker, G.: The expansive phase of magnetospheric substorms. I. Development of the auroral electrojets and auroral arc configuration during a substorm. J. Geophys. Res. **79**, 972–984, 1974

Kisabeth, J.L., Rostoker, G.: Modelling of the three-dimensional current systems associated with magnetospheric substorms. Geophys. J.R. Astrom. Soc. **49**, 655–683, 1977

Klumpar, D.M.: Relationship between auroral particle distributions and magnetic field perturbations associated with field-aligned currents (accepted for publication). J. Geophys. Res. 1979

Klumpar, D.M., Burrows, J.R., Wilson, M.D.: Simultaneous observations of field-aligned currents and particle fluxes in the post-midnight sector. Geophys. Res. Lett. **3**, 395–398, 1976

Küppers, F., Untiedt, J., Baumjohann, W., Lange, K., Jones, A.G.: A two-dimensional magnetometer array for ground-based observations of auroral zone electric currents during the International Magnetospheric Study (IMS). J. Geophys. **46**, 429–450, 1979

Lui, A.T.Y., Venkatesan, D., Anger, C.D., Akasofu, S.-I., Heikkila, W.J., Winningham, J.D., Burrows, J.R.: Simultaneous Observations of particle precipitations and auroral emissions by the ISIS 2 satellite in the 19–24 MLT sector. J. Geophys. Res. **82**, 2210–2226, 1977

Maurer, H., Theile, B.: Parameters of the auroral electrojet from magnetic variations along a meridian. J. Geophys. **44**, 415–426, 1978

McPherron, R.L.: Growth phase of magnetospheric substorms. J. Geophys. Res. **75**, 5592–5599, 1970

McPherron, R.L.: Magnetospheric substorms. Rev. Geophys. Space Phys. **17**, 657–681, 1979

McPherron, R.L., Russel, C.T., Aubry, M.P.: Satellite studies of magnetospheric substorms on August 15, 1968. 9. Phenomenological model for substorms. J. Geophys. Res. **79**, 3131–3149, 1973

Mozer, F.S.: Origin and effects of electric fields during isolated magnetospheric substorms. J. Geophys. Res. **76**, 7595–7608, 1971

Oguti, T., Fukunishi, H., Tohmatsu, T., Nagata, T : H_β-emissions during auroral breakup. Memoirs of National Inst. of Polar Research, Special Issue No. 3, Proceedings of antarctic review meeting, 10–20, 1974

Opgenoorth, H.J.: Vergleich von Strukturen im Polarlicht und im Magnetfeld gleichzeitiger Ionosphärenströme. Diplomarbeit Physik, Inst. f. Geophys. Univ. Münster, FRG, 1978

Pellinen, R.J.: Induction model and observations of onset of magnetospheric substorms. PhD Thesis, University of Helsinki, Finland 1979

Pellinen, R.J., Heikkila, W.J.· Energization of charged particles to high energies by an induced substorm electric field within the magnetotail. J. Geophys. Res. **83**, 1544–1550, 1978a

Pellinen, R.J., Heikkila, W.J.: Observations of auroral fading before breakup. J. Geophys. Res. **83**, 4207–4217, 1978b

Pike, C.P., Whalen, J.A.: Satellite observations of auroral substorms. J. Geophys. Res. **79**, 985–1000, 1974

Potemra, T.: Current systems in the earth's magnetosphere. Rev. Geophys. Space Phys. **17**, 640–656, 1979

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

Ranta, H.: The onset of an auroral absorption substorm. J. Geophys. Res. **83**, 3893–3899, 1978

Rostoker, G.: Macrostructure of geomagnetic bays. J. Geophys. Res. **73**, 4217–4229, 1968

Rostoker, G., Boström, R.: A mechanism for driving the gross Birkeland current configuration in the auroral oval. J. Geophys. Res. **81**, 235–244, 1976

Saito, T.: Geomagnetic pulsations. Space Sci. Rev. **10**, 319–412, 1969

Sharber, R.J., Heikkila, W.J.: Fermi acceleration of auroral particles, J. Geophys. Res. **77**, 3397–3410, 1972

Stauning, P.: Compilation of Ionlab riometer data for IMS-workshop in Bad Lauterberg, FRG in October 1978. Ionlab report R51, Danish Metereol. Institute, Techn. Univ., Lungby, Denmark, 1978

Swift, D.W.: Auroral mechanisms and morphology. Rev. Geophys. Space Phys. **17**, 681–696, 1979

Untiedt, J., Pellinen, R.J., Küppers, F., Opgenoorth, H.J., Pelster, W.D., Baumjohann, W., Ranta, H., Kangas, J., Czechowsky, P., Heikkila, W.J.: Observations of the initial development of an auroral and magnetic substorm at magnetic midnight. J. Geophys. **45**, 41–65, 1978

Vallance Jones, A.: Aurora. Dordrecht, Boston: D. Reidel Publ. Comp. 1976

Wedeken, U., Hillebrand, O., Krenzien, E., Ranta, A., Ranta H., Voelker, H.: Cosmic noise absorption events and geomagnetic pulsations activity during substorms. J. Geophys. **46**, 249–259, 1979

Wescott, E.M., Stenbaek-Nielsen, H.C., Davis, T.N., Murcray, W.B., Peek, H.M., Bottoms, P.J.: The L=6.6 Oosik Barium plasma injection experiment and magnetic storm on March 7, 1972. J. Geophys. Res. **80**, 951–966, 1975

Whalen, J.A.: Auroral oval plotter and nomograph for determining corrected geomagnetic local time, latitude, and longitude for high latitudes in the northern hemisphere. Environmental Res. Papers No. 327, Air Force Cambridge Res. Lab., Bedford, Mass., 1970

Received December 12, 1979; Revised Version March 17, 1980; Accepted March 17, 1980

J. Geophys. 48, 116–120, 1980

Journal of Geophysics

Simultaneous Maxima of Geomagnetic Pulsations and Riometer Absorption on the Afternoon Side of the Auroral Zone

Yu.P. Novikov[1], Yu.A. Kopytenko[2], O.M. Raspopov[3], T. Bösinger[4], J. Kangas[4], and H. Ranta[5]

[1] Department of Physics, State University, Leningrad, USSR
[2] Leningrad Branch Izmiran, Leningrad, USSR
[3] Polar Geophysical Institute, Kola Branch of Academy of Sciences, Apatity, USSR
[4] University of Oulu, Department of Physics, Oulu, Finland
[5] Geophysical Observatory, Sodankylä, Finland

Abstract. Bursts of irregular long-period magnetic pulsations have been observed in the afternoon sector of the auroral zone when a substorm was developing in the midnight sector. The maximum amplitude of these pulsations is coincident both in space and time with the maximum of riometer absorption. Short-period magnetic pulsations, usually an IPDP event, appear also at the same location at the same time. These observations are discussed for modelling the afternoon Pi2-like magnetic pulsations. Similarities with the midnight Pi2's are noted.

Key words: Pi2 magnetic pulsations – Substorm – Plasmapause.

Introduction

Irregular long-period magnetic pulsations called Pi2 pulsations occur typically at the onset of a substorm. Pi2's are believed to be an important manifestation of substorm processes (see, as general reviews, Saito 1969; Orr 1973). It has been confirmed by Raspopov et al. (1972), Stuart (1974), Fukunishi (1975), and Saito et al. (1976) that on the night side of the magnetosphere the amplitude of Pi2 waves has the main maximum (high latitude maximum) at the auroral oval and a secondary maximum (midlatitude maximum) inside the plasmapause. Phase reversals of about 180° are reported to occur at the high latitude maximum and at the minimum (subauroral minimum) between the two maxima (Björnsson et al. 1971; Saito et al. 1976).

Kopytenko et al. (1977) made an analysis of magnetic pulsations during the substorm in the local midnight sector using a dense network of magnetometers spaced by about 1° from $\phi = 53°$ to $\phi = 65°$ They reported an interesting result: an intermediate amplitude maximum of long-period magnetic pulsations appears between the high- and midlatitude maxima. It was concluded that this new maximum occurs at the plasmapause projection onto the ground. This was concluded earlier also for the subauroral minimum (Saito 1976), which, however, according Kopytenko et al. (1977) is located still 2–3° south of the plasmapause projection.

It is further known that Pi2 waves at the auroral oval are closely associated with the auroral electrojet (Olson and Rostoker 1975) and electron precipitation into the high atmosphere (Stuart et al. 1977). Also high-frequency pulsations are superimposed on Pi2 waves at high latitudes (Saito 1969).

In this paper bursts of large-amplitude irregular magnetic pulsations in the Pi2 period range on the afternoon side of the magnetosphere are studied. These bursts occur during the substorm in the longitude sector of the positive bay in the auroral zone. Such pulsations seem to be common according to Jacobs and Sinno (1960). A special interest here has been to give a detailed picture of the association of these long-period pulsations with the substorm development and with the electron precipitation events and high-frequency magnetic pulsations in the same local time sector using a dense network of riometers and magnetometers across the auroral zone and plasmapause region.

Observations

Long-period magnetic pulsations ($T \sim 10$–200 s) analysed here have been recorded by magnetostatic magnetometers of the same stations as those used by Kopytenko et al. (1977). The time resolution is about 4 s. The geomagnetic coordinates of the stations are given in Table 1 (Evans et al. 1969). A north-south chain of four 27.6 MHz riometers was operated by the Geophysical Observatory of Sodankylä. The time resolution is of the order of 2 min. A north-south chain of 3 induction coil magnetometers was operated by the University of Oulu. These detect short-period magnetic pulsations ($T \sim 1.0$–10 s) recorded on analog tapes and displayed by sonagrams with a time resolution of about 2 min. The geomagnetic coordinates of the Finnish stations are given in Table 2 (corrected geomagnetic coordinates calculated after Gustafsson 1970).

In this report, data from March 2 and 3, 1974 will be presented. The general magnetic activity during the events was characterized by the Kp value 4.

Table 1. Location of Russian magnetometer stations

Station	L	Geomagn. coord. (corrected)	
		Latitude	Longitude
Lovozero	5.1	63.8°	118°
Kniazhaya Guba	4.7	62.4°	116°
Angozero	4.5	61.4°	116°
Kem'	4.2	60.7°	116°
Muezerka	4.0	60.0°	113°
Suckozero	3.75	59.0°	111°
Suisaari	3.6	58.0°	114°

0340-062X/80/0048/0116/$01.00

Table 2. Locations of Finnish riometers and pulsation magnetometers

Station	L	Geomagn. coord. (corrected)	
		Latitude	Longitude
Kevo[a]	6.0	66.2°	111°
Ivalo[a]	5.5	65.0°	110°
Sodankylä	5.1	63.9°	109°
Oulu	4.3	61.8°	107°
Nurmijärvi	3.3	57.0°	103°

[a] Only riometer

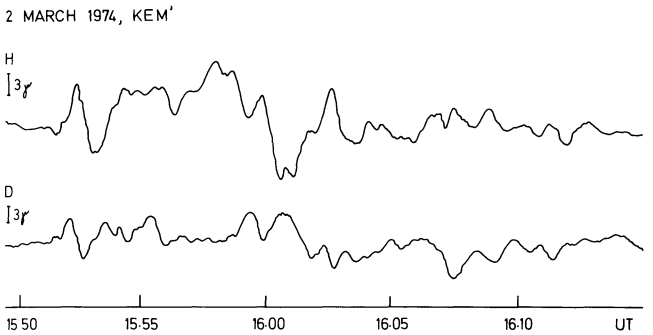


Fig. 1. Long-period magnetic pulsations registered in Kem' at T_2 = 15.52–16.10 UT on March 2, 1974

Events on March 2, 1974

Two trains of intense long-period magnetic pulsations appear at T_1 = 15.22–15.38 UT (~18.30 MLT) and at T_2 = 15.52–16.10 UT (~19.00 MLT). An example of the pulsations at T_2 is given in Fig. 1. Note that for the analysis we have always taken that part of the pulsation train that can be identified at each station as belonging to the same burst (a train usually consists of several bursts). The maximum amplitude of waves – derived from the H-component – occurs at $L \sim 4.7$ for the first event (T_1) and at or beyond $L \sim 4.2$ for the second event (T_2), as shown in the amplitude/time plots in the top half of Fig. 2a. It should be noted that data at $L > 4.2$ for T_2 are not available, and that 'amplitude' here is defined as the variation range between maximum and minimum in the records.

In the lower half of Fig. 2a the latitudinal variation of riometer absorption is presented. At the end of T_1 riometer absorption starts to increase at Sodankylä ($L \sim 5.1$), reaching its maximum shortly after T_1. Later at T_2 a strong absorption event has been recorded at Oulu ($L \sim 4.3$).

Dynamic spectra of high-frequency pulsations from two stations located at about the same meridian as the Russian magnetometers are illustrated in Fig. 2b. At the auroral zone station (Sodankylä) unstructured Pc1–2 waves (see Heacock 1971a) appear at T_1 which are most intense at the time of the strong absorption event at Sodankylä. Later, at T_2 typical Pi1 activity appears in the Sodankylä recordings. At a more southern station (Oulu) Pc1–2 pulsations have been observed at T_1 but at T_2 an IPDP pulsation event with a steep increase of the frequency is the prominent characteristic of short-period magnetic pulsations.

Figure 2c, d show the development of magnetic disturbances in the auroral zone and at low latitude, respectively. It is seen in Fig. 2c that negative magnetic bays start at Chelyuskin at 15.20 UT (~22.25 MLT) and at Dikson at 15.50 UT (~21.35 MLT). Both of these onsets at 15.20 and 15.50 UT can be related with the increase of the magnetic field at the stations in the afternoon sector, namely at Kiruna (~17.55, 18.25 MLT) and at Loparskaya (~18.20, ~18.50 MLT).

At low latitude stations at Guam and Lunping at T_1 and at Hyderabad and Alibag at T_2 an increase in the H-component has been measured as shown in Fig. 2d. This has occurred in the longitude sector of the negative bay in the auroral zone (cf. Fig. 2c). The local magnetic midnight of the various stations is denoted by a dot in Fig. 2c, d.

Events on March 3, 1974

A similar set of data as for the events of March 2 is shown for an event on March 3, 1974 in Fig. 3a–d. A burst of intense irregular long-period magnetic pulsations was observed at 13.47–14.05 UT (~16.50 MLT) with maximum at $L \sim 4.2$ as shown in the upper half of Fig. 3a. At the same time an intense absorption event appears in riometer recordings at the Oulu station (~16.20 MLT) at $L \sim 4.3$ as shown in the lower half of Fig. 3a. According to Fig. 3b strong IPDP pulsations have also been observed at Oulu.

A magnetic substorm starts at Dikson at 13.30 UT (~19.10 MLT) which is associated with a positive bay in the afternoon sector at Kiruna (~15.20 MLT) and Loparskaya as illustrated in Fig. 3c. At the same time an increase of magnetic field intensity is seen at low latitudes (Fig. 3c) at the meridian of the high latitude negative bay as shown in Fig. 3d.

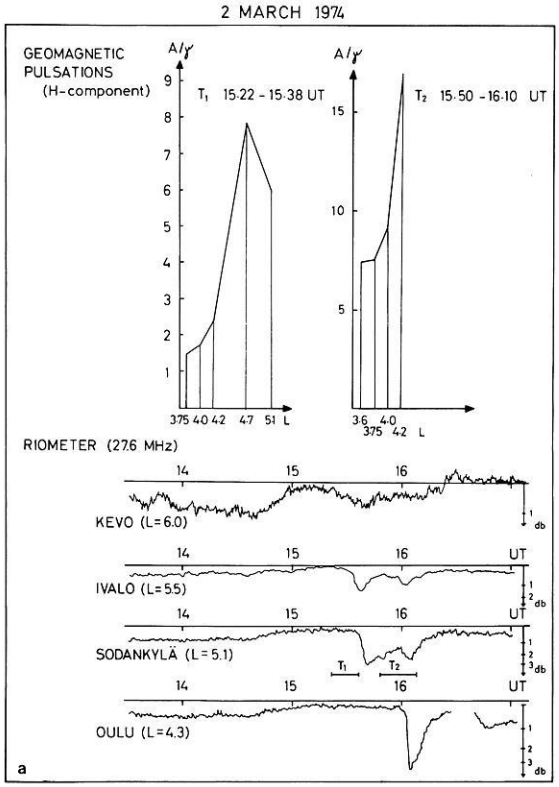
Summary and Discussion

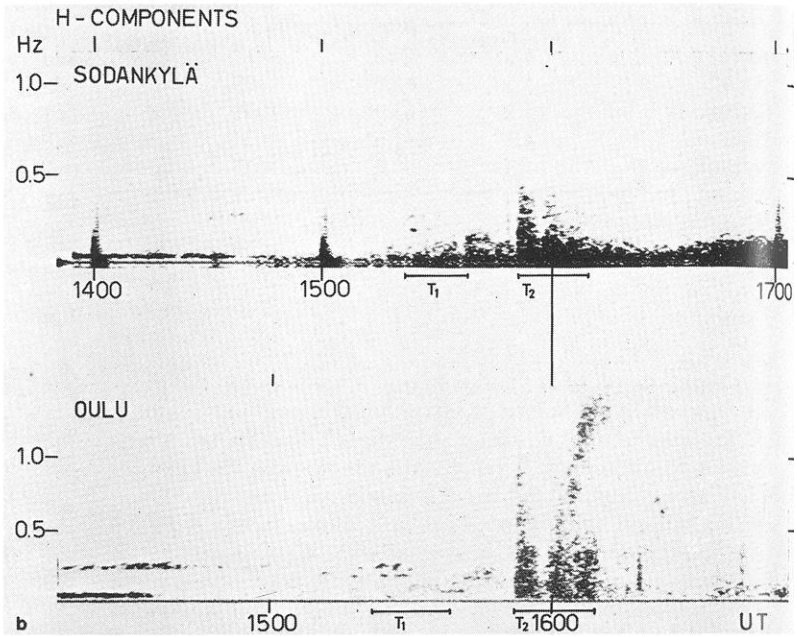
Meridional profiles of riometer absorption and occurrence of short-period magnetic pulsations have been analysed in the afternoon sector at the time when a burst of irregular longperiod magnetic pulsations has been observed during the positive bay at about the same meridian. Our observations can be summarized as follows:

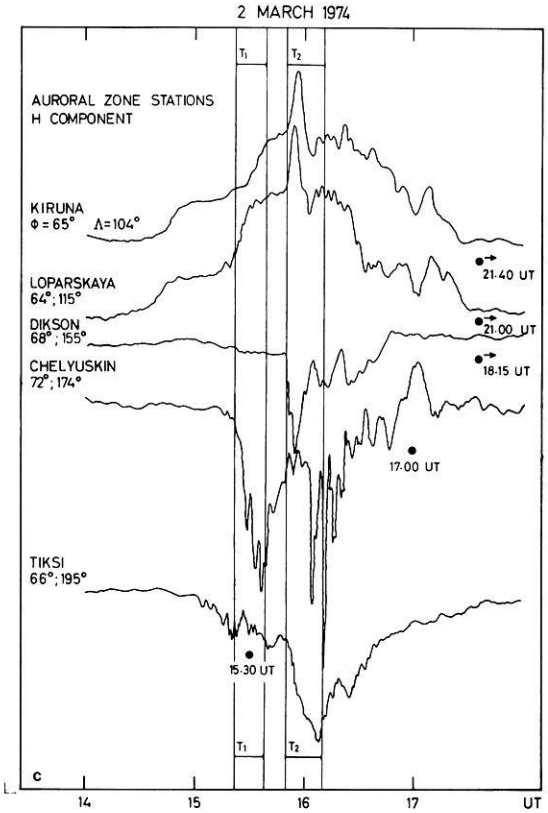
1. The bursts of irregular long-period magnetic pulsations are associated with the substorm developing to the east from the meridian of pulsation magnetometers.
2. The maximum amplitude of pulsations occurs at about the same latitude and at about the same time as the riometer absorption maximum.
3. Short-period magnetic pulsations typed unstructured Pc1–2 or IPDP are also observed at the same latitude.

Note that the longitudinal difference of the Russian and Finnish meridians can be neglected as regards Pi2, IPDP and riometer absorption events, since each of them is reported to occur simultaneously over a large longitudinal extent (Sutcliffe 1979; Heacock 1971a, b; Berkey et al. 1974).

The comparison of the auroral zone and low latitude magnetograms with the observed Pi2-like pulsations shows that our afternoon Pi2's monitor the substorm development in the same way as Pi2's recorded in the midnight sector. We therefore propose that the afternoon Pi2's are due to night-time westward propagating Pi2 waves which succeed in penetrating into the afternoon side of the magnetosphere. This is consistent with Sutcliffe's (1979)







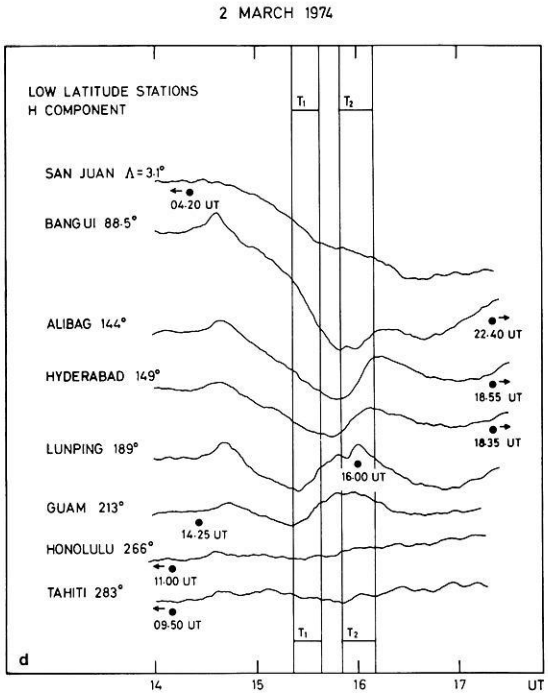
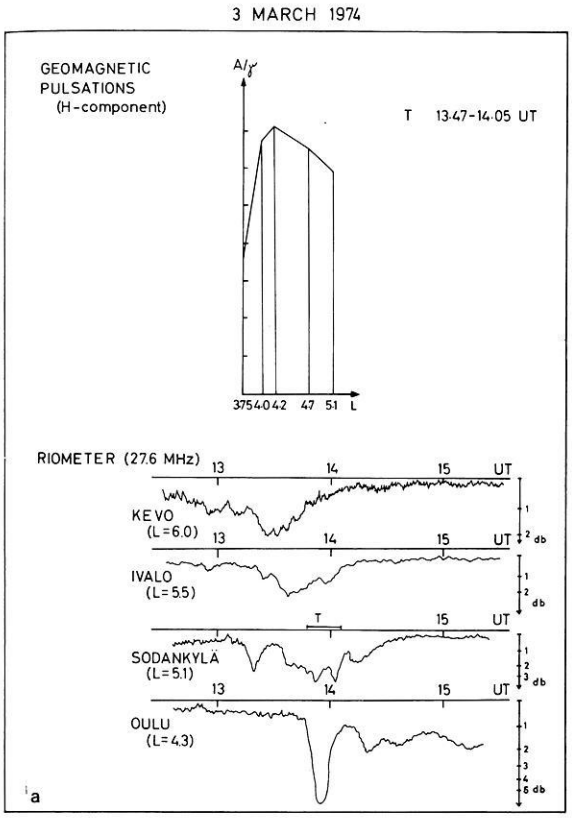
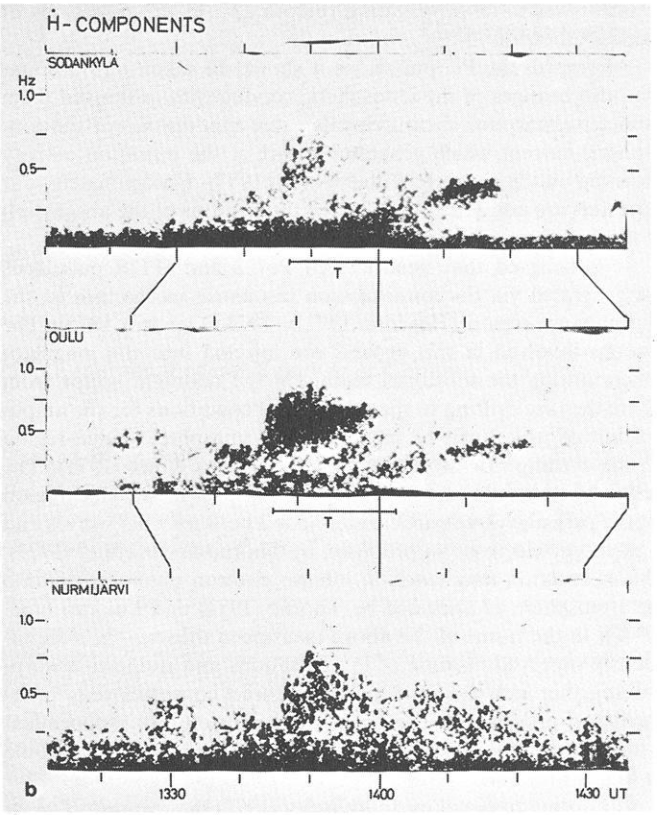
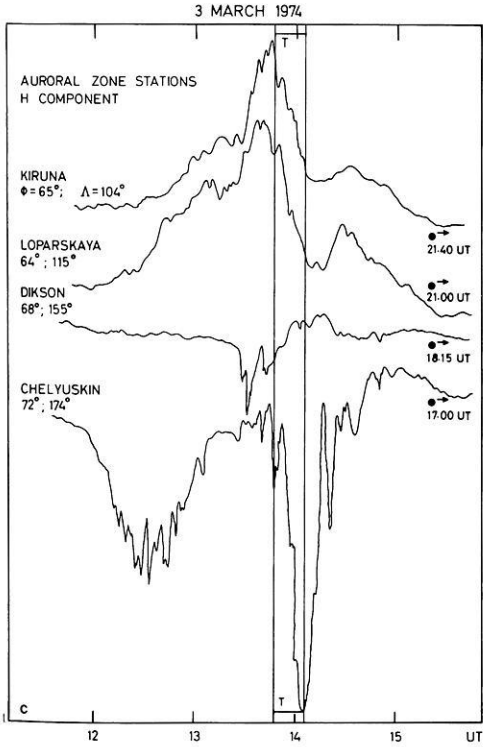


Fig. 2. **a** Latitude profiles of the maximum amplitude of long-period magnetic pulsations at $T_1 = 15.22–15.38$ UT and $T_2 = 15.52–16.10$ UT and riometer absorption in the local afternoon sector of the auroral zone on March 2, 1974. **b** Dynamic spectra of short-period magnetic pulsations recorded at Sodankylä and Oulu on March 2, 1974. **c** Magnetic recordings at some auroral zone stations on March 2, 1974. Local magnetic midnight is denoted by a *dot*. **d** Magnetic recordings at some low-latitude stations on March 2, 1974. Local magnetic midnight is denoted by a *dot*







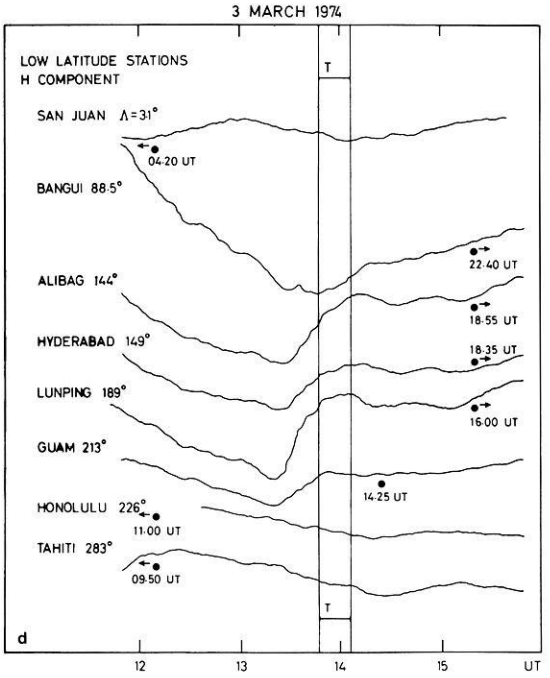


Fig. 3. a Latitude profiles of the maximum amplitude of long-period magnetic pulsations at $T = 13.47–14.02$ UT and riometer absorption in the local afternoon sector of the auroral zone on March 3, 1974. **b** Dynamic spectra of short-period magnetic pulsations at Sodankylä, Oulu and Normijärvi on March 3, 1974. **c** Same as in Fig. 2c but for March 3, 1974. **d** Same as in Fig. 2d but for March 3, 1974

result of westward propagating (before 23 LT) Pi2 pulsations at auroral latitudes.

As regards the Pi2 pulsations it should be taken into account that also changes of the ionospheric conductivity as derived from riometer absorption measurements cause modulation of the ionospheric current which generates a part of the pulsation activity observed on the ground (Wilhelm et al. 1977). Unfortunately, our riometers are not able to resolve the fluctuations of the absorption in the frequency range of the observed magnetic pulsations.

It is believed that unstructured Pc1–2 and IPDP pulsations are generated via the ion-cyclotron resonance mechanism in the plasmapause region (Heacock 1971b, 1973; Gendrin 1975). The protons involved in this process are injected into the magnetosphere during the substorm, mainly in the midnight sector from where they are drifting to the west. Good conditions for the amplification of waves can be met in the plasmaspheric bulge region on the evening side as shown by Perraut and Roux (1975). The same region in the magnetosphere is also a favourable region for the parasitic wave-particle interaction between ULF-waves and high-energy electrons as proposed by Thorne and Kennel (1971). This interaction may result in intense electron precipitation into the atmosphere as discussed by Thorne (1974) and Lukkari et al. (1977). In the frame of the above mentioned theories the simultaneously observed maxima of Pi2 pulsations and riometer absorption together with the IPDP pulsations are a consequence of interconnected processes in the plasmapause region. An independent measurement of the location of the plasmapause, therefore, would be highly desirable.

According to Novikov and Kiselev (1977) the increase of magnetic field intensity in the equatorial plane of the magnetosphere during substorms leads (among other agents like inward motion of a generation region and energy dispersion of westward drifting protons) to an increase of the wave frequency of IPDP events. Thus our display of low latitude magnetograms does not only allow to localize in time and space the substorms but also shows a possible connection between IPDP and equatorial magnetic field intensification.

Summarizing we can say that all three events presented here show a close connection between the pulsation amplitude and the riometer absorption maximum at the statistical position of the plasmapause projection onto the ground (Chappell 1972). Simultaneously observed IPDP events further support the idea of an interconnection of all three phenomena at the plasmapause where most favourable conditions for the development of ion and electron cyclotron instabilities exist.

References

Berkey, F.T., Driatskiy, V.M., Henriksen, K., Hultqvist, B., Jelly, D.H., Schuka, T.I., Theander, A., Yliniemi, J.: A synoptic investigation of particle precipitation dynamics for 60 substorms in IQSY (1964–1965) and IASY (1969). Planet. Space Sci. **22**, 255–307, 1974

Björnsson, A., Hillebrand, O., Voelker, H.: First observational results of geomagnetic pi2 and pc5 pulsations on a north-south profile through Europe. Geophys. **37**, 1031–1042, 1971

Chappell, C.R.: Recent satellite measurements of the morphology and dynamics of the plasmapause. Rev. Geophys. Space Phys. **10**, 951–979, 1972

Evans, J.E., Newkirk, L.L., McCormac, B.M.: DASA Lockheed Palo Alto Res. Laboratory, 2347, 1969

Fukunishi, H.: Polarization changes of geomagnetic Pi2 pulsations associated with the plasmapause. J. Geophys. Res. **80**, 98–110, 1975

Gendrin, R.: Waves and wave-particle interactions in the magnetosphere: A review. Space Sci. Rev. **18**, 145–200, 1975

Gustafsson, G.: A revised corrected geomagnetic coordinate system. Ark. Geofys. **5**, 595–617, 1970

Heacock, R.R.: Types Pc1–2 and IPDP micropulsations at the auroral zone. Report of Geophysical Institute University of Alaska, AFCRL-71-0097, 1971a

Heacock, R.R.: Spatical and temporal relations between Pi bursts and IPDP micropulsation events. J. Geophys. Res. **76**, 4494–4504, 1971b

Heacock, R.R.: Type IPDP magnetospheric plasma wave events. Nature **246**, 93–96, 1973

Jacobs, J.A., Sinno, K.: World wide characteristics of geomagnetic micropulsations. Geophys. J. **3**, 333–352, 1960

Kopytenko, Yu.A., Raspopov, O.M., Koshelevsky, U.K., Novikov, Yu.P.: Plasmasphere structure obtained from geomagnetic pulsation data. J. Atmos. Terr. Phys. **39**, 165–172, 1977

Lukkari, L., Kangas, J., Ranta, H.: Correlated electron precipitation and magnetic IPDP events near the plasmapause. J. Geophys. Res. **82**, 4750–4756, 1977

Novikov, Yu.P., Kiselev, B.V.: On the relation between slow variations of the earth magnetic field at equatorial stations and the spectra of IPDP pulsations (in Russian). Report of the Department of the Physics of the Earth of the State University of Leningrad, USSR 1977

Olson, J.V., Rostoker, G.: Pi2 pulsations and the auroral electrojet. Planet. Space Sci. **23**, 1129–1139, 1975

Orr, D.: Magnetic pulsations within the magnetosphere: A review. J. Atmos. Terr. Phys. **35**, 1–50, 1973

Perraut, S., Roux, A.: Respective role of the cold and warm plasma densities on the generation mechanism of ULF-waves in the magnetosphere. J. Atmos. Terr. Phys. **37**, 407–418, 1975

Raspopov, O., Kochelevsky, V., Troitskaya, V., Baransky, L., Belenkaya, B., Roquet, I., Fambiyakoye, O.: Characteristique des spectres de pulsations du type Pi2 suivant un profil meridieu. C.R. Acad. Sci. (Paris) Sect. B.: **275**, 465–468, 1972

Saito, T.: Geomagnetic pulsations. Space Sci. Rev. **10**, 319–412, 1969

Saito, T., Sakurai, T., Koyama, Y.: Mechanism of association between Pi2 pulsation and magnetospheric substorm. J. Atmos. Terr. Phys. **38**, 1265–1277, 1976

Stuart, W.F.: A mechanism of selective enhancement of Pi2's by the plasmasphere. J. Atmos. Terr. Phys. **36**, 851–859, 1974

Stuart, W.F., Green, C.A., Harris, T.J.: Correlation between modulation of the intensity of precipitating electrons in the auroral zone and a coincident Pi2. J. Atmos. Terr. Phys. **39**, 631–635, 1977

Sutcliffe, P.R.: Some aspects of Pi2 morphology. Proceedings of the international Workshop on selected topics of magnetospheric physics, Tokyo, March 1979. pp. 94–98, 1979

Thorne, R.M., Kennel, C.F.: Relativistic electron precipitation during magnetic storm main phase. J. Geophys. Res. **76**, 4446–4453, 1971

Thorne, R.M.: A possible cause of dayside relativistic electron precipitation events. J. Atmos. Terr. Phys. **36**, 635–645, 1974

Wilhelm, K., Münch, I.W., Kremser, G.: Fluctuations of the auroral zone current system and geomagnetic pulsations. J. Geophys. Res. **82**, 2705–2716, 1977

Received June 13, 1979; Revised Version January 31, 1980

J. Geophys. 48, 121–123, 1980

Journal of Geophysics

Numerical Investigation of the Spectral Resolving Power of Burg's Maximum Entropy Method*

F. Wenzel and W. Zürn
Geophysikalisches Institut, Universität Karlsruhe, Hertzstr. 16, D-7500 Karlsruhe, Federal Republic of Germany

Abstract. The maximum entropy method (MEM) proposed by Burg (1967) for power spectral analysis has often been cited for superior resolving power (Ulrych and Bishop 1975). The resolving power of classical techniques is limited to $\Delta f \geqq T^{-1}$, where T is the length of the time series or the maximum lag, for the periodogram and the power spectrum respectively. The limits of MEM have, to our knowledge, never been explored. Using signals consisting of two and three decaying sinusoids, with random noise superimposed, we tried to find criteria for the limits of resolution with MEM. The best resolution (in the sense that the real frequency separation in the time series is verified by the analysis with relatively small errors) we found was $\Delta f_{SYN} \cdot T \approx 0.2$; but this is not a well defined limit.

Key words: Maximum entropy method – Spectral resolution.

Introduction

Numerous papers (Ulrych and Bishop 1975 for references) have dealt with the theory and the assumptions of the maximum entropy method (MEM). Numerical experiments have shown, and applications to geophysical time series have indicated, increased spectral resolution of MEM as compared with the classical methods, the periodogram and power spectrum. However, Chen and Stegen (1974) pointed out that MEM has disadvantages also, the most severe being that the spectra may exhibit peaks with no physical meaning (artificial splitting). Another problem they found is the shifting of the locations of spectral peaks (LOSP), an effect which can be rather strong for short signal durations T, and depends on the initial phases φ_i [see Eq. (1)] of the signals. Amplitudes and line shapes are not as readily interpreted as in classical methods; this was recognized in the first publications on MEM (e.g. Lacoss 1971).

Most of the authors who investigated the properties of MEM using synthetic signals confined their interest to the range where the resolution needed is slightly better than classical methods can achieve, i.e., $\Delta f_{SYN} \cdot T \geqq 0.27$ (Table 1). In all these cases MEM was clearly able to resolve the lines. Here we explored the range $\Delta f_{SYN} \cdot T$ between $5 \cdot 10^{-7}$ and 10 in the input series, in order to find the limitations of MEM and possibly some criteria for future applications for spectral resolution. In order to suppress the shifting of the LOSP we used only signals with duration T longer than about 20 times the signal period.

* Contribution No. 225, Geophysikalisches Institut, Universität Karlsruhe

Table 1. Frequency separations in synthetic time series used in the literature to show the resolving power of MEM. Δf is the frequency separation, T the length of the series, $\Delta f \cdot T$ indicates the resolving power: if $\Delta f \cdot T$ is less than 1 the resolution is better than $1/T$

Author	Δf	$1/T$	$\Delta f \cdot T$
Lacoss (1971)	0.15	0.1	1.50
Ulrych et al. (1973)	0.04	0.05	0.80
Ulrych and Bishop (1975)	0.03	0.025	1.20
Radoski et al. (1975)	1.0	0.78	1.28
Baggeroer (1977)	0.5	0.42	1.19
Newman (1977)	0.04	0.067	0.60
Bowling (1977)	2.0	4.0	0.50
Frost (1977)	0.016	0.030	0.52
	0.012	0.030	0.41
	0.010	0.030	0.33
	0.0060	0.015	0.39
	0.0042	0.015	0.27

Numerical Experiments

Synthetic Signals

Synthetic signals x_k of the following type

$$x_k = \sum_{i=1}^{L} A_i \cdot \exp(-\delta_i \cdot k \cdot \Delta T) \cdot \cos(2\pi f_i \cdot k \cdot \Delta T - \varphi_i) + n_k$$
$$k = 1, 2, \ldots, M \qquad (1)$$

were analyzed using Burg's MEM algorithm. The frequency separations in the time series are:

$$\Delta f_{SYN} = |f_i - f_{i+1}|$$

$L = 2, 3$ is the number of decaying harmonic signals in the series; A_i, δ_i, f_i, φ_i are amplitude, attenuation factor, frequency and initial phase of the i-th harmonic; ΔT is the sampling interval and $M \cdot \Delta T$ is the length of the time series. The noise n_k consists of normally distributed random numbers, the signal-to-noise ratio (SNR) is specified in dB and defined as the ratio of the power in the sinusoids to the power in the random numbers.

These signals could simulate a record of free vibrations of the earth after narrow band-pass filtering. The attenuation fac-

0340-062X/80/0048/0121/$01.00

Table 2. List of parameters for the time series used in this paper. For definitions see Eq. (1)

Series	M	ΔT	T	SNR dB	i	A_i	δ_i	f_i	$\varphi_i(°)$	Remarks
B	500	0.001	0.5	5	1	1	5	100	0	Δf varied from 10 to 5×10^{-7}, the corresponding filter length, where resolution is achieved, from 60 to 400
					2	1	5	$100+\Delta f$	180	
F	100	0.01	1.0	5	1	1	1	20	0	Δf varied from 10 to 0.001, the filter length from 20 to 80. In order to see the influence of the SNR and the phase shift, we also operated with SNR = 0 and 30 dB and $\varphi_2=45°$. The results remained unchanged
					2	1	1	$20+\Delta f$	90	
H	100	0.01	1.0	30	1	1	0	20	0	Two undamped sinusoids. Δf varied from 10 to 0.01, the filter length from 20 to 95
					2	1	0	$20+\Delta f$	45	
K	100	0.01	1.0	5	1	1	0	20	0	Three undamped sinusoids. Δf varied from 10 to 0.01, the filter length from 20 to 90
					2	1	0	$20+\Delta f$	45	
					3	1	0	$20+2\Delta f$	90	

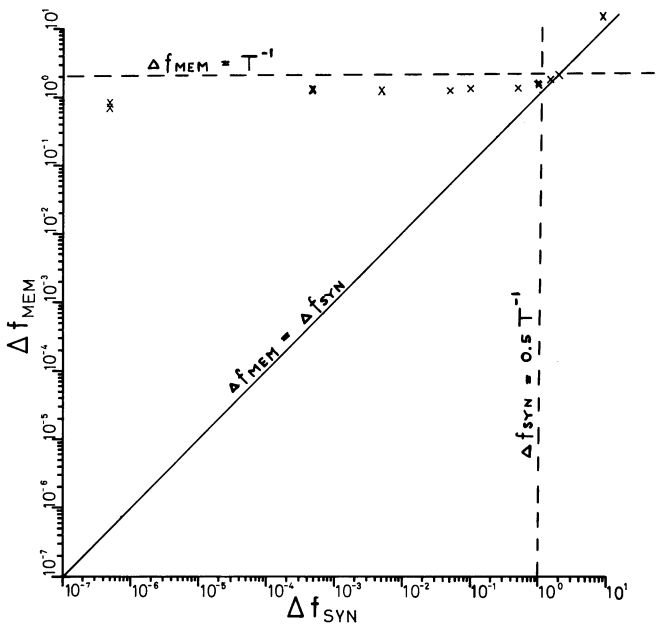


Fig. 1. Plot of Δf_{MEM} versus Δf_{SYN} for resolved peaks in the MEM spectra of time series *B*. Δf are measured in units of reciprocal time in the time series (sampling interval $\Delta T=1$ unit)

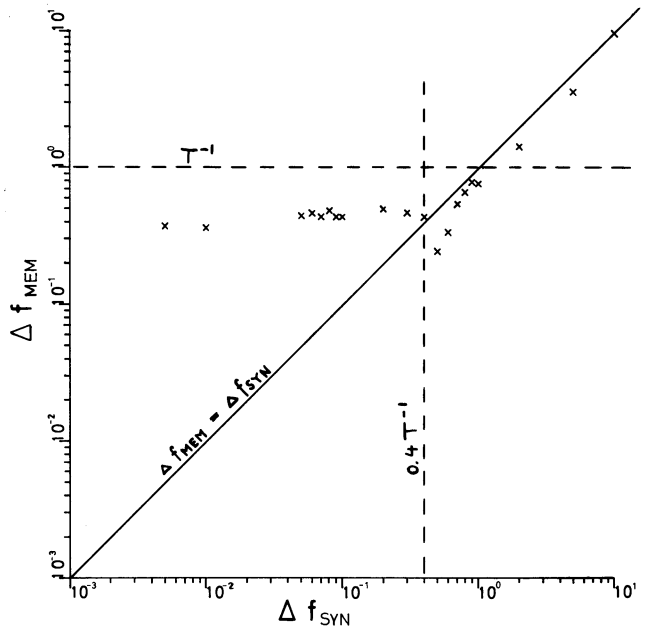


Fig. 2. Plot of Δf_{MEM} versus Δf_{SYN} for resolved peaks in the MEM spectra of time series *F*

tors were in fact chosen to be of the order of magnitude observed for free oscillations of the earth.

The frequency differences $\Delta f_{\mathrm{SYN}}=|f_i-f_{i+1}|$ were varied over a wide range. The signal parameters and the prediction error filter lengths (NPEF) used are listed in Table 2.

Results

MEM spectra were computed for all these time series and in each case the NPEF was increased until L peaks appeared in the spectrum. The resulting frequency separations Δf_{MEM} were plotted in Figs. 1–3 versus the frequency separation in the input time series Δf_{SYN}. When less or more than L distinct peaks were found, the results were discarded. Ideally the data points should follow a straight line with a slope of 45 degrees, a result which cannot be expected for very small Δf_{SYN}. For classical techniques no data points would occur for Δf_{SYN} smaller than about T^{-1}. It was surprising to find separated peaks with MEM for Δf_{SYN} as low as $5\cdot10^{-7}\cdot T^{-1}$ for $L=2$. However, the frequency separation obviously cannot be considered reliable below a certain threshold $f_{\mathrm{SYN}}\cdot T=c$. For the three time series *B*, *F*, and *H* we found c to be between 0.4 and 0.5. This threshold does not appear to depend on NPEF, SNR or φ_i.

For the *K* series, where $L=3$, however, the three constituents were never resolved when $\Delta f_{\mathrm{SYN}}\cdot T$ was less than 1. In this case therefore, MEM is only moderately superior to a periodogram as far as resolution is concerned.

It should be mentioned here, that when higher resolution is required than $\Delta f\cdot T\approx1$, for more complicated signals, and when

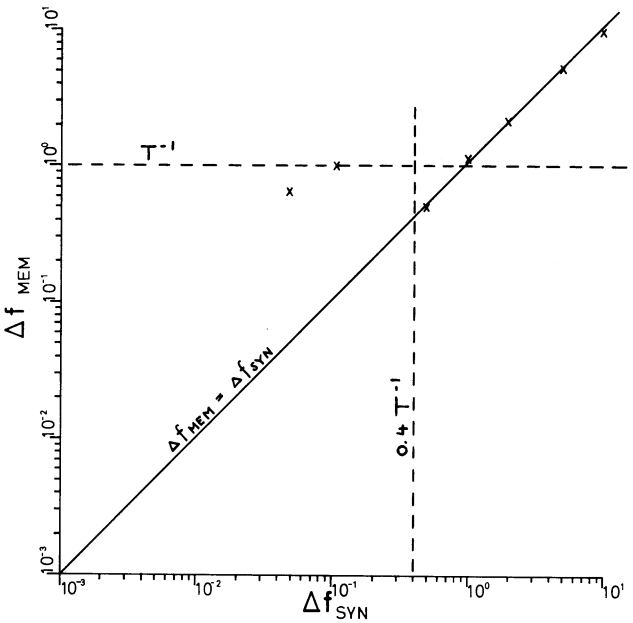


Fig. 3. Plot of Δf_{MEM} versus Δf_{SYN} for resolved peaks in the MEM spectra of time series *H*

NPEF is large, artificial splitting may occur and may complicate, to say the least, the interpretation of real data (e.g. Wenzel 1978).

Discussion

Our experiments demonstrate that for two decaying sinusoids MEM always shows two peaks, but does not give the correct frequency separation if $\Delta f \cdot T$ is less than 0.4. For two undamped sinusoids we found a similar threshold, while for three undamped sinusoids no resolution occurred below $\Delta f \cdot T \approx 1$.

Because of the complexity and nonlinearity of MEM we cannot assess analytically the threshold found, even in the simplest case. Table 1 and our results show that for the simple case of two sinusoids, decaying or not, the resolution of MEM is markedly better than a periodogram's (by a factor of 5 to 10). This property deteriorates as signals become more complicated, as in the free oscillations of the earth, for example (Wenzel 1978).

MEM attenuates the effects of truncation of a time series by prediction filtering and thus virtually increases its length and consequent resolution (Smylie et al. 1973). The quality of the prediction filter found from the data depends on the type of time series under investigation. The best performance of MEM must be expected when autoregressive processes are being analyzed (Ulrych and Bishop 1975).

Acknowledgements. We are greatly indebted to Dr. K.-O. Millahn for helpful discussions and constructive criticisms. We also want to thank Drs. K. Fuchs, G. Müller, and E. Rietsch for critically reviewing the paper and Ingrid Hörnchen for typing the manuscript.

References

Baggeroer, A.V.: Confidence intervals for maximum entropy estimates. In: Aspects of signal processing, Part 2, G. Tacconi, ed.: Dordrecht, Holland: Reidel Publ. Comp., 617–630, 1977

Bowling, S.B.: Linear prediction and maximum entropy spectral analysis for radar application. M.I.T. Lincoln Laboratory Project Rep. RMP-122, 1977

Burg, J.P.: Maximum entropy spectral analysis. Paper presented at 37th Annual Meeting, Int. Soc. Exploration Geophysicists, Oklahoma City, Oklahoma, USA 1967

Chen, W.J., Stegen, G.R.: Experiments with maximum entropy power spectra of sinusoids. J. Geophys. Res. **79**, 3019–3022, 1974

Frost, O.L.: Power spectrum estimation. In: Aspects of signal processing, Part 1, G. Tacconi, ed.: Dordrecht, Holland: Reidel Publ. Comp., 125–162, 1977

Lacoss, R.T.: Data adaptive spectral analysis methods. Geophysics **36**, 661–675, 1971

Newman, W.I.: Extension of the maximum entropy method. IEEE Trans. Inform. Theory IT-23, 89–93, 1977

Radoski, H.R., Fougere, O.F., Zawalick, E.J.: A comparison of power spectral estimates and applications of the maximum entropy method. J. Geophys. Res. **80**, 619–625, 1975

Smylie, D.E., Clarke, G.K., Ulrych, T.J.: Analysis of the irregularities in the Earth's rotation. In: Methods of computational physics, B. Bolt, ed.: pp 391–430. New York: Academic Press 1973

Ulrych, T.J., Bishop, T.N.: Maximum entropy spectral analysis and autoregressive decomposition. Rev. Geophys. Space Phys. **13**, 183–200, 1975

Ulrych, T.J., Smylie, D.E., Jensen, O.G., Clarke, G.K.: Predictive filtering and smoothing of short records by using maximum entropy. J. Geophys. Res. **78**, 4959–4964, 1973

Wenzel, F.: Hochauflösende Spektralanalyse-Verfahren mit Anwendung der Maximum Entropie Methode auf das Kolumbienbeben 1970. Diplomarbeit, Geophysikalisches Institut, University of Karlsruhe, 1978

Received November 8, 1979; Accepted March 11, 1980

J. Geophys. 48, 124–125, 1980

Journal of Geophysics

Letter to the Editor

Errors in Matrix Element Expressions for the Reflectivity Method

D.R. Baumgardt

ENSCO, INC., Data and Computer Sciences Division
P.O. Box 2578, 1930 Highway A1A, Indian Harbour Beach, Florida 32937, USA

The purpose of this note is to correct some long standing errors in expressions for the bottom half space delta matrix used in the computation of the propagator matrix of the transition zone in the reflectivity method. The expressions were first given by Fuchs (1968) in his Sect. 6.3, and the same erroneous expressions have reappeared in Kind (1976).

Using the notation of Fuchs (1968), the lower half space matrix for an $n-2$ layer transition zone between 2 half spaces is:

$$T_n^{-1} = F \cdot \begin{pmatrix} -j2\mu_n k \nu_n \nu'_n & +j\mu_n l_n \nu'_n & -\nu_n \nu'_n & -k\nu'_n \\ -j\mu_n l_n \nu_n & -j2\mu_n k \nu_n \nu'_n & -k\nu_n & +\nu_n \nu'_n \\ -j2\mu_n k \nu_n \nu'_n & -j\mu_n l_n \nu'_n & -\nu_n \nu'_n & +k\nu'_n \\ +j\mu_n l_n \nu_n & -j2\mu_n k \nu_n \nu'_n & +k\nu_n & +\nu_n \nu'_n \end{pmatrix} \tag{1}$$

where

$$F = \frac{\beta_n^2}{2\mu_n \nu_n \nu'_n \omega^2}. \tag{2}$$

The elements of the delta matrix of the lower half space are the 2×2 subdeterminants of the T_n^{-1} matrix. The correct expressions are:

$$\begin{aligned}
&\hat{t}^n|^{12}_{12} = \hat{t}|^1_1 = -\frac{\beta_n^4}{4\omega^4} \cdot \left(4k^2 + \frac{l_n^2}{\nu_n \nu'_n}\right) \\
&\hat{t}^n|^{12}_{13} = \hat{t}^n|^1_2 = \frac{j\beta_n^2}{4\mu_n \nu'_n \omega^2} \\
&\hat{t}^n|^{12}_{14} = \hat{t}^n|^1_3 = \hat{t}^n|^{12}_{23} = \hat{t}^n|^1_4 = \frac{-j\beta_n^4}{4\mu_n \omega^3 c} \cdot \left(2 + \frac{l_n}{\nu_n \nu'_n}\right) \\
&\hat{t}^n|^{12}_{24} = \hat{t}^n|^1_5 = -\frac{j\beta_n^2}{4\mu_n \nu_n \omega^2} \\
&\hat{t}^n|^{12}_{34} = \hat{t}^n|^1_6 = -\frac{1}{4\rho_n^2 \omega^4} \cdot \left(1 + \frac{k^2}{\nu_n \nu'_n}\right).
\end{aligned} \tag{3}$$

These are the only elements of the 6×6 delta matrix for the lower half space that are required for computing reflection and transmission coefficients of the transition zone. Equivalent expressions for the half space delta matrix elements are also given by Watson (1970). They differ from the expression in (3) because Watson (1970) uses the dilatational and rotational wave solutions given by Haskell (1953) rather than the potential solutions given by Fuchs (1968).

The expressions given by Fuchs (1968) and Kind (1976) were apparently obtained by evaluating the 2×2 subdeterminants of the matrix in the brackets in (1) and multiplying by the common factor F in (2). In fact, the subdeterminants of the matrix in brackets should have been multiplied by F^2. Therefore, all delta matrices $m|^{ij}_{kl}$ of the Haskell propagator matrix M computed with the expressions in Fuchs (1968) and Kind (1976) would have a multiplicative error of F.

It should also be pointed out that the definitions of the vertical wavenumbers ν_i and ν'_i given by Kind (1976) are incorrect. The correct definitions are given in (12) and (13) in Fuchs (1968).

Fortunately, none of the results of the application of the reflectivity method to date are invalidated (Fuchs and Müller 1971; Kennett 1972; Kind and Müller 1975, to name a few). This is because these applications have required only the computation of topside reflection coefficients from the transition zone. For example, the reflection coefficients for incident P waves on the top of the transition zone are (Fuchs 1968; Červený 1974):

$$\begin{aligned}
R_{PP} &= R_{11} = m|^{12}_{23}/m|^{12}_{12} \\
R_{PS} &= R_{12} = -m|^{12}_{13}/m|^{12}_{12}.
\end{aligned} \tag{4}$$

Since the numerators and denominators are in error by a factor of F, the errors are removed by division. The error does not divide out for other coefficients, however. For the case of an incident P wave from the top, the top to bottom transmission coefficients are (Fuchs 1968; Červený 1974):

$$\begin{aligned}
T_{PP} &= R_{13} = m_{31} \cdot R_{PP} + m_{32} \cdot R_{PS} + m_{33} \\
T_{PS} &= R_{14} = m_{41} \cdot R_{PP} + m_{42} \cdot R_{PS} + m_{43}
\end{aligned}$$

The m_{ij} are elements of the Haskell matrix M. For these expressions, the errors remain in the elements of the M matrix and are not divided out. From Červený (1974), the expressions for R_{13}, R_{14}, R_{23}, R_{24}, R_{31}, R_{32}, R_{41}, and R_{42} would also contain the multiplicative error whereas in the expressions for R_{21}, R_{22}, R_{33}, R_{34}, R_{43}, and R_{44}, the error is removed by division.

References

Červený, V.: Reflection and transmission coefficients for transition layers. Stud. Geophys. Geod. **18**, 59–68, 1974

Fuchs, K.: Das Reflexions- und Transmissionsvermögen eines geschichteten Mediums mit beliebiger Tiefenverteilung der elastischen Moduln und der Dichte für schrägen Einfall ebener Wellen. Z. Geophys. **34**, 389–413, 1968

0340-062X/80/0048/0124/$01.00

Fuchs, K., Müller, G.: Computation of synthetic seismograms with the reflectivity method and comparison with observations. Geophys. J. R. Astron. Soc. **35**, 243–264, 1971

Haskell, N.A.: The dispersion of surface waves on multilayered media. Bull. Seismol. Soc. Am. **43**, 17–34, 1953

Kennett, B.L.N.: Seismic waves in laterally inhomogeneous media. Geophys. J. R. Astron. Soc. **27**, 301–325, 1972

Kind, R., Müller, G.: Computation of SV waves in realistic earth models. J. Geophys. **41**, 149–172, 1975

Kind, R.: Computation of reflection coefficients for layered media. J. Geophys. **42**, 191–200, 1976

Watson, T.H.: A note on fast computation of Rayleigh wave dispersion in the multilayered elastic halfspace. Bull. Seismol. Soc. Am. **65**, 765–766, 1970

Received December 21, 1979; Revised Version March 11, 1980

Comment of the Editor

The three authors of the Reflectivity Method (K. Fuchs, G. Müller, R. Kind) acknowledge that D.R. Baumgardt has pointed out an important error in the reflectivity method for the case of transmission.

J. Geophys. 48, 127–138, 1980

Journal of Geophysics

Magnetometer Array Observations of a Giant Pulsation Event

K.-H. Glaßmeier
Institut für Geophysik der Universität Münster, Gievenbecker Weg 61, D-4400 Münster, Federal Republic of Germany

Abstract. On 19 November 1976 a rather well developed giant pulsation event occurred at about 0600 MLT over northern Scandinavia. Results of observations obtained with a dense network of magnetometer stations are reported. The method of analysis is based on the concept of the analytical signal. The disturbance region was strongly localised. The amplitudes of the horizontal components show a pronounced difference: the amplitude maximum of the N-S component is clearly moving to the west while that of the E-W component is nearly fixed relative to the array. The frequency of the event decreased with time. Both the frequency decrease and the drift of the amplitude maximum are interpreted in terms of bounce – resonance instability. The spatial polarisation pattern found is in agreement with the fieldline-resonance theory in that the sense of rotation of the horizontal disturbance vector is as predicted by the theory. At the region of maximum amplitudes the pulsation is linearly polarised in E-W direction, as observed by other workers. This does not agree with the fieldline-resonance model. Apparent azimuthal wave numbers of the horizontal components show that the event analysed is a westward-travelling wave. The wave numbers are as large as predicted by Olson and Rostoker's (1978) empirically derived relation between the E-W wave number and the frequency of pc 4–5 pulsations.

Key words: Giant pulsation – pc 4 and 5 Pulsations – Bounce resonance instability – Fieldline-resonance model – Apparent EW wave numbers – Analytical signal.

Introduction

Among the various types of geomagnetic micropulsations, giant pulsations (pg) are striking because of their sinusoidal form and their large amplitudes. According to the definition given by Annexstad and Wilson (1968), (1) they occur simultaneously at conjugate points; (2) their average period is about 100 s; (3) their spatial distribution is confined to a narrow range of latitudes in the auroral zone; (4) they occur predominantly in the early morning hours (Harang 1936); and (5) they are well modulated and have amplitudes of several tens of nT (Jacobs 1970). Another well-known feature of pgs is the change of their period with time. This may be positive or negative and typically is of the order 10 s/h (Eleman 1967; Annexstad and Wilson 1968; Hillebrand 1974; Green 1979).

As a possible cause of pgs, Obayashi and Jacobs (1958) suggested fieldline-guided hydromagnetic waves. This idea was supported by Annexstad and Wilson's (1968) study of the conjugate behaviour of pgs, and by Ol's (1963) observation of a clear relationship between the latitude of a pg region and the main period of the event.

Hillebrand (1976), Green (1979), and Rostoker et al. (1979) recently showed that giant pulsations exhibit some features that are otherwise known from pc 4–5 micropulsation studies (e.g. Samson et al. 1971; Green 1978). A region exists where the disturbance amplitude reaches a clear relative maximum. North of this demarcation or resonance region the horizontal disturbance vector shows a clockwise rotation (if viewed in the direction of the geomagnetic fieldline); south of the resonance region the opposite sense of rotation is found. Starting from these observations and from the fieldline-resonance theory as described by Southwood (1974) and Chen and Hasegawa (1974), Green (1979) suggested that pgs are westward travelling waves.

To prove this, and the applicability of the fieldline resonance model to giant pulsations, and also to derive a more detailed picture of the spatial amplitude, phase and polarisation behaviour, observations by a dense magnetometer array seem to be necessary.

Since summer 1976 the University of Münster has been operating such an array of up to 36 magnetometers of an improved Gough-Reitzel type (Gough and Reitzel 1967; Küppers and Post in press 1980) in northern Scandinavia (Küppers et al. 1979). In this paper we will give results of the observation of a pg event by means of the Scandinavian Magnetometer Array and of a detailed data analysis.

Instrumentation

The sites of the magnetometers used in this study are shown in Fig. 1. The instruments are located along five N-S profiles nearly parallel to geomagnetic meridians (profile 1: FRE-HAS, profile 2: AND-LYC, profile 3: MIK-PIT, profile 4: SOY-SAU, profile 6: BER-SKO). The spacing between magnetometers along these profiles and between the profiles themselves varies 100–150 km. All magnetometers observe the magnetic field variations with 10 s temporal resolution. A detailed description of this magnetometer array has been given by Küppers et al. (1979).

The coordinate system indicated in Fig. 1 was introduced by Küppers et al. (1979) and has been named the Kiruna System. It is a cartesian system obtained by a stereographic projection of the globe onto a tangential plane centered at Kiruna, Sweden (67.8° N, 20.4° E). The y_{KI} axis of the system, whose origin is

0340-062X/80/0127/$02.40

situated at Kiruna (KIR, see Fig. 1), has been chosen as the tangent to the projection of the line $\phi_c = \phi_c$ (KIR) = 64.8° with ϕ_c denoting the revised corrected geomagnetic latitude as given by Gustafsson (1970). The x_{KI} axis points approximately 12° west of geographic north of KIR.

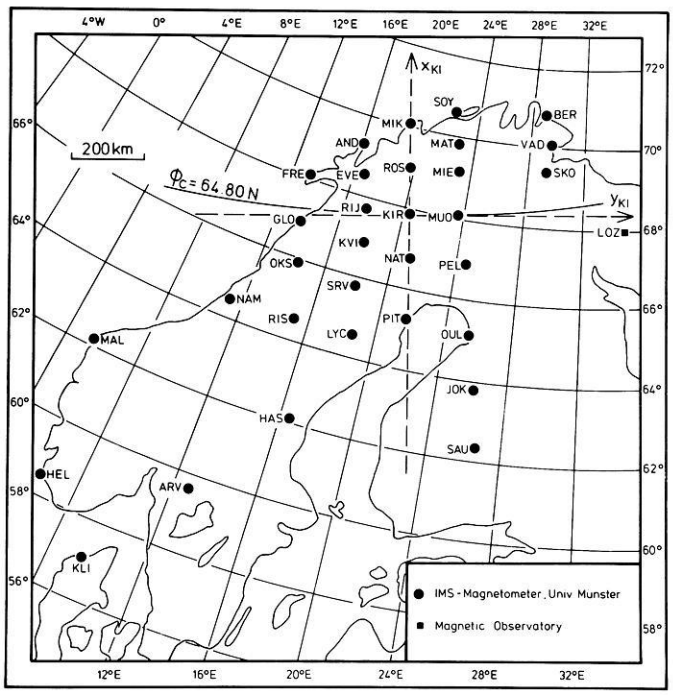


Fig. 1. Station map in geographic coordinates. Also indicated are the axes of the Kiruna system (see text) and the line of constant revised corrected geographic latitude $\phi_c = 64.80$ N (after Gustafsson 1970)

For pulsation studies the amplitude and phase response of the magnetometers are of some importance. Typical response curves may be found in Küppers and Post (1980). Here, we restrict ourselves to pointing out that, for periods around 100 s, no amplitude correction is necessary whereas phase distortion is of the order of 10°. The uncertainty in phase determination caused by a time resolution of 10 s is more severe and, for a period of 100 s, may amount to ±20°. The amplitude resolution is about 2 nT (Küppers et al. 1979).

Description of Data and Method of Analysis

On 19 November 1976 a rather well developed giant pulsation event was observed between 0330 UT and 0450 UT (about 0550–0710 MLT) by the 20 northernmost instruments of the Scandinavian Magnetometer Array at a time of moderate magnetospheric activity ($K_p = 3+$). As an example, the magnetic variations of the A component (magnetic deflection parallel to the x_{KI} axis), the B component (magnetic deflection parallel to the y_{KI} axis) and the Z component are given for the northernmost stations of profile 4 relative to the averaged quiet day level on 18 November 1976, 0800–1200 UT (Fig. 2). The magnetograms show the known quasi-monochromatic nature of pgs as reported by others (Rolf 1931; Annexstad and Wilson 1968; Hillebrand 1976; Green 1979). The pulsation event occurred during the recovery phase of a small magnetic deflection seen in the A and Z components, while the low frequency part of the B component remained nearly constant. To isolate the pulsation train with a main period of about 100 s, all records were band-pass filtered with a filter of trapezoidal shape (Schmucker 1978) and cut-off frequencies at 4 and 16 mHz.

Some of the band-pass-filtered A and B data are given in Fig. 3 for a north-south profile (SOR-PEL) and east-west line (FRE-VAD). The strong localisation of the disturbance is remarkable. Between MIE and MUO the amplitude diminishes to

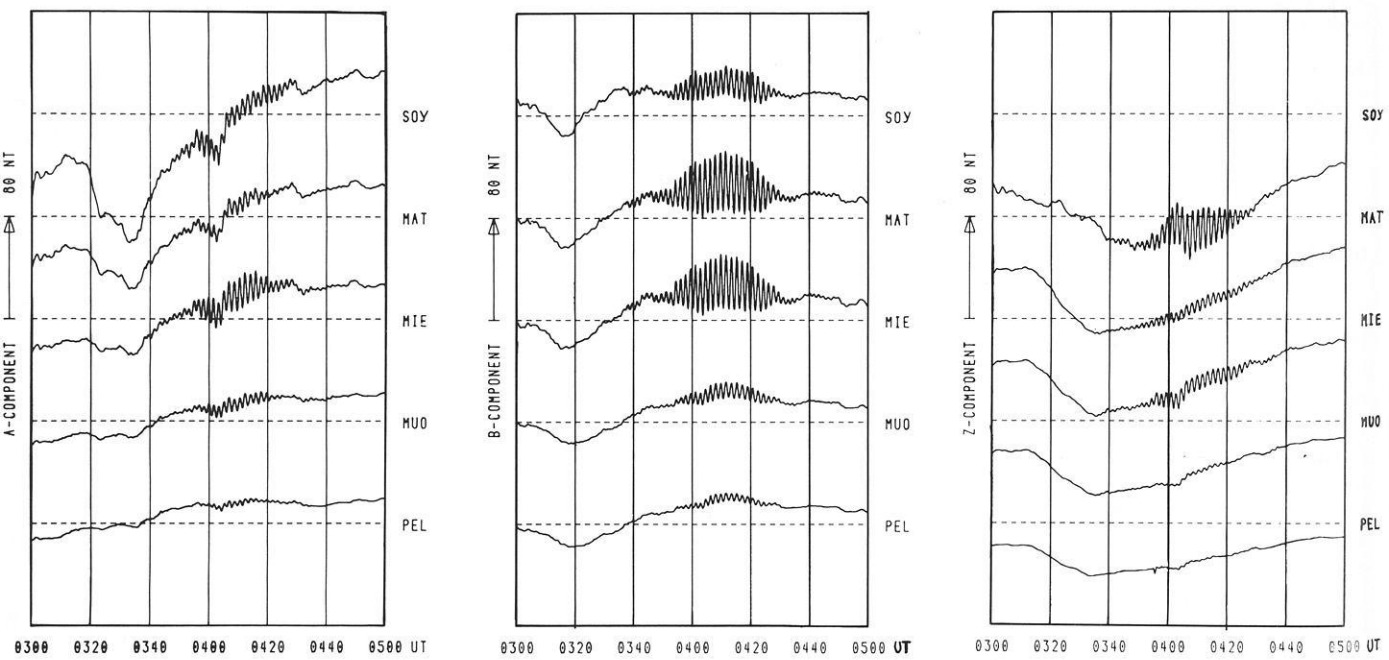


Fig. 2. Magnetograms of the unfiltered A, B, and Z components for profile 4 on 19 November 1976, 0300–0500 UT. The A and B components are defined parallel to the x_{KI} and y_{KI} axis, respectively (see Fig. 1)

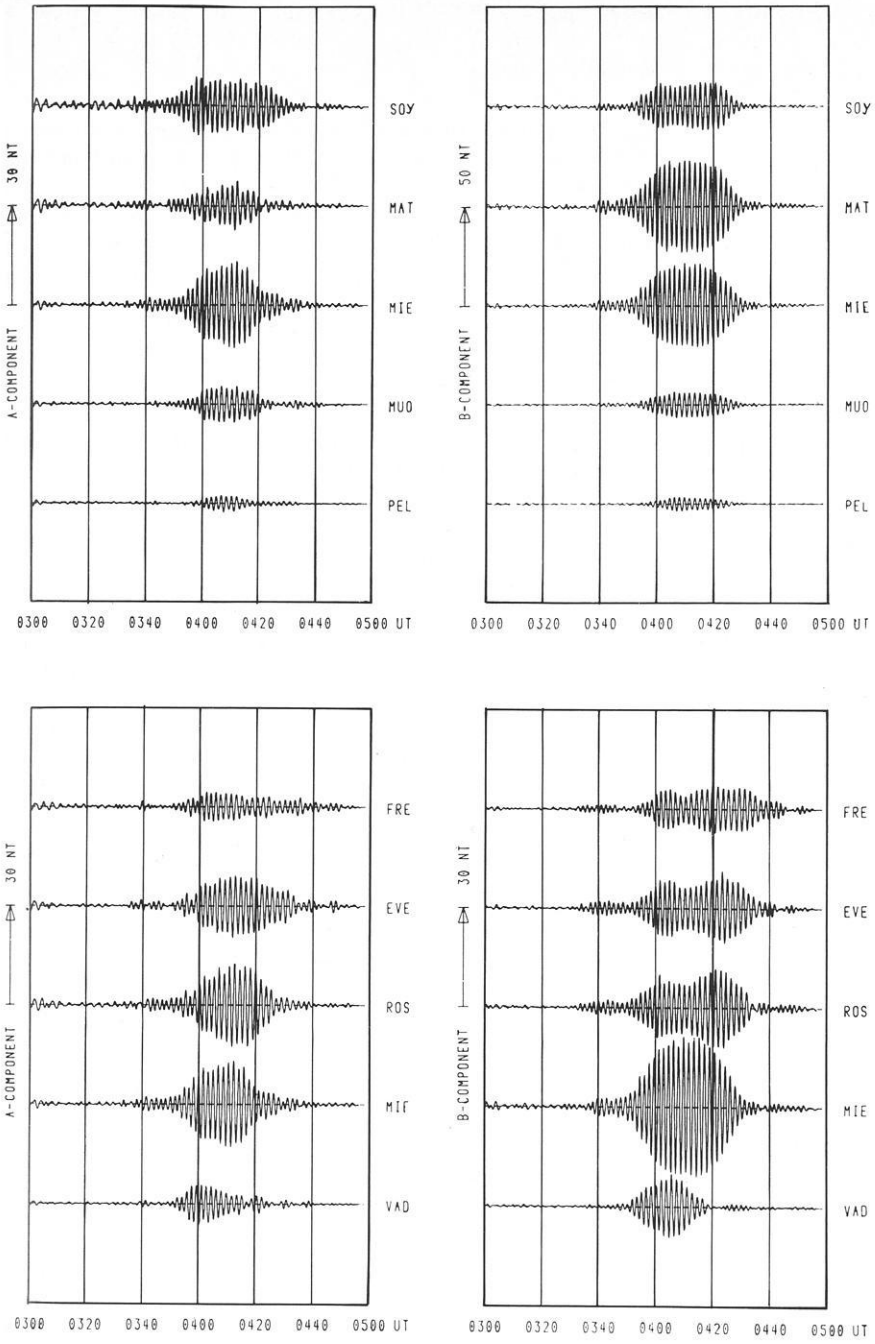


Fig. 3. Magnetograms of band-pass filtered A and B components for profile 4 (*upper part*) and for an E-W profile from FRE to VAD (*lower part*) (see Fig. 1). Note also the different scale in the upper right part of this figure

about one-half over a north-south distance of about 100 km. On the profile (FRE-VAD) the stations west of MIE show a clear double-peak structure in B, whereas all other stations show a simple wave-packet. In the following no regard is paid to the Z component because, in the pulsation frequency band, there are marked induction effects in this component over Scandinavia, especially near the northern coast (Jones and Olafsdottir 1979; Küppers et al. 1979).

During further analysis of the data we considered the nearly sinusoidal character of pgs and regarded them as amplitude- and phase-modulated quasi-monochromatic wave trains. We considered the techniques of statistical frequency analysis to be inapplicable to our data because of the obvious non-stationary character of the event. To derive amplitude and phase information from the data we used the complex or analytical signal method as first described by Gabor (1946).

For a given amplitude and phase modulated monochromatic signal $x(t)=A(t)\cos(\omega t+\varphi(t))$ a complex or analytical signal $z(t)$ may be defined by

$$z(t)=x(t)+iy(t) \tag{1}$$

where $y(t)$ is the Hilbert transform of $x(t)$

$$y(t)=-\frac{1}{\pi}\int_{-\infty}^{+\infty}\frac{x(\tau)}{t-\tau}\,d\tau. \tag{2}$$

For the instantaneous amplitude $A(t)$, the instantaneous phase

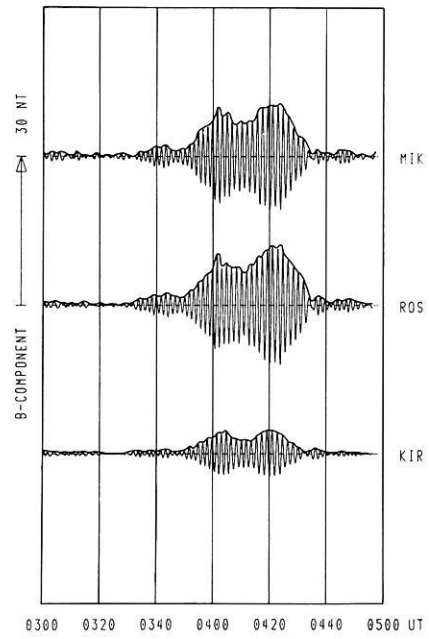


Fig. 4. Magnetograms of the band-pass filtered B components of some selected stations together with the curves of the instantaneous amplitude computed by the analytical signal method

$\varphi(t)$, and the instantaneous frequency $f(t)$, the following relations hold:

$$A(t)=\sqrt{x^2(t)+y^2(t)}$$

$$\varphi(t)=\arctan(y(t)/x(t))$$

$$f(t)=\frac{d}{dt}\varphi(t). \tag{3}$$

The idea of the analytical signal is based on certain advantages shown by operations with exponentials $\exp(i\xi)$ as compared to operations with $\sin\xi$ or $\cos\xi$ functions (Gabor 1946). The appearance of the Hilbert transform may be understood by remembering that the sin function is the Hilbert-transform of the cos function. To provide for the computation of the Hilbert transform of our records all band-pass filtered data (see the examples shown in Fig. 3) were set to zero before 0300 UT and after 0500 UT. More detailed descriptions of this demodulation technique are given by Fischer (1969), Farnbach (1975), and Kodera et al. (1977).

As an example of what the above-described method may yield, in Fig. 4 the band-pass filtered B records of the stations MIK, ROS and KIR are shown together with the amplitude curves as computed by the analytical signal method. The amplitude curves are just the envelopes of the records, and even very minute changes in the envelopes are resolved.

Assuming the amplitude error, δA, to be as large as the magnetic resolution of 2 nT (Küppers et al. 1979), we may estimate the phase error, $\delta\varphi$, and the frequency uncertainty, δf, from Eq. (3) to be $\delta\varphi\approx 11°$ and $\delta f\approx 2$ mHz for an amplitude A $=10$ nT.

With the amplitude and phase values of two components, A and B for example, it is also possible to compute instantaneous polarisation parameters in the way given by Born and Wolff (1975, p. 28).

Observations

Amplitude Behaviour

The geographical distribution of the instantaneous amplitudes of the horizontal components (Fig. 5a, b) shows, at all times, an ellipsoidal shape, i.e., strongly localised, but with slightly different half-widths, in both main directions. In the N-S direction the half-width of the A and B distributions may be estimated to be about 300 km, as already found by Hillebrand (1974, 1976). In the E-W direction it amounts to about 500 km for the A component and 300 km for the B component. For most of the time there is a pronounced amplitude maximum in the B component situated on profile 4. Pulsation magnetograms from the Soviet Union station Lovozero (LOZ, see Fig. 1) still show the pg event but with an amplitude smaller than 2 nT.

The most striking feature in Fig. 5 is the completely different behaviours which the disturbance maxima of the A and B components show in time. The A maximum is indicated by the dashed circles in Fig. 5a. It clearly moves to the west along the line MIE-FRE (65.7° geomagnetic latitude). Only very late within the event (0430 UT, Fig. 5b) do we see a movement of the B maximum to the west. For the A component a small movement of the maximum to the south may also be noticed (Fig. 5a), but this observation is of doubtful significance.

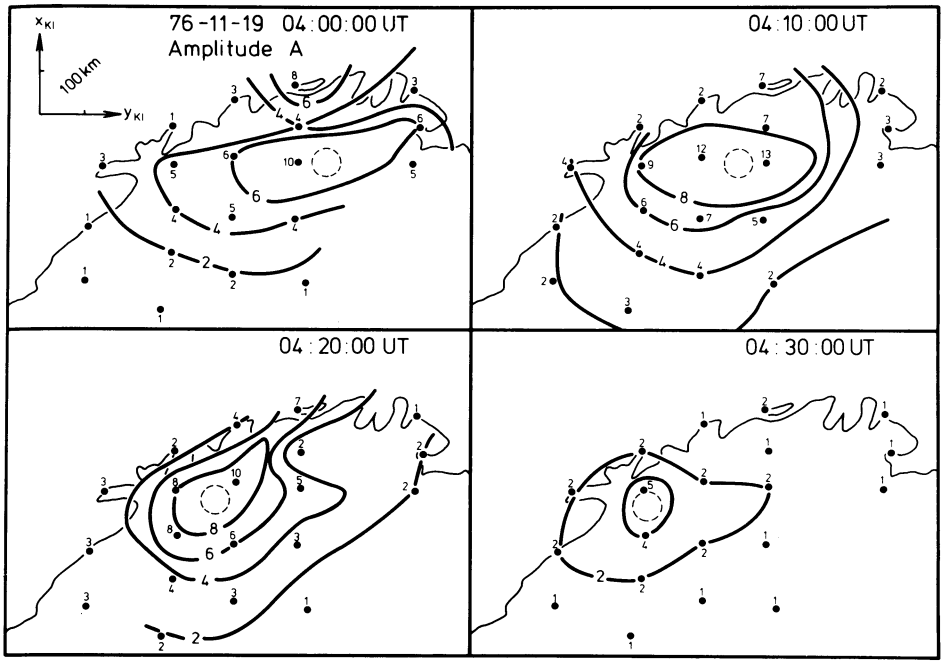
The E-W drift velocity of the A maximum is $v=-(0.18\pm 0.02)$ km/s, where the negative sign indicates a drift to the west. If we transform this value to an apparent angular velocity around the earth's axis by virtue of $\omega_A=v/R_E\cdot\cos\Phi$, with R_E the earth's radius and Φ the geomagnetic latitude, we get $\omega_A=-(6.8\pm 0.7)\cdot 10^{-5}$ rad/s which is, in magnitude, remarkably close to the earth's angular velocity of rotation of $7\cdot 10^{-5}$ rad/s. Altogether this means that the A maximum is nearly fixed with respect to the magnetosphere throughout the event.

Phase Behaviour

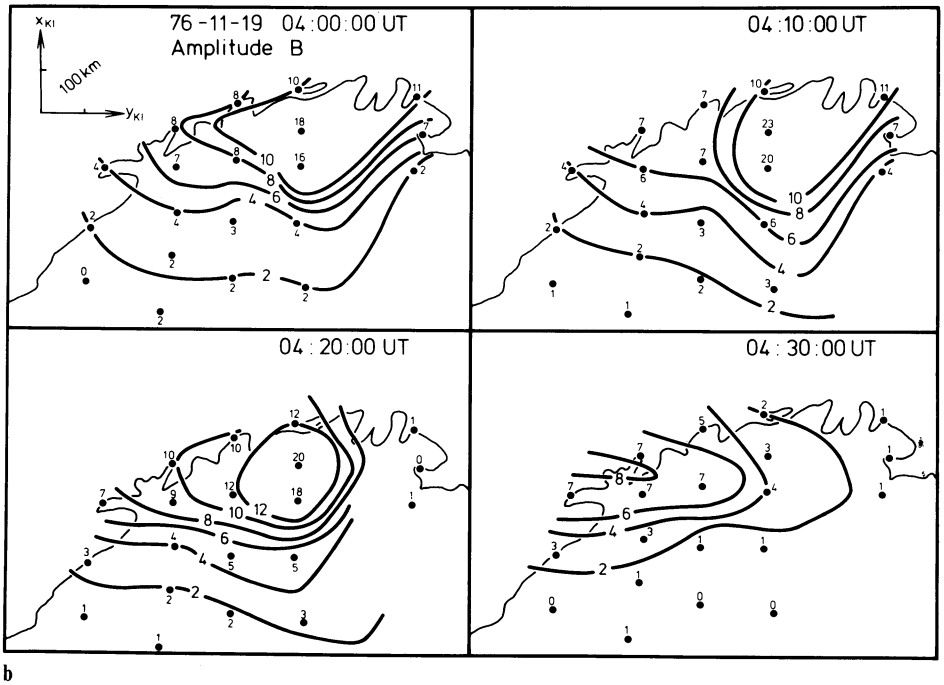
Again isolines are used to demonstrate the phase behaviour of the event analysed (Fig. 6a, b). For the A component the isolines are very roughly straight, N-S directed lines up to the latitude of KIR. Further to the north large-phase differences between the coastal stations and the neighbouring inland stations are observable. For example, at 0410 UT this difference reaches 177° between the stations SOY and MAT over a distance of only about 100 km.

For the B component no such sudden phase changes are visible. The isolines are strongly aligned along the N-S direction. As for the A component, the eastern stations lead in phase which means phase propagation to the west. This was suggested earlier by Green (1979) from pg observations at Kiruna and Tromsö using the predictions of the fieldline resonance theory (Southwood 1974), and was found observationally by Rostoker et al. (1979).

For three E-W profiles (AND, MAT, VAD along about 66.6° geomagnetic latitude, FRE, EVE, ROS, MIE, SKO along about 66.0°, and GLO, RIJ, KIR, MUO along about 64.8°) we computed apparent azimuthal wave numbers, m, for the B component at three different times (see Table 1). For the A component azimuthal wave numbers have only been computed on the last two profiles because the northernmost E-W profile (AND, MAT, VAD) is already in the region where large N-S spatial phase-gradients are found. The given wave numbers are



a



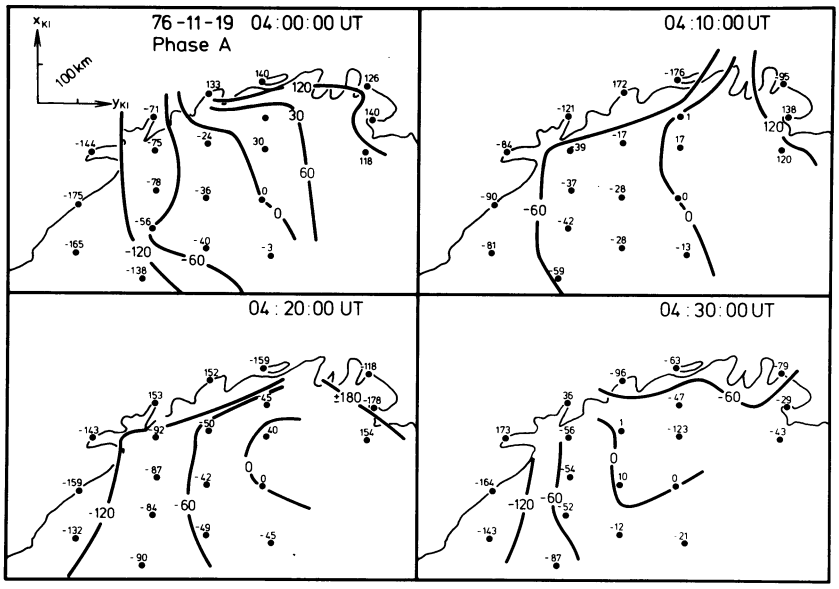
b

Fig. 5. a Maps of the instantaneous amplitude (in nT) of the A component at four different times. Linear interpolation has been used to construct the isolines. The *dashed circles* indicate the region of the A disturbance maximum
b As Fig. 5a, but for the B component

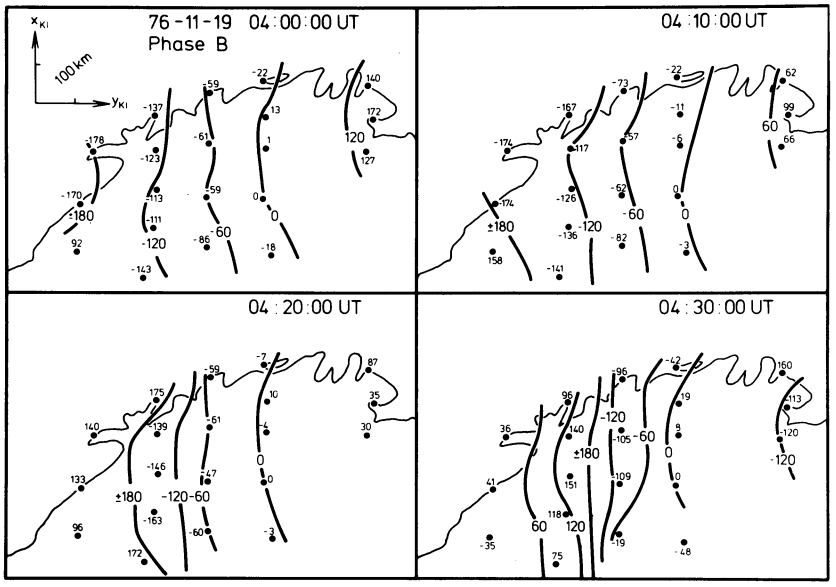
mean values as computed from the phase differences between successive station pairs. An apparent wave length, λ, is defined by $\lambda = \frac{2 \cdot \pi \cdot R_E \cdot \cos \Phi}{m}$ (Green 1976). As may be seen from Table 1 there is no systematic difference between the wave numbers for the various times and geomagnetic latitudes of the profiles considered. Within the uncertainty of each m value there is no significant difference between the wave number of the A and B components.

Polarisation

As examples, the polarisation ellipses in the A-B plane of the event analysed, at two times, are given in Fig. 7. In every case, two perpendicular straight lines may be recognised dividing the pg region into four quadrants, S_1-S_4 (Fig. 8). One line, in the E-W direction, separates two regions in which the senses of rotation of the horizontal disturbance vectors are different. Along this line, nearly linear polarisation in the direction of the



a



b

Fig. 6. a Maps of the instantaneous phase (in degrees) of the A component, φ_A, relative to the station MUO: $\varphi_A = \varphi_{ST} - \varphi_{MUO}$, where φ_{MUO} is the phase of the reference-station and φ_{ST} is the phase of a station as computed by the analytical signal method
b As Fig. 6a, but for the B component

Table 1. Apparent azimuthal wave-numbers, m, at different times, measured along the three EW-profiles (AND, MAT, VAD), (FRE, EVE, ROS, MIE, SKO) and (GLO, RIT, KIR, MUO). λ denotes the corresponding apparent azimuthal wave-length

Time (UT)	A component				B component					
	FRE		GLO		AND		FRE		GLO	
	m	λ (km)	m	λ (km)	m	λ (km)	m	λ (km)	m	λ (km)
4:00	24.0 ±5.8	678 ±130	20.5 ±6.5	831 ±180	29.9 ±2.9	534 ±50	23.5 ±4.1	693 ±100	21.7 ±4.7	785 ±120
4:10	17.9 ±11.9	909 ±350	10.2 ±5.6	1,670 ±600	25.5 ±4.8	626 ±100	19.5 ±3.8	834 ±100	22.7 ±8.0	751 ±200
4:20	24.1 ±9.9	675 ±200	19.4 ±1.8	878 ±80	36.2	441	27.5 ±4.4	591 ±80	29.1 ±13.8	585 ±180

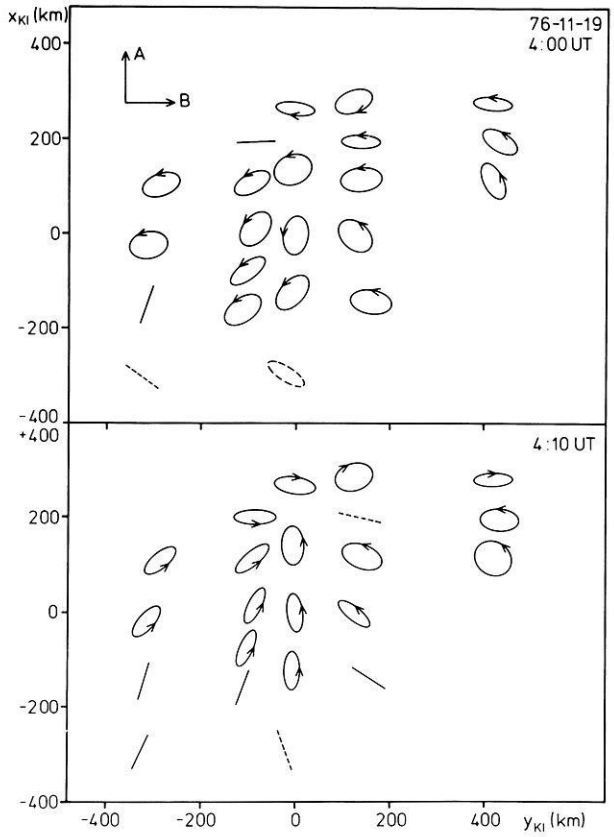


Fig. 7. This figure shows the polarisation state of the pg event analysed, for two different times. The ellipticity, the azimuth angle, the main axis, and the sense of rotation are averaged over two periods (4:00:00–4:03:20, 4:10:00–4:13:20). At all stations the horizontal disturbance vector is normalised to the main axis so that the polarisation ellipses shown all are of the same length. The direction of the main axis is given correctly, whereas different ellipticities have only been indicated. The sense of rotation is given by the *arrows*. *Dashed ellipses* are ambiguous

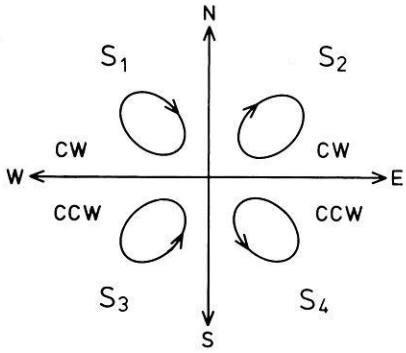


Fig. 8. Schematic representation of the observed polarisation state of the pg analysed. S_1 to S_4 denote the different sectors. The line $\overline{WE}$ represents the demarcation line and $\overline{NS}$ is the meridional separator line (for a definition of these lines see text)

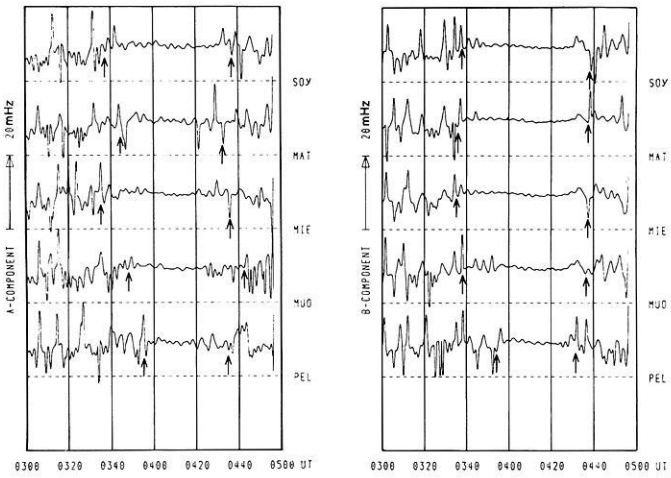


Fig. 9. Curves of the instantaneous frequency of the A (*left*) and B (*right*) component for the northernmost stations of profile 4. The *arrows* indicate the beginning and end of the pulsation event judged from the band-pass filtered magnetograms (see Fig. 3)

B component is found and therefore this may be called the demarcation line, according to the definition of Samson et al. (1971). The second line, in the N-S direction, which we would like to call the meridional separator line, divides two regions in which the inclination of the main axis of the ellipses with respect to the northern direction is positive or negative, respectively. In agreement with the prediction of the fieldline-resonance model for a westward travelling wave (as in our case) in the northern sectors S_1 and S_2 the rotation is clockwise (when viewed in the direction of the geomagnetic field), and counterclockwise in the southern sectors S_3 and S_4. The change of the sense of rotation of the horizontal disturbance vector at the station BER between 0400 UT and 0410 UT may be explained by a drift of the pg region, and therefore also of the demarcation line, to the south. Such a drift is also indicated in Fig. 5a, as mentioned above. The observation of an equatorward drift of the pg region is in contrast to the observation of Rostoker et al. (1979) who reported a poleward motion.

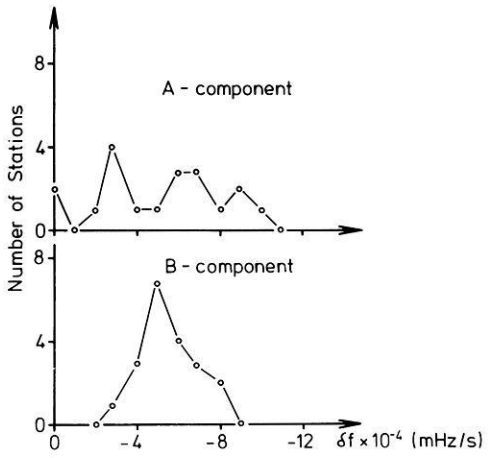


Fig. 10. Distribution of the rate of the frequency change with time as computed by a least-squares fit to the curves of the instantaneous frequencies between the two times marked by arrows in Fig. 9, at the 20 northernmost stations of the array for the A and B components

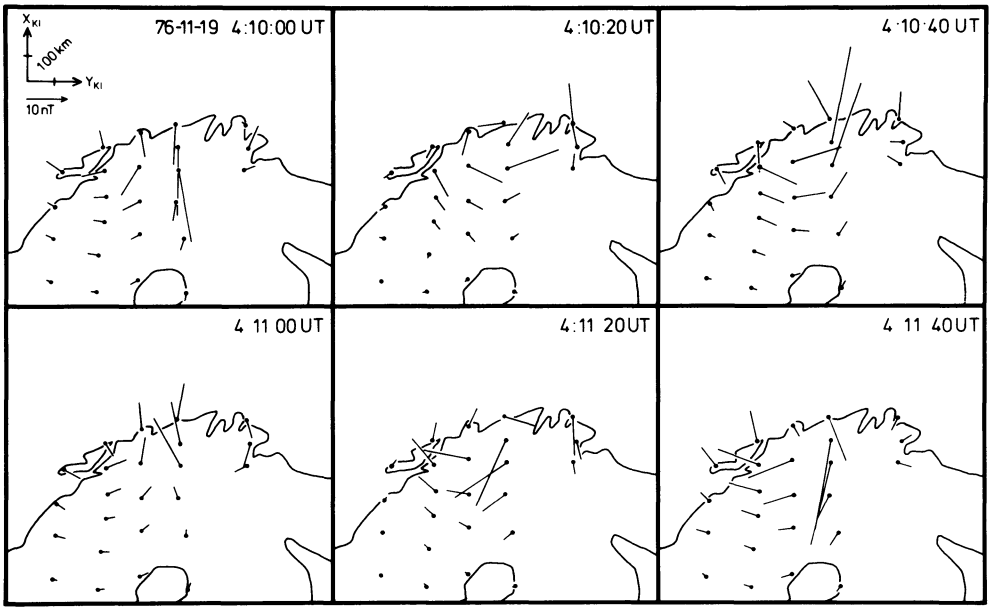


Fig. 11. Equivalent current vectors (in nT) at selected instants of time within the northern part of Scandinavia. The current vectors have their origin at the stations where the corresponding magnetic disturbances have been recorded

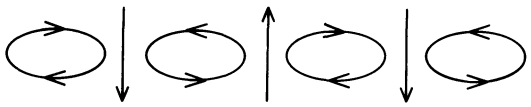


Fig. 12. Schematic representation of an equivalent current system whose periodicity in space may produce the observed temporal variations if the system drifts (after Obertz and Raspopov 1968)

Instantaneous Frequency

Curves of the computed instantaneous frequency show, on the average during the event, a decrease (Fig. 9). Between the start and end points of the event, which are indicated by arrows in Fig. 9, a linear least-squares fit to the frequency curves was made. The distribution of the calculated frequency gradients for A and B is given in Fig. 10. Again, as for the amplitude, the two horizontal components show different behaviour. For the B component we see a narrowly peaked distribution around the value $-5\ 10^{-4}$ mHz/s whereas the A component shows a wide distribution between $-2\ 10^{-4}$ and $-10\ 10^{-4}$ mHz/s. The mean frequency gradient from both components is $-5.5 \cdot 10^{-4}$ mHz/s corresponding to an increase of the period of about 20 s/h, a value already found by others (e.g., Annexstad and Wilson 1968; Green 1979). No dependence of the frequency gradient on the location of the magnetometer station has been observed.

System of Equivalent Current Vectors

Equivalent overhead current systems have been used before to represent magnetic micropulsations (e.g., Jacobs and Sinno 1960; Wilson 1966; Obertz and Raspopov 1968). For the event analysed, at 0410 UT the equivalent current vectors, i.e., horizontal magnetic disturbance vectors rotated 90 degrees clockwise if viewed from above, seem to show two current vortices (Fig. 11). The one to the west is well developed while that to the east is only suggested. The following diagrams in Fig. 11 demonstrate the time behaviour of the equivalent currents during one period (100 s). It appears that a system of current vortices with the currents flowing against each other (see Fig. 12) is moving to the west. The extent of a single current vortex is about 10° in longitude (see Fig. 11), and the mean velocity of the whole system is estimated from Fig. 11 to be (8.3 ± 2.7) km/s.

Discussion and Interpretation

The question of a drift of the pg region has already been discussed by Eleman (1967). He states – in contrast to the corresponding result of the present paper – that the pulsation region remains rather fixed relative to the earth. We suppose that the magnetometer network which he used was not as dense as is necessary for an exact determination of the location of the disturbance maximum. Furthermore we like to point out Eleman's (1967) observation that the events which he analysed were apparent not only in Scandinavia but also at Leirvogur, Iceland, at about 1800 km distance west of KIR, and occurred there up to 1 h later. This may correspond to a westward drift of the disturbance center of the order of 0.5 km/s, which is comparable to our present observation.

Interpreting the measured drift velocity as the azimuthal group velocity of a pg signal, a possible explanation for its value may be found in terms of a drift instability like the bounce resonance mechanism (Southwood et al. 1969; Southwood 1976). The resonance condition for such an instability is

$$\omega - l\omega_D = N\omega_B \tag{4}$$

where ω, ω_D, and ω_B are the wave frequency, the drift frequency and the bounce frequency, respectively, of a particle trapped in the earth's magnetic field, l is the magnetospheric azimuthal wave number of the disturbance, and N is an integer.

Our results about the phase behaviour of the pg event considered imply phase velocity towards the west and therefore $l<0$ in a coordinate system in which Λ, the azimuthal coordinate, increases towards the east. As Southwood et al. (1969) showed for a westward travelling wave, protons with an inward-

directed density gradient are a possible population of resonant particles. Using the center angular frequency of the pg event analysed, $\omega = 20 \cdot \pi$ mHz, in (4) we get two resonance energies, a low and a high one (Southwood et al. 1969). For the lower energy the disturbance frequency approximately equals the bounce frequency of the protons, and we have for the energy, W, of the resonant particles $W \sim 13$ keV.

Using an expression for ω_D as given by Hamlin et al. (1961) and ignoring a slight dependence of ω_D on the equatorial pitch angle (see Southwood et al. 1969), we have

$$\omega_D \approx -\frac{2 \cdot L \cdot W}{q \cdot B_0 \cdot R_E^2}. \tag{5}$$

Here L, W, and q are the McIlwain parameter, the energy, and the electric charge, respectively, for the particles under consideration; B_0 is the equatorial geomagnetic flux density at the earth's surface; and R_E is the earth's radius. ω_D as calculated by Hamlin et al. (1961) involves only the gradient and the curvature drifts. For such low energy particles as the resonant protons (see above) the drift velocity is not governed by gradient and curvature drifts alone, but we also have to take into account an $\mathbf{E} \times \mathbf{B}$ drift because of the dawn-dusk magnetospheric electric field (Schulz and Lanzerotti 1974, p. 6f). This drift is to the east on the morning side and therefore lowers the westward drift velocity as computed by Hamlin et al. (1961). For the angular velocity $\omega_{E \times B}$ that is related to the $\mathbf{E} \times \mathbf{B}$ drift it follows

$$\omega_{E \times B} = \frac{E}{B_0} \frac{L^2}{R_E} \text{ [rad/s]} \tag{6}$$

with E denoting the magnitude of the dawn-dusk field. A typical value is $E \sim 5 \cdot 10^{-4}$ V/m (Brice 1967). With $L = 6.6$ (this corresponds, e.g., to the L value of MAT) and putting $\omega_{TOT} = \omega_D + \omega_{E \times B}$, where ω_{TOT} is now the drift velocity of a trapped particle because of gradient and curvature drifts as well as the $\mathbf{E} \times \mathbf{B}$ drift. From (5) and (6) and putting in the above values we get $\omega_{TOT} \sim -2.6 \cdot 10^{-5}$ rad/s which agrees roughly with the measured apparent angular velocity, $\omega_A = -6.8 \ 10^{-5}$ rad/s. On the other hand, now putting $\omega_{TOT} \sim \omega_A$, we may estimate in the same way the dawn-dusk field to have been $3.1 \cdot 10^{-4}$ V/m.

Therefore, starting from the resonace condition (4) one finds a clear relation between the observed drift velocity of the disturbance center in the A component and the drift of the trapped protons if not only gradient and curvature drifts are considered but an $\mathbf{E} \times \mathbf{B}$ drift is also taken into account. The idea that the observed signal moves with the same velocity as the resonant particles may also explain the regularity and large amplitudes of pgs because then the disturbance has time to build up.

No good explanation can be given for the different drift behaviour of the B component. We suggest that some unknown influence of the ionosphere on a magnetospheric pulsation signal might be a reason for that difference. How the relative motion of the pulsation source against the ionosphere is of importance still has to be investigated.

The period drift of a giant pulsation (see Fig. 9) may also be discussed in terms of the bounce resonance mechanism. We propose that a change of the resonance condition (4) is responsible for the period change as is probably the case for IPDPs (*I*ntervals of *P*ulsations with *D*iminishing *P*eriods, see review by Jacobs 1970). IPDPs are believed to be caused by ion cyclotron instabilities, and a change of the resonance condition is caused, for example, by an inward convection of the resonating protons which results in an increase of the particle gyration frequency (Heacock et al. 1976).

A bounce resonance between protons and a magnetic disturbance results in a radial diffusion of the resonant particles. As Dungey (1964) showed, a meridional magnetic field disturbance, b_N, results in a zig-zag motion of the bouncing particles (see Dungey's Fig. 1), and the particles experience a change of their L value with time (Dungey, 1964),

$$\dot{L} = \frac{b_N \cdot v_{\parallel} \cdot \cos^3 \Theta \cdot L^3}{B_0 R_E} \tag{7}$$

where $v_{\parallel}$ is the particle's velocity parallel to the ambient field and Θ is the colatitude of the particle's gyration center. As Southwood et al. (1969) showed for a westward travelling wave, as observed in our case, there must be an inward-directed gradient of the proton density for the instability to arise. This means that L has to increase during the event because the protons cannot diffuse against the density gradient. With $b_N = 10$ nT and an equatorial pitch angle of, for example, 80° one finds for 13 keV protons (such protons have a bounce period of about 100 s, see above): $\dot{L} \approx 4 \cdot 10^{-3}$ s^{-1}. This corresponds to a frequency decrease of $\delta f \sim 4 \cdot 10^{-2}$ mHz/s, a value which is much larger than the observed one. A more refined discussion of the diffusion described and a consideration of non-linear effects may result in better coincidence of theoretical and observed values. In agreement with Green (1979), who suggested that a radial motion of the source region might be the cause of the frequency change, we propose that radial diffusion of the resonant particle distribution might be the reason for such a radial motion and a possible explanation of the observed period drift.

The results of the polarisation analysis in the present paper (compare Fig. 8) and the earlier works of Hillebrand (1976) and Green (1979) show that the fieldline resonance theory as described by Southwood (1974) and Chen and Hasegawa (1974) is strictly not applicable to pgs because at the resonance region they show linear polarisation in the direction of the B component and not in that of the A component as predicted by the theory (Southwood 1974; Walker et al. 1979) and as observed, for example, for pc 4–5 pulsations (Green 1978). The fieldline-resonance model also does not predict the existence of the meridional separator line. In agreement with Rostoker et al. (1979) we believe the observation that the disturbance is largest in the B component and is strongly localised in space (Fig. 5a, b) to be best described by the guided poloidal mode (Orr and Matthew 1971; Radoski 1967). However, as Chen and Hasegawa (1974) pointed out, this mode does not have a resonant coupling with, for example, a surface wave at the magnetopause. This also shows the restricted applicability of the fieldline resonance theory. Only the sense of rotation of the horizontal disturbance vector on both sides of the demarcation line is as predicted by the model. For a westward travelling wave we find clockwise rotation north of the resonance and counterclockwise rotation south of it (Southwood 1974; Fig. 8 of the present paper).

There is another possible mode in the magnetospheric plasma having fieldline-resonance-like behaviour (Southwood 1977). For the ring-current plasma, Southwood showed the existence of a compressional mode which is strongly localised. The disturbance vector of this mode is mainly in the meridional direction if phase propagation occurs in E-W direction. Considering the ionospheric rotation of a magnetospheric signal as

described by Inoue (1973) and Hughes (1974), the observed larger amplitudes of the B component as compared to the A component, combined with a purely westward phase propagation, and the strong localisation in space are in agreement with the properties of the Southwood mode. The strong localisation of this mode is an effect of the magnetic and pressure changes being in counterphase. Just this anticorrelation was found observationally by Hughes et al. (1979) for pg-like disturbances seen by ATS 6.

The strong localisation in E-W direction is of some importance for statistical studies of pgs. For example, an analysis of the occurrence of pgs and of the distribution of the pulsation period should not be done in relation to local time of a magnetometer station (Harang 1936; Green 1979) but in relation to local time of the pg center. To find this center a dense magnetometer array is necessary. The localisation may also explain the different results of Annexstad and Wilson (1968) and Green (1979) concerning conjugate behaviour. While the events that Green (1979) analysed were of small amplitude and showed odd-mode characteristics, Annexstad and Wilson (1968) reported on pgs with amplitudes up to 40 nT as seen also for the event reported in this paper, and even-mode characteristics. Therefore the present author believes that Green (1979) measured the conjugage behaviour somewhere in the periphery of the pg region where properties may be different from those in the center region.

In recent years apparent azimuthal wave numbers have been derived from observations by several workers (Green 1976; Olson and Rostoker 1978; Mier-Jedrzejowicz and Southwood 1979). From a study of a large number of pc 4–5 pulsation events Olson and Rostoker (1978) derived the following relation between the apparent azimuthal wave number, m, and the frequency, f, of a pulsation event: $m=(1.4\pm0.4)\cdot f+0.26$, where f is measured in mHz. The apparent azimuthal wave numbers as measured for the present event (see Table 1) are in accord with the above-mentioned relation which shows that there is no significant difference between the wave numbers of normal pc 4–5 pulsations and the giant pulsation analysed in this paper. A corresponding result was found recently by Rostoker et al. (1979) using observations of four pg events.

Concerning the meaning of apparent azimuthal wave numbers, Green (1976) believes them to give us information about magnetospheric properties of a pulsation event. An apparent azimuthal wave number, $m\sim20$ for example, corresponds to an apparent horizontal phase velocity on the ground $v_{ph}\sim8$ km/s, at auroral latitudes. Phase velocities of the same order of magnitude have also been derived theoretically for phase propagation in the ionosphere, perpendicular to the earth's magnetic field, of waves in the extraordinary mode, for example (Prince and Bostick 1964). As Hughes (1974) showed, the behaviour of a magnetic disturbance in the atmosphere is quasi-static and the phases on the ground are a direct image of the phases in the ionosphere. We therefore suppose apparent azimuthal wave numbers, as measured in the present work, and associated apparent horizontal phase velocities to be controlled by ionospheric rather than by magnetospheric conditions.

Equivalent current systems (Fig. 11) have earlier been used by Obertz and Raspopov (1968) to explain, for example, the phase differences in the N-S component of pc 5 pulsations on an N-S profile. From Fig. 11 we see that the equivalent currents of the pg analysed may be interpreted as a system of moving current vortices as proposed by Obertz and Raspopov (1968) (see Fig. 12). If the periodic change of the disturbance vectors in time is explained by a moving periodic structure in space then, for a period of 100 s and a longitudinal extent of a single current vortex of about 10° (see above), one arrives at a velocity of 8 km/s for the motion of the periodic current system. The measured velocity is (8.3 ± 2.7) km/s.

A system of current vortices as shown in Fig. 12 may be generated by moving field-aligned currents. As Fukushima (1976) shows in an ionosphere with laterally homogeneous electric conductivity, the equivalent ionospheric currents of an incident field-aligned current and the Pedersen current cancel each other in their magnetic effect at ground, so that one only measures the curl-like Hall currents at the earth's surface. Whether such moving field aligned currents periodically incoming or outflowing in E-W direction, are realistic needs more investigation by means of magnetospheric and ionospheric measurements. However, there is still one argument which restricts the applicability of a current model as proposed by Obertz and Raspopov (1968) to pgs: From Fig. 3 we know that the event lasted about 30 periods in time which means a longitudinal extent of the whole current system of about 300°. This seems to be in contrast to the observed localisation.

Conclusions

One of the main results of the present paper is the observation of a westward drift of the disturbance center in the A component. This observation is in contrast to the work of Eleman (1967) and shows the advantage of a dense two-dimensional magnetometer array in pulsation studies.

The evidence for a bounce-resonance instability as a possible mechanism for explaining the drift and also the measured increase in period is suggestive but not categorial. For further decision one needs more information about the magnetospheric plasma distribution during a pg event. To clarify the result that the drift velocity of the disturbance center in the A component agrees roughly with the drift velocity of the resonant protons, one needs a more thorough analysis of the dynamical behaviour of the proposed instability, as has been done in the case of the ion-cyclotron instability by Jacobs and Watanabe (1966) to see whether it is a convective or a non-convective one.

As Southwood et al. (1969) pointed out, a bounce resonance, for example, requires an even-mode oscillation of the fieldline and therefore the conjugate behaviour of giant pulsations is of some interest. Because of the different results of Green's (1979) and Annexstad and Wilson's (1968) analysis of conjugate behaviour of pgs this question must be treated again with magnetometer stations showing better conjugacy and lying under or very close to the pulsation center and not somewhere at the periphery of the pg region.

The observation of a moving pulsation region may be of some interest regarding the question of the influence of the ionosphere on a magnetospheric signal. The motion of the magnetospheric radiation source relative to the ionosphere may explain the different behaviour of the two horizontal components. In future analyses a separation of both the drifting and stationary parts of the observed signal should be studied to get more insight into the dynamical properties of the pulsation center.

Localised compressional waves in the ring current region as described by Southwood (1977) show some features observed with the present pg event, namely the strong localisation and the predominance of the B component at the resonance region.

More information about the spatial polarisation characteristics of this mode is needed for a detailed comparison. For example, it would be interesting to know about the existence of a meridional separator line within this mode as was observed in the polarisation pattern of the present event (Figs. 7 and 8).

Acknowledgements. I am greatly indebted to those past and present members of the magnetometer group at the University of Münster, who were involved in collecting the magnetic data. The observations were performed in cooperation with Aarhus University, the Department of Plasma Physics of the Royal Institute of Technology at Stockholm, the Finnish Meteorological Institute at Helsinki, the University of Bergen, the Geophysical Observatory Sodankylä of the Finnish Academy of Science and Letters, the Kiruna Geophysical Institute, the University of Oulu and the University of Tromsö. Thanks are due to these institutions for their permanent support. Dr. Loginov, Polar Geophysical Institute at Murmansk, made available magnetic data from the Observatory at Lovozero. I am also grateful to the pulsation study group of the University of Göttingen, especially Professor M. Siebert and to Drs. C.A. Green, Edinburgh, and W.J. Hughes, Boston, for helpful discussions. For critical and helpful comments during preparation of this paper I thank Professor J. Untiedt, Mr. W. Baumjohann, and Dr. A.G. Jones. Finally, for their encouragement throughout all stages of this work I thank Dr. F. Küppers and H. Jödicke.

The magnetometer array observations were supported by grants from the Deutsche Forschungsgemeinschaft.

References

Annexstad, J.O., Wilson, C.R.: Characteristics of pg micropulsations at conjugate points. J. Geophys. Res. **73**, 1805–1818, 1968

Born, M., Wolff, E.: Principles of optics. 5th edn. Oxford: Pergamon Press 1975

Brice, N.M.: Bulk motion in the magnetosphere. J. Geophys. Res. **72**, 5193–5211, 1967

Chen, L., Hasegawa, A.: A theory of long-period magnetic pulsations, 1. Steady state excitation of fieldline resonance. J. Geophys. Res. **79**, 1024–1032, 1974

Dungey, J.W.: Effects of electromagnetic perturbations on particles trapped in the radiation belts. Space Sci. Rev. **4**, 199–222, 1964

Eleman, F.: Studies of giant pulsations, continuous pulsations and pulsation trains in the geomagnetic field. Ark. Geofys. **5**, 231–282, 1967

Farnbach, J.S.: The complex envelope in seismic signal analysis. Bull. Seismol. Soc. Am. **65**, 951–962, 1975

Fischer, F.A.: Einführung in die statistische Übertragungstheorie. Mannheim, Zürich: Bibliographisches Institut 1969

Fukushima, N.: Ground magnetic effect of field-aligned currents connected with ionospheric currents. Geophysics Research Laboratory, Univ. Tokio, UT-GRL-76-01, 1976

Gabor, D.: Theory of communication. J. Inst. Elec. Eng. **93**, 429–457, 1946

Gough, D.I., Reitzel, J.S.: A portable three-component magnetic variometer. J. Geomagn. Geoelectr. **19**, 203–215, 1967

Green, C.A.: The longitudinal phase variation of midlatitude pc 3–4 micropulsations. Planet. Space Sci. **24**, 79–85, 1976

Green, C.A.: Meridional characteristics of a pc 4 micropulsation event in the plasmasphere. Planet. Space Sci. **26**, 955–967, 1978

Green, C.A.: Observations of pg pulsations in the northern auroral zone and at lower latitude conjugate regions. Planet. Space Sci. **27**, 63–77, 1979

Gustafsson, G.: A revised corrected geomagnetic coordinate system. Ark. Geofys. **5**, 595–617, 1970

Hamlin, D.A., Karplus, R., Vik, R.C., Watson, K.M.: Mirror and azimuthal drift frequencies for geomagnetically trapped particles. J. Geophys. Res. **66**, 1–4, 1961

Harang, L.: Oscillations and vibrations in the magnetic records at high-latitude stations. Terr. Magn. Atmos. Electr. **41**, 329–336, 1936

Heacock, R.R., Henderson, D.J., Reid, J.S., Kivinen, M.: Type IPDP pulsation events in the late evening-midnight sector. J. Geophys. Res. **81**, 273–280, 1976

Hillebrand, O.: Registrierung von Riesenpulsationen in Skandinavien und Analyse der Daten. Diss. Math.-Nat. Fak. Univ. Göttingen, 1974

Hillebrand, O.: Spatial characteristics of giant pulsations. J. Geophys. **42**, 257–269, 1976

Hughes, W.J.: The effect of the atmosphere and ionosphere on long-period magnetospheric microspulsations. Planet. Space Sci. **22**, 1157–1172, 1974

Hughes, W.J., McPherron, R.D., Barfield, J.N., Mauk, B.H.: A compressional pc 4 pulsation observed by three satellites in geostationary orbit near local midnight. Planet. Space Sci. **27**, 821–840, 1979

Inoue, Y.: Wave polarisations of geomagnetic pulsations observed in high latitudes on the earth's surface. J. Geophys. Res. **78**, 2959–2976, 1973

Jacobs, J.A.: Geomagnetic micropulsations. Berlin, Heidelberg, New York: Springer 1970

Jacobs, J.A., Sinno, K.: World-wide characteristics of geomagnetic micropulsations. Geophys. J.R. Astron. Soc. **3**, 333–353, 1960

Jacobs, J.A., Watanabe, T.: Amplification of hydromagnetic waves in the magnetosphere by a cyclotron instability process with applications to the theory of hydromagnetic whistlers. J. Atmos. Terr. Phys. **28**, 235–253, 1966

Jones, A.G., Olafsdottir, B.: Geomagnetic induction studies in northern Scandinavia (Abstract). Geophys. J.R. Astron. Soc. **57**, 265, 1979

Kodera, K., Gendrin, R., de Villedary, C.: Complex representation of a polarised signal and its application to the analysis of ULF-waves. J. Geophys. Res. **82**, 1245–1255, 1977

Küppers, F., Post, H.: A second generation Gough-Reitzel magnetometer. J. Geomagn. Geoelectr. in press, 1980

Küppers, F., Untiedt, J., Baumjohann, W., Lange, K., Jones, A.G.: A two-dimensional magnetometer array for ground-based observations of auroral zone electric currents during the International Magnetometer Study (IMS). J. Geophys. **46**, 429–450, 1979

Mier-Jedrzejowicz, W.A.C., Southwood, D.J.: The east-west structure of mid-latitude geomagnetic pulsations in the 8–25 mHz band. Planet. Space Sci. **27**, 617–630, 1979

Obayashi, T., Jacobs, J.A.: Geomagnetic pulsations and the earth's outer atmosphere. Geophys. J.R. Astron. Soc. **1**, 53–63, 1958

Obertz, P., Raspopov, O.M.: Study of the spatial characteristics of type pc 5 geomagnetic pulsations. Geomagn. Aeron. (USSR) **8**, 424–427, 1968

Ol', A.L.: Long periodic gigantic geomagnetic field pulsations. Geomagn. Aeron. (USSR) **3**, 90–95, 1963

Orr, D., Matthew, J.A.D.: The variation of geomagnetic

micropulsation periods with latitude and the plasma-pause. Planet. Space Sci. **19**, 897–905, 1971

Olson, J.V., Rostoker, G.: Longitudinal phase variations of pc 4–5 pulsations. J. Geophys. Res. **83**, 2481–2488, 1978

Prince, C.E., Bostick, F.X.: Ionospheric transmission of transeversely propagated plane waves at micro-pulsation frequencies and theoretical power spectrums. J. Geophys. Res. **69**, 3213–3234, 1964

Radoski, H.R.: Highly asymmetric MHD resonances: The guided poloidal mode. J. Geophys. Res. **72**, 4026–4027, 1967

Rolf, B.: Giant micropulsations at Abisko. Terr. Magn. Atmos. Electr. **36**, 9–14, 1931

Rostoker, G., Lam, H.-L., Olson, J.V.: Pc 4 giant pulsations in the morning sector. J. Geophys. Res. **84**, 5153–5166, 1979

Samson, J.C., Jacobs, J.A., Rostoker, G.: Latitude-dependent characteristics of high-latitude pc 4 and pc 5 micropulsations. J. Geophys. Res. **76**, 3675–3683, 1971

Schmucker, U.: Auswerteverfahren Göttingen. In: Elektromagnetische Tiefenforschung, Kolloquium Neustadt/Weinstraße, pp. 163–188. Hrsg.: Haak, V., Institut für Geophysikalische Wissenschaften, Berlin und Homilius, J., Niedersächsisches Landesamt für Bodenforschung Hannover, 1978

Schulz, M., Lanzerotti, L.J.: Particle diffusion in the radiation belts. Berlin, Heidelberg, New York: Springer-Verlag 1974

Southwood, D.J.: Some features of fieldline resonances in the magnetosphere. Planet. Space Sci. **22**, 483–491, 1974

Southwood, D.J.: A general approach to low-frequency instability in the ring current plasma. J. Geophys. Res. **81**, 3340–3348, 1976

Southwood, D.J.: Localised compressional hydromagnetic waves in the magnetospheric ring current. Planet. Space Sci. **25**, 549–554, 1977

Southwood, D.J., Dungey, J.W., Ethington, R.J.: Bounce resonant interaction between pulsations and trapped particles. Planet. Space Sci. **17**, 349–361, 1969

Walker, A.D.M., Greenwald, R.A., Stuart, W.F., Green, C.A.: Stare auroral radar observations of pc 5 geomagnetic pulsations. J. Geophys. Res. **84**, 3373–3388, 1979

Wilson, C.R.: Conjugate three-dimensional polarisation of high-latitude micropulsations from a hydromagnetic wave-ionospheric current model. J. Geophys. Res. **71**, 3233–3249, 1966

Received February 22, 1980; Revised Version April 8, 1980
Accepted April 11, 1980

J. Geophys. 48, 139–147, 1980

Journal of Geophysics

Palaeomagnetic Secular Variation Curves Extending Back to 13,400 years B.P. Recorded by Sediments Deposited in Lac de Joux, Switzerland

Comparison with U.K. Records

K.M. Creer[1], T.E. Hogg[1,3], P.W. Readman[1,4], and C. Reynaud[2]

1 Department of Geophysics, University of Edinburgh, Edinburgh EH9 3JZ, Scotland
2 Département de Géologie et de Paléontologie, Université de Genève, Switzerland
3 Now at Imperial Metals Industries, Birmingham, England
4 Now at Geophysics Section, Dublin Institute for Advanced Studies, Ireland

Abstract. Palaeosecular variation records have been obtained from three cores from Lac de Joux, Switzerland (46°37′N, 6°16′E) on which a parallel biostratigraphy has been constructed. All three cores show a well developed Late-glacial sequence which extends to beyond 13,400 years B.P. and the recorded palaeomagnetic declination and inclination variations can be correlated between these cores. The longer core (no. 3) also shows a well developed Post-glacial sequence and the declination and inclination records can be correlated in detail with the independently dated United Kingdom records which extend back to about 10,000 years B.P. Ages along Lac de Joux core 3, obtained by palaeomagnetic correlation with U.K. cores, are compared with ages based on the local palynology and thus a timescale has been attached to the Lac de Joux record back to beyond 13,400 years B.P. during which time there is no evidence of any geomagnetic excursion or short event.

Key words: European geomagnetic secular variation – Lake sediments – Palaeomagnetism – Palynology.

1. Introduction

In the summer of 1975 several lakes in the region of the Alps were cored as part of a research project to investigate the behaviour of the geomagnetic field through the Holocene. The field work was carried out by a team from the Department of Geophysics, University of Edinburgh, and from the Department of Mineralogy, University of Geneva. This paper deals with results obtained from Lac de Joux.

Lac de Joux is situated just to the north of Lake Geneva at an altitude of 1,014 m and lies between latitudes of 46°41′N and 46°38′N and longitudes 6°17′E and 6°20′E. A map of the lake, which has only two small rivers flowing into it, is shown in Fig. 1. Four cores were taken using a pneumatically controlled Mackereth corer (Mackereth 1958). The bottom sediments were very hard and the longest core (no. 3) was 5.5 m long but the average length was only 4 m due to the hardness of the sediment penetrated.

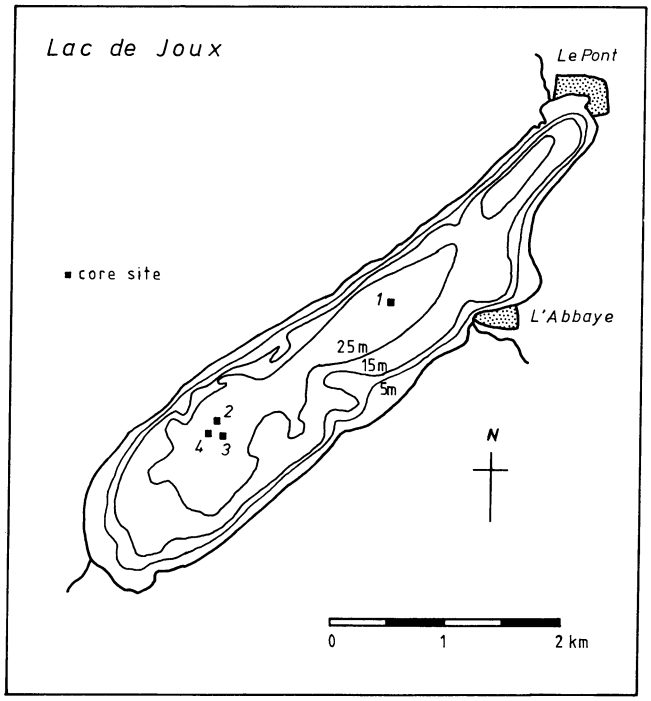


Fig. 1. Sketch map of Lac de Joux (46.6° N, 6.3° E) showing coring sites. Contours show water depths in the lake

The cores were transported to Edinburgh for study. They were split in half lengthwise and samples were withdrawn for palaeomagnetic study in square shaped polystyrene boxes of side 17 mm, at about 2 cm intervals, from one of the halves of each core. The other halves were sent to Geneva for palynological study.

2. Pollen Studies

Pollen analyses were carried out in the limnogeological laboratory of the University of Geneva and supported financially by the

0340-062X/80/0139/$01.80

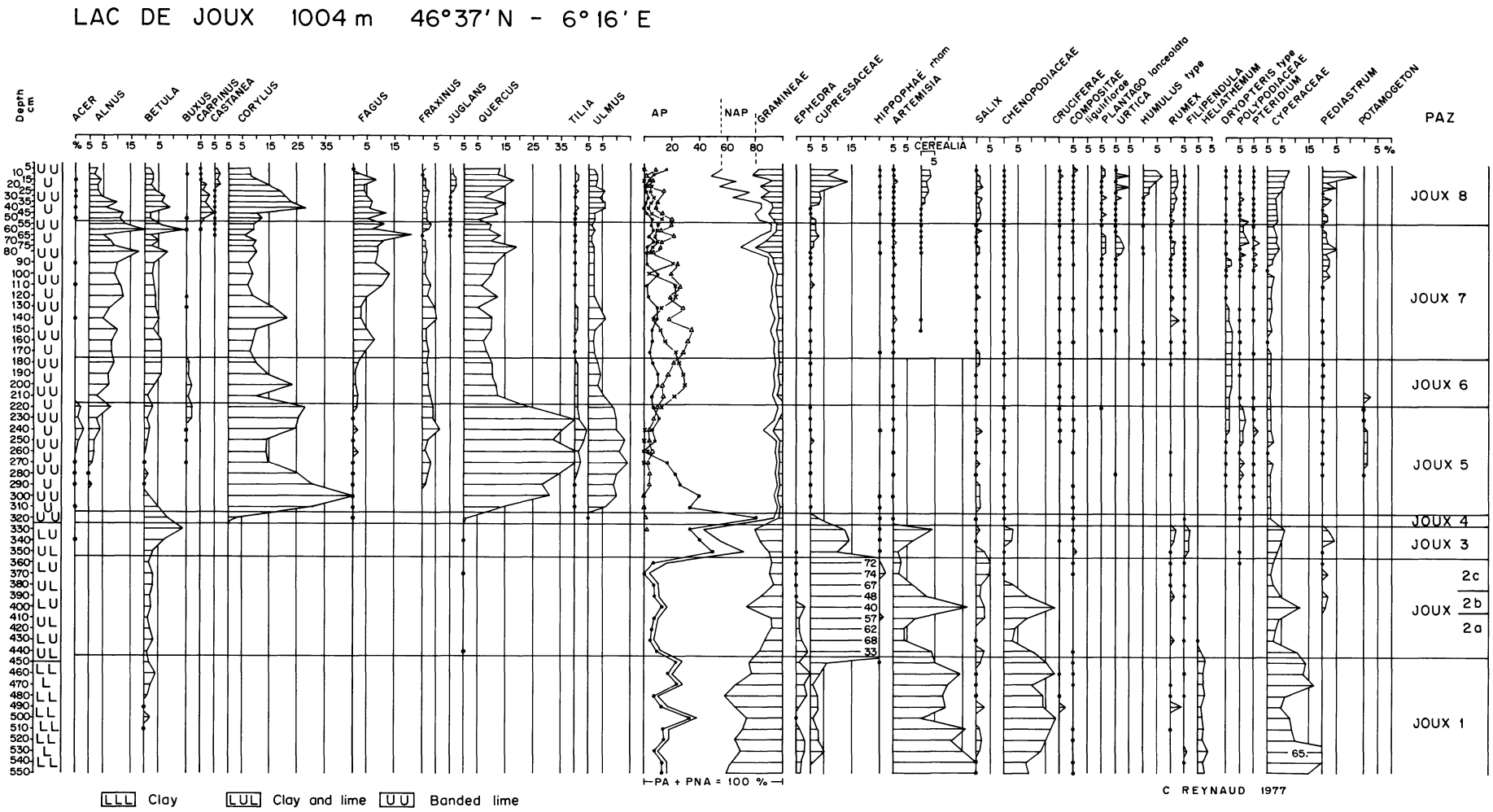


Fig. 2. Pollen diagrams for Lac de Joux core 3 presenting the most important taxa. The biozones Joux 1–Joux 8 are described in Sect. 2.1. *AP* = arboreal pollen; *NAP* = non-arboreal pollen

'Fonds national de la Récherche scientifique' (Switzerland). Laboratory treatment followed the recommendation of Faegri and Iversen (1964). One cm^3 of fresh sediment was treated with concentrated HCl and 40% HF and an acetolysis was performed on the Post-glacial sediments with rich organic content. In core no. 3, the sediments analysed consisted of clay up to 450 cm, of clay and calcaereous mud up to 325 cm and of banded lime in the upper part of the core which, in total, was 550 cm long. The results are presented (Fig. 2) in the form of conventional relative pollen diagrams in which the total pollen arboreal (AP) plus non-arboreal (NAP) = 100%. Cyperaceae, spores, and other microfossils have been excluded from the NAP and expressed as percentages of the pollen sum.

2.1. Local Biozone Descriptions (Fig. 2)

Joux 1: Herbs Zone, 550–445 cm. The relative AP values are very low: less than 30% of the pollen sum. Among the herbs the dominating taxa are Gramineae, *Artemisia*, Chenopodiaceae, *Helianthemum, Thalictrum*, and *Plantago alpina*. Shrubs are also present with *Ephedra, Salix* and *Juniperus*, each being less than 5%.

Joux 2a: Juniperus Zone, Juniperus Sub-zone, 445–405 cm. The AP values are still low and are represented as in Joux 1 by pine (20%) and birch (5%). *Juniperus* dominates all the other taxa with high values: between 33 and 74%. A few *Hippophaë rhamnoides* are also present.

Joux 2b: Juniperus Zone, Juniper-Herbs Sub-zone, 405–385 cm. Juniper values are lowered to less than 50% and there is a sudden increase of steppic elements such as *Artemisia* and Chenopodiaceae.

Joux 2c: Juniperus Zone, Juniper-Pine Sub-zone, 385–355 cm. Juniper first and then pine increase again. The AP values are much higher, up to 70% of the total pollen sum.

Joux 3: Pine-Herbs Zone, 355–325 cm. The sudden relative increase of pine (to 50%) is depressed by the rising values of the herbs: Gramineae, *Artemisia*, Chenopodiaceae, *Helianthemum, thalictrum*, and *Plantago alpina*. In the level 340 cm many spores of fungi were found.

Joux 4: Pine-Birch Zone, 325–315 cm. All the herbs are depressed to values of less than 5%. Birch and pine (80%) are the main components of the AP curve.

Joux 5: Oak-Hazel-Elm Zone, 315–215 cm. The thermophilous trees appear suddenly and are dominated by oak. Pine decreases to less than 10%.

Joux 6: Fir-Spruce Zone, 215–175 cm. Spruce present in the former biozone increases gently but fir dominates while oak and elm register a corresponding decline. Low percentages of beech (1%–3%) are present.

Joux 7: Spruce-Fir-Beech Zone, 175–55 cm. Fir decreases while spruce and beech are the dominating species. Among the components of the *Quercetum mixtum* only oak and hazel are important.

Joux 8: Spruce-Oak-Herbs Zone, 55–5 cm. The total AP shows a decline while herbs and juniper increase to 40%. In this biozone consistent amounts of common walnut, hornbeams, chestnut, and cerealia are present.

2.2. Discussion

According to Wegmüller (1966) the pollen stratigraphy throughout the southeastern part of the Jura mountains starts with a Late Würmian herbs zone (Fig. 3). It is followed by a shrubs zone dominated by *Juniperus* with some *Hippophaë rhamnoides* present. It suggests a treeless landscape although some immigrating and isolated birches might have been present close to the area investigated.

Later on, about 12,300 years B.P., park-forest invaded the Jura and this is clearly indicated by the birch zone. Birch was then replaced by pine and it is in the sediments bearing pine pollen assemblage that ashes emitted by the Laacher Volcano are found; it is dated at 10,980 years B.P. (Wegmüller and Welten 1973).

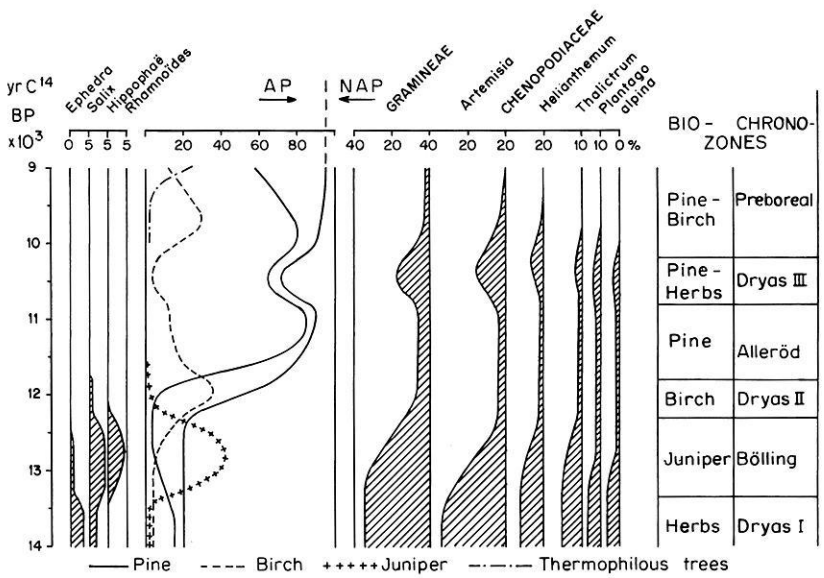


Fig. 3. Schematic diagram for the Late-glacial biostratigraphy in the Jura mountains

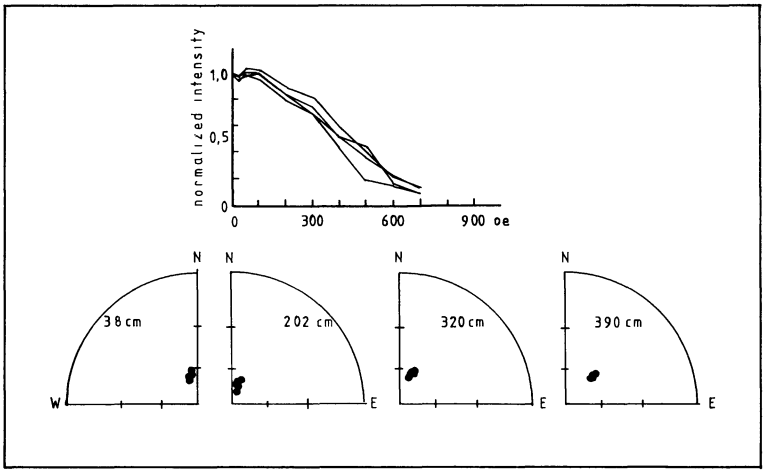


Fig. 4. Alternating field demagnetization plots of pilot samples from core 3. Depths of samples in cm from the top of the core are indicated

At about 10,800 years B.P., herbs increase at the expense of pine and this is interpreted as a climatic deterioration during youngest Dryas (DR III). The Late Würmian/Early Flandrian transition can be distinguished by the re-advance of pine and birch forest and the appearance of thermophilous trees.

The schematic pollen diagram in Fig. 3 is based on the results of Wegmüller (1966) from the Cruilles bog (1,035 m) situated alongside Lac de Joux and on the pollen investigations of Matthey (1971) in the central part of the Juras. When the Würmian Late-glacial sequence of Lac de Joux is compared with Fig. 3 it is clear that both birch and pine zones, ascribed classically to the Older-Dryas-Alleröd complex, are missing in the Lac de Joux core 4.

There is also a hiatus in the biostratigraphy at around 330 cm marked by a very short flip in the pine curve and the sudden appearance of the thermophilous pollen. This may be linked to the rather complicated process of ice retreat in the Joux valley as described by Aubert (1937). This gap in the sedimentation possibly represents a time span of about 1,000 years.

The Post-glacial vegetational development in the Joux valley described above is in agreement with the results of Wegmüller (1966). After a short appearance of the mixed oak forest close to the site, spruce and fir developed around the lake. Land clearance partly destroyed the natural forest (from 75 cm upwards) and this can be dated to the Bronze Age. The progressive thinning of the forest is marked by the relative increase of both herbs and juniper pollen and the abundance of *Cerealia* pollen.

3. Palaeomagnetic Results

3.1 Method

Natural remanent magnetizations (NRM) were measured with a Digico fluxgate magnetometer (Molyneux 1971) and susceptibilities were measured with a Digico bridge. First, pilot samples were progressively demagnetized in alternating magnetic fields (Creer 1958, 1959). Demagnetization characteristics are illustrated for four typical samples in Fig. 4 which shows that a weak viscous component was removed during the first two demagnetization steps of 30 and 60 Oe. The directions of RM of all pilot samples showed very little change on cleaning in alternating fields of up to 700 Oe (peak) and the median destructive field was typically between 400 and 450 Oe. The viscous component could alternatively be removed by storage in zero field for several months, and this procedure was adopted before measuring the NRM of the samples withdrawn from the cores.

3.2. Susceptibility and NRM Intensities Along Core No. 3

The palaeomagnetic logs are shown in Fig. 5. We discuss first the intensity and susceptibility logs. Susceptibility shows an overall correlation with the colour of the mud. Lowest values (< 5 μG/Oe) are encountered in a white Post-glacial mud between 2 m and 3 m depth most of which has been placed in the Joux 5 pollen zone (Sect. 2) Susceptibilities lie mainly in the range 5–10 μG/Oe and attain maximum values in the Joux 4 pollen zone part of the brown mud, i.e., at the base of the Post-glacial. The lower part of this brown mud unit is Late-glacial (Joux 3 pollen zone). There is a sharp discontinuity in susceptibility and an even greater one in intensity at the Post-glacial/Late-glacial boundary (Joux 4/Joux 3).

The shapes of the intensity and susceptibility logs are similar through the cream Late-glacial clay where the Q-ratios (intensity/susceptibility) are low, of the order of unity, but through the Post-glacial sediments where the Q-ratios are high, the shapes of the logs are quite different. In fact, in the banded lime of the Joux 5 pollen zone, Q-ratios attain values of 20–50 suggesting that the magnetic fraction of the detrital grains must have been rather more efficiently aligned than in most lake sediments we have previously studied.

3.3. Directional Variations Along Core No. 3

The declination log recovered from the Post-glacial sediments show several oscillations of about 15° amplitude. In the Late-glacial sediments the oscillations are of larger amplitude and are systematically displaced to the east relative to the Post-glacial ones.

The inclination log, below about 160 cm depth is characterized by four pronounced features, rounded on the high inclination sides with cusps on the low inclination sides. The top part of the inclination log shows less regular variations. In particular, low values between about 90 and 105 cm in the middle of feature ε, separating it into two smaller maxima, labelled ε1 and ε3, are

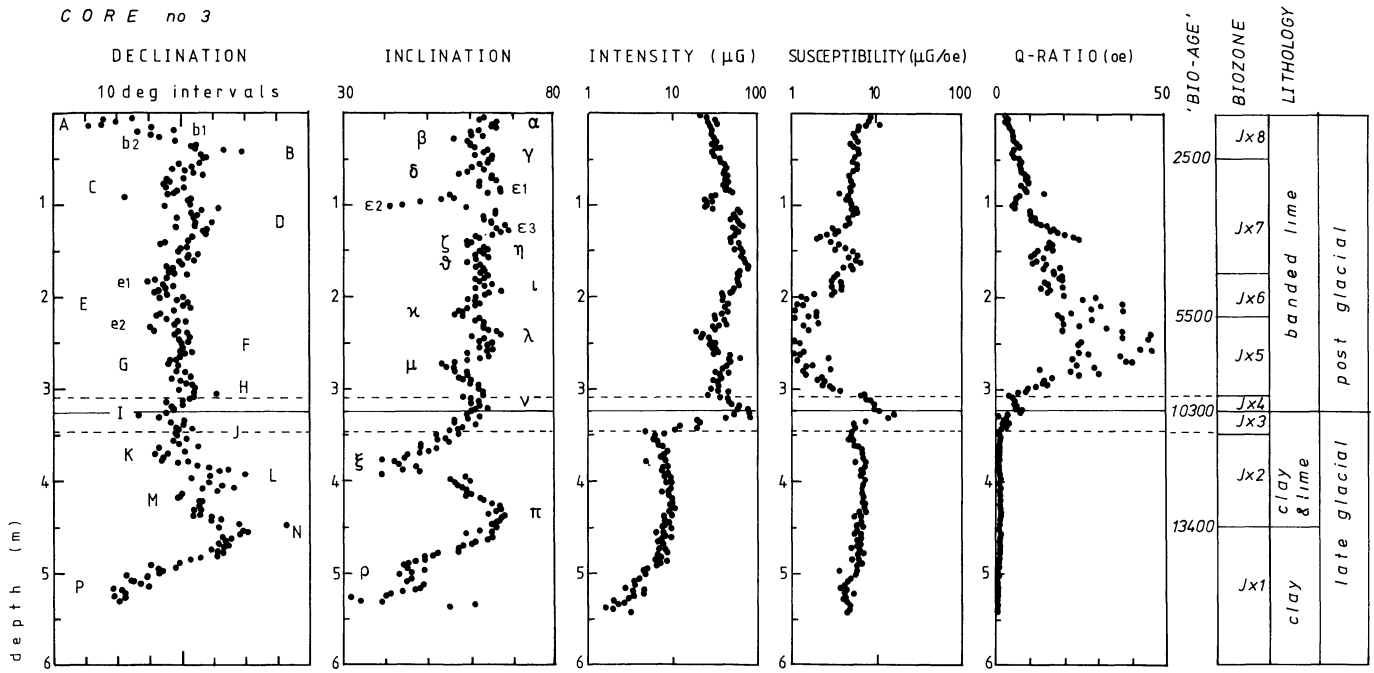


Fig. 5. Declination (D), inclination (I), remanence intensity (J), susceptibility (k) and Q ratio (J/K) logs for core 3, Lac de Joux. Units of J are μG, of k, μG/Oe and of Q, Oe. Declination and inclination features are labelled A–P and α–ρ respectively. Lithological and biozone boundaries shown at right of figure. *Solid line* running across plots represents Post-glacial/Late-glacial boundary. *Broken lines* represent upper limit of Joux 4 and lower limit of Joux 3 biozones respectively

accompanied by low intensities of NRM. One possible explanation is that the sediments at this level have been physically disturbed causing partial disorientation of the magnetic grains since corresponding susceptibility values are similar to those measured just above and below. However, the declinations accompanying the scattered low inclinations are not themselves particularly scattered as they should be if the sediment had in fact been disturbed so that the band of low inclinations recorded in the sediment could possibly reflect low geomagnetic inclinations.

Both logs can be correlated well with United Kingdom (U.K.) palaeomagnetic declination and inclination logs from Lake Windermere (Mackereth 1971; Creer et al. 1972; Thompson and Turner 1979) and with similar, though shorter logs from Poland (Creer et al. 1979). The labels attached to the logs (Fig. 5) illustrate this correlation: the declination labels are the same as previously used for the U.K. logs (Creer et al. 1976b, 1979), and the inclination labels are the same as used for the U.K. logs (Thompson and Turner 1979) and for the Polish logs (Creer et al. 1979). The Lac de Joux record from core 3 extends the European secular variation records by some 3,500 years into the Late-glacial.

3.4. Results from Core No. 4

Palynological studies on the samples taken along this rather short core (3.25 m) showed the Late-glacial/Post-glacial boundary to be at 1.5 m depth. The Late-glacial unit consisted of cream clay and carbonate. The Post-glacial deposits consisted of pale grey clay with some varve like stratifications, with carbonate above 1.25 m. The uppermost 12 cm consisted of brown clay carbonate with organic matter.

The palaeomagnetic logs are shown in Fig. 6. The top of the Late-glacial unit is marked by a sharp increase in intensity and by a marked increase in Q-ratio as in core no. 3.

Turning now to the directional logs, starting from the Post-glacial/Late-glacial boundary as marker horizon we can identify in core 4 (Fig. 6) declination features *H–N* and inclination features ν–π which were seen in the logs from core 3. Directions are only poorly recorded in the Post-glacial unit which seems to be rather thin and disturbed at this coring site.

3.5. Results from Core No. 2

Palynological study of seven samples from this 3.3 m long core show that it consists of a very well developed Late-glacial sequence of cream carbonate and clay overlain by a very short (27 cm) Post-glacial unit consisting of light yellow carbonate-clay, yellow-banded lime and brown clay-carbonate with organic matter.

The declination and inclination records carried by core 2 (Fig. 7) are not so well defined as for cores 3 and 4, particularly so between about 50 and 120 cm depth. Even so, a correlation can be made with cores 3 and 4 and we conclude that core 2 has penetrated into slightly older Late-glacial clay than either of the other two cores, additional features *P* (declination) and σ (inclination) being observed.

4. Comparison of Swiss and U.K. Records

Second generation lake sediment records from U.K. (Thompson and Turner 1979) show finer detail than the original records from Lake Windermere (Mackereth 1971; Creer et al. 1972). They have

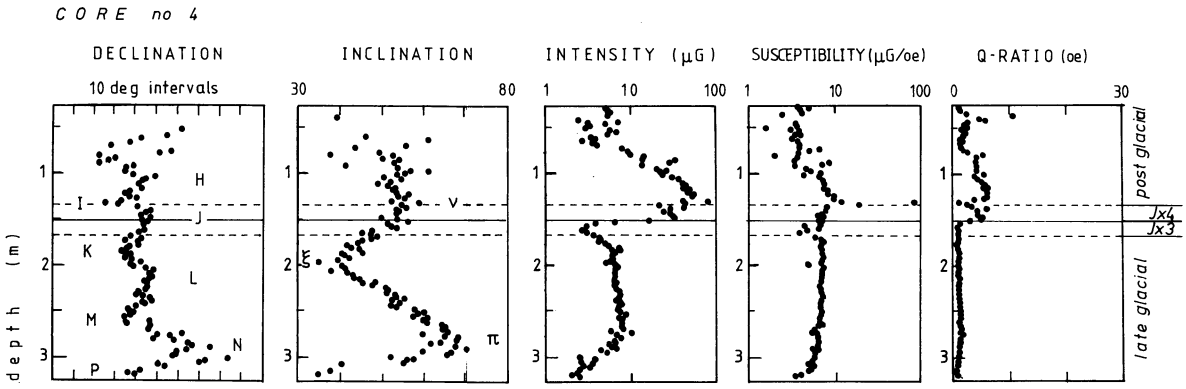


Fig. 6. Palaeomagnetic logs for Lac de Joux core 4. Declination and inclination features are labelled as for core 3. *Solid line* across plots represents Late-glacial/Post-glacial boundary. Jx3 and Jx4 represent biozones

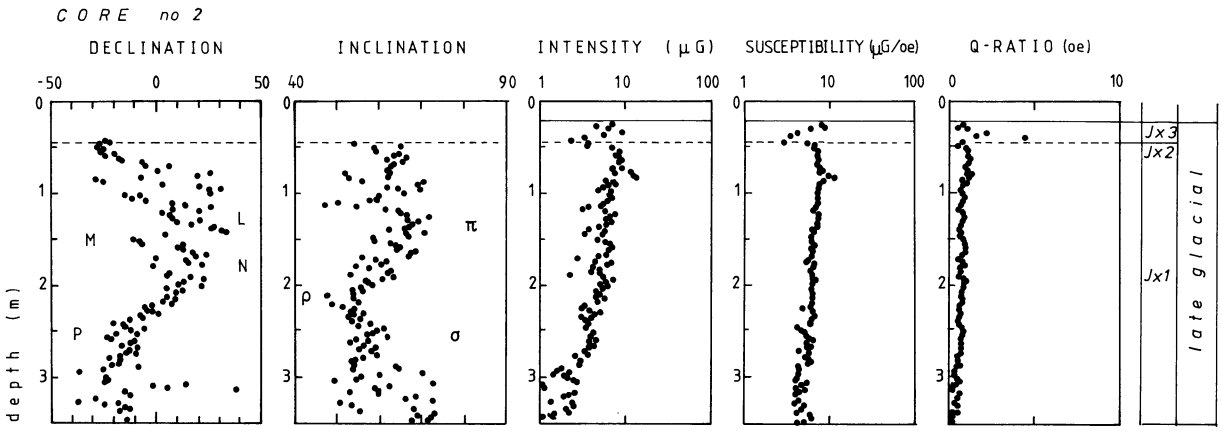


Fig. 7. Palaeomagnetic logs for core 2 Lac de Joux which mainly covers the Late-glacial. Declination and inclination features are labelled as for core 3. *Solid line* across plots represents Late-glacial/Post-glacial boundary. Jx1, Jx2, and Jx3 represent biozones

been dated by radiocarbon and have thus been shown to extend back for 10 millenia (Lake Windermere, NE England) and for 7 millenia (Loch Lomond, Scotland and Lake Geirionydd, Wales). Practical use of palaeomagnetic declination and inclination variation curves for regional age correlation is limited by variations in amplitude and detail of shape of the individual features as recorded from place to place caused by the somewhat imperfect natural magnetic recording process and by spacial variations in the temporal behaviour of the geomagnetic field. Thus we should not necessarily expect to find such good correlation between the Lac de Joux and the U.K. lakes as is observed between the U.K. lakes themselves.

We now compare the Lac de Joux record from core 3 with the record from a Lake Windermere core supplied by Dr. Turner. As noted in Sect. 3.3, most of the features shown by the Lake Windermere records can be identified in the Lac de Joux records. To illustrate the correlation better, the Lac de Joux and Lake Windermere inclination and declination records have been plotted against time rather than sediment depth in Fig. 8a and b respectively in which the time scale is based on the ages determined by Turner (1979) for the U.K. inclination features. This assumes that the geomagnetic variations occurred synchronously in the U.K. and Switzerland. In this respect, we note that westward drift at the historically observed rate of about 0.2 Oe/yr would cause a phase lag between Switzerland and U.K. of only about 30 years which is negligible compared with the precision of absolute dating.

We now compare these 'magnetic' ages, obtained for Lac de Joux core 3, with the independent set of ages based on the observed depths and published dates of biozone boundaries for the Joux Valley. The latter are illustrated in Fig. 9 by a time-depth plot consisting of five straight line portions, together with a dashed line taking account of the suspected hiatus of about 1,000 years at 330 cm depth (Sect. 2.2).

We note that there is some discrepancy between 'magnetic' and 'pollen' ages through the upper 2 m of core comprising the Joux 6, 7, and 8 biozones. While the 'magnetic' ages of the declination and inclination features indicate a steady rate of deposition of about 0.39 mm/yr throughout the whole of the last 5,500 years, the ages of 2,500 years B.P. attached to the Joux 7/8 biozone boundary implies a more rapid depositon rate of about 0.53 mm/yr through the Joux 6 and 7 biozones and a slower rate of about 0.22 mm/yr through the Joux 8 biozone. The ages of the younger declination (*A* and *B*) and inclination (α, β, γ and δ) features are controlled by observatory or archaeological data and therefore are more secure than the ages of the older features which are controlled by radiocarbon and are hence, in the case of lake sediments, subject to systematic error. However, contamination by

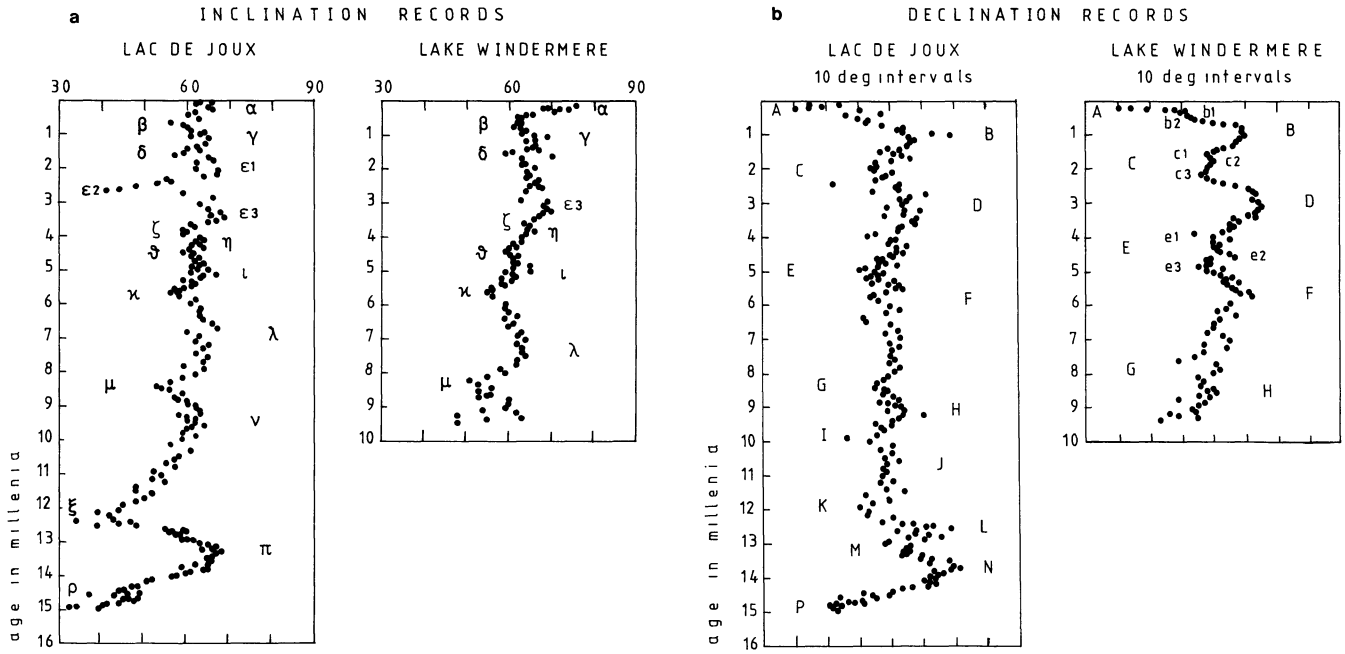


Fig. 8. a Inclination and **b** declination secular variation curves plotted as a function of time for Lac de Joux and Lake Windermere. Time scale constructed using dates of inclination features α–μ as determined for Lake Windermere (Turner 1979). Age control of Late-glacial part of Lac de Joux core is from dated palynology of the Joux valley

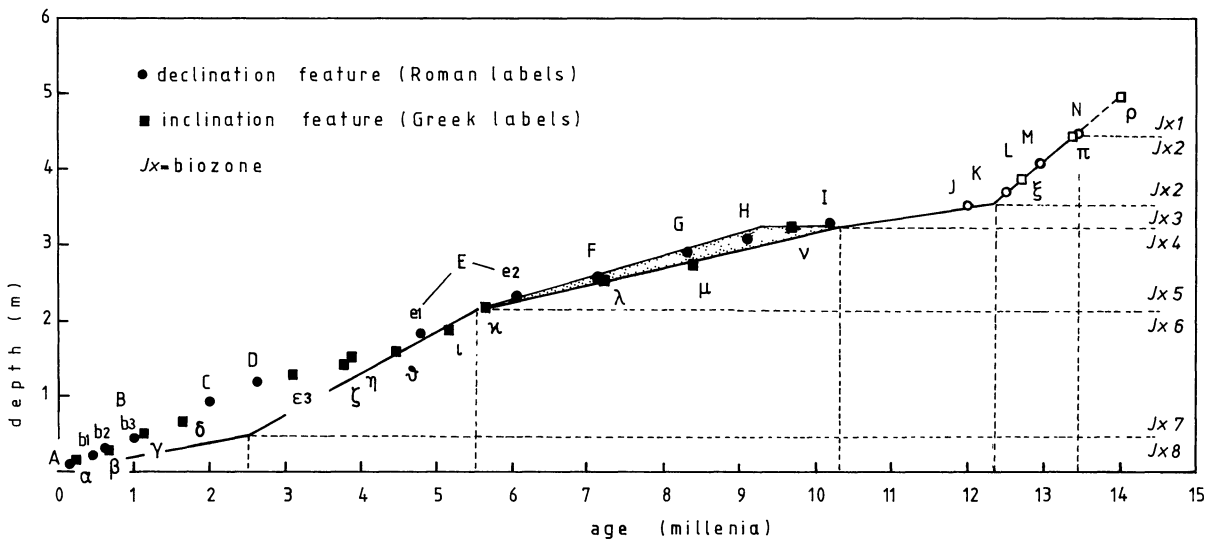


Fig. 9. Sediment accumulation (time vs depth) curve deduced for Lac de Joux core 3. The solid line is constructed from the biostratigraphy (see Figs. 2 and 3). The upper solid line running between 5,500 and 10,300 years B.P. takes into account a suspected hiatus lasting about 1,000 years at about 10,000 years B.P. *Dots* and *squares* represent 'magnetic' ages of declination and inclination features respectively, dated by correlation with U.K. curve

'old' carbon makes radiocarbon ages of lake sediments too old whereas, in this particular case, the 'magnetic' ages are younger than those indicated by the palynologically deduced time-depth curve. Thus it is difficult to account for the age discrepancy by errors in the 'magnetic' ages.

Turning now to sediments of the Joux 5 biozone, it appears that the 'magnetic' ages of the declination (*F, G, H,* and *I*) and inclination (κ, λ, μ and ν) features appear consistent with the existence of the suggested hiatus at 330 cm (Sect. 2.2), since their representative points fall within the shaded triangle bounded by the upper line which supposes a 1,000 years hiatus and the lower line which supposes no hiatus.

The dated U.K. palaeomagnetic declination and inclination curves only extend back to 10,000 years B.P. so we cannot use the depths at which the older declination features (*K–P*) and inclination features (ξ, π and ρ) occur as an independent check of the date

Table 1. Ages (in conventional radiocarbon years) of Late-glacial declination and inclination features

Feature label		Assigned age (year B.P.)
Declination	Inclination	
J		11,400
K		12,250
	ξ	12,400
L		12,600
M		13,000
	π	13,250
	ρ	14,200
N		14,250
P		14,500

Dates younger than 13,400 year B.P. assume steady deposition rate through Joux 1, 2 and 3 biozones. Older dates obtained by extrapolating this rate back in time

of 13,400 years B.P. for the Joux 1/2 biozone boundary (Sect. 2.2). Rather we use the biochronological time-depth curve to assign provisional dates to the declination and inclination features as shown in Table 1. If we extrapolate the average deposition rate of 0.39 mm/yr between 10,300 and 13,400 years B.P. for which we have biostratigraphic control back to the bottom of the core we date the bottom of the core at some 15,500 years B.P.

5. The Late-Glacial Record

The Late-glacial inclination records from Lac de Joux show two features ξ and ρ dated at 12,400 and 14,200 years B.P. respectively where values as low as 40° are reached. However, we note that no negative inclinations are encountered: there is no suggestion of a short lived geomagnetic reversal (excursion) as has been reported to be recorded by Late-glacial sediments from Sweden (Mörner et al. 1971) and from Lake Erie, N. America (Creer et al. 1976a). The former, or so-called Gothenberg excursion, dated by Mörner at 12,350 years B.P. has been attributed to a slumping of the sediment in the stratigraphic level corresponding to the Fjard interstadial by Thompson and Berglund (1976). Since the latter (Erieau) excursion is not seen in nearby records from Lake Michigan (Creer et al. 1976b), its reality as a geomagnetic excursion is highly doubtful.

6. Conclusions

The pattern of secular variations in geomagnetic direction recorded by sediments deposited in Lac de Joux shows a remarkable similarity to the pattern recorded by U.K. lake sediments through Post-glacial time. The Lac de Joux sediments are highly stable to alternating field demagnetization and it is probable that the main source of error of the palaeomagnetic logs arises from imperfections in the magnetic recording mechanism rather than remagnetization by geomagnetic fields after acquisition of the primary remanence. Physical disturbance of the sediment due to natural causes such as slumping or due to the coring process has spoilt the palaeomagnetic record at some levels: for example the inclination record in core 3 at about 1 m depth and the top metre or so of cores 2 and 4.

The estimated duration of the minor features of the declination and inclination logs which can be correlated between Switzerland and U.K. is a few hundred years (Fig. 9) and this places an upper limit on the duration of the magnetization process which is thought to be post-depositional.

The largest uncertainty in defining logs of variations in palaeomagnetic directions is in the independent dating control. Radiocarbon dates can be up to several thousand years too old (Mackereth 1971; Creer et al. 1979), so that dating a core to a precision of a few hundred years, the duration of the briefest palaeomagnetic features, is not generally possible. This means that it is not possible to deduce whether there is a phase lag between the U.K. and Swiss based records as one would expect if the field had been systematically drifting to the west throughout Post-glacial time as through the last few centuries: the phase lag expected over the 6° longitude difference would amount only to about 30 years assuming a drift rate of 0.2°/yr. The discrepancy between the magnetic ages of declination features *A–E* and inclination features α–ι with the time-depth curve based on the biostratigraphy (Fig. 9) is more likely due to errors in radiocarbon dating than to a phase difference between the geomagnetic field variations at the respective sites.

Acknowledgements. The field and laboratory work was carried out as part of the project funded by Grant GR3/2238 from the U.K. Natural Environment Research Council which also supported Dr. Readman. Dr. Hogg was in receipt of a NERC Research Studentship. The vehicle and boat were funded by a grant from the Royal Society of London. Pollen analyses were carried out in the Limnological Laboratory of the University of Geneva. We express our appreciation to Dr. J.-J. Wagner of the Petrophysics Laboratory of the University of Geneva for his help in the field and with the planning of the project which extended to several other lakes in Switzerland and France, the results of which will be published shortly.

References

Aubert, D.: Les glaciers quaternaires d'un bassin fermé: La vallee de Joux (Canton de Vaud). Bull. Soc. Vaudpise Sci. Nat. **60**, 117–130, 1937

Creer, K.M.: Preliminary palaeomagnetic measurements from South America. Ann. Géophys. **14**, 373–390, 1958

Creer, K.M.: A.C. demagnetization of unstable Triassic Keuper Marls from S.W. England. Geophys. J. R. Astron. Soc. **2**, 261–275, 1959

Creer, K.M., Anderson, T.W., Lewis, C.F.M.: Late Quaternary geomagnetic stratigraphy recorded in Lake Erie sediments. Earth Planet. Sci. Lett. **31**, 37–47, 1976a

Creer, K.M., Gross, D.L., Lineback, J.: Origin of regional magnetic variations recorded by Wisconsin and Holocene sediments from Lake Michigan, USA and Lake Windermere, England. Bull. Geol. Soc. Am. **87**, 531–540, 1976b

Creer, K.M., Hogg, T.E., Malkowski, Z., Mojski, J.E., Niedziolka-Krol, E., Readman, P.W., Tucholka, P.: Palaeomagnetism of Holocene lake sediments from north Poland. Geophys. J.R. Astron. Soc. **59**. 287–314, 1979

Creer, K.M., Mackereth, F.H., Molyneux, L., Thompson, R.: Geomagnetic secular variation recorded in the stable magnetic remanence of recent sediments. Earth Planet. Sci. Lett. **14**, 115–127, 1972

Faegri, K., Iversen, J.: Textbook of pollen analysis. p. 237. Copenhagen: Munksgaard 1964

Mackereth, F.J.H.: A portable core sampler for lake deposits, Limnol. Oceanogr. **3**, 181–191, 1958

Mackereth, F.J.H.: On variations in directions of the horizontal component of remanent magnetization in lake sediments. Earth Planet. Sci. Lett. **12**, 332–338, 1971

Matthey, F.: Contribution à l'étude de l'Évolution Tardi- et Postglaciaire de la végétation dans le Jura Central. Matériaux pour le levé géobotanique de la Suisse, Fasc. 53, p. 86. Berne: Hans Huber 1971

Molyneux, L.: A complete results magnetometer for measuring the remanent magnetization of rock. Geophys. J.R. Astron. Soc. **32**, 479–481, 1971

Mörner, N.A., Lanser, J., Hospers, J.P.: Late Weichselian palaeomagnetic reversal. Nature Phys. Sci. **234**, 173, 1971

Thompson, R., Berglund, B.: Late Weichselian geomagnetic 'reversal' as a possible example of the reinforcement syndrome. Nature **263**, 490–491, 1976

Thompson, R., Turner, G.M.: British geomagnetic master curve 10,000–0 years B.P. for dating European sediments. Geophys. Res. Lett. **6**, 249–252, 1979

Turner, G.M.: Geomagnetic investigations of some recent British sediments 204 pp. University of Edinburgh: Ph.D. Thesis, 1979

Wegmüller, S.: Über die spät- und postglaziale Vegetationsgeschichte des südwestlichen Jura. Beitr. Geobot. Landesaufnahme Schweiz **48**, 144, 1966

Wegmüller, S., Welten, M.: Spätglaziale Bimstufflagen des Laacher Vulkanismus im Gebiet der westlichen Schweiz und der Dauphiné (F). Eclogae Geol. Helv. **66/3**, 533–541, 1973

Received February 25, 1980; Revised May 5, 1980;
Accepted May 10, 1980

J. Geophys. 48, 148–152, 1980

Journal of Geophysics

Anhysteretic Remanent Magnetization of Small Multidomain Fe_3O_4 Particles Dispersed in Various Concentrations in a Non Magnetic Matrix

E. Schmidbauer and R.J. Veitch
Institut für Allgemeine und Angewandte Geophysik der Universität München, Theresienstr. 41, D-8000 München 2, Federal Republic of Germany

Abstract. The anhysteretic remanent magnetization (ARM) and hysteresis parameters of synthetic, small, multidomain particles (0.2–0.4 μ) of Fe_3O_4, dispersed at various concentrations in a non magnetic matrix (packing factors $p=0.01$–0.3), have been measured. The coercive force H_c is weakly dependent on p. At higher concentrations the shape of the sample has a strong influence on the magnetization. There appears to be an additional concentration dependence of the ARM. However, the independence of partial ARMs also holds for larger p.

Key words: Rock magnetism – Small particles – Anhysteretic remanent magnetization – Fe_3O_4.

Introduction

In recent years, in rock magnetism, a series of studies have been devoted to the problem of determining the palaeointensity of the geomagnetic field from the magnitude of the natural remanent magnetization (NRM) of rocks. In the models and methods used for this purpose it is suggested that the NRM is a thermoremanent magnetization (TRM). In order to determine palaeointensities, the method proposed by Thellier and Thellier (1959), or one based on it with minor or major modifications, is generally applied. For this procedure the samples are given a TRM in the laboratory, i.e., they have to be heated, so chemical alteration might occur.

For some years investigations of palaeointensity have been performed with the aim of replacing the laboratory TRM by the anhysteretic remanent magnetization (ARM). In order to produce an ARM, a sample is subjected to the combined action of a small, steady (D.C.) and an alternating (A.F.) magnetic field. The A.F. field initially saturates the sample during each cycle and decreases slowly to zero, with the D.C. field remaining constant. The resulting magnetization is called anhysteretic magnetization (AM). After reducing the D.C. field to zero the remanence is the ARM. Both the TRM and the ARM represent some kind of magnetic equilibrium state, although the two states may not be the same. This similarity, which may apply preferentially to single-domain particles (SDP), helps explain why both remanences show some similar properties: (1) Similar stability against A.F. demagnetization (Rimbert 1959; Levi and Merrill 1976): (2) Both are linearly proportional to a weak, inducing magnetic field (Rimbert 1959): (3) The addition law is valid for partial ARM and TRM (Patton and Fitch 1962; Dunlop and West 1969). On the other hand, there is a major difference between the states: the TRM arises predominantly at high temperatures whereas the ARM is usually produced at room temperature (Dunlop et al. 1975).

Up to 1969, the AM and the ARM were explained, in particular for SDP, in terms of static models including magnetostatic interaction between grains in the specimens (review by Wohlfarth 1964; Kneller 1968). Jaep (1969, 1971) introduced a kinetic analysis for the ARM of SDP based on the concept of fluctuation phenomena, as developed by Néel (1949). In this model the analogy between ARM and TRM becomes clear and, the attempt to replace TRM by ARM measurements is justified in principle. In this picture the initial anhysteretic susceptibility, $\chi_{AM}=\lim_{H\to 0} dI_{AM}/dH$ (I_{AM}=anhysteretic magnetization, H=steady magnetic field), remains finite even for highly diluted SDP, contrary to the old models. Using Jaep's results, Banerjee and Mellema (1974) derived formulae for the ratio ARM/TRM of SDP, including interactions between particles. Shcherbakov and Shcherbakova (1977) presented similar formulae. Kneller and Köster (1977) investigated the relation of χ_{AM} to the static magnetic parameters of SDP. For the ARM of multidomain particles (MDP), a microscopic theory like that developed for SDP is still missing. In this case the magnetization process is complicated by the presence of domain walls, whose motion is determined by factors such as lattice imperfections (e.g., stacking faults) and surface irregularities. A phenomenological theory of ARM was given by Gillingham and Stacey (1971).

In many rocks, the magnetization is carried by magnetic grains consisting of Fe_3O_4 or spinels which are close to Fe_3O_4 in composition. The Fe_3O_4 grains in rocks are frequently above single-domain (SD) size (300 Å–500 Å, Morrish and Yu 1955; Dunlop 1973; Butler and Banerjee 1975) and below multidomain (MD) size (<20 μ, Stacey 1963; Parry 1965. These grains are small MDP which show neither pure SD nor pure MD behaviour. In rock magnetism they are sometimes called pseudo-SDP. These grains can occur in quite high concentrations in oxidized titanomagnetite intergrowths (Evans and Wayman 1974; Davis and Evans 1976).

We have investigated the ARM and hysteresis properties of synthetic equidimensional Fe_3O_4 grains in the grain size range 0.2–0.4 μ, in relation to their concentration in specimens with volume fractions (packing factors) $p=0.01$–0.30.

The ARM of small Fe_3O_4 MDP in low concentrations has been investigated in the past by Rimbert (1959), Patton and Fitch (1962), Gillingham and Stacey (1971) and Johnson et al. (1975) The last mentioned authors found no difference between the demagnetization properties of the ARM carried by dispersed SDP and small MDP. Levi and Merrill (1976, 1978), in their study

0340-062X/80/0148/$01.00

of ARM and TRM characteristics of small Fe_3O_4 grains, gave curves of the ratio ARM/TRM with varying grain size. Their Fe_3O_4 grains, in the small MD size range were diluted to less than 1 wt.%, in a non-magnetic Al_2O_3 matrix. Sugiura (1979) investigated the concentration dependence of ARM in small Fe_3O_4 MDP with packing factors $p = 3 \cdot 10^{-6}$–$2 \cdot 10^{-2}$. He had a mixture of SDP and small MDP in his samples. Surprisingly he found a marked concentration dependence of the ARM even for these high dilutions.

Sample Characterisation

The theoretical domain wall thickness in bulk Fe_3O_4 is approximately 0.15 μ (Morrish and Yu 1955; Butler and Banerjee 1975). Therefore Fe_3O_4 grains in the size range of 0.2–0.4 μ contain only one or a few domain walls. This and surface effects prevent the development of MD characteristics. In a detailed model one has to consider furthermore whether the grains are equidimensional, of predominantly cubic or spherical form, or irregularly shaped. In all three cases the thickness, the number and the structure of the domain walls may be different.

The magnetic structure of such two- or more-domain particles, shaped as parallelepipeds or spheres, has been investigated theoretically by Amar (1957, 1958a, b) and Stapper (1969), respectively. For Fe_3O_4 grains, Morrish and Yu (1965), Stacey (1963), Dunlop (1973), Stacey and Banerjee (1974) and Butler and Banerjee (1975) have estimated the size for the transition from SDP to MDP and the possible structure of small MDP. In the last mentioned two papers the authors have emphasized the effect of surface terminations of domain walls. They conclude that SD-like moments occur at the surface of a grain. According to their model, the magnetization of a grain can be described by a superposition of SD and MD behaviour caused by different parts of the grain. Dunlop et al. (1974) found that equidimensional Fe_3O_4 grains of 0.22 μ are two-domain grains and those of less than 0.2 μ (still above SD size) consist essentially of one domain wall which behaves like one SDP. Shcherbakov (1978), assuming a particle to have the shape of a parallelepiped and using Amar's formulae, estimated that Fe_3O_4 grains of 0.2–0.4 μ would be two- and three-domain grains. According to this concept the ratio of the saturation remanent magnetization (I_{RS}) to the saturation magnetization (I_S) should be $I_{RS}/I_S = 0.2$–0.6 for two-domain grains and $I_{RS}/I_S = 0.1$–0.18 for three-domain grains. A similar estimation concerning the number of domain walls has been carried out by Moskowitz and Banerjee (1979). A circular configuration of spins for small, spherical MDP, as discussed by Morrish and Yu (1955), implies very low saturation remanent magnetization.

Effective Field in a Sample

It has been found that the ARM for heterogeneous samples is related to internal magnetic interaction effects. For this reason it is suited to the study of structural properties. The ARM depends also on the external demagnetizing factor of the sample. The latter effect is significant when the magnetic material is dispersed in high concentration. Both effects may possibly be separated when the ARM is measured as a function of the sample shape and of the concentration.

In order to describe magnetic interaction and demagnetizing effects one can aplly the 'effective field method' In this model one considers the internal magnetic field H_i acting on each individual grain. H_i can be thought to be approximately composed of the following terms

$$H_i = H_{ex} + H_d^o + H_{Lor}^o + H_{Lor}^i + H_d^{gr}.$$

H_{ex} = External magnetic field

H_d^o = Demagnetizing field of the sample. When the sample is macroscopically homogeneous and has the outer form of an ellipsoid, H_d^o is uniform and is given by $H_d^o = -N_g I$. N_g = demagnetizing tensor of the sample, I = average magnetization of the sample

H_{Lor}^o = Field inside a spherical cavity centered on the grain in question (Lorentz sphere). When the magnetization of the sample, I, is macroscopically homogeneous and a sufficient large sphere is taken, the effect of the grains outside the cavity can be described by a uniform field $H_{Lor}^o = +4\pi I/3$ within the sphere

H_{Lor}^i = Field of the grains inside the Lorentz sphere acting on the central grain. This field depends on the properties and distribution of the grains and on the interaction between them. It is called the local or interaction field

H_d^{gr} = Demagnetizing field due to the magnetization of the grain itself. This field will vary in some complicated way inside a small MDP

The concept of Lorentz sphere is, in general, valid if the grains are arranged isotropically and their orientations are statistically random.

Experimental Method

Several commercially available Fe_3O_4 powders, with grain sizes of 1–5 μ, were checked for purity using magnetization data and Mössbauer spectroscopy. With the help of the latter it was found that the powders contained up to about 20% of either γ-Fe_2O_3 or a solid solution of Fe_3O_4–γ-Fe_2O_3. Similar observations concerning commercial Fe_3O_4 powders have been made by other workers (Levi and Merrill 1978).

For this reason, and to reduce magnetic interaction which may hinder the dispersal of magnetic grains in a non-magnetic matrix, hematite (α-Fe_2O_3) was taken as the starting material of the samples. After the dispersal α-Fe_2O_3 was reduced to Fe_3O_4. Commercial α-Fe_2O_3 powder (Fluka AG) and ZrO_2 (Fluka AG), which was previously milled to a grain size $< 1\mu$, were dispersed and mixed ultrasonically in acetone. The mixing was continued in a slowly rotating plastic cylinder, containing a few hundred small glass spheres, while the acetone evaporated. Under the optical microscope the α-Fe_2O_3 grains appeared to be well dispersed (magnification × 1,100). The mixture was pressed into pellets (2–3 kbar). The reduction of the α-Fe_2O_3 to Fe_3O_4 was attempted at first using hydrogen at 300° C. In addition to the Fe_3O_4, however, the samples contained a certain amount of either Fe_2O_3 or FeO and metallic Fe. Better Fe_3O_4 was obtained using a mixture of approximately 90% N_2 and 10% H_2 at 500° C for about 40 min. The quality of Fe_3O_4 grains, prepared in this way, has been studied by Weisweiler and Alavi (1977). By examining some dozens of specimens, thus produced, a series of samples was found to contain fairly pure Fe_3O_4. Others had the correct saturation magnetization but were found to be contaminated by FeO and Fe (Mössbauer spectroscopy). These findings proved the samples to be too critically dependent on the gas mixture and on temperature variations. Finally, for sample preparation, a mixture of H_2 and H_2O vapour

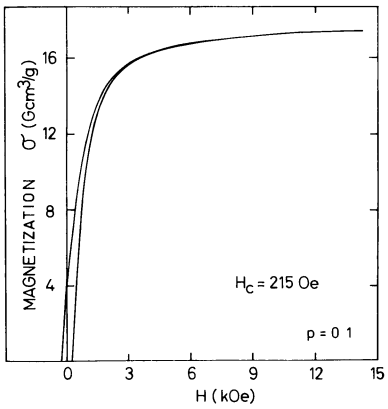


Fig. 1. Upper branch of a hysteresis loop for a spherical sample with $p=0.1$, measured at room temperature

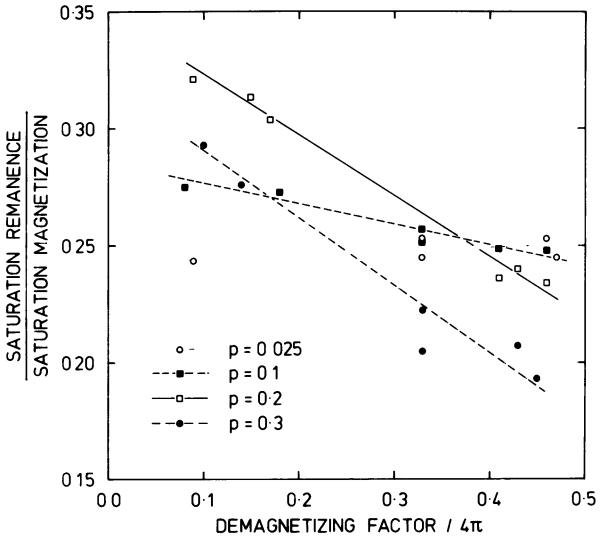


Fig. 2. The ratio I_{RS}/I_S vs the geometrical demagnetizing factor N_g of samples with various packing factors p

was used at 400° C (ratio of partial pressures $p_{H_2}/p_{H_2O} \simeq 12$). The Fe_3O_4 prepared in this way was reproducible and the samples were homogeneous. The specimens were cooled down to room temperature in 5–10 min by pulling them into a cold extension of the furnace.

It is possible, in view of the low reduction temperature, that the Fe_3O_4 grains were clusters of still smaller grains arising from the reduction process. However, no evidence is seen from scanning electron micrographs (SEM). The SEM showed that the grain sizes were distributed between 0.2 and 0.4 μ. No larger grains were seen. On the other hand, it was not possible to rule out the presence of grains of much less than 1 μ (magnification × 50,000). Both the Fe_2O_3 and Fe_3O_4 grains are almost spherical.

Small cylinders and spheres were cut from the reduced, pressed pellets and fixed in polyester resin to protect them from physical damage and oxidation. The cylinders were 5–7 mm long with length to diameter ratios between 2 and 6; the spheres had an average diameter of 3 mm.

The ARM was generated using an air-cored coil with a maximum A.F. amplitude of 1,500 Oe. A superimposed magnetic D.C. field up to ≃80 Oe was produced parallel to the A.F. field by an additional coil. All remanences were measured with a Digico spinner magnetometer.

Hysteresis properties were determined with a PAR vibrating sample magnetometer. The I_s was ca. 5% lower than the nominal values. I_s was accurate to 1%–2%.

Results

In Fig. 1 the upper branch of a hysteresis loop, measured on a spherical sample, is presented; it is typical of all concentrations used. The coercive force H_c varies between 210 and 235 Oe with a tendency to the lower values for increasing packing factors p. The relative isothermal saturation remanent magnetization I_{RS}/I_S is plotted in Fig. 2 for varying p and geometrical demagnetizing factors N_g of the samples. The magnetometric demagnetizing factors (Zijlstra 1967) were taken for the cylinders. Values of N_g in Fig. 2 correspond to measurements on spheres and on cylinders

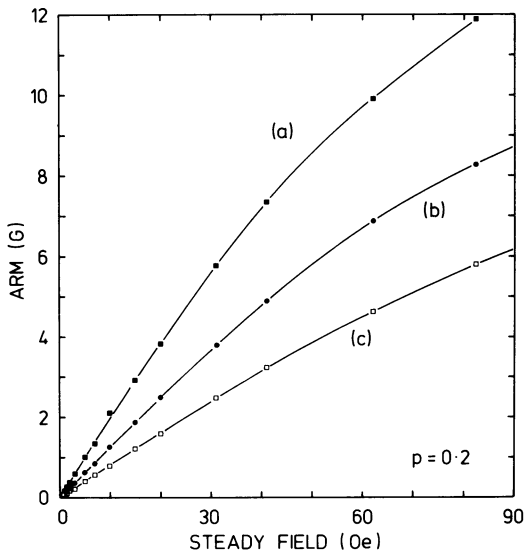


Fig. 3. Dependence of the anhysteretic remanent magnetization (ARM) on the steady field H for samples with $p=0.2$ and three values of the demagnetizing factor N_g: (*a*) $N_g/4\pi=0.08$, (*b*) $N_g/4\pi=0.33$, (*c*) $N_g/4\pi=0.46$

magnetized along and perpendicular to their axes. The points are connected, in a first approximation, by straight lines. Samples with high p show a clear dependence of I_{RS}/I_S on N_g. There is no dependence on N_g for the most diluted samples.

Figure 3 shows how the ARM depends on the applied steady field H for a specimen with $p=0.2$ and three values of N_g. Up to about 15 Oe the ARM varies linearly with H. The concentration dependence of the ARM in a field of 2 Oe for specimens with different N_g is shown in Fig. 4. The dependence of the ratio ARM (2 Oe)/I_S on N_g for high p is much more pronounced than for I_{RS}/I_S. As the law of the variation is not known, an extrapolation to $N_g \to 0$ cannot be made. A.F. demagnetization curves of I_{RS} and ARM, carried by spheres with $p=0.01$ and 0.2, are plotted in Fig. 5. The samples wth high p have lower median destructive

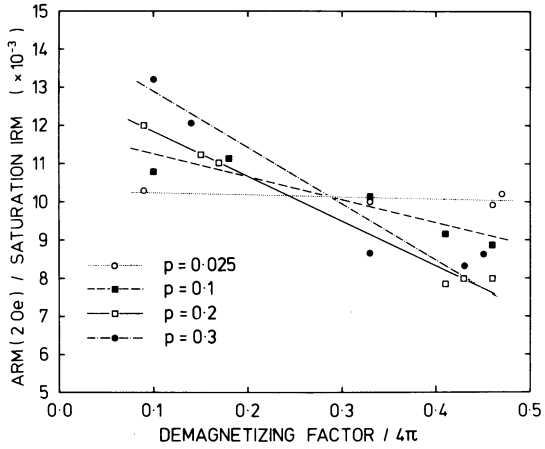


Fig. 4. ARM (2 Oe)/saturation magnetization vs the geometrical demagnetizing factor N_g for samples with various packing factors p

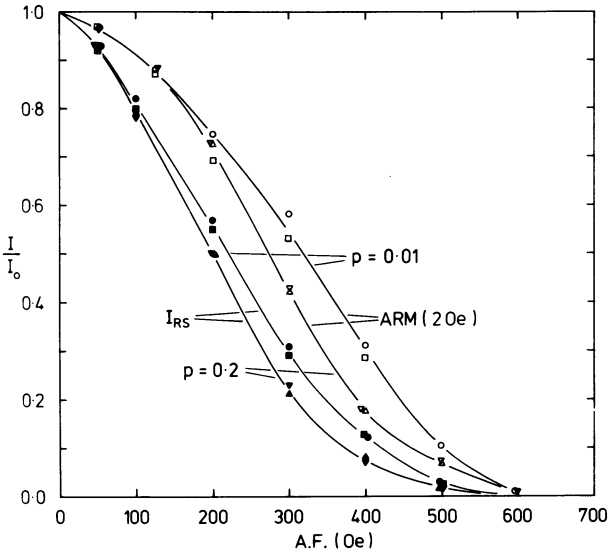


Fig. 5. A.F. demagnetization curves of the ARM and I_{RS} of spherical samples with $p=0.01$ and 0.2

fields $H_{1/2}$ ($H_{1/2}$ = peak field for which half of the initial remanence is removed) for both I_{RS} and ARM. In each case $H_{1/2}$ is smaller for I_{RS} than ARM (2 Oe). Figure 6 shows the relation between the acquisition and the demagnetization of the ARM for a sample with $p=0.2$ and three values of N_g. The partial ARM acquired in a peak alternating field ($\tilde{H}$) is plotted against the magnetization remaining after the full ARM is demagnetized in the same oscillating field. For all three N_g the points of varying alternating fields are on straight lines with slopes $\simeq -1$.

Discussion

The weak dependence of H_c on p is in agreement with the findings of Morrish and Yu (1955) and of Levi and Merrill (1978) for small MDP of Fe_3O_4. For completely spherical SDP, H_c is determined by the crystalline anisotropy constant $K_1 \cdot H_c \simeq 2\, K_1/I_s$, which for Fe_3O_4, is, about 145 Oe, whereas for shape anisotropy

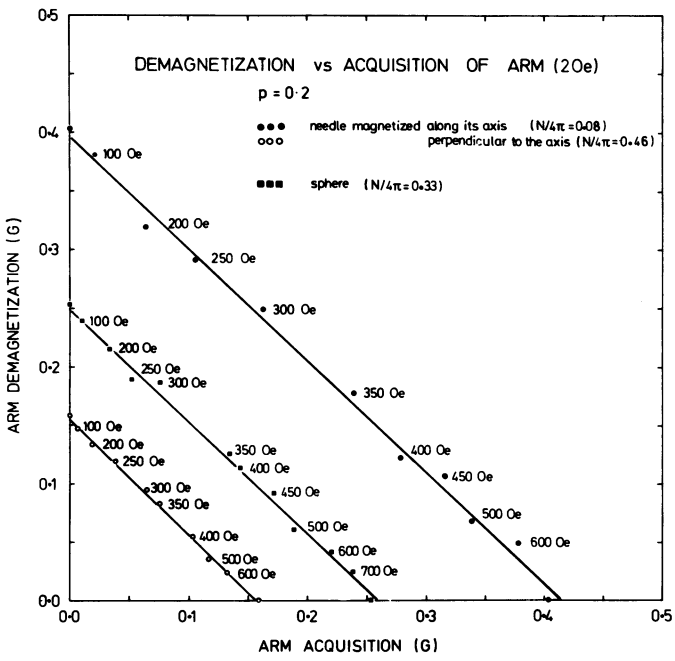


Fig. 6. ARM acquisition vs demagnetization of specimens with $p=0.2$ and different demagnetizing factors. The magnetic fields given at the points are the A.F. peaks taken for the generation of the ARM and for the demagnetization of the full ARM

H_c is much larger. On the other hand, for MDP H_c is always found to be smaller than for SDP. Therefore, from H_c measured for highly diluted concentrations of Fe_3O_4 grains the conclusion may be drawn that, on average, a slight deviation of the grains from spherical shape might occur. The values of H_c are compatible with those measured by Morrish and Watt (1958). Dunlop (1973) found for $\simeq 0.22\ \mu$ monocrystalline, cubic Fe_3O_4 grains, at room temperature, $H_c \simeq 100$ Oe, which is much lower. A possible reason for increased H_c might be pores or voides in the grains, strain and imperfect conversion of α-Fe_2O_3 to Fe_3O_4 (exchange anisotropy).

For $p=0.025$, $I_{RS}/I_S \simeq 0.22$. This value is probably much higher than one would expect from any circular spin configuration in the grains. It is close to the one predicted by Shcherbakov (1978) for two-domain grains of parallelepipedal or cubic shape. Therefore, in the spherical grains a similar arrangement of domains may be present. For higher p one has to consider agglomeration of grains in pairs or, to a lesser extent, in bridges. These clusters may have a magnetic structure different from the individual grains.

H_d^o, H_{Lor}^o, and H_{Lor}^i will all depend on the concentration of the magnetic grains, and any concentration dependence of the magnetic properties derives both from the outer shape of the measured sample and from the distribution and nature of the magnetic particles. The lack of a full model of the process of anhysteretic magnetization makes it difficult to subtract the shape effect and determine the interactions which occur between neighbouring particles. If the local field H_{Lor}^i is zero or can be approximated by a linear dependence on the average magnetization ($H_{Lor}^i = \alpha\, I$) there should be some shape of the sample such that $(-N_g + 4\pi/3 + \alpha)\, I = 0$ and the static field acting on an individual grain would be independent of p. Figures 2 and 4 show that, within the probable experimental uncertainties, such a crossing

point may occur (perhaps with the exception of the $p=0.3$ sample) at a demagnetizing factor of about 0.3. This suggests that the interaction effect between near neighbours is small for p less than 0.3. Due to the lack of data this conclusion is uncertain.

The difference between ARM and I_{RS} A.F. demagnetization curves for the highly diluted Fe_3O_4 powder (Fig. 5) had been observed in a similar form by Dunlop et al. (1973) on dispersed 0.1 μ Fe_3O_4 grains and by Johnson et al. (1975) on SDP and small MDP of Fe_3O_4 which may have contained a certain amount of γ-Fe_2O_3.

The linearity of the plots in Fig. 6 shows that, although shape-demagnetizing and structural effects influence strongly the intensity of remanence, they do not, in this case, invalidate palaeointensity techniques using partial demagnetization and remagnetization. This is, as mentioned earlier, important for rocks which bear highly concentrated Fe_3O_4 grains in titanomagnetite intergrowths and for archaeomagnetic specimens.

The acquisition of an anhysteretic remanent magnetization may be divided into two parts; the induction of a magnetization in an applied field (superposition of fields) and the partial self-demagnetization when the external field is removed. These processes depend on the demagnetization field of the sample and complicate the interpretation of the remanence data. For SDP the difference between AM and ARM is thought to be negligible, whereas for small MDP, due to domain wall motion, it cannot be neglected.

Acknowledgements. The authors are indebted to Professor G. Angenheister, Head of the Institut für Allgemeine und Angewandte Geophysik, for support and advice. They would like to thank the members of the rock and mineral magnetism group for help and discussions.

References

Amar, H.: On the width and energy of domain walls in small multidomain particles. J. Appl. Phys. **28**, 732–733, 1957

Amar, H.: Size dependence of the wall characteristics in a two-domain iron particle. J. Appl. Phys. **29**, 542–543, 1958a

Amar, H.. Magnetization mechanism and domain structure of multidomain particles. Phys. Rev. **111**, 149–153, 1958b

Banerjee, S.K., Mellema, J.P.: A new method for the determination of paleointensity from the ARM properties of rocks. Earth Planet. Sci. Lett. **23**, 177–184, 1974

Butler, R.F., Banerjee, S.K.: Theoretical single-domain grain size range in magnetite and titanomagnetite. J. Geophys. Res. **80**, 4049–4058, 1975

Davis, P.M., Evans, M.E.: Interacting single-domain properties of magnetite intergrowths. J. Geophys. Res. **81**, 989–994, 1976

Dunlop, D.J.: Superparamagnetic and single-domain threshold sizes in magnetite. J. Geophys. Res. **78**, 1780–1793, 1973

Dunlop, D.J., Bailey, M.E., Westcott-Lewis, M.F.: Lunar paleointensity determination using anhysteretic remanence (ARM): A critique. Proc. Lunar Sci. Conf. **6**th, 3063–3069, 1975

Dunlop, D.J., Hanes, J.A., Buchan, K.L.: Indices of multidomain magnetic behavior in basic igneous rocks: Alternating-field demagnetization, hysteresis, and oxide petrology. J. Geophys. Res. **78**, 1387–1393, 1973

Dunlop, D.J., Stacey, F.D., Gillingham, D.E.W.: The origin of thermoremanent magnetization: Contribution of pseudo-single-domain magnetic moments. Earth Planet. Sci. Lett. **21**, 288–294, 1974

Dunlop, D.J., West, G.F.: An experimental evaluation of single-domain theories. Rev. Geophys. Space Phys. **7**, 709–757, 1969

Evans, M.E., Wayman, M.L.: An investigation of the role of ultrafine titanomagnetite intergrowths in palaeomagnetism. Geophys. J.R. Astron. Soc. **36**, 1–10, 1974

Gillingham, D.E.W., Stacey, F.D.: Anhysteretic remanent magnetization (ARM) in magnetic grains. Pure Appl. Geophys. **8**, 160–165, 1971

Jaep, W.F.: Anhysteretic magnetization of an assembly of single-domain particles. J. Appl. Phys. **40**, 1297–1298, 1969

Jaep, W.F.: Role of interactions in magnetic tapes. J. Appl. Phys. **42**, 2790–2794, 1971

Johnson, H.P., Lowrie, W., Kent, D.V.: Stability of anhysteretic remanent magnetization in fine and coarse magnetite and maghemite. Geophys. J.R. Astron. Soc. **41**, 1–10, 1975

Kneller, E.: Magnetic interaction effects in fine – particle assemblies and thin films. J. Appl. Phys. **39**, 945–955, 1968

Kneller, E., Köster, E.: Relation between anhysteretic and static magnetic tape parameters. IEEE Trans. Magn. **Mag-13**, 1388–1390, 1977

Levi, S., Merrill, R.T.: A comparison of ARM and TRM in magnetite. Earth Planet. Sci. Lett. **32**, 171–184, 1976

Levi, S., Merrill, R.T.: Properties of single-domain, pseudo-single-domain and multidomain magnetite. J. Geophys. Res. **83**, 309–323, 1978

Morrish, A.H., Watt, L.A.K.: Coercive force of iron oxide micro-powders at low temperatures. J. Appl. Phys. **29**, 1029–1033, 1958

Morrish, A.H., Yu, S.P.: Dependence of the coercive force on the density of some iron oxide powders. J. Appl. Phys. **26**, 1049–1055, 1955

Moskowitz, B.M., Banerjee, S.K.: Grain size limits for pseudosingle domain behavior in magnetite: Implications for paleomagnetism. IEEE Trans. Magn. **Mag-15**, 1241–1246, 1979

Néel, L.: Théorie du traînage magnétique des ferromagnétiques en grains fins avec applications aux terres cuites. Ann. Geophys. **5**, 99–136, 1949

Parry, L.G.: Magnetic properties of dispersed magnetite powders. Philos. Mag. **11**, 303–312, 1965

Patton, B.J., Fitch, J.L.: Anhysteretic remanent magnetization in small steady fields. J. Geophys. Res. **67**, 307–311, 1962

Rimbert, J.: Contribution à l'étude de l'action de champs alternatifs sur les aimantations rémanents des roches, Applications géophysiques. Rev. Inst. Fr. Pet. **14**, 123–155, 1959

Shcherbakov, V.P.: Theory concerning the magnetization properties of pseudo-single domain grains. Bull. Acad. Sci. USSR, Earth Phys. **14**, 356–362, 1978

Shcherbakov, V.P., Shcherbakova, V.V.: Calculation of thermoremanence and ideal magnetization of an ensemble of interacting single-domain grains. Bull. Acad. Sci. USSR, Earth Phys. **13**, 413–421, 1977

Stacey, F.D.: The physical theory of rock magnetism. Adv. Phys. **12**, 45–133, 1963

Stacey, F.D., Banerjee, S.K. The physical principles of rock magnetism. Amsterdam, London, New York: Elsevier 1974

Stapper, C.H.: Micromagnetic solutions for ferromagnetic spheres. J. Appl. Phys. **40**, 798–802, 1969

Sugiura, N.: ARM, TRM and magnetic interactions: Concentration dependence. Earth Planet. Sci. Lett. **32**, 451–456, 1979

Thellier, E., Thellier, O.: Sur l'intensité du champ magnétique terrestre dans le passé historique et géologique. Ann. Geophys. **15**, 285–376, 1959

Weisweiler, W., Alavi, M.: Reduktion von alpha-Eisenoxid zu aktivem Magnetit und die unmittelbare Oxidation zu phasenreinem gamma-Eisenoxid. Ber. Dtsch. Keram. Ges. **54**, 268–271, 1977

Wohlfarth, E.P.: A review of the problem of fine-particle interactions with special reference to magnetic recording. J. Appl. Phys. **35**, 783–790, 1964

Zijlstra, H.: Measurements of magnetic quantities. In: Experimental methods in magnestism, E.P. Wohlfarth, ed. Amsterdam: North-Holland Publishing Company 1967

Received January 29, 1980; Revised Version June 3, 1980
Accepted June 3, 1980

J. Geophys. 48, 153–157, 1980

Journal of Geophysics

Stirred Remanent Magnetization: A Laboratory Analogue of Post-Depositional Realignment

P. Tucker

University of Edinburgh, Department of Geophysics, James Clerk Maxwell Building, Mayfield Road, Edinburgh EH 9 3JZ, Scotland

Abstract. The remanent magnetization acquired by a slurry stirred in a magnetic field was measured as a function of the applied field, stirring rate and water content. The experimental results were fitted by a theoretical model in which the stirring process was approximated as a periodic randomization of the grains. The acquired remanence was proportional to the applied magnetic field and independent of the stirring rate only for weak fields (<160 A/m) and slow stirring rates (<10 rad/s). The remanent intensity decreased with decreasing water content. The implications for the laboratory modelling of post-depositional remanent magnetization are discussed.

Key words: Rock magnetism - Sediments - Post depositional remanent magnetization - Modelling.

1. Introduction

The settling of magnetic grains through water and their immediate interaction with the sediment interface produces a statistical grain alignment which is determined by the ambient magnetic field at the time of deposition. Although the declination record preserved in the sediment is generally a good representation of the field declination, the recorded inclinations may be too shallow according to the mean shape of the grains (King 1955) or dependent on the physical nature of the sediment/water interface. Theoretically, Stacey (1972) showed that, for a distribution of grains with moments up to m_{max}, the intensity (M) varies with field as $\frac{1}{x}\ln\left(\frac{\sinh x}{x}\right)$; $x=\frac{m_{max}H}{kT}$ which is linear for geophysically realistic fields.

Analyses of both natural and laboratory deposited sediments (see Verosub 1977) have shown that the inclination error is often less than predicted and that the intensity record often bears no simple relationship to the applied field strength. Post-depositional models (Irving and Major 1964; Lovlie 1976) in which grain re-alignment occurs after the initial deposition, have been invoked to account for these features. It is convenient to classify the post-depositional effects into two categories: (i) when no external perturbations are present and (ii) where such perturbations do occur. The first category has been previously discussed (Tucker 1980): this paper is concerned with the second effect.

In the absence of external perturbations, the internal constraints on grain movement (void size and rigidity of the sediment, entrapped gas, cohesive forces and friction etc) may be sufficient to largely inhibit realignment by any magnetic torque acting on the remanence carriers. It has been shown (Tucker 1980) that for realistic field strengths (<200 A/m) only a small fraction of the carriers may be susceptible to realignment in this way. For the larger-scale realignments which have been proposed in order to account for the natural remanences of many fine-grain sediments, it is necessary to postulate the presence of additional time-dependent disturbances to the sediment. These may include local 'stirring' by for example bioturbation, shaking via earth movements or gross movement of the sediment during slumping. These mechanisms may temporarily reduce or remove the constraints on grain movement just as heating or the application of high alternating fields reduce or remove the effective barriers to domain-wall movement (or domain rotation) in TRM or ARM acquisition respectively.

In the laboratory, a convenient way of simulating the natural disturbances is simply by stirring a slurry in an applied field. Kent (1973) demonstrated that sediments prepared in this way could acquire a stable remanence whose intensity was approximately proportional to the applied field strength. Games (1977) found that sun-dried adobe bricks became magnetized at the time when the wet clay was 'thrown' into the mould. He successfully simulated the process by the 'throwing' and stirring of 'stiff' slurries. Verosub et al. (1979) noted that slurries stirred and dried in a field were more intensely magnetized than those stirred in zero field then field dried. The feature common to these experiments is the close relationship between the time at which the magnetization is acquired and the stirring process itself.

2. Stirring Experiments

In order to isolate the effects of stirring from those imposed by the drying process, the remanence was directly monitored using a fluxgate gradiometer, after and during the stirring process

0340-062X/80/0153/$01.00

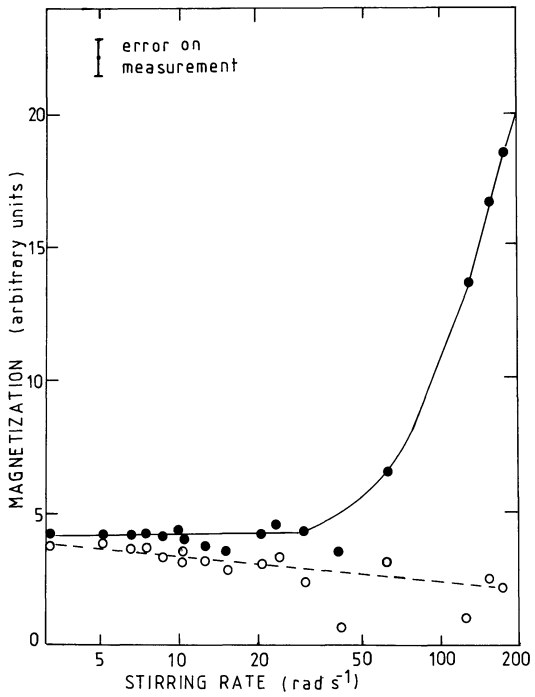


Fig. 1. The magnetization acquired by a synthetic sediment (slurry 1) during stirring in a magnetic field of 80 A/m (1 Oe) (*dotted line*) and the maximum remanence reached after stirring (*solid line*)

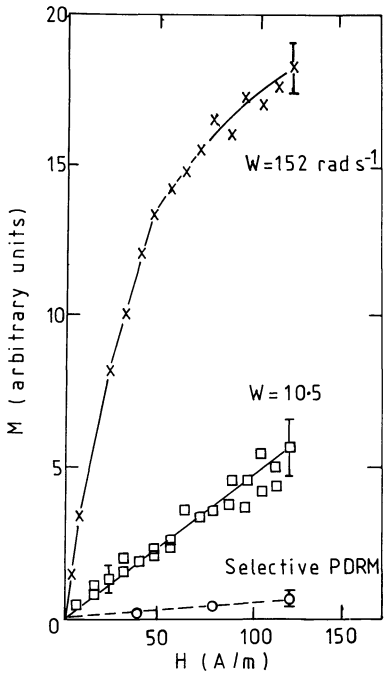


Fig. 3. The stirred remanent magnetization acquisition curves for fast and slow stirring respectively. The maximum intensity achieved without stirring (after Tucker 1980) is plotted for comparison

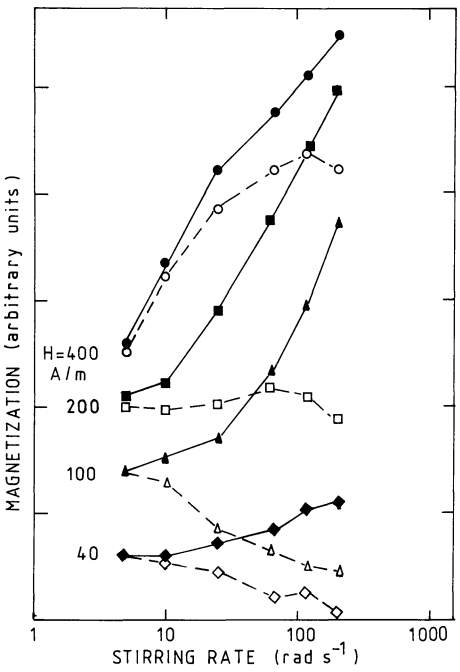


Fig. 2. The magnetization acquired by a synthetic sediment (slurry 2) during stirring in a magnetic field (*dotted line*) and the maximum remanence reached after stirring (*solid line*). The curves are plotted for a range of applied field strengths

itself. The laboratory field was cancelled and a field applied in the horizontal direction by a set of Helmholtz pairs. A mechanical stirrer, removed from the vicinity of the deposition apparatus, drove a nonmagnetic propeller via a cord and pulley system. The magnetization was measured along the field axis. After stirring the samples were left to dry, subsampled and measured on a spinner magnetometer.

The sediments were artificially prepared from a 2% concentration of natural magnetite of grain size 1–32 μm in a silica matrix (1–35 μm). The water content was set initially to 75%–76% which corresponded to a saturated slurry. The intensities were measured whilst stirring and 30 s after stirring, by which time visual signs of gross movement and the observed increase in magnetization had ceased.

For an inducing field of 80 A/m (Fig. 1) and slow stirring rates, the sediments retained a magnetization in the field direction during the stirring process itself. The intensity rose only marginally after the disturbance was removed. For fast stirring rates, the magnetization during stirring was correspondingly smaller, however this rose dramatically after the stirring finished. These general features were common to all the weak-field curves (Fig. 2) with the onset of the pronounced rise occurring at progressively lower stirring frequencies as the inducing field increased. At the higher inducing fields, however, the remanence during stirring initially increased with stirring rate and reached a maximum before falling away at the highest stirring speeds. The intensity measured 1 minute after a slow stir was proportional to the strength of the applied field for fields up to at least 120 A/m (slurry 1, Fig. 3) and up to 100 A/m for slurry 2.

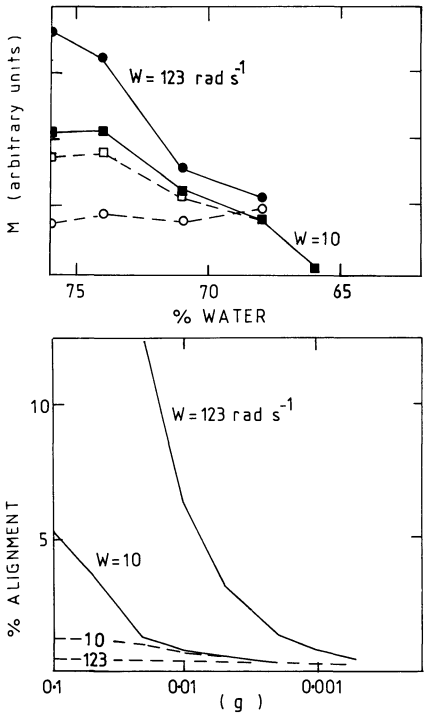


Fig. 4. The effect of progressive deatering on the alignment capacity of a synthetic sediment. The *dotted* and *solid lines* refer to during and after stirring respectively. The model predictions for the same stirring rates are shown for comparison. Here the parameter g is used as a measure of water content

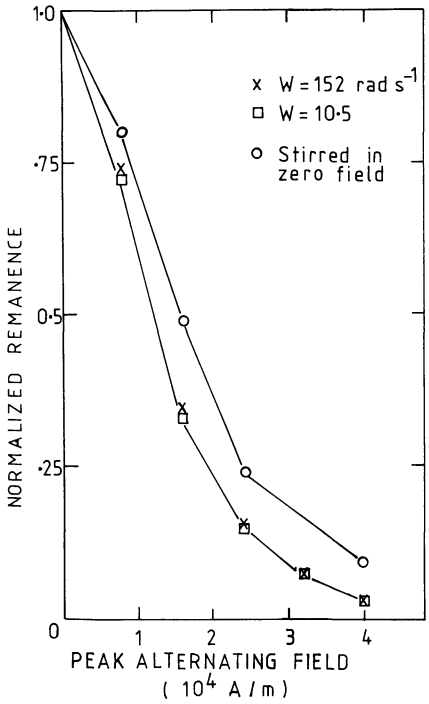


Fig. 5. The AF demagnetization characteristics of a synthetic sediment prepared by stirring in an 80 A/m field followed by field drying. The same sediment stirred in zero field then field dried is noticeably harder emphasising the selective activation of the high-coercivity fraction when field aided external perturbations are absent

For a fast stir, this relationship no longer held. The stirring-aided realignments were larger by a factor of 10 or more than the alignment achieved in the absence of any external perturbations (Fig. 3). The maximum realignment achieved decreased as the samples were progressively dried out (Fig. 4). For water contents of less than 65%, the matrix became too rigid for laboratory stirring. The reproducibility of the stirred remanent magnetization was better than $\pm 15\%$.

The intensity reached after stirring was stably preserved for several tens of hours and presumably would be for even longer times. Following a reversal in the direction of the applied field, the specimen remained magnetized in the original direction, the intensity decreasing by only 10%–20%.

Alternating field demagnetization of the dried out samples showed little difference between the coercivity fractions activated by each stirring rate, whereas the slurry stirred in zero field then field dried was noticably magnetically harder (Fig. 5).

The dramatic rise in post-stirring intensity for the high frequency disturbances may be because a greater number of grains are liberated for realignment, perhaps from the breaking up of grain clusters. According to the coercivity data these extra carriers would have to be of approximately the same size distribution as those previously activated. A second alternative is that the grains, activated, individually achieve greater realignment after the more rapid stirring. It will be shown below that the second alternative on its own is sufficient to account for the observed effects.

3. A Model of Stirring

Consider a spherical magnetic grain moving freely, under the influence of a magnetic torque, in a circular path. Its equation of motion is

$$f(\psi)+I\ddot{\theta}+\lambda\dot{\theta}+\mu_0 mH\sin\theta=0 \tag{1}$$

where θ is the angle between the grain moment (m) and the applied field (H), I the moment of inertia of the particle and $\lambda\dot{\theta}$ the viscous drag. $f(\psi)$ is the contribution of the circular motion. To solve the above equation it is necessary to make certain simplifying assumptions. In general the inertial term is small and may be neglected (Collinson 1965). The rotation terms contained in $f(\psi)$ can be treated by the following extremes.

(i) The stirring process serves to periodically randomize the grains without causing bodily rotation (in which case $f(\psi)$ can be omitted).

(ii) Where the particle moves in a circular path at the frequency ($w=\dot{\psi}$) of the stirring. The particle is then allowed to rotate about its own axis due to the magnetic torque.

With most designs of stirrers, for relatively thick slurries, the first effect is dominant for all but the most rapid of stirring rates; even with vigourous stirring, the particle's angular velocity may be very much lower than that of the stirrer.

Solving Eq. (1) for case (i) gives

$$\tan\frac{\theta(t)}{2}=\tan\frac{\theta_0}{t}\exp(-\mu_0 Hmt/\lambda). \tag{2}$$

As we are concerned with the component of remanence in the applied field direction (i.e., $m\cos\theta$), Eq. (2) can be transformed into

$$\cos\theta(t)=\left[\frac{1-\tan(\theta_0/2)\,e^{-2kt}}{1+\tan(\theta_0/2)\,e^{-2kt}}\right] \tag{3}$$

or

$$\cos\theta=F(\theta_0,k,t) \tag{4}$$

where $k=\mu_0 Hm/\lambda$ and θ_0 is the original magnetization direction. A periodic randomization every π/w seconds is assumed to produce a uniform distribution of θ_0. The grains will then realign according to Eq. (4) until $t=\pi/w$ or $t=\tau$ where τ is a characteristic time over which the grains remain mobile. At any instant, a fraction of the grains will have been randomized over the preceding τ seconds whilst the remainder would have reached their final orientation. The total magnetization of the sample is thus given by

$$M=m\sum_{\theta_0}\left(\sum_{t=0}^{\tau}F(\theta_0,k,t)+\sum_{t=\tau}^{\pi/w}F(\theta_0,k,\tau)\right) \tag{5}$$

for $\tau<\pi/w$ and

$$M=m\sum_{\theta_0}\sum_{t=0}^{\pi/w}F(\theta_0,k,t) \tag{6}$$

for $\tau>\pi/w$.

The above can be solved numerically once a further assumption is made about the nature of τ. The characteristic time would be expected to decrease as the sediment become more rigid and should be largest for the greatest magnitude of disturbance. Visual observation of the slurries showed that for $w\approx 300$ rad/s, internal movement appeared to have stopped within approximately 5 s after stirring had finished, and the increase in remanence was virtually complete within 10–20 s. For slower rotation rates these times were correspondingly reduced. A first approximation to the value of τ may thus be made as $\tau=gw$ where g is a 'stiffness' constant, w being taken as a measure of the magnitude of disturbance. On the above considerations, g would be of the order 0.05. The predicted results do not, in fact, critically depend on the exact form of expression; any function of τ which monotonically increases with w would suffice. The approximation chosen is thought to be least valid at the very lowest rotation rates.

Direct calculation of the constant k, for say a 1–10 μm particle, with $m\sim 10^3-10^4\ r^3$ Am^2, $\lambda=3\times10^{-2}\ r^3$ for a field $H=80$ A/m gives $k\sim 5$–50. This is thought to be an overestimation as the calculated drag coefficient (λ) only strictly applies to an isolated particle: in a concentrated slurry λ may be very much higher.

Equations (5) and (6) were numerically integrated for $k=0.1$, 1, and 10 with $g=0.01$, 0.1, and 1. The model curves are shown in Fig. 6.

After stirring is complete, the grains would continue to move until a time τ had elapsed since they had last been disturbed. The maximum remanence (M) achieved would therefore be

$$M=m\sum_{\theta_0}F(\theta_0,k,\tau) \tag{7}$$

The predicted curves are shown in Fig. 6.

On comparison of Figs. 1 and 2 with Fig. 6, it is seen that there is good agreement between experiment and theory. As the stirring frequency increases, the remanence during stirring decreases for the low k (weak field) curves and peaks for the higher k (high field) curves. After stirring, little further realignment is possible with a low stirring rate but with rapid stirring a large increase in magnetization is expected. The onset of the rapid rise occurs at progressively higher frequencies as $k(H)$ decreases.

The effect of an increase in characteristic time (τ) is to shift the solid curves, in Fig. 6, to the left (i.e., give saturation at lower frequencies) and to increase the low frequency values of the dotted curves. A decrease in water content would show as a decrease in τ (modelled by a decrease in g; Fig. 4). Progressive drying out may also effect the microscopic viscosity (λ) and

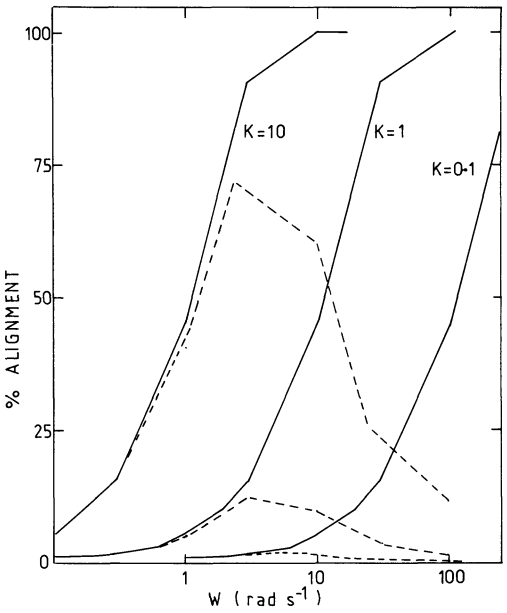


Fig. 6. The model predictions for alignment during stirring (*dotted line*) and the maximum alignment reached after stirring (*solid line*) as a function of stirring rate. The curves are drawn for a 'stiffness' constant $g=0.1$

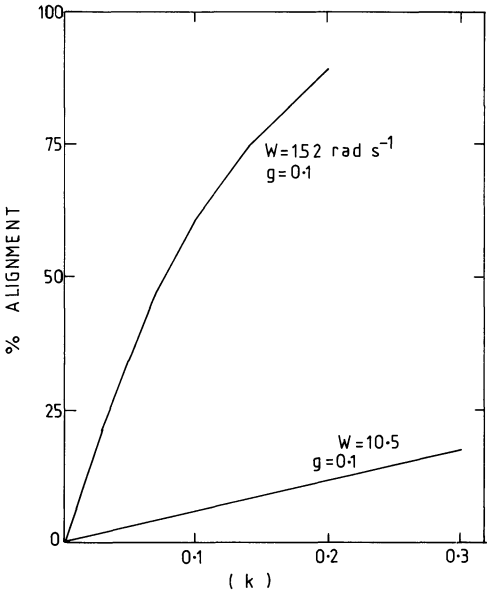


Fig. 7. The predicted stirred remanent magnetization acquisition curves, for high and low stirring rates respectively. k is a parameter proportional to the applied field strength

could be expected to further accentuate the fall in alignment capability on dewatering.

The 'best' fit to the data of Fig. 1 is given by $k=0.1$, $g=0.1$. Using the value $g=0.1$, the field ($\propto k$) dependence of the stirred remanent magnetization was calculated as a function of stirring rate (Fig. 7). The model predicts a linear relationship, with low stirring rates, for fields up to $k=0.3$ (~ 250 A/m) whereas with a high rate of stirring the linear relationship no longer holds. This again is in agreement with experiment (Fig. 3).

Allowing for a circular motion of the particle, the particle velocity ($\dot{\psi}$) may influence its alignment rate ($\dot{\theta}$). The extremes would be (i) θ independent of ψ and (ii) an additional change in θ of $\dot{\theta}'$ ($\propto\dot{\psi}$). The first leads to saturation at all frequencies and the second to a rapid fall off in remanence (whilst stirring) at moderately low particle velocities and, except for high τ, a similar fall in remanence after stirring. Neither extreme was evident for the sediments investigated.

4. Conclusions and Implications

Although many simplifying assumptions were made, the periodic randomization model of stirring does provide a solution which is consistent with all the experimental results to date. The main features would equally apply to other types of periodic disturbance and indeed to a 'one-shot' disturbance. The model relies heavily on the following two important concepts:

(i) In order to achieve realignment, constraining forces must be broken.

(ii) After the disturbance, these forces reassert themselves after a characteristic time (τ) at which point the particle is locked into place again preserving its current alignment.

It follows that the characteristic time will depend on the physical properties of the sediment (i.e., 'stiffness' or rigidity which depend on the water content, depth of burial, particle size and type) and on the magnitude of the disturbance.

Stirring does liberate a wide spectrum of grain coercivities for prospective realignment. The coercivity spectrum involved seems to be virtually independent of stirring rate. It should be borne in mind, however, that it may be dangerous to generalise this last point to include different types of disturbance. Games (1977) has shown that a 'throwing' of a sediment may activate a different coercivity spectrum to the stirring of the same slurry.

A rapidly stirred sediment (and a sediment poured in a field) do not show a linear relationship between the acquired magnetization and the applied field whereas a gentle disturbance (slow stirring or gentle tapping) does induce a remanent magnetization that is linear with the applied field, for any geophysically realistic field. For slow stirring rates, the final intensity reached is virtually independent of the stirring rate.

If these conclusions also apply to natural sediments, then it is evident that the magnitude of the PDRM may depend critically on the type and scale of disturbance. In order to model the PDRM process in the laboratory, it is essential that the measurements are made self-consistent. This would mean working with low frequency or small-scale disturbances and a saturated, or near-saturated, slurry where any fluctuations in stirring rate or differences in water content would have least effect. Further, with such low stirring rates, the remanent intensity assumes a linear relationship with the applied field and may provide an effective normalizing parameter for palaeointensity determination. Indeed, it may well be that the above conditions are the closest analogue to the natural PDRM acquisition conditions. Because the stirred remanent magnetization is acquired within a few tens of seconds after the disturbance is applied, there may be no need for recourse to prolonged drying. In situ measurements within a cryogenic magnetometer would permit a high throughput of results.

Acknowledgements. The work was carried out whilst in receipt of a Natural Environment Research Council fellowship.

References

Collinson, D.W.: Depositional remanent magnetization in sediments, J. Geophys. Res. **70**, 4663–4668, 1965

Games, K.P.: The magnitude of the palaeomagnetic field: a new non-thermal non detrital method using sun-dried bricks, Geophys. J. R. Astron. Soc. **48**, 315–329, 1977

Irving, E., Major, A.: Post-depositional remanent magnetization in a synthetic sediment. Sedimentology **3**, 135–143, 1964

Kent, D.V.: Post-depositional remanent magnetization in a deep-sea sediment. Nature **246**, 32–34, 1973

King, R.F.: Remanent magnetization of artificially deposited sediments. Mon Not. R. Astron. Soc. [Geophys. Suppl.] **7**, 115–134, 1955

Løvlie, R.: The intensity pattern of post-depositional remanence acquired in some marine sediments deposited during a reversal of the external magnetic field. Earth Planet. Sci. Lett. **30**, 209–214, 1976

Stacey, F.D.: On the role of Brownian motion in the control of detrital remanent magnetization of sediments. Pure Appl. Geophys. **98**, 139–145, 1972

Tucker, P.: A grain mobility model of post-depositional realignment, Geophys. J.R. Astron. Soc. in press, 1980

Verosub, K.L.: Depositional and post-depositional processes in the magnetization of sediments. Rev. Geophys. Space Phys. **15**, 129–143, 1977

Verosub, K.L., Ensley, R.A., Ulrick, J.S.: The role of water content in the magnetization of sediments. Geophys. Res. Lett. **6**, 226–228, 1979

Received February 15, 1980; Revised Version March 18, 1980
Accepted March 19, 1980

J. Geophys. 48, 158–160, 1980

Journal of Geophysics

High Precision Measurement of the Frequency of Mode $_0S_0$ *

W. Zürn[1], L. Knopoff[2], and P.A. Rydelek[2]

1 Geowissenschaftliches Gemeinschaftsobservatorium, Universitäten Karlsruhe/Stuttgart, Heubach 206, D-7620 Wolfach, Federal Republic of Germany

2 Institute of Geophysics and Planetary Physics, University of California at Los Angeles, Los Angeles, California 90024, USA

Abstract. Mode $_0S_0$ was excited by the Indonesian earthquake of 19 August 1977 and recorded at the South Pole and at UCLA. Phase estimates from 16 independent time lapse spectra were used to determine the frequency of this mode to high precision. The result is $f(_0S_0)=2.932851$ cph with a standard deviation of 11 ppm.

Key words: Frequency – $_0S_0$ – Time lapse spectra – Sumbawa earthquake.

Introduction

The frequencies of the earth's free vibrations form an important set of data and are used to infer the global average internal structure of the earth. The precision with which these eigenfrequencies can be measured depends in part on the length of seismic records available for spectral analysis; these depend in turn on the initial amplitude of a mode after an earthquake, its specific attenuation factor Q and the noise level of the instrument and/or station at the time. Recorded noise may be due to instrumental disturbances, aftershocks or earthquakes at other locations and to meteorological sources. The influence of changes in atmospheric mass on recordings by a gravimeter, for example, has been demonstrated by Warburton and Goodkind (1977) and Slichter et al. (1979) for tidal periods; these processes must certainly be important in the free mode band of periods as well.

The measurement of the frequency of mode $_0S_0$ is favored by its high Q; on the other hand it is excited only by very large earthquakes with a usable signal-to-noise ratio. Derr (1969) reports four observations of this mode; the present-day standard of measurement of the frequency of $_0S_0$ is (Slichter 1967)

$$f(_0S_0)=2.9324\pm0.0003 \text{ cph}$$

The Sumbawa earthquake of 19 August 1977 has provided an additional opportunity to measure the periods of this and other low order modes for the earth (Buland et al. 1979; Knopoff et al. 1978; Knopoff et al. 1979; Linton et al. 1979, Riedesel et al. 1979).

* Contribution No. 229, Geophysikalisches Institut, Universität Karlsruhe, Hertzstr. 16, D-7500 Karlsruhe, Federal Republic of Germany and Publication No. 2071, Institute of Geophysics and Planetary Physics, University of California, Los Angeles

Measurement of the Frequency of $_0S_0$

Knopoff et al. (1979) determined Q of $_0S_0$ from recordings of the Sumbawa earthquake using ultralong period seismographs at South Pole (SPA) and Los Angeles (LMS). For the determination of Q for this mode, we used amplitude information from time lapse spectra while phase information was ignored. Details about the data and the data reduction can be found in the above paper. The phases for the different portions of the records are listed in Table 1.

A fast Fourier transform of the longest clean record available from the two stations (9,900 min starting at minute 332,480.0 in 1977 from SPA) gave the following spectral amplitudes in the vicinity of $_0S_0$:

$f_{483}=2.927273$ cph $\quad A_{483}=1.2540$
$f_{484}=2.933333$ cph $\quad A_{484}=7.3647$
$f_{485}=2.939393$ cph $\quad A_{485}=1.7700$

Table 1. Phase estimates for $f=2.933333$ cph as a function of lapse time of the start of each record (in minutes from 0000 hours U.T., 1 January 1977). The origin time of the earthquake was 331568.92 min. Each record was 45.0 h long. Corrections t_{ci} were measured with stop-watch against time signals. The sample denoted with a question mark was deleted in the final analysis. Reference time $t_0=t_1$

t_i (min)	t_i-t_1 (h)	Station	t_{ci} (s)	Phase (°)	Phase (cycles)
332480	0.00	SPA	0.4	−164.567	−0.4571
334880	40.00	LMS	− 1.5	73.670	0.2046
337320	80.67	SPA	0.3	− 13.654	−0.0379
339840	122.67	LMS	−12.5	− 84.945	−0.2360
343520	184.00	SPA	0.2	− 49.177	−0.1366
344160	194.67	LMS	−13.5	−141.703	−0.3936
347160	244.67	SPA	0.1	− 16.274	−0.0452
351360	314.67	LMS	−16.0	−122.898	−0.3414
353616?	352.27	SPA	0.0	−147.189	−0.4089
355680	386.67	LMS	−17.0	−170.921	−0.4748
360000	458.67	SPA	− 0.1	122.919	0.3414
360240	462.67	LMS	−18.5	−131.794	−0.3661
362880	506.67	SPA	− 0.1	177.294	0.4925
362945	507.75	LMS	−19.0	130.097	0.3614
365760	554.67	SPA	− 0.2	− 64.893	−0.1803
367200	578.67	LMS	−20.5	146.084	0.4058
370080	626.67	LMS	− 0.5	−125.273	−0.3480

0340-062X/80/0158/$01.00

where the subscripts are the estimate number in the Fourier spectrum of the record and the amplitudes, A, are in relative units. From the near symmetry of these amplitudes we conclude that the frequency of ${}_0S_0$ must be very close to f_{484} and is bounded by

$$2.936363 \text{ cph} > f({}_0S_0) > 2.930303 \text{ cph}$$

where the two limiting values are the midfrequencies between the three spectral estimates. If the phase of the mode is not affected by earthquakes subsequent to the start, t_1, of the first lapsed record, then the phase information from the time lapse spectra can be used to measure $f({}_0S_0)$ more accurately.

The phases in the spectra of the records are shifted with respect to those in the actual ground motion, due to instrumental phase shifts that include the effects of the free mode filters (Nakanishi et al. 1976). At both stations the frequency responses are sufficiently close to identical that we assume these phase shifts to be the same for all spectra; no corrections for instrumental effects have been applied.

The quartz clocks that control the digital sampling of the data drift at both stations. This drift was checked regularly against time signals and appropriate corrections (Table 1, column 4) have been applied.

A zero order approximation to the phases of ${}_0S_0$ in the lapsed records is given by the phases for the frequency $f_0 = 2.933333$ cph. The phases for f_0 differ from those of ${}_0S_0$ by an amount which depends on the difference between f_0 and $f({}_0S_0)$, the Q of the mode, and the length of the time series. Since all these factors are the same for all spectra, no corrections needed to be applied for these effects.

The phase p_i (in cycles) of the mode is defined to decrease linearly with starting time, t_i, for each segment of the seismogram, namely

$$p_i = -f({}_0S_0) \cdot (t_i - t_0) + p_0 \tag{1}$$

where (t_0, p_0) are a reference time and phase. Because phase is uncertain by an integer number of cycles, the results are only describable as p_i (modulo 1.0). The straight line of Eq. (1) is therefore mapped onto a sawtooth function with unknown slope (Fig. 1). There are many cycles of ${}_0S_0$ between the different starting times t_i of the record segments. Without *a priori* bounds on $f({}_0S_0)$ and in the presence of noise, it would be difficult to find the sawtooth-function that fits the data in a maximum likelihood sense. There is an obvious analogy to the aliasing problem of spectral analysis.

The *a priori* knowledge gleaned from the spectral analysis of the 9,900 minute record described above permits us to "demodulate" the sawtooth function. If we choose the "carrier-frequency" to be close enough to $f({}_0S_0)$, then the demodulated phases p_i' will span a phase interval of less than one cycle over all the record segments, $i = 1,2,\ldots$ We choose the estimate $f_0 = 2.933333$ cph from the above analysis as the demodulating frequency and let $\delta f = f({}_0S_0) - f_0$. Equation (1), as modified, is,

$$p_i' = -\delta f \cdot (t_i - t_0) + p_0 \tag{2}$$

where $p_i' = [p_i + f_0 \cdot (t_i - t_0)]$ (modulo 1.0). In other words, the value of f_0 allows us to estimate the number of full cycles in the interval $(t_i - t_0)$. If this estimate is accurate, then p_i' will satisfy the straight line equation within one cycle, except for noise. The demodulated phases are plotted in Fig. 2 and we note that f_0 is sufficiently close to $f({}_0S_0)$ that the p_i' span less than one cycle, for all values of i.

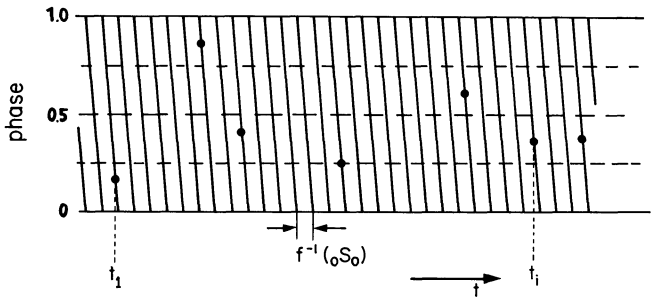


Fig. 1. Sketch of phase as a function of lapse time before demodulation. *Solid circles* denote phase estimates up to an arbitrary integer number of cycles

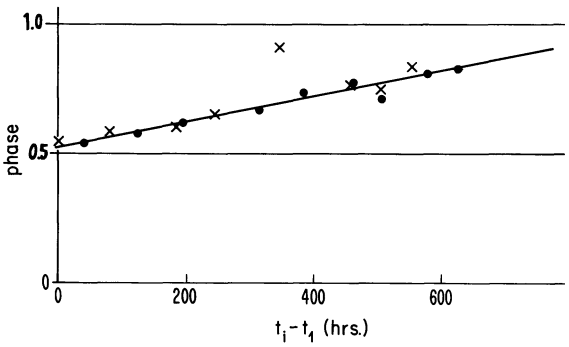


Fig. 2. Phase estimates (in cycles) as function of lapse time after demodulation with $f = 2.9333333$ cph. *Crosses* and *solid circles* are data from SPA and LMS respectively. *Straight line* is the least-squares fit to the estimates

The statistics of phase estimation for sinusoidal signals in the presence of Gaussian noise are close enough to normal for the noise levels in our spectra (Knopoff et al. 1979) that a least squares solution to Eq. (2) is justified (Middleton 1968, Fig. 9.4). This is in contrast to the Ricean statistics we were obliged to use to study spectral amplitudes. Because noise-free ${}_0S_0$ is in phase for both stations and because instrumental and data processing phase shifts are identical, we can analyze the phase data for both stations jointly, after the clock corrections have been applied.

As can be seen in Fig. 2, the phase estimate labelled with a question mark in Table 1 deviates conspicuously from the general trend of the other estimates. Inspection of the relevant time series revealed that a small earthquake was recorded during this segment; this may have produced an erratic phase estimate. This particular time series did not produce any more scatter in the spectral amplitudes than the other data and was not rejected in the amplitude analysis performed in Knopoff et al. (1979). In the present analysis this point has been deleted and we obtain, from the least squares solution for the remaining 16 time lapse spectra, the value

$$f({}_0S_0) = (2.932851 \pm 0.000031) \text{ cph}$$
$$= 2.932851 \text{ cph} \pm 11 \text{ ppm}.$$

Using seismograms from the IDA-network (and a different method of analysis) Riedesel et al. (1979), obtained

$$f({}_0S_0) = (2.932794 \pm 0.000015) \text{ cph}$$
$$= 2.932794 \text{ cph} \pm 5 \text{ ppm}.$$

These two results are in disagreement at the 1 s.d. level but agree at the 2 s.d. level. The effects of mode conversion, especially from nearby mode $_0S_5$, are probably small.

Acknowledgements. The superlative skills of Dr. Tadashi Yogi in operating the ultralong period seismometer at the South Pole during the year November 1976 to November 1977 made this research possible. Many discussions with Dr. Erhard Wielandt are appreciated. Drs. Dieter Emter, Karl Fuchs, and Gerhard Müller read the manuscript critically. We are grateful to Professor Karl Fuchs, Geophysical Institute, University of Karlsruhe, Germany for supporting the participation of Walter Zürn in this program. This work was supported by the National Science Foundation under grant numbers DPP 76-17234 A01 and EAR 78-23841. The Office of Polar Programs provided invaluable support for the program at the South Pole.

References

Buland, R., Berger, J., Gilbert, F.: Observations from the IDA network of attenuation and splitting during a recent earthquake. Nature **277**, 358–362, 1979

Derr, J.. Free oscillation observations through 1968. Bull. Seismol. Soc. Am. **59**, 2079–2099, 1969

Knopoff, L., Countryman, R.L., Rydelek, P.A., Smythe, W.D., Yogi, T., Zürn, W.: Free oscillations of the earth and tilt at the South Pole. Antarct. J.U.S., **XIII**, 4, 41–43, 1978

Knopoff, L., Zürn, W., Rydelek, P.A., Yogi, T.: Q of mode $_0S_0$. J. Geophys. **46**, 89–95, 1979

Linton, J.A., Smylie, D.E., Jensen, O.G.: Gravimeter observations of free modes excited by the August 19, 1977 Indonesia earthquake. Bull. Seismol. Soc. Am. **69**, 1445–1454, 1979

Middleton, D.: An introduction to statistical communication theory. New York: McGraw-Hill Book Co. 1968

Nakanishi, K.K., Knopoff, L., Slichter, L.B.: Observation of Rayleigh wave dispersion at very long periods. J. Geophys. Res. **81**, 4417–4421, 1976

Riedesel, M., Berger, J., Gilbert, F.: Frequency and attenuation of $_0S_0$ (Abstract). E.O.S., Trans. Am. Geophys. Union **60**, 379, 1979

Slichter, L.B. Spherical oscillations of the earth. Geophys. J. R. Astron. Soc. **14**, 171–177, 1967

Slichter, L.B., Zürn, W., Syrstad, E., Knopoff, L., Smythe, W.D., Uffelman, H.: Long period gravity tides at the South Pole. J. Geophys. Res. **84**, 6207–6212, 1979.

Warburton, R.J., Goodkind, J.M.: The influence of barometric pressure variations on gravity. Geophys. J. R. Astron. Soc. **48**, 281–292, 1977

Received March 6, 1980; Accepted May 28, 1980

J. Geophys. 48, 161–172, 1980

Journal of Geophysics

The Solution of Dynamic Problems of Elastic Wave Propagation in Inhomogeneous Media by a Combination of Partial Separation of Variables and Finite-Difference Methods*

A.S. Alekseev, B.G. Mikhailenko

Computing Center, Siberian Branch, Academy of Sciences of the USSR, Prospekt Nauki 6, 630090 Novosibirsk, USSR

Abstract. A new method for the calculation of theoretical seismograms, which is suitable for a wide class of inhomogeneous media, is suggested. The method is based on a combination of partial separation of variables with finite-difference techniques. Different variants of the method, based on the application of the Fourier-Bessel transform, finite integral transforms, expansions in Legendre polynomials, etc, are discussed in detail. Examples of theoretical seismograms for various simple structures are presented.

Key words: Theoretical seismograms – Finite differences – Partial separation of variables – Non-ray effects – Finite integral transforms.

1. Introduction

One of the basic problems of theoretical seismology and seismic prospecting is the computation of complete seismograms for inhomogeneous media. At present various methods can be used to compute theoretical seismograms. These methods have been applied successfully to the solution of many important problems in seismology. A brief review of these methods can be found in Červený et al. (1977). Most of these methods, however, give only incomplete theoretical seismograms, corresponding, to body waves or surface waves, for example. The most general method, at the present time, is the method of finite differences. The method is quite universal, but it is rather time consuming, and requires large amounts of computer store.

In this paper we will describe a new method for the calculation of theoretical seismograms, which is suitable for a wide class of inhomogeneous media, including vertically inhomogeneous media with block structures. The method is based on the combination of separation of variables with finite difference techniques. The basic principle of the method is the separation of the spatial variables (e.g., the coordinate corresponding to the epicentral distance). After this separation, the equation has reduced dimensionality, but remains hyperbolic. This equation can then be solved efficiently by finite differences.

We start with the application of this method to a vertically inhomogeneous halfspace and then proceed to more complicated cases.

* Presented at the Workshop Meeting on Seismic Waves in Laterally Inhomogeneous Media, Liblice, ČSSR, 27 February–3 March, 1978

2. Method of Solution

2.1. *Vertically Inhomogeneous Halfspace (Lamb's Problem)*

The physical problem that we are going to solve is the following: to determine the motion of the free surface and the interior of an inhomogeneous halfspace when a source of the normal-force type is located on the free surface. Since the geometry of the system has axial symmetry, it is convenient to use cylindrical coordinates (r, z).

The equation of motion of an inhomogeneous elastic medium is given by

$$(\lambda+2\mu)\,\mathrm{grad}\,\mathbf{U}-\mu\,\mathrm{rot}\,\mathrm{rot}\,\mathbf{U}+\mathrm{div}\,\mathbf{U}\,\mathrm{grad}\,\lambda +2(\mathrm{grad}\,\mu\cdot E)=\rho\frac{\partial^2\mathbf{U}}{\partial t^2}, \tag{1}$$

with the boundary conditions

$$\tau_{zz}|_{z=0}=-f(t)\,r^{-1}\,\delta(r), \qquad \tau_{rz}|_{z=0}=0, \tag{2}$$

and the initial values

$$\mathbf{U}|_{t=0}=\left.\frac{\partial\mathbf{U}}{\partial t}\right|_{t=0}=0, \tag{3}$$

where the following notation is used: $\mathbf{U}=\begin{pmatrix}U_r\\U_z\end{pmatrix}$ is the displacement vector, $\rho(z)$ is the density, $\lambda(z)$, $\mu(z)$ are Lamé's constants, E is the deformation tensor, the function $f(t)$ represents the time variation of the source, τ_{zz} and τ_{rz} are normal and tangential stresses, respectively.

(a) *The First Modification.* We seek a solution in the form of Fourier-Bessel integrals:

$$U_z=\int_0^\infty R(z,k,t)\,J_0(kr)\,dk, \tag{4}$$

$$U_r=\int_0^\infty S(z,k,t)\,J_1(kr)\,dk. \tag{5}$$

As a result, we obtain the following boundary value problem of reduced dimensionality for the functions $R(z, k, t)$ and $S(z, k, t)$:

0340-062X/80/0161/$02.40

$$\frac{\partial^2 \mathbf{G}}{\partial z^2} + A(z, k)\frac{\partial \mathbf{G}}{\partial z} + B(z, k)\,\mathbf{G} = C(z)\frac{\partial^2 \mathbf{G}}{\partial t^2}, \tag{6}$$

$$\left(\frac{\partial \mathbf{G}}{\partial z} + D(z, k)\,\mathbf{G}\right)\Big|_{z=0} = \varepsilon, \tag{7}$$

$$\mathbf{G}|_{t=0} = \frac{\partial \mathbf{G}}{\partial t}\Big|_{t=0} = 0, \tag{8}$$

where

$$\mathbf{G} = \begin{pmatrix} S(z, k, t) \\ R(z, k, t) \end{pmatrix}, \qquad \varepsilon = \begin{pmatrix} 0 \\ -kf(t) \end{pmatrix},$$

and A, B, C, D are known matrices. To solve this problem we use the method of finite differences. We must solve our problem for different values of k and then calculate the integrals (4) and (5) numerically. The system of Eqs. (6)-(8) must be solved for the values of k, which are knots of the quadrature formula, to allow the calculation of these integrals. If the values of k are small we can use an explicit scheme. Equation (6), in the form of an explicit scheme, enables us to compute the functions $R(z, k, t)$ and $S(z, k, t)$ at each spatial grid point, at the time step $(j+1)$, exclusively in terms of the values at the two previous time steps j and $(j-1)$. If the values of k are large, one has to use either a smaller time step Δt for the explicit scheme, or pass to an implicit scheme in order to make the computation stable.

The convergence of the Fourier-Bessel integrals is determined by the behaviour of the functions $R(z, k, t)$ and $S(z, k, t)$ as $k \to \infty$. For the case of an impulsive SH-torque source for a homogeneous model, we have analyzed the analytical solution for the boundary-value problem obtained after the separation of variables. The behaviour of $R(z, k, t)$ and $S(z, k, t)$ has been found to depend on the smoothness of the function $f(t)$ in the boundary condition. If $f(t)$ is a discontinuous function, one can speak of convergence of the integrals (4), (5) only in general terms. For smooth finite functions $f(t)$ the integrands $R(z, k, t)$ and $S(z, k, t)$ decrease exponentially as the parameter k increases, the integrals (4), (5) converge and therefore can be calculated numerically.

Replacing the integrals (4), (5) by those over a finite interval, we note that they are integrals of strongly oscillating functions. Therefore, to compute them we follow Filon's method: we consider $J_1(kr)$, $J_0(kr)$ as weight functions and substitute for $R(z, k, t)$ and $S(z, k, t)$ using an interpolation polynomial. Spline-interpolation of the second order is used.

The error in the method as a whole can be determined by comparison with the exact solution for the problem of an impulsive SH-torque source, for the case of a homogeneous model. The error in the displacement $U(r, z, t)$ does not exceed two or three per cent at distances up to $30\,\lambda_0$ (here λ_0 is the dominant wavelength radiated by the source). For details see Alekseev and Mikhailenko (1976), Mikhailenko (1973, 1974).

(b) *The Second Modification.* This method is based on combining finite-integral transformations (see Koshlyakov et al. 1970) with finite-difference methods. The use of finite-integral transformations considerably increases the possibilities of the method. Numerical integration of the rapidly-oscillating integrals (4), (5) is no longer necessary. To compute them one has only to solve a great number of one-dimensional problems at fixed values of the parameters k.

Consider problem (1)-(3) again and introduce new boundary conditions:

$$U_r(r, z, t)|_{r=r_0} = 0, \qquad U_z(r, z, t)|_{r=r_0} = 0. \tag{9}$$

We have thus introduced a reflecting surface at the distance $r = r_0$ from the origin. We select a sufficiently large distance r_0 and consider the wave field up to $t = T$, where T is the time taken for propagation of the wavefront up to the reflecting surface.

Let us apply finite Hankel integral transformations (see Sneddon 1951) along the coordinate r,

$$R_1(z, k_i, t) = \int_0^{r_0} r\,U_z(r, z, t)\,J_0(k_i r)\,dr, \tag{10}$$

$$S_1(z, \tilde{k}_i, t) = \int_0^{r_0} r\,U_r(r, z, t)\,J_1(\tilde{k}_i r)\,dr, \tag{11}$$

$$U_z(r, z, t) = \frac{2}{r_0^2}\sum_{i=1}^{\infty} R_1(z, k_i, t)\frac{J_0(k r_i)}{[J_1(r_0 k_i)]^2}, \tag{12}$$

$$U_r(r, z, t) = \frac{2}{r_0^2}\sum_{i=1}^{\infty} s_1(z, \tilde{k}_i, t)\frac{J_1(r\tilde{k}_i)}{[J_1'(r_0 \tilde{k}_i)]^2}, \tag{13}$$

where k_i are the roots of the equation $J_0(k_i r_0) = 0$, (14)

and $\tilde{k}_i$ are those of the equation $J_1(\tilde{k}_i r_0) = 0$. (15)

The boundary value problem of reduced dimensionality is:

$$\frac{\partial^2 \mathbf{Q}}{\partial z^2} + \tilde{A}(z, k_i^*)\frac{\partial \mathbf{Q}}{\partial z} + \tilde{B}(z, k_i)\,\mathbf{Q} = \tilde{C}(z)\frac{\partial^2 \mathbf{Q}}{\partial t^2}, \tag{16}$$

$$\frac{\partial \mathbf{Q}}{\partial z} + \tilde{D}(z, k_i^*)\,\mathbf{Q} = \tilde{\varepsilon} \quad \text{at} \quad z = 0, \tag{17}$$

$$\mathbf{Q}|_{t=0} = \frac{\partial \mathbf{Q}}{\partial t}\Big|_{t=0} = 0, \tag{18}$$

where $\mathbf{Q}(z, k_i^*, t)$ is the vector with the components $R_1(z, k_i, t)$ and $S(z, \tilde{k}_i, t)$, the coefficients $\tilde{A}$, $\tilde{B}$, $\tilde{C}$, $\tilde{D}$, $\tilde{\varepsilon}$ are known matrices, and the vector $k_i^* = \begin{pmatrix} k_i \\ \tilde{k}_i \end{pmatrix}$ contains the roots of Bessel's equations. Having solved problem (16)-(18) numerically for various roots of the Bessel's Eqs. (14), (15), we can find the components of the displacement vector by (12), (13).

The convergence of the series (12), (13) depends on the smoothness of the function $f(t)$. It can be shown that, if the function $f(t)$ satisfies the Dirichlet conditions, then as $k_i^* \to \infty$ the series converge rapidly. The above method allows one to calculate complete theoretical seismograms on the computer BESM-6 up to distances exceeding $100\,\lambda_0$ (λ_0 is the dominant wavelength generated by the source).

The total error in the determination of the displacement vector, caused by inaccuracy of the difference schemes and the truncation of the series, does not exceed 3 %-4 %. The method is easily generalized for anisotropic media. In this case only the coefficients in the difference equations are changed.

Other details can be found in Alekseev and Mikhailenko (1978).

2.2. Half-Space Inhomogeneous in the Horizontal Direction

Consider the application of the method to the calculation of wave fields in media where the elastic parameters are arbitrary functions of epicentral distance. In the system of Cartesian coordinates (x, z) propagation of SH waves from a line source is described by the equation

$$\frac{\partial}{\partial x}\left(\mu\frac{\partial U}{\partial x}\right)+\mu\frac{\partial^2 U}{\partial z^2}=\rho\frac{\partial^2 U}{\partial t^2}-\rho\,\delta(x)\,\delta(z-h)\,f(t), \tag{19}$$

with the boundary condition

$$\left.\frac{\partial U}{\partial z}\right|_{z=0}=0, \tag{20}$$

and the initial values

$$U|_{t=0}=\left.\frac{\partial U}{\partial t}\right|_{t=0}=0, \tag{21}$$

where $\mu(x)$ is Lamé's constant, and $\rho(x)$ is density.

For application of the finite integral transformation we introduce the boundary condition

$$\left.\frac{\partial U}{\partial z}\right|_{z=z_0}=0, \tag{22}$$

and select z_0 sufficiently far from the origin. The wave field is then considered up to the time $t=T$, where T is the time taken for wave propagation to the reflecting surface at $z=z_0$.

We employ the cosine transformation with finite limits

$$S(x,\eta,t)=\int_0^{z_0} U(x,z,t)\cos\frac{\eta\pi z}{z_0}\,dz, \tag{23}$$

$$U(x,z,t)=\frac{1}{z_0}S(x,0,t)+\frac{2}{z_0}\sum_{\eta=1}^{\infty}S(x,\eta,t)\cos\frac{\eta\pi z}{z_0}. \tag{24}$$

Multiplying Eq. (19) by $\cos\dfrac{\eta\pi z}{z_0}$ and integrating from zero to z_0, making use of conditions (20), (22), we obtain

$$\frac{\partial}{\partial x}\left(\mu\frac{\partial S}{\partial x}\right)-\mu\frac{\eta^2\pi^2}{z_0^2}S=\rho\frac{\partial^2 S}{\partial t^2}-\rho\,\delta(x)\cos\frac{\eta\pi h}{z_0}f(t), \tag{25}$$

$$S|_{t=0}=\left.\frac{\partial S}{\partial t}\right|_{t=0}=0. \tag{26}$$

Problem (25), (26) is solved by finite-difference methods for fixed values of η. Summing the values $S(x,\eta,t)$ for $\eta=0,1,2,\ldots,N$, according to formula (24), we find $U(x,z,t)$.

2.3. Radially Symmetric Medium

In the spherical system of coordinates r, Θ, φ

$(0<r\leqq a,\ 0\leqq\Theta\leqq\pi,\ 0\leqq\varphi\leqq 2\pi)$,

we consider a radially symmetric elastic medium where Lamé's constants λ, μ and density ρ are arbitrary, piecewise-continuous functions of the coordinate r. At the point $r=d$, $\Theta=0$, a vertical force is applied

$$\mathbf{F}=\delta(r-d)\frac{\delta(\Theta)}{d^2\sin\Theta}f(t)\,\mathbf{e}_r. \tag{27}$$

This force has axial symmetry and generates the displacement field

$$U=U_r(r,\Theta,t)\,\mathbf{e}_r+U_\Theta(r,\Theta,t)\,\mathbf{e}_\Theta. \tag{28}$$

The components $U_r(r,\Theta,t)$ and $U_\Theta(r,\Theta,t)$ are defined from the system of equations of the dynamic elasticity theory,

$$(\lambda+2\mu)\,\mathrm{grad}\,\mathrm{div}\,\mathbf{U}-\mu\,\mathrm{rot}\,\mathrm{rot}\,\mathbf{U}+\mathrm{div}\,\mathbf{U}\,\mathrm{grad}\,\lambda$$
$$+2(\mathrm{grad}\,\mu\cdot E)+\rho\mathbf{F}=\rho\frac{\partial^2\mathbf{U}}{\partial t^2}, \tag{29}$$

for zero initial values

$$\mathbf{U}|_{t=0}=\left.\frac{\partial\mathbf{U}}{\partial t}\right|_{t=0}=0, \tag{30}$$

and with boundary conditions at the free surface

$$\tau_{rr}|_{r=a}=0,\qquad \tau_{\Theta r}|_{r=a}=0, \tag{31}$$

where E is the deformation tensor, and τ_{rr}, $\tau_{\Theta r}$ are vertical and tangential stresses, respectively.

If, at the depth $r=d$, an impulsive SH-torque is applied

$$\mathbf{F}=2\,\delta(r-d)\frac{\delta(\Theta)}{d^3\sin^2\Theta}f(t)\,\mathbf{e}_\varphi, \tag{32}$$

then only SH-waves arise in the medium. The component $U_\varphi(r,\Theta,t)$ is determined from the equation

$$\frac{\mu}{r^2}\frac{\partial}{\partial\Theta}\left(\frac{\partial U_\varphi}{\partial\Theta}-U_\varphi\cot\Theta\right)+\frac{\partial}{\partial r}\left[\mu\left(\frac{\partial U_\varphi}{\partial r}-\frac{U_\varphi}{r}\right)\right]+3\frac{\mu}{r}\left(\frac{\partial U_\varphi}{\partial r}-\frac{U_\varphi}{r}\right)$$
$$+\frac{2\mu}{r^2}\left(\frac{\partial U_\varphi}{\partial\Theta}-U_\varphi\cot\Theta\right)\cot\Theta+\rho F=\rho\frac{\partial^2 U_\varphi}{\partial t^2}, \tag{33}$$

with the initial values

$$U_\varphi|_{t=0}=\left.\frac{\partial U_\varphi}{\partial t}\right|_{t=0}=0, \tag{34}$$

and the boundary condition at the free surface

$$\mu\left(\frac{\partial U_\varphi}{\partial r}-\frac{U_\varphi}{r}\right)\Bigg|_{r=a}=0. \tag{35}$$

If the Earth's core is liquid, then a similar boundary condition is introduced at the boundary of the core at $r=r_0$.

We seek a solution of Eqs. (29)–(31) in the form

$$U_r(r,\Theta,t)=\sum_{\eta=0}^{\infty}R(r,\eta,t)\,P_\eta(\cos\Theta), \tag{36}$$

$$U_\Theta(r,\Theta,t)=\sum_{\eta=0}^{\infty}S(r,\eta,t)\frac{\partial P_\eta(\cos\Theta)}{\partial\Theta}, \tag{37}$$

where the $P_\eta(\cos\Theta)$ are Legendre polynomials. To determine the functions $R(r, \eta, t)$ and $S(r, \eta, t)$ a new boundary value problem,

$$\frac{\partial}{\partial r}\left\{\lambda\left[\frac{\partial R}{\partial r}+\frac{2}{r}R-\frac{\eta(\eta+1)}{r}S\right]+2\mu\frac{\partial R}{\partial r}\right\}$$
$$+\frac{\mu}{r^2}\left\{4\frac{\partial R}{\partial r}r-4R+\eta(\eta+1)\left(3S-R-r\frac{\partial S}{\partial r}\right)\right\}+\rho F_r=\rho\frac{\partial^2 R}{\partial t^2}, \quad (38)$$

$$\frac{1}{r}\lambda\left[\frac{\partial R}{\partial r}+\frac{2}{r}R-\frac{\eta(\eta+1)}{r}S\right]+\frac{\partial}{\partial r}\left[\mu\left(\frac{\partial S}{\partial r}-\frac{S}{r}+\frac{R}{r}\right)\right]$$
$$+\frac{\mu}{r^2}\left[5R+3R\frac{\partial S}{\partial r}-S-2\eta(\eta+1)S\right]=\rho\frac{\partial^2 S}{\partial t^2}, \quad (39)$$

must be solved.

The boundary conditions at the free surface at $r=a$ are

$$(\lambda+2\mu)\frac{\partial R}{\partial r}+\frac{2\lambda}{r}R-\frac{\lambda}{r}\eta(\eta+1)S=0$$

and

$$\frac{\partial S}{\partial r}+\frac{1}{r}R-\frac{1}{r}S=0. \quad (40)$$

The initial conditions take the form

$$R|_{t=0}=\left.\frac{\partial R}{\partial t}\right|_{t=0}=0, \quad S|_{t=0}=\left.\frac{\partial S}{\partial t}\right|_{t=0}=0. \quad (41)$$

Problem (38)–(41) is solved by finite-difference methods for different values of the parameter η which is the summation index in formulae (36), (37). We will not discuss the peculiarities of the numerical solution of problem (38)–(41); they are described in Alekseev and Mikhailenko (1977). Summing the functions $S(r, \eta, t)$ and $R(r, \eta, t)$ according to formulae (36), (37), we determine the displacement components $U_r(r, \Theta, t)$ and $U_\Theta(r, \Theta, t)$.

The number of terms of the series (36) and (37) depends on the smoothness of the function $f(t)$, i.e., on the frequency of the signal from the source. This number increases linearly with frequency. Theoretical seismograms for waves with periods longer than 5–10 s can be calculated on the computer BESM-6. For other details see Alekseev and Mikhailenko (1977).

2.4. *Diffraction by a Wedge in an Inhomogeneous Medium*

In a cylindrical system of coordinates $0\leqq r<\alpha$, $0\leqq\varphi<2\pi$, consider the wave equation with the variable velocity $v_p(r)$ and a line source located outside a wedge, at the point (r_0, φ_0):

$$\frac{\partial^2 U}{\partial r^2}+\frac{1}{r}\frac{\partial U}{\partial r}+\frac{1}{r^2}\frac{\partial^2 U}{\partial\varphi^2}$$
$$=\frac{1}{v_p^2(r)}\frac{\partial^2 U}{\partial t^2}-\frac{2\pi}{r_0}\delta(\varphi-\varphi_0)\,\delta(r-r_0)\,f(t). \quad (42)$$

On the boundaries of the wedge the conditions

$$U|_{\varphi=0}=0, \quad U|_{\varphi=\alpha}=0, \quad (43)$$

are fulfilled, where α is the angle of the wedge ($\pi<\alpha<2\pi$). The initial values are

$$U|_{t=0}=0, \quad \left.\frac{\partial U}{\partial t}\right|_{t=0}=0. \quad (44)$$

Applying the sine transformation with finite limits,

$$R(r, \eta, t)=\int_0^\alpha U(r, \varphi, t)\sin\frac{\eta\pi\varphi}{\alpha}\,d\varphi, \quad (45)$$

$$U(r, \varphi, t)=\frac{2}{\alpha}\sum_{\eta=1}^{\infty}R(r, \eta, t)\sin\frac{\eta\pi\varphi}{\alpha}, \quad (46)$$

we obtain a new problem

$$\frac{\partial^2 R}{\partial r^2}+\frac{1}{r}\frac{\partial R}{\partial r}-\frac{\eta^2\pi^2}{r^2\alpha^2}R=\frac{1}{v_p^2(r)}\frac{\partial^2 R}{\partial t^2}-\frac{4\pi}{r_0}\sin\frac{\eta\pi\varphi}{\alpha}\delta(r-r_0)\,f(t), \quad (47)$$

$$R|_{t=0}=\left.\frac{\partial R}{\partial t}\right|_{t=0}=0. \quad (48)$$

Problem (47), (48) is solved by finite-difference methods at fixed values of η. The displacement $U(r, \varphi, t)$ is then found from (46). If the boundary conditions on the sides of the wedge are of the form

$$\left.\frac{\partial U}{\partial r}\right|_{\varphi=0}=0, \quad \left.\frac{\partial U}{\partial r}\right|_{\varphi=\alpha}=0, \quad (49)$$

then a cosine transformation with finite limits is applied.

Changing the velocity and the angle of the wedge, sufficiently complicated models of the media can be obtained.

2.5. *Other Diffraction Problems: Block Structures*

Without significant modifications the method is applied both to diffraction by a cylinder in an inhomogeneous medium and diffraction of a spherical wave by an opaque cone. In the first case we use the Fourier series with respect to the angular coordinate $\varphi(0\leqq r<\infty, 0\leqq\varphi<2\pi)$; in the second case the finite-integral Legendre transformation with respect to the angular coordinate $\Theta(0\leqq r<\infty, 0\leqq\Theta<\alpha)$ is used.

Employing finite transformations within the limits 0 to a and b, and connecting the solutions with the help of the finite-difference method, one can solve problems for inhomogeneous media with block structure, each block having its own parameters $\lambda(z)$, $\mu(z)$ and density $\rho(z)$.

At present, the method has been applied to construct theoretical seismograms for certain types of visco-elastic media, porous media and pre-stressed media. If Lamé's constants λ, μ and density ρ are arbitrary functions of some space variables, the classical separation of variables does not take place. In this case, employing finite integral transformations can be suggested as a rather effective method for calculation of theoretical seismograms (Mikhailenko 1978, 1979).

3. Certain Dynamical Peculiarities of Seismic Waves in Inhomogeneous Media: Non-Ray Effects

In this section, we will present examples of computations of theoretical seismograms for some simple models. These examples are not of methodological character, but they demonstrate

certain results which are of great interest in seismology and in seismic prospecting.

In all the examples presented, we consider a source with a source time function $f(t)$ given by the formula:

$$f(t)=\begin{cases} C\cdot[\sin(2\pi t/T)-0.5\sin(4\pi t/T)], & 0\leqq t\leqq T, \\ 0, & t<0,\ t>T. \end{cases} \tag{50}$$

3.1. *Vertically Inhomogeneous Media*

First Example. First we will consider a classical seismological model: a homogeneous layer overlaying a homogeneous halfspace, with a point source located at arbitrary depth (see Fig. 1). We select an explosive-type source, with unidirectional radiation characteristic. The parameters of the model under consideration are as follows:

$v_{p1}/v_{p2}=0.5$, $v_{s1}/v_{s2}=0.5$, $\rho_1/\rho_2=0.8$. The source is located at a depth of 2 km, the interface at a depth of 3 km. A number of papers have been devoted to this problem. For example, Alterman and Karal (1968) have solved this problem by the method of finite differences (with explicit schemes).

The theoretical seismograms of horizontal displacement U_r and vertical displacement U_z for six distances are shown in Fig. 2.

An interesting effect can be observed in Fig. 2, in the theoretical seismograms of the vertical component U_z, for the epicentral distance $R=0$. According to geometric ray theory the converted *PS* wave (reflected from the interface between the layer and the halfspace) should vanish for $R=0$. In Fig. 2, however, the *PS* wave following the *PP* wave can be seen clearly, even for $R=0$. The amplitude of the *PS* wave at $R=0$ depends on the distance of the source from the interface and on the elastic parameters of both media. The form of the *PS* wave at $R=0$ is the integral of the incident *P*-signal. As the epicentral distance increases, the form of the *PS* wave changes successively; at larger epicentral distances it is the same as the form of the *PP* wave, with opposite polarity. Note that the source time-function, given by $f(t)$, is differentiated in the explosive-type source). It should be noted that, for the source depth chosen (about 2 wavelengths), the Rayleigh wave is too weak to be seen in the seismograms.

Second Example. In this example, we consider a homogeneous halfspace with a free surface. A point source of the vertical-force type is located at a depth of 2 km. Two waves are seen clearly in the seismograms (Fig. 3) – the *P* wave, immediately followed by the *S* wave. The pulse shape of the vertical component of the *S* wave is the integral of $f(t)$ at small epicentral distances and changes successively to $f(t)$ as the epicentral distance increases. For angles of incidence larger than $\Theta=\sin^{-1}(v_s/v_p)$, the *S* wave changes its polarization from linear to elliptic.

Third Example. In this example, we consider a high-velocity layer located in a homogeneous halfspace (see Fig. 4). We consider a point *SH*-torque source located at the free surface. The wave field connected with high-velocity layers has a number of interesting peculiarities. A considerable amount of attention has been devoted by various authors to the so-called 'interference head waves' and to the 'tunnel waves'. It is of great importance to investigate the properties of these waves as a function of l/λ_2 (where l is the thickness of the layer and λ_2 the wavelength within the layer).

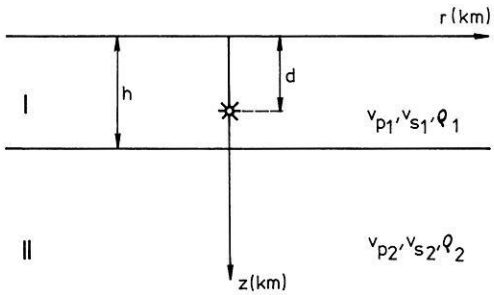


Fig. 1. Layered halfspace with an explosive point source

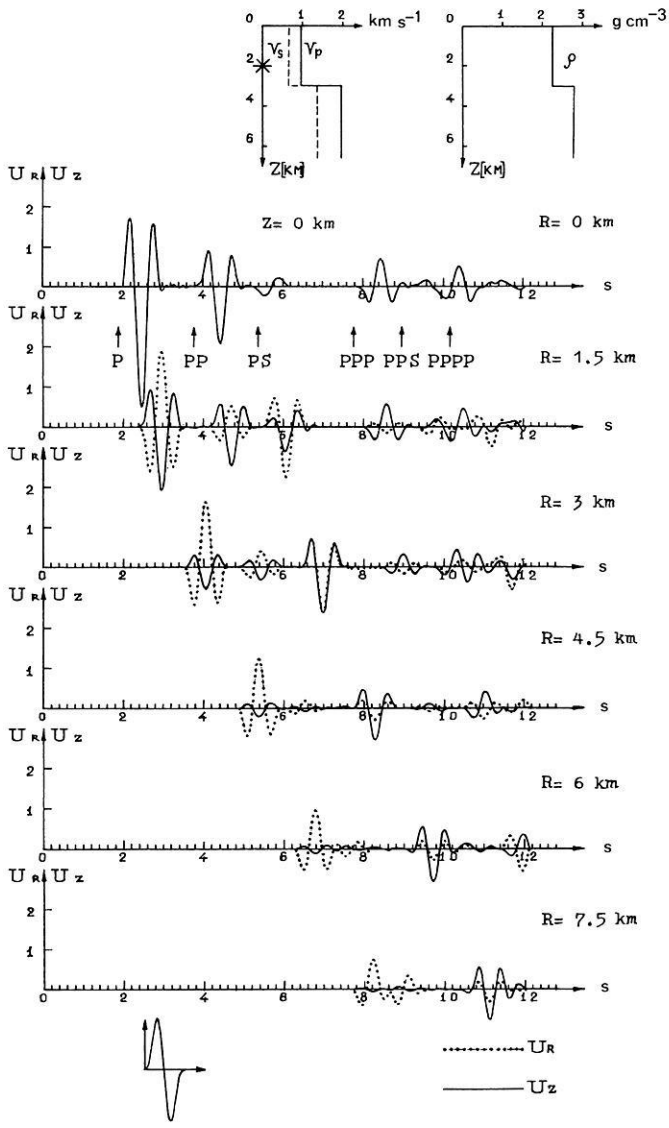


Fig. 2. Horizontal and vertical displacements U_r *(dotted curve)* and U_z *(solid curve)* on the surface of a layered halfspace ($z=0$) for different epicentral distances R. A point source of explosive type is located at a depth of 2 km, the interface at a depth of 3 km ($v_{p1}/v_{p2}=0.5$, $v_{s1}/v_{s2}=0.5$, $\rho_1/\rho_2=0.8$). The source time-function with a duration of 1 s is shown at the bottom-left. At the top are shown *P* and *S* wave velocity-depth graphs and the density-depth graph for the model used for computations

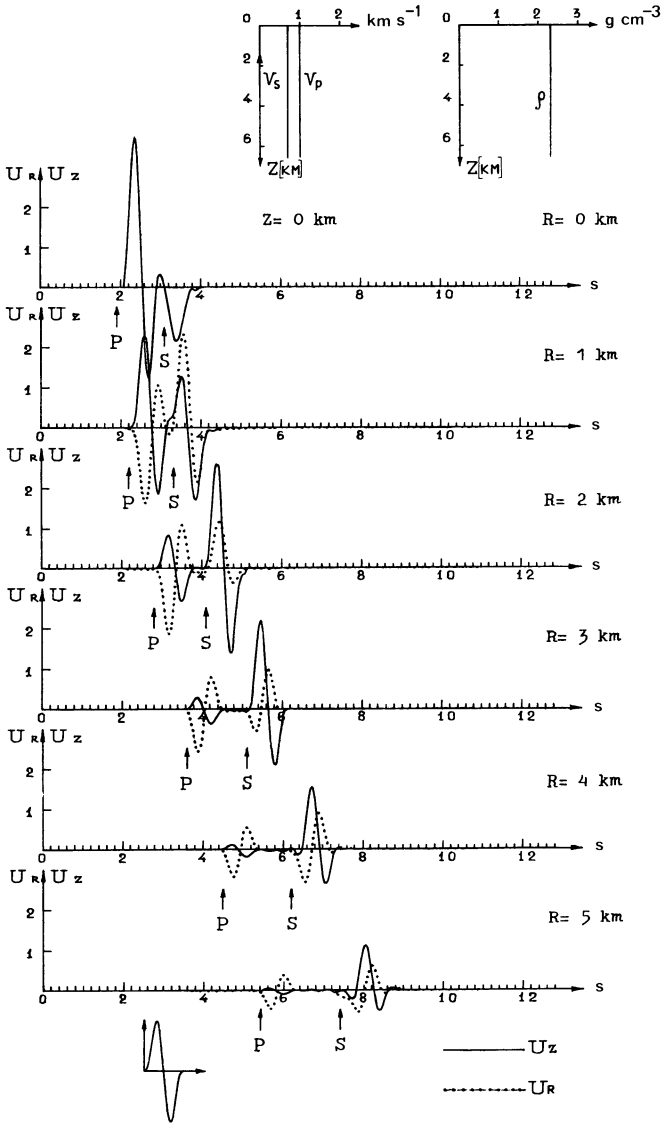


Fig. 3. Horizontal and vertical displacements U_r *(dotted curve)* and U_z *(solid curve)* on the surface of a homogeneous halfspace ($z=0$) for different epicentral distances R. An impulsive vertical-force point-source is located at a depth of 2 km. The source time-function with a duration of 1 s is shown at the bottom-left. At the top are shown P and S wave velocity-depth graphs and the density-depth graph for the model used for computations

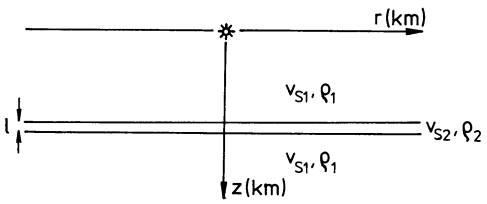


Fig. 4. Thin layer embedded in a halfspace model of Fig. 4

The process generating the interference head-wave along a thin high-velocity layer is closely connected with the interference group of multiply-reflected waves propagating within the layer. This interference group of multiply-reflected waves forms a distinct wave. Computations show that the amplitudes of interference head-waves depend not only on l/λ_2, but also on the velocities inside and outside the layer. For a larger velocity change (say, $v_{s1}/v_{s2}=0.5$), the maximum amplitudes have been observed for $l/\lambda_2 \sim 0.1$. The velocity of propagation of the interference head-wave is approximately 6%–7% lower than the velocity of head-wave propagation along an interface between two halfspaces. For a smaller velocity change ($v_{s1}/v_{s2}>0.7$) the amplitudes of interference head-waves decrease with decreasing l/λ_2. The change of velocity is not observed in this case.

Another interesting non-ray wave, connected with the high-velocity layer, is the tunnel wave (also called the screened wave). The generation of the tunnel wave cannot be explained by ray theory, it does not propagate through the layer along a ray-path,

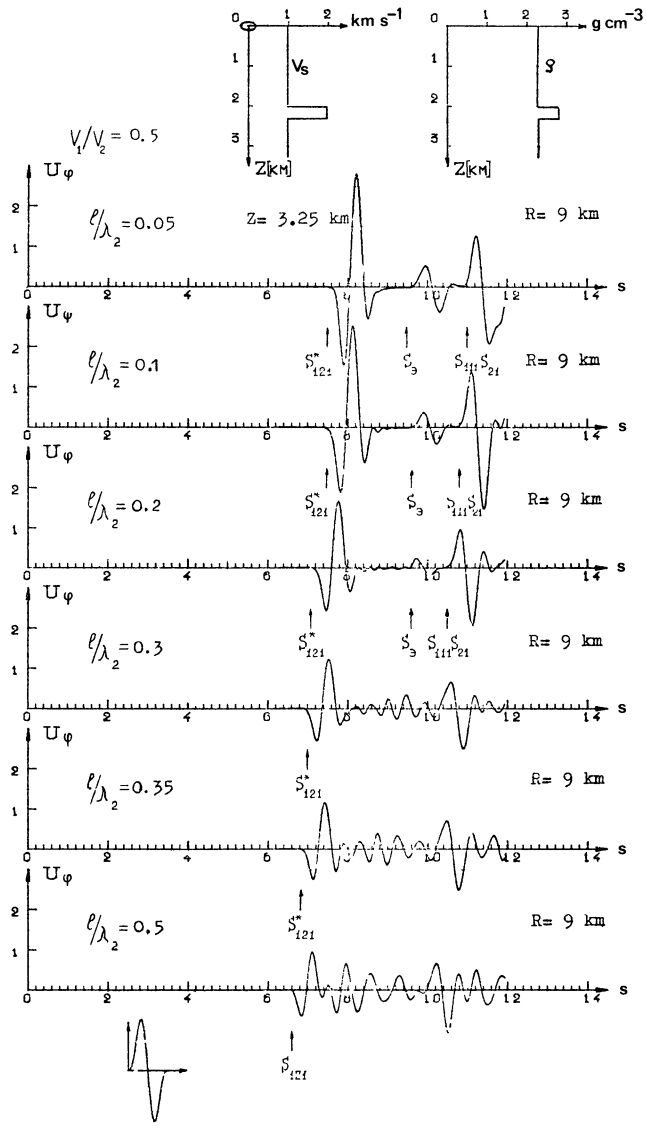


Fig. 5. Horizontal displacement U within the model of Fig. 4 a depth of 3.25 km and at the epicentral distance 9 km for different l/λ_2, where l is the thickness of the layer and λ_2 is the wavelength within the layer. A point SH-torque source is located at the free surface. The source time-function with a duration of 1 s is shown at the bottom-left.
At the top are shown S wave velocity-depth and density-depth graphs for the model used for computations

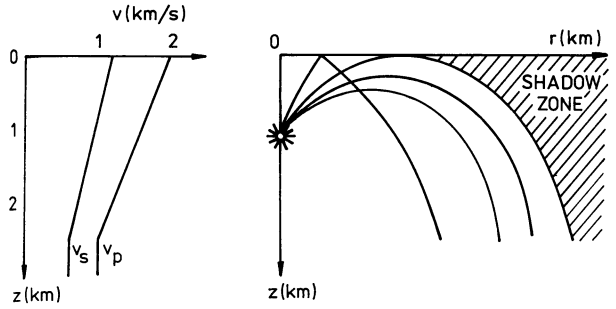


Fig. 6. Model of an anti-waveguide and corresponding rays

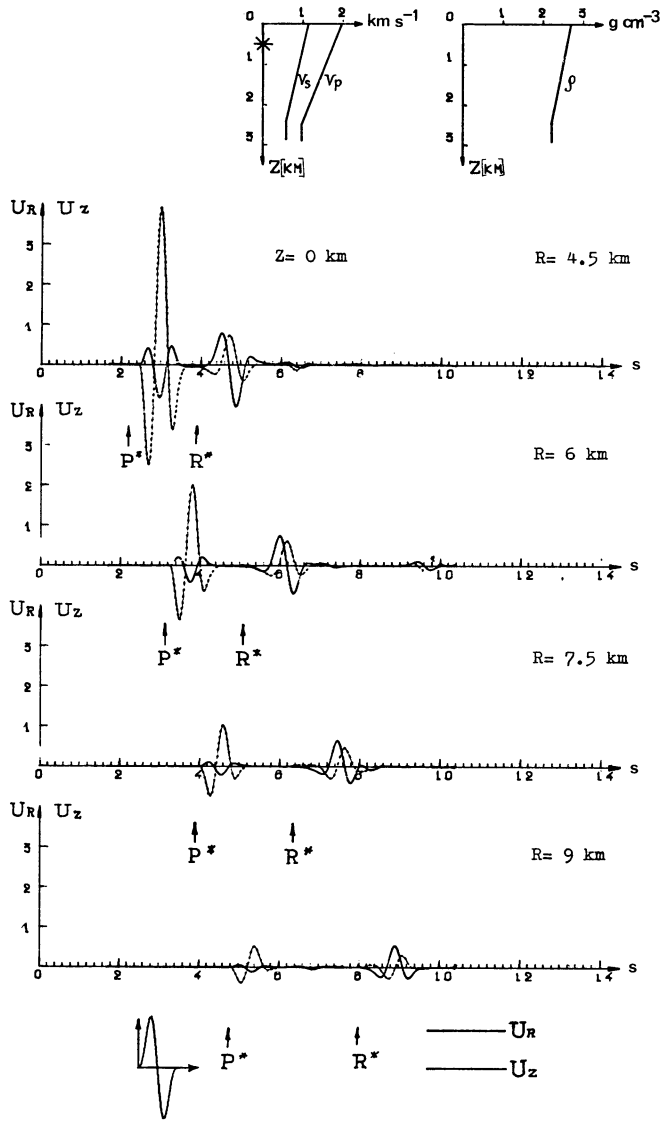


Fig. 7. Horizontal and vertical displacements U_r *(dotted curve)* and U_z *(solid curve)* at the surface of an inhomogeneous half-space (Fig. 6) for different epicentral distances R. An explosive-type point-source is located within the inhomogeneous anti-waveguide at a depth of 1 km. The source time-function with a duration of 1 s is shown at the bottom-left.
At the top are shown P and S wave velocity-depth graphs and the density-depth graph for the model used for computations

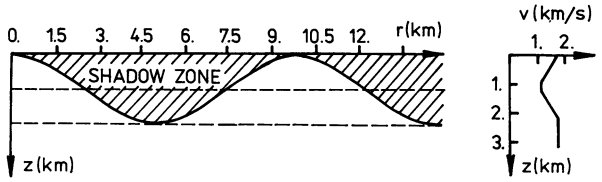


Fig. 8. Model of a waveguide and corresponding limiting ray

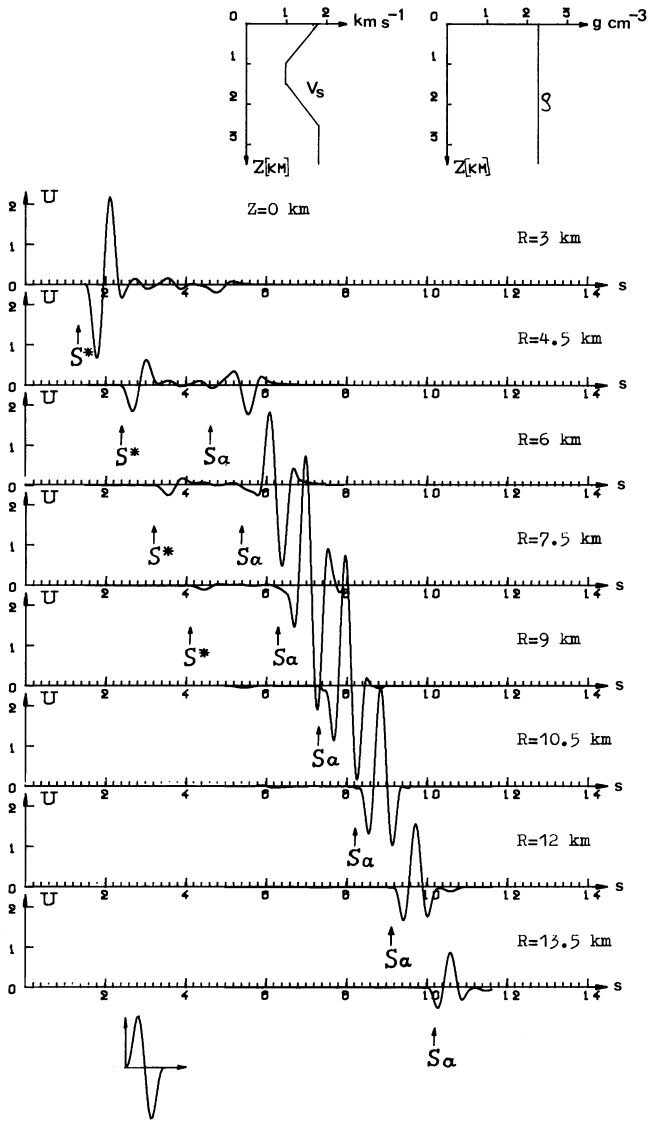


Fig. 9. Horizontal displacement U_φ at the free surface of the waveguide (Fig. 8) for different epicentral distances R. A point SH-torque source is located at the free surface. S_a is a channel wave. The source time-function with a duration of 1 s is shown at the bottom-left. At the top are shown S wave velocity-depth and density-depth graphs for the model used for computations

it 'tunnells'. Only lower frequencies are tunnelled, higher frequencies are screened. Thus, the dominant frequencies of the tunnel wave decrease. The amplitudes of tunnel waves depend on l/λ_2; they decrease with increasing l/λ_2.

Theoretical seismograms for various values of l/λ_2, for the epicentral distance $R=9$ km are shown in Fig. 5. The source is

located on the free surface, the receiver, at a depth of 3.25 km below the high-velocity layer ($v_{s1}/v_{s2}=0.5$). The wave observed as the first arrival (S^*_{121}) is the interference head wave. The tunnel wave S_3 does not have a distinct onset, its pulse shape corresponds approximately to the integral of the incident pulse.

Fourth Example. Now we will consider a model with an anti-waveguide (see Fig. 6). A point source of explosive type is located within the inhomogeneous anti-waveguide, at a depth of 1 km. Theoretical seismograms for four epicentral distances in the shadow zone are shown in Fig. 7. The wave reflected from

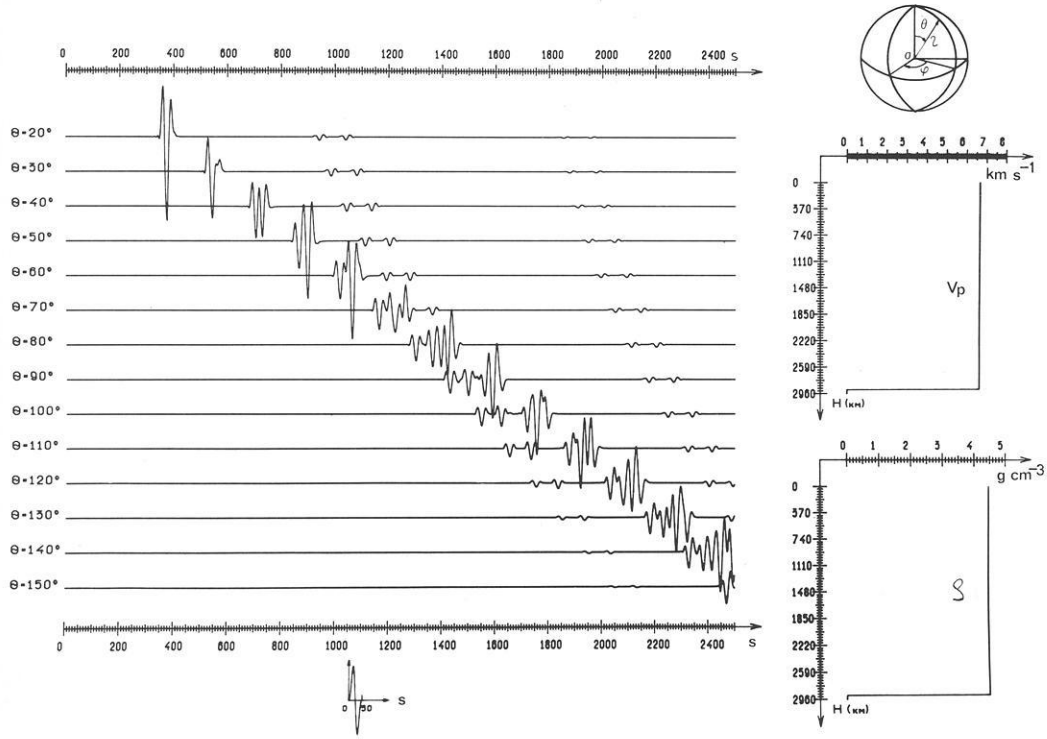


Fig. 10. Theoretical seismograms for a homogeneous spherical Earth model ($v_s=6.24$ km/s, $\rho=4.46$ g/cm^3) with a liquid core for different epicentral distances θ. A torque-type source is located at a depth of 300 km and generates a pulse with a duration of 50 s. The source time-function is shown at the bottom-left. At the right-hand side are P wave velocity-depth (*top*) and density-depth (*bottom*) graphs for the model used for computations

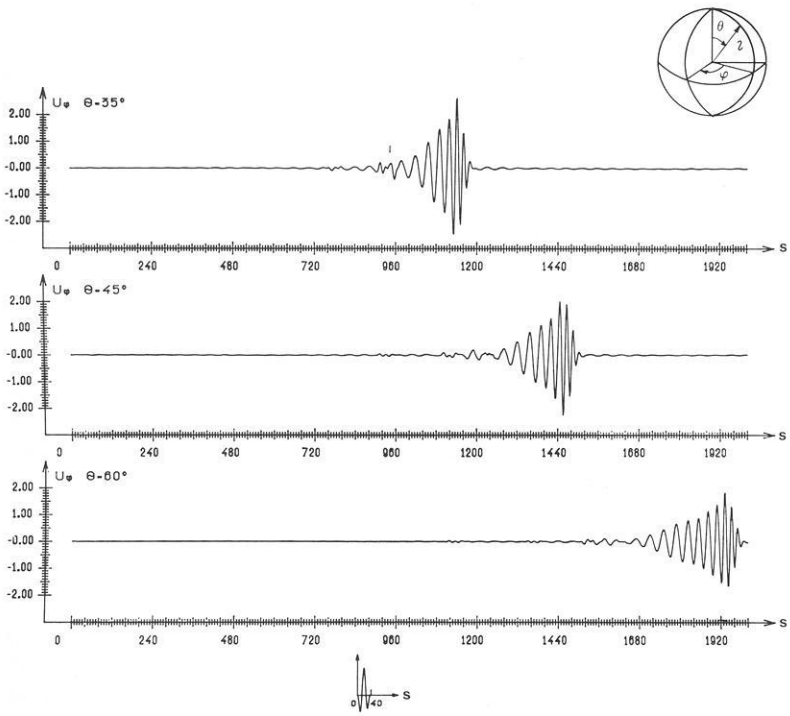


Fig. 11. Theoretical seismograms for the Gutenberg model for different epicentral distances Θ. A torque-type source is located at a depth of 28 km and generates a pulse with a duration of 40 s. The source time-function is shown at the bottom-left

the interface of the second order at the depth of 2.5 km is very weak. As shown in Fig. 7, the shape of the diffracted wave P^* changes with increasing epicentral distance. The amplitude spectrum of the diffracted wave is shifted to lower frequencies. The prevailing frequency of the diffracted wave is about 30% lower than the prevailing frequency of the direct wave in the illuminated region. In Fig. 7, one can also observe the Rayleigh wave (denoted by R^*). The displacement vector of the R^* wave rotates in the rz plane. This rotation can be seen clearly in Fig. 7.

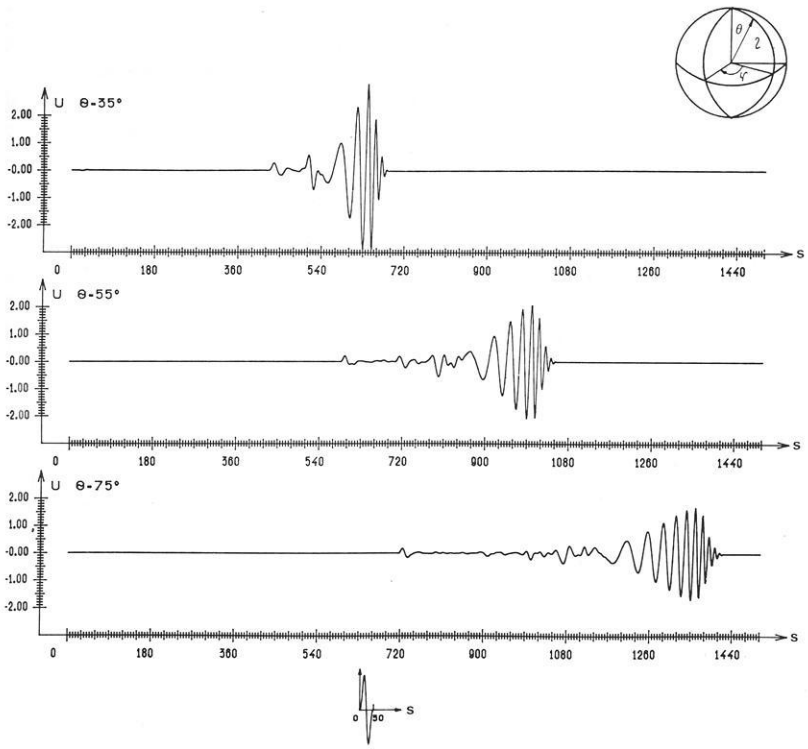


Fig. 12. Theoretical seismograms of P waves for a liquid Earth model similar to the Jeffreys-Bullen model for different epicentral distances Θ. An explosive-type source is located at the free surface and generates a pulse with a duration of 30 s. The source time-function is shown at the bottom-left

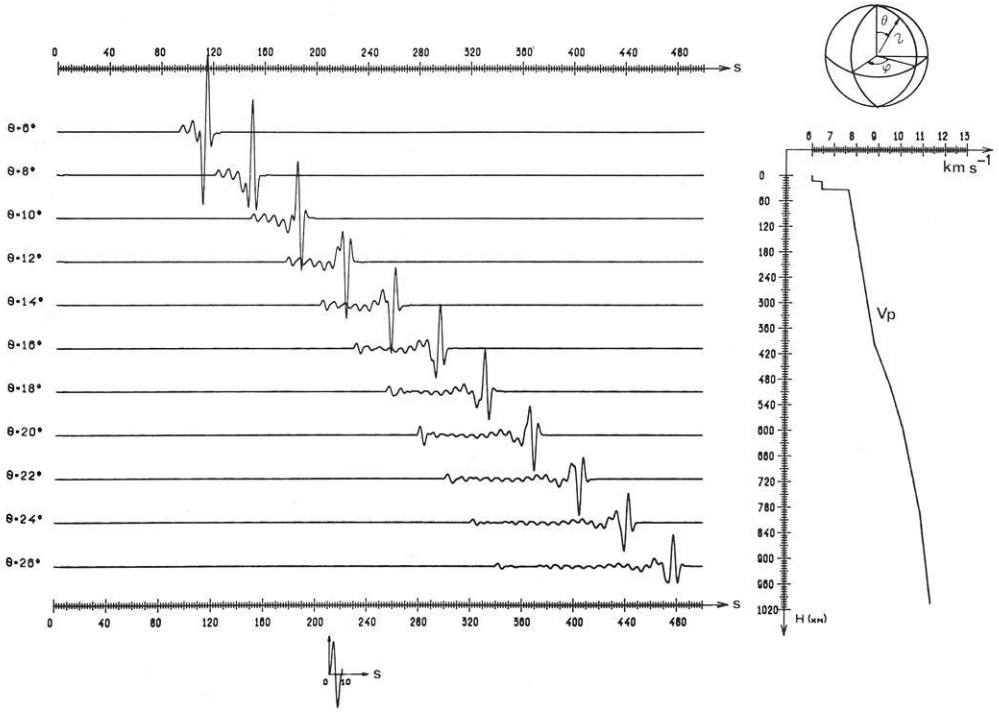


Fig. 13. Theoretical seismograms of P waves for a liquid Jeffreys-Bullen model for different epicentral distances θ. An explosive-type source is located at the free surface and generates a pulse with a duration of 10 s. The source time-function is shown at the bottom-left. At the right-hand side is a P wave velocity-depth graph for the model used for computations

Fifth Example. Now we will consider a symmetrical sub-surface waveguide. The model and corresponding rays are shown in Fig. 8. The source radiates *SH* waves and is located on the free surface. For this model large shadow-zones are formed on the free surface. The shadow zones are separated by illuminated regions or by regions where rays are focussed. The zones of maximum energy correspond to regions where the boundary rays touch the free surface (see Fig. 8). Theoretical seismograms for this model are shown in Fig. 9. The characteristic peculiarity of these seismograms is the disappearance of first arrivals and the increase in amplitudes of second arrivals, S_a (the channel waves).

3.2. Radially Symmetric Medium

First Example. We consider a homogeneous spherical Earth model ($v_s = 6.24$ km/s, $\rho = 4.46$ g/cm^3) with a liquid core. A torque-type source is located at a depth of 300 km and generates a pulse with a period $T = 50$ s. The corresponding theoretical seismograms, at the Earth's surface, for epicentral distances from $\Theta = 20°$ to $\Theta = 150°$ are shown in Fig. 10.

Second Example. In Fig. 11, theoretical seismograms are given for the Earth's surface for epicentral distances of 35°, 45°, and 60° for the Gutenberg model. A torque-type source is located at a depth of 28 km and generates a pulse with a period of 40 s. The amplitudes of the direct *SH* waves are small compared to those of the Love wave and hence are not well seen in the seismograms.

Third Example. In Fig. 12, theoretical seismograms of *P* waves are presented for the Earth's surface for epicentral distances of 35°, 45°, and 60° for an Earth model similar to the Jeffreys-Bullen model. An explosion-type source is located at the free surface and generates a pulse with a period of 30 s. The mantle-refracted wave is recorded as the first arrival.

Fourth Example. In Fig. 13 theoretical seismograms of *P* waves are presented for epicentral distances from 6° to 26° for the Jeffreys-Bullen model. An explosion-type source is located at the free surface and generates a pulse with a period of 10 s. In the Jeffreys-Bullen model the refracted wave is clearly distinguished as the first arrival for all distances. This wave has maximum amplitudes at 20°.

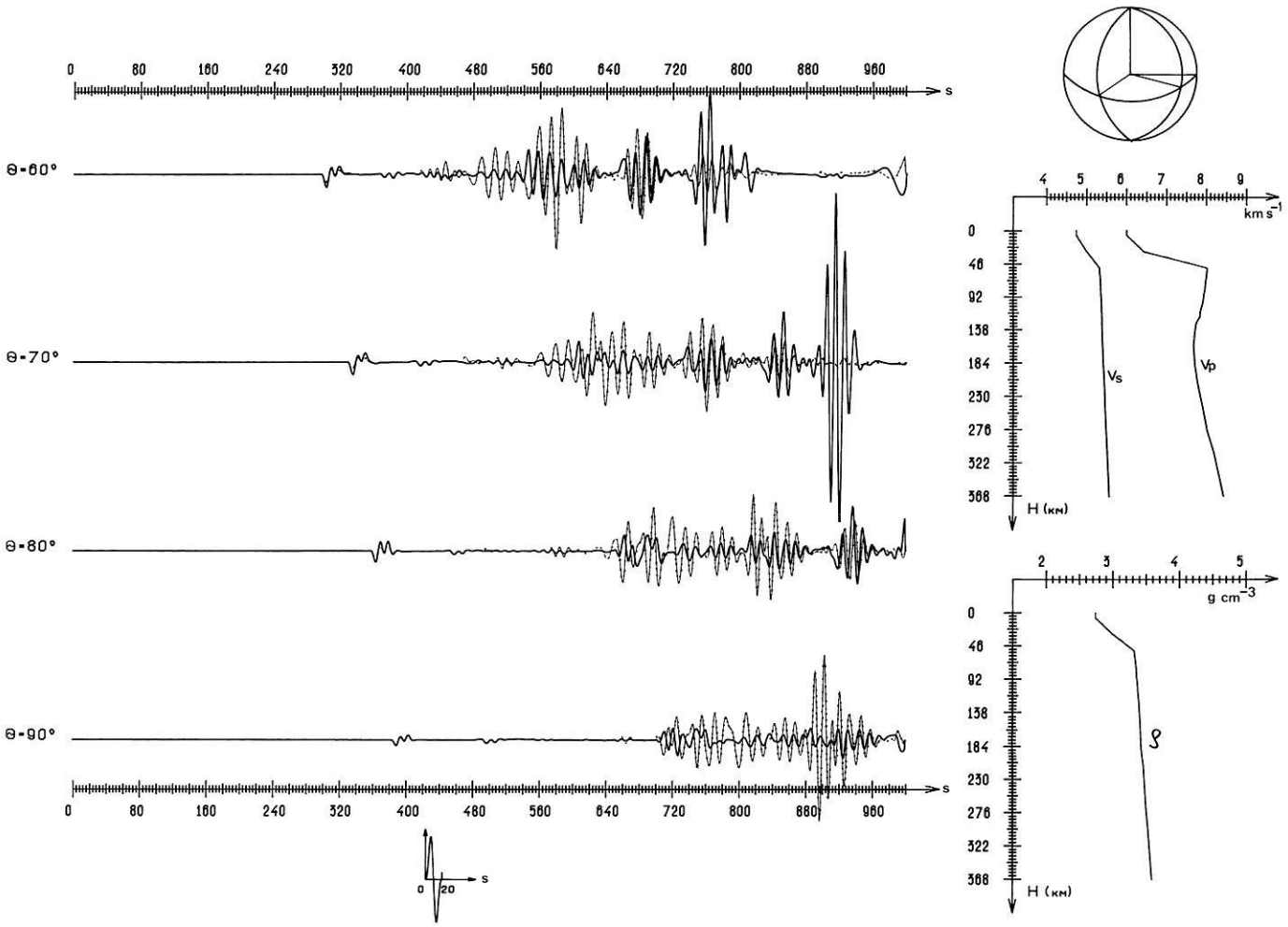


Fig. 14. Theoretical seismograms for the components U_r (*solid curve*) and U_Θ (*dotted curve*) at the free surface of the model on the right for different epicentral distances θ. A vertical-force-type source is located at a depth of about 60 km and generates a pulse with duration of 20 s. The source time-function is shown at the bottom-left. At the right-hand side are *P* and *S* wave velocity-depth graphs (*top*) and density-depth graph (*bottom*) for the model used for computations

In Fig. 14 theoretical seismograms are presented for epicentral distances of 60°, 70°, 80°, and 90° for a vertical-force-type source located at a depth of about 60 km, generating a pulse with a period of 20 s.

3.3. *Diffraction by a Wedge and a Cylinder*

First Example. In Fig. 15, theoretical seismograms for the diffraction of a cylindrical wave by a wedge are presented. An explosion-type source is located at the point ($z=2\lambda$, $\varphi_0=45°$) where λ is the dominant wavelength. The seismograms have been computed for 10 points located on a semicircle with radius $R=2\lambda$ at angular intervals of 20°, beginning with 90° (the first point). The angle of the wedge is $\alpha=315°$.

Second Example. In Fig. 16, the computed wave-field for the diffraction of a cylindrical wave by a cylinder is presented. An explosion-type source is located at the point $d=5\lambda$. The three-dimensional representation of the wave field is constructed in the square abcd outside the cylinder, for a certain instant of time. As can be seen, focussing of the wave-field takes place within the limits ab, as the velocity in the cylinder is less than that outside.

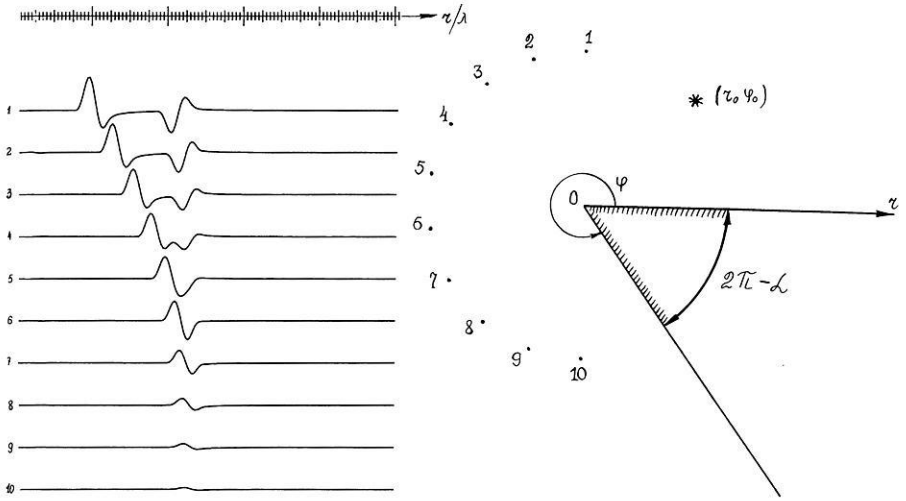


Fig. 15. Theoretical seismograms for 10 points located on a semicircle with radius 2λ (λ is the dominant wavelength) for the diffraction of a cylindrical wave by a wedge. The angle α of the wedge is 305°. The source is located at the point denoted by the asterisk

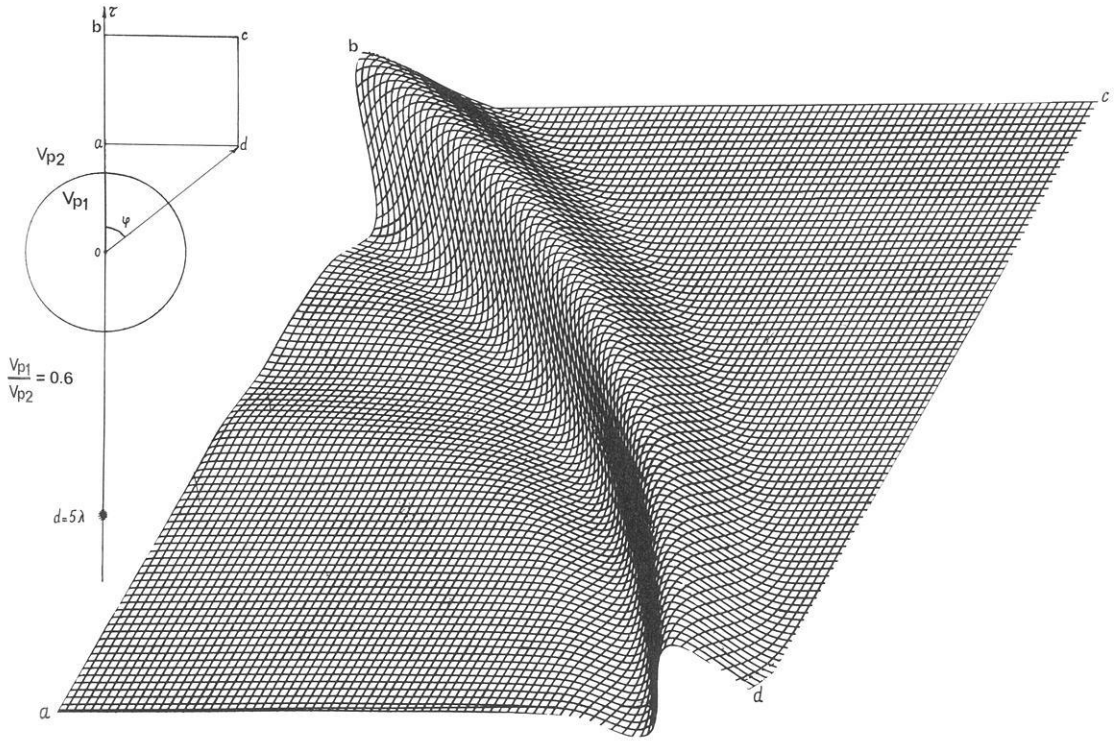


Fig. 16. Three-dimensional wave-field constructed in the square abcd for a certain instant of time, for the diffraction of a cylindrical wave by a cylinder. A explosive-type source is located outside the cylinder at the point 5λ

Conclusion

The method suggested has some advantages compared to known methods of calculation of theoretical seismograms. It does not require large amounts of computer store and is much more efficient and accurate than ordinary finite-difference methods, applicable to the solution of plane and axially-symmetric elasticity-theory problems with coefficients dependent on one space variable. It is not very difficult to apply the method to the calculation of theoretical seismograms for inhomogeneous, anisotropic, viscoelastic media as well as to porous and prestressed media, complete theoretical seismograms being computed in each case.

At present the method has been developed further for the calculation of theoretical seismograms for media whose parameters are arbitrary functions of two or three space variables.

References

Alekseev, A.S., Mikhailenko, B.G.: The solution of Lamb's problem for a vertically inhomogeneous elastic half-space. Izv. Akad. Nauk SSSR, Fiz. Zemli **12**, 11–25, 1976

Alekseev, A.S., Mikhailenko, B.G.: Numerical modelling of seismic waves propagating in a radially inhomogeneous Earth's model. Dokl. Akad. Nauk SSSR **235**, 46–49, 1977

Alekseev, A.S., Mikhailenko, B.G.: The method of calculation of theoretical seismograms for complex media models. Dokl. Akad. Nauk SSSR **240**, 1062–1065, 1978

Alterman, Z., Karal, F.C.: Propagation of elastic waves in layered media by finite difference methods. Bull. Seismol. Soc. Am. **58**, 367–398, 1968

Červený, V., Molotkov, I.A., Pšenčík, I.: Ray method in seismology. Univerzita Karlova: Praha, 1977

Koshlyakov, N.S., Gliner, E.B., Smirnov, M.M.: Partial differential equations of mathematical physics. Moscow: Vysshaya shkola 1970

Mikhailenko, B.G.: Numerical solution of Lamb's problem for inhomogeneous half-space. In: Mathematical problems of geophysics, Vol. 4, M.M. Lavrentiev and A.S. Alekseev, eds: pp. 273–297. Novosibirsk: Computing Center Acad. Sci. USSR (Sib. Division) 1973

Mikhailenko, B.G.: Numerical simulation of elastic wave propagation in the thin layer. In: Mathematical problems of geophysics, Vol. 5/1, M.M. Lavrentiev and A.S. Alekseev, eds.: pp. 187–194. Novosibirsk: Computing Center Acad. Sci. USSR (Sib. Division) 1974

Mikhailenko, B.G.: Calculation of theoretical seismograms for multi-dimensional medium models. In: Uslovno-korrektnye zadachi matematicheskoy fiziki v interpretatsii geofizicheskikh nablyudenii. Izd. VTs SO AN USSR 1978

Mikhailenko, B.G.: The method of solution of dynamic seismic problems for two-dimensional inhomogeneous media models. Dokl. Acad. Nauk USSR **246**, 47–51, 1979

Sneddon, I.: Fourier transforms. New York, Toronto, London 1951

Received August 1, 1978; Revised Version July 15, 1979
Accepted March 17, 1980

J. Geophys. 48, 173–180, 1980

Journal of Geophysics

A Characteristic Method for Numerical Solution of the Inverse Kinematic Seismic Problems*

M.E. Romanov and A.S. Alekseev

Computing Center, Siberian Branch, Academy of Sciences of the USSR, Prospect Nauki 6, 630090 Novosibirsk, USSR

Abstract. The problem of determination of a multi-dimensional velocity function, supposed smoothly dependent on coordinates, from the observed travel-times is considered. An accurate mathematical formulation of this problem is obtained by formulating an inverse problem for a Hamilton-Jacobi-type differential equation. A numerical algorithm is constructed for a medium with velocity increasing monotonically with depth and slightly different from a linear function within any small domain of the medium.

In seismic investigations a problem arises in correcting the initial model of the medium, on the basis of comparison of the model with the observed data. An approach to the solution of this problem employing a linearized formulation of the inverse kinematic problem and a numerical method for the solution of some integral geometry problems are considered.

Questions of solvability, stability and practical applicability of the methods developed are discussed.

Key words: Hamilton formalism – Ray method – Inverse problems – Laterally inhomogeneous media – Numerical solution.

Introduction

At present solution of inverse kinematic problems with the assumption of a one-dimensional law of velocity-distribution in the medium is widely used. Such models of real geological media are, from the present-day viewpoint, imperfect. Some progress has been made in the application of mathematical modelling methods to seismic wave-propagation in complicated media. Inverse problems are solved by many of these methods in two steps:

(a) determination of an initial approximation to the medium;

(b) refinement of the model by optimization methods.

The paper deals with formulations and a numerical method for the solution of these problems in the multidimensional case. Although the problems and the method of solution considered here are of interest in themselves, they are presented under the assumption that the procedures proposed can be applied to seismic investigations[1].

An algorithm is described for the determination of a two dimensional velocity-function, supposed smoothly dependent on coordinates, from the observed travel-times of refracted waves. An accurate mathematical formulation of this problem is equivalent to the problem of the determination of unknown functions in a Hamilton-Jacobi-type differential equation, describing the propagation of refracted waves in an inhomogeneous medium, on the basis of information available about this equation.

In the multidimensional case, the mathematical study of the inverse kinematic problem is one of the basic problems in the theory of improperly-posed problems of mathematical physics (Lavrentiev 1967; Lavrentiev et al. 1970; Romanov 1974a). The accurate mathematical formulation of the inverse kinematic problem considered here was first presented by Belonosova and Alekseev (1967). This paper presents a generalized statement of this problem including the case of three-dimensional medium with smooth interfaces.

Inverse Problems for Refracted Waves

Let us consider a two-dimensional medium, where the travel-velocity distribution for seismic waves is described by the function $v(\xi, \eta)$ depending continuously on the horizontal variable ξ and monotonically increasing with the vertical variable η, which characterizes the depth (Fig. 1). Let us assume that, within any small domain of the medium, the velocity function $v(\xi, \eta)$ differs slightly from the linear function $\bar{v}(\xi, \eta) = v_0 + v_1 \xi + v_2 \eta$, where v_0, v_1, v_2 are constants dependent on the size of the domain and on the properties of the medium in this domain. In other words, the assumption of a sufficiently-smooth change of the medium properties is introduced.

1. As is generally well-known, the travel-time of a refracted wave between the points A and B in the medium equals the integral value

$$I = \int_A^B \frac{dS}{v(\xi, \eta)} \tag{1}$$

* Contributed to the Workshop of the Commission on Controlled Source Seismology (International Association of Seismology and Physics of the Earth's Interior), Karlsruhe, Federal Republic of Germany, August 1–6, 1977

[1] Space does not allow presentation of all mathematical calculations and proofs, therefore presentation of some ideas is schematic. But we hope that this will not present difficulties for advanced readers

0340-062X/80/0173/$01.60

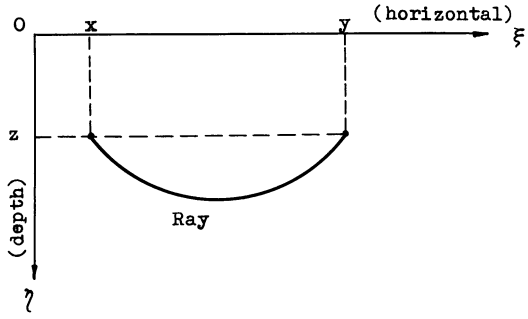


Fig. 1. Orientation of axes and labelling of ray-path in formulation of the inverse problem

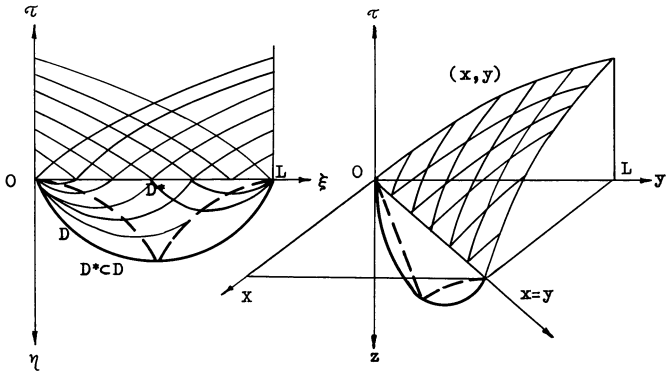


Fig. 2. Travel-time curves and their representation $\varphi(x, y)$, x and y are source and receiver coordinates. D is the lightened domain. D^* is the domain of stability of the inverse-problem solution

taken along the seismic ray-path connecting these points. Let A and B lie on some straight line $\eta = z$ parallel to the axis $O\xi$ and have the coordinates (x, z), (y, z) respectively. Different points A and B on different lines correspond to different values of the travel-time between them, or, in other words, the value of integral (1) is a function of the coordinates of the ends of the ray-path:

$$I = \tau(x, y, z).$$

2. Now let a complete system of travel time curves of refracted waves be given for the profile of observations of length L (or within the interval $[0, L]$ of the medium surface), i.e., the function $\tau(x, y, 0) = \varphi(x, y)$ is known (Fig. 2). The principle of reciprocity of the receiver-point and the shotpoint gives $\varphi(x, y) = \varphi(y, x)$, so one can limit oneself to the case $x \leqq y$. So, in consideration of the travel time-curve system, if $(x, 0)$ are the shotpoint coordinates, then $(y, 0)$ are the receiver coordinates. Assuming monotonic velocity increase with depth for the observational system considered, the seismic rays fill-in the domain D within the medium, limited by the boundary interval $[0, L]$ and the seismic ray joining the boundary points of the profile with the coordinates $(0, 0)$ and $(0, L)$. This domain is often called the domain "lightened" by the given system of travel-time-curves.

The Gaussian curvature of a manifold whose metric is given in the form $ds^2 = v^{-2}(\xi, \eta)(d\xi^2 + d\eta^2)$ equals $K = v^2 \Delta \ln v$. Suppose that everywhere in D, $K \leqq 0$. Then, making use of the Gauss-Bonnet theorem, one can easily show the absence of adjoint points (i.e., intersections of geodesics coming from the same point or caustics of rays). Condition $\partial v/\partial \eta > 0$ provides the return of the ray to the line $\eta = 0$.

3. Let us set the inverse kinematic problem, i.e., determine the velocity function $v(\xi, \eta)$ in the domain D using the observed travel times $\varphi(x, y)$ of refracted waves.

If the point A in (1) has the coordinates (x, z_1), the point B has the coordinates (y, z_2), then $I = T(x, z_1, y, z_2)$ is a function of coordinates of A and B. With A or B fixed, one finds eikonal equations

$$\left(\frac{\partial T}{\partial y}\right)^2 + \left(\frac{\partial T}{\partial z_2}\right)^2 = v^{-2}(y, z_2), \quad \left(\frac{\partial T}{\partial x}\right)^2 + \left(\frac{\partial T}{\partial z_1}\right)^2 = v^{-2}(x, z_1).$$

Consider the functions $z_1(z) = z$, $z_2(z) = z$, $\tau(x, y, z) = T[x, z_1(z), y, z_2(z)]$. Evidently, $\partial\tau/\partial x = \partial T/\partial x$, $\partial\tau/\partial y = \partial T/\partial y$, $\partial\tau/\partial z = \partial T/\partial z_1 + \partial T/\partial z_2$.

Now, expressing the right-hand side of the latter equality from the eikonal equations, having made the necessary substitutions and taken $\partial\tau/\partial z < 0$ into consideration, with $x < y$, we arrive at the equation satisfied by the function $\tau(x, y, z)$:

$$\frac{\partial \tau}{\partial z} + \sqrt{f_1^2(x, z) - \left(\frac{\partial \tau}{\partial x}\right)^2} + \sqrt{f_2^2(y, z) - \left(\frac{\partial \tau}{\partial y}\right)^2} = 0. \tag{2}$$

Here f_1 and f_2 are related to the slowness at the points (x, z), (y, z):

$$f_1(x, z) = -v^{-1}(x, z), \qquad f_2(y, z) = v^{-1}(y, z).$$

An accurate mathematical formulation of the inverse kinematic problem is given as the problem of determining the unknown functions f_1, f_2 in the Hamilton-Jacobi-type Eq. (2), if

$$\tau(x, y, 0) = \varphi(x, y), \qquad 0 \leqq x \leqq y \leqq L < \infty. \tag{3}$$

4. With the above assumptions about the velocity, the characteristic method is applied for the numerical solution of the inverse problem as formulated, whose essence is as follows: with f_1 and f_2 given [that is, given the function $v(\xi, \eta)$], the solution of Cauchy problem (2)–(3) is equivalent to the solution of the characteristic system of ordinary differential equations

$$\frac{dx}{dz} = -\frac{p}{\sqrt{f_1^2 - p^2}}, \qquad \frac{dy}{dz} = -\frac{q}{\sqrt{f_2^2 - q^2}},$$

$$\frac{dp}{dz} = -\frac{f_1 \dfrac{\partial f_1}{\partial x}}{\sqrt{f_1^2 - p^2}}, \qquad \frac{dq}{dz} = -\frac{f_2 \dfrac{\partial f_2}{\partial y}}{\sqrt{f_2^2 - q^2}} \tag{4}$$

with the initial conditions

$$x(0) = x^0, \qquad y(0) = y^0, \quad 0 \leqq x^0 \leqq y^0 \leqq L < \infty,$$

$$p(0) = p^0 = \left.\frac{\partial \varphi}{\partial x}\right|_{\substack{x = x^0 \\ y = y^0}}, \qquad q(0) = q^0 = \left.\frac{\partial \varphi}{\partial y}\right|_{\substack{x = x^0 \\ y = y^0}}. \tag{5}$$

A pair of functions $x(z)$, $y(z)$ defines a seismic ray (its ascending and descending branches respectively) joining the points (x, z), (y, z). The travel-time $\tau\,[x(z), y(z), z]$ is found from the equation

$$\frac{d\tau}{dz} = -\frac{f_1^2}{\sqrt{f_1^2 - p^2}} - \frac{f_2^2}{\sqrt{f_2^2 - q^2}} \tag{6}$$

with the initial condition

$$\tau(0)=\tau^0=\varphi(x^0, y^0). \tag{7}$$

Therefore the inverse problem is to determine the functions $f_1(x, z)$, $f_2(y, z)$, which are part of the right-hand sides of system (4), (6), using the data (5), (7). The domain P_k will be determined in the form of a band, cut out of the domain D by the straight lines $\eta=kh$ and $\eta=(k+1)h$, where h is the value of the numerical integration-step of system (4), (6), $k=0, 1, 2, \ldots$

5. The numerical method for solution of the inverse problem uses successive recalculation of initial data (5), (7) along the characteristics (rays), determined by the system of Eqs. (4), by local determination of the functions f_1, f_2 in the domains P_k using the following algorithm.

(a) Take $k=0$ (i.e., the given band P_0, where $0\leqq z\leqq h$). Let us determine within the domain $0\leqq x\leqq y\leqq L<\infty$ with $\eta=z=0$ a set $M^{(k)}$ of discrete points (x^0, y^0) rather densely distributed. For example, let (x^0, y^0) coincide with the values of coordinates of shotpoints and receivers respectively on the profile; the values p^0, q^0, τ^0 for all the points of the set $M^{(k)}$ are calculated with regard to (5), (7).

(b) Let us specify a certain number δ and take two sets $M_\delta^{(k)}$ and $M_L^{(k)}$ from $M^{(k)}$ with the condition that the point (x^0, y^0) belongs to $M_\delta^{(k)}$ if $y^0-x^0\leqq\delta$ and it belongs to $M_L^{(k)}$ if $y^0-x^0>\delta$ (δ is chosen so that the points of $M_\delta^{(k)}$ are uniformly distributed along the profile and the above assumption, that $|v-\bar{v}|$ is small for all the points of $M_\delta^{(k)}$, is satisfied.

(c) Let us consider the respective Cauchy problem (4) to (7) for the points $M_\delta^{(k)}$. With the given assumptions of locality and the choice of $M_\delta^{(k)}$, in the vicinity of each of (x^0, y^0), $v(\xi, \eta)$ differs little from $\bar{v}(\xi, \eta)$. Therefore, setting $v=\bar{v}$ in (4), (6), relations expressing the parameters v_0, v_1, v_2 in terms of p^0, q^0, τ^0, x^0, y^0 are found (Romanov 1972), i.e., in the vicinity of each point $(x^0, y^0)\in M_\delta^{(k)}$ on the plane $z=kh$, parameters v_0, v_1, v_2 of local approximations $\bar{v}(\xi, \eta)$ to the unknown velocity function $v(\xi, \eta)$ are determined within the band P_k.

(d) Applying a method of smoothing by spline functions we sew the local approximations $\bar{v}(\xi, \eta)$ into the smooth function $v(\xi, \eta)$ within the band P_k.

A stable solution to the inverse problem can be obtained only as follows. Within the domain (Fig. 2) where the solution of the inverse problem is sought, we define the domain D^* as limited by the set of points of maximum depth on each ray on the given observational profile. The function $v(\xi, \eta)$ can be stably determined [according to items (c), (d)] only on the intersection of the band P_k and D^*. Therefore we call D^* the domain of stability of the inverse problem solution.

(e) To determine $v(\xi, \eta)$ outside the domain D^* we use extrapolation or some additional relations for characteristics (see Sect. 3).

(f) Substituting the function obtained, $v(\xi, \eta)$, in the right-hand sides of Eqs. (4) and (6), on the band P_k, the Cauchy problem is solved numerically, with (5) and (7) determined for each point of $M_L^{(k)}$. As a result we have values of the functions $x(z)$, $y(z)$, $p(z)$, $q(z)$, $\tau(z)$ with $z=(k+1)h$. Now a new set $M^{(k+1)}$ is formed by the points $x(z)$, $y(z)$ of the three-dimensional space $\{x, y, z\}$. When solving the Cauchy problem (4) to (7) some characteristics cross the plane given by the equation $x=y$ in three-dimensional space. As a rule these are characteristics originating from points in $M_\delta^{(k)}$. In further calculations those characteristics or rays, whose depth of maximum penetration does not exceed $z=(k+1)h$, do not take part, therefore $M^{(k+1)}$ contains a smaller number of points than $M^{(k)}$.

(g) Changing k into $k+1$, one should come back to item (b) if $M^{(k+1)}$ contains at least one point.

Thus, the algorithm described realizes a recurrent process of successive determination of $v(\xi, \eta)$ within the domain D of uniqueness of the solution of the inverse problem. Here a stable function $v(\xi, \eta)$ is generated in the domain D^*; the accumulation of resultant errors is generally caused by the necessity of using unstable extrapolation procedures beyond the stability domain D^*.

Nevertheless, note that in actual observational systems (especially in seismic prospecting) the range of observational profiles exceeds the maximum distance between the source and receiver in this system: $\max\limits_{(x^0, y^0)}|y^0-x^0|\ll L$. Then the domain of instability $D\setminus D^*$ is significant for determination of the velocity in D^*. Therefore the algorithm with linear extrapolation of the velocity function from D^* to $D\setminus D^*$ is to be applied, for the interpretation of the data from such observational systems.

Method of Refinement for Models of the Medium

The inverse kinematic seismic problem is nonlinear, since seismic ray-paths are to be determined along with the velocity function. In some cases the linear inverse problem can be formulated.

In seismic studies an approximate model of the inhomogeneous medium investigated, or some velocity function $v_0(\xi, \eta)$ can be given on the basis of some a priori data or by approximate methods employing different information in the observed wave field. In particular, the inverse kinematic problem of the determination of the velocity function $v_0(\xi, \eta)$ making use of the observed travel time curves of refracted waves $\varphi(x, y)$ has been considered (see Sect. 1).

Application of the characteristic method results in the function $v_0(\xi, \eta)=v(\xi, \eta)-\Delta v(\xi, \eta)$ where $\Delta v(\xi, \eta)$ is a velocity variation – smaller than $v_0(\xi, \eta)$.

Now consider the problem of determination of $\Delta v(\xi, \eta)$ from the difference $\Delta\varphi(x, y)$ of the observed travel time curves $\varphi(x, y)$ and calculated travel time curves $\varphi_0(x, y)$. This problem was considered in Romanov (1974a) and in Lavrentiev and Romanov (1966) and is reduced to the so-called "problem of integral geometry", if the velocity function $v_0(\xi, \eta)$ is such that the seismic ray $\gamma(x^0, y^0)=\gamma^0$ connecting the points $(x^0, 0)$ and $(y^0, 0)$ is strictly inside that part of the medium that is limited by the ray connecting $(x^1, 0)$ and $(y^1, 0)$ with the condition $0\leqq x^1<x^0<y^0<y^1\leqq L$. This is a linear inverse problem since γ^0 is given. It can be solved numerically by the characteristic method (Romanov 1975).

In fact, let us (similarly to Sect. 1) introduce the function

$$\psi(x, y, z)=\int\limits_{\gamma^0}\Delta n(\xi, \eta)\,ds, \qquad \Delta n=-\frac{\Delta v}{v_0(v_0+\Delta v)}$$

where γ^0 is the seismic ray in the medium with the velocity $v_0(\xi, \eta)$ connecting the points with the coordinates (x, z), (y, z).

Then we arrive at Hamilton-Jacobi equation (see Romanov 1975).

$$\frac{\partial\psi}{\partial z}+\cot\Theta_1\frac{\partial\psi}{\partial x}+\cot\Theta_2\frac{\partial\psi}{\partial y}-\frac{f_1}{\sin\Theta_1}-\frac{f_2}{\sin\Theta_2}=0 \tag{9}$$

satisfied by the function $\psi(x,y,z)$. Here Θ_1, Θ_2 are the angles formed by the axis $O\xi$ and the tangents at the points (x,z), (y,z) of the ray γ^0.

The functions

$$f_1(x,z)=-\Delta n(x,z), \quad f_2(y,z)=\Delta n(y,z)$$

are to be determined by the given function

$$\psi(x,y,0)=\Delta\varphi(x,y) \tag{10}$$

and the rays γ^0 determining the angles Θ_1, Θ_2.

Thus we have an inverse problem similar to that considered in Sect. 1 with the only difference being that the rays in (8) for the model v_0 are given. Hence one can apply the above characteristic method for determination of the velocity model correction $\Delta v(\xi,\eta)$.

Note that initial condition (10) of this problem is approximate if $\psi(x,y,z)$ is introduced in accordance with (8). Condition (10) is determined to within small values of the order of $(\Delta v)^2$ (Romanov 1974a). Therefore, considering $v_1(\xi,\eta)=v_0(\xi,\eta)+\Delta v(\xi,\eta)$ as a new approach to the real velocity function, a new correction can be determined. In other words the method of refinement described for the model of the medium can become the basis of the method of successive approximations for the solution of a multidimensional inverse kinematic problem.

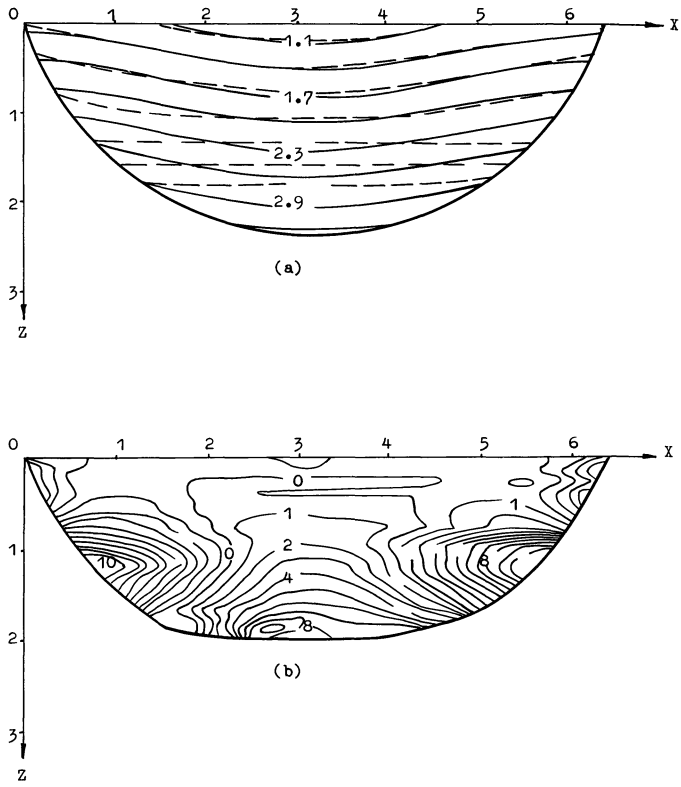


Fig. 3a and b. *Test I.* **a** Isolines of the velocity function $v(x,z)=1.1+0.2\cos(x)+z$ (*solid lines*) and regenerated function (*dashed lines*). **b** Isolines of the error function (in %) show the domain D^*

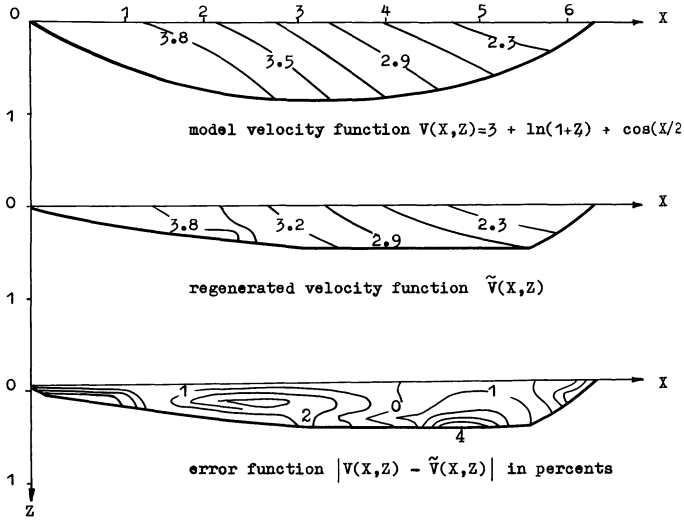


Fig. 4. *Test II:* Isolines of test functions

Discussion of the Results

1. The investigation of the multidimensional inverse kinematic problem is closely connected to that of inverse problems for differential equations and a number of papers of theoretical character have been devoted to the determination of conditions for the uniqueness of solutions. From the point of view of practical applications, the statement of the inverse kinematic problem with initial data on that part of the boundary of the domain where solutions are being sought is most important. It is in this problem that there are some principal difficulties, since the inverse problem here belongs to the class of improperly-posed problems.

In a case where the initial data are given on the whole boundary of the domain of solution, as recent investigations have shown (Mukhometov 1975), the inverse problem appears correct. Geophysical interpretation of this result is that, if the travel-times of a seismic wave between any pairs of points of the boundary of domain, where solution is being sought, are known, the velocity function is determined, stable and unique if it is such that the corresponding family of rays have no caustics, $\Delta \ln v \leqq 0$.

First results on the multidimensional inverse kinematic seismic problem were obtained (Lavrentiev and Romanov 1966; Romanov 1974a) making use of a linearized method. The velocity distribution of seismic waves on the Pamir-Baikal profile was investigated on this basis (Alekseev et al. 1971). Uniqueness of the solution in the class of analytic functions was proved in Anikonov (1969, 1971). In Romanov (1974b) a fairly wide class of functions with unique solutions has been presented, that is the class of functions $n(x,y)\in C^3(D)$, satisfying inequalities $a\leqq n(x,y)\leqq b$, $a\leqq -n'_y(x,y)\leqq b'$ in D and presented in the form

$$f[n(x,y)]=\sum_{k=1}^{N}\varphi_k(x)\psi_k(y)$$

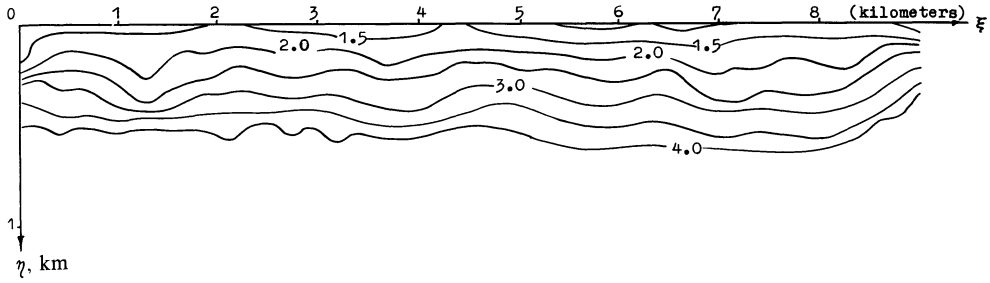


Fig. 5.
Example of the velocity cross-section based on seismic prospecting data and regenerated by the characteristic method (values of isolines in km/s)

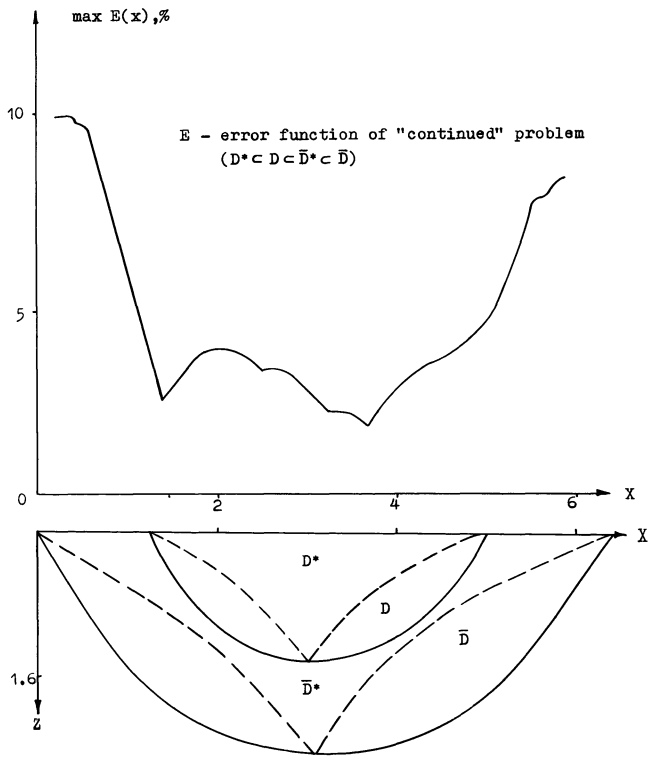


Fig. 6. Results of the regularization experiment for *Test I*. *Dashed lines* correspond to boundary of stability domains, D, D^* are uniqueness and stability domains of the original problem, $\bar{D}$, $\bar{D}^*$ are the same domains of the continued problem. $\bar{D}^*$ contains D, therefore the error-function in D of the continued problem is less than the same function of the original problem (see Fig. 3b)

where

$n(x, y) = v^{-1}(x, y), \quad x = (x_1, \ldots, x_m),$

$D = \{(x, y): |x| < \infty, 0 \leqq y \leqq H\},$

$\varphi_k \in C^3(R^m), \quad \psi_k \in C^3[0, H], \quad a', b'.$

a, b, H are positive constants, $f(z) \in C^3[a, b]$, $|f'(z)| > 0$, the variable y being the depth.

Another approach to the investigation of uniqueness of solution of multidimensional inverse problems is described in Anikonov (1971). For the case where the travel time curves satisfy some differential equation, then the velocity function satisfies a corresponding differential equation. On the basis of this approach a number of particular solutions of the multidimensional inverse kinematic problem are obtained (Anikonov and Shasheva 1971; Anikonov 1974). In Jobert (1973) a method for an approximate inversion of the travel-time curves is presented for the three-dimensional case where the surfaces of equal velocity are planes with an infinitesimal dip. In this case our method gives a complete solution for any dips.

2. In our opinion the approach considered here for the solution of the multidimensional inverse kinematic problem has possibilities applicable to the creation of efficient algorithms. A set of programs in ALGOL-60, implementing the above characteristic method for processing real data, obtained from observational systems, was created in Novosibirsk Computing Center, Siberian Branch of the USSR Academy of Sciences. In Figs. 3 and 4 results of test calculations by the characteristic method are shown. An example of isolines of velocity functions generated from the real velocity data are shown in Fig. 5. Data from prospecting profiles, of average length 130 km, have been processed by the characteristic method. The results obtained were used for estimating the velocity parameters of the upper part of the cross-sections of Siberian platform.

3. In the course of numerical experiments it was found that in the uniqueness domain D of the inverse problem the stability domain D^* is distinctly determined. This domain is the set of deeper points of rays in the given finite interval $[0, L]$ of observations $\varphi(x, y)$. Outside the domain D^* the solution of the problem is generally unstable.

If the initial data $\varphi(x, y)$ for problem are given on the whole boundary of the halfspace $\eta \geqq 0$, then $D = \{(\xi, \eta): \eta \geqq 0\}$ and $D^* = D$. In this case the problem is correct (Mukhometov 1975), since the instability domain is absent. But in our problem the data are given on the segment $[0, L]$ thus giving rise to the instability domain $D \setminus D^*$. Therefore our problem belongs to the class of improperly-posed problems. In order to obtain the solution of the improperly-posed problem one should employ a proper regularization method.

However, taking the remark at the end of Sect. 1 into account we can limit ourselves by linear extrapolation of the values f_1 (or f_2) from D^* to $D \setminus D^*$.

The problem can be regularized in three ways. Firstly, if the initial data are continued smoothly beyond the interval of observations, then the stability domain of the continued problem contains the uniqueness domain of the original problem (Fig. 6). In this case an implicit connection between velocity values in D^* with velocity values outside is introduced. Numerical analysis (Romanov 1972) has shown that stability in D can be essentially increased in this way. Secondly, fixing the class of velocity functions where the solution is being sought, one may try to find integrals of Hamilton system (4), (6), thus introducing

additional relations on characteristics, determining an explicit connection between the values f_1, f_2 in D. Thirdly, one should consider additionally the variation along the rays of the value

$$r=\frac{\partial\tau}{\partial z}=-\sqrt{f_1^2-p^2}-\sqrt{f_2^2-q^2}.$$

If the value $r(z)$ on the ray is known then the value f_2 is determined from this relation making use of the value f_1 in D^* (or, on the contrary, making use of the value f_2 in D^*, f_1 is determined outside D^*). In this case the Hamilton-Jacobi equation itself is the additional relation which connects the values f_1 and f_2 at the ends of rays.

Stability is also affected by errors arising in numerical implementation of the method: these are errors in obtaining local approximations by the linear functions $\bar{v}(\xi,\eta)$, errors in numerical integration of Eqs. (4), (6) and errors caused by inaccurate initial data (5), (7). However, their estimations require a special discussion and here we note only that an error introduced in local approximation is easily estimated by the method described in Sect. 2. This error is proportional to h^2 and depends on estimations of higher derivatives of the velocity function. To decrease this error in the program for solution of the inverse problem, two blocks are provided: a spline-smoothing block and a block for integration of Eqs. (4), (6) at $v\equiv\bar{v}$ with the step $0.1\times h$.

The numerical method described for solving the inverse problem is known to have an error of second order in approximation.

4. As compared to other approximate methods employed in seismic practice the methods considered are based on the accurate statement of the inverse problem under conditions providing uniqueness of the inverse-problem solution, i.e., under the a priori assumption of a velocity increase with depth. It is essential that the function of two variables $\varphi(x,y)$ given in the domain $P=\{(x,y)\colon 0\leqq x\leqq y\leqq L<\infty\}$ is used as the data. The characteristic method considered is based on the discrete set of points $(x^0,y^0)\in P$ where the values $\varphi(x^0,y^0)$, $\partial\varphi(x^0,y^0)/\partial x$, $\partial\varphi(x^0,y^0)/\partial y$ should be given. Therefore, a dense distribution of the points (x^0,y^0) is necessary both for qualitative approximation of $\varphi(x,y)$ and for the determination of the derivatives $\partial\varphi/\partial x$, $\partial\varphi/\partial y$ and the details of the velocity law $v(\xi,\eta)$.

The greater the density of points (x^0,y^0) in the observational system, the better the results.

5. The inverse problem can be also solved when the function $\varphi(x,y)$ is ambiguous (in the presence of loops). In this case, from some point $(x^0,y^0)\in P$, where $\varphi(x,y)$ is ambiguous, rays originate at various angles [different pairs p^0, q^0 in (5)].

6. Sections 1 and 2 describe concrete applications of the theory of Hamilton formalism to the solution of problems in the case of a two-dimensional medium. This formalism can also be applied in the three-dimensional case. In this case the vector-functions

$$x(z)=[x_1(z),x_2(z)],\ y(z)=[y_1(z),y_2(z)],\ p(z)=[p_1(z),p_2(z)],$$

$q(z)=[q_1(z),q_2(z)]$ should be considered instead of $x(z)$, $y(z)$, $p(z)$, $q(z)$.

Equation (2) also describes the kinematics of reflected waves (see Sect. 4) and the kinematics of waves in a medium with curvilinear interfaces. Therefore the development of this approach to kinematic problems opens possibilities for the development of computerized systems of real seismic-data-processing in for example reflection and refraction profiling and seismic deep sounding.

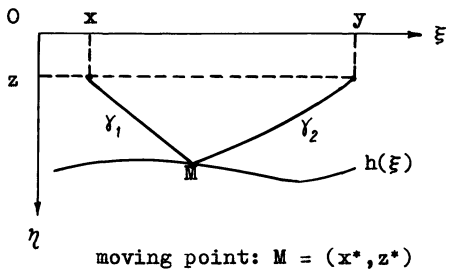


Fig. 7. Orientation of axes and labelling of ray-path and velocity-interface in formulation of the variational problem

The Characteristic Method – an Application of Hamilton Formalism to Kinematic Problems for Multi-Dimensional Media

Let us consider some general ideas on the application of Hamilton formalism to the kinematic seismic problems. For the sake of simplicity we will restrict ourselves to a two-dimensional medium with one smooth interface, since all the results are valid for three-dimensional and multi-interfaced media.

Let (ξ,η) be a point in the half-space $\eta\geqq 0$ (Fig. 7), $v(\xi,\eta)$—a velocity function with discontinuities on the interface $\eta=h(\xi)$. Consider a curve γ connecting those points in the medium with the coordinates (x,z), (y,z) and consisting of two branches $\gamma_1=\{(\xi,\eta)\colon\ \xi=x(\eta),\ 0\leqq z\leqq\eta\leqq z^*\}$ and $\gamma_2=\{(\xi,\eta)\colon\ \xi=y(\eta),\ 0\leqq z\leqq\eta\leqq z^*\}$. For all $\eta\in[z,z^*]$, $x(z^*)=y(z^*)=x^*$, i.e., (x^*,z^*) are the coordinates of the intersection of the branches. Let us define a plane G in the three-dimensional space of the variables $\{x,y,z\}$ by the equation $x=y$ and the interface $\Gamma=\{(x,y,z)\colon\ x=y=\xi,\ z=h(\xi)\}$. Then all the kinematic seismic problems connected with waves refracted and/or reflected from Γ lead to the consideration of a variational problem of minimizing the functional $J(\gamma)$ with a moving point (x^*,x^*,z^*):

$$J(\gamma)=\int_{z^*}^{z}L(x(\eta),y(\eta),\dot{x}(\eta),\dot{y}(\eta),\eta)\,d\eta \tag{11}$$

where

$$L=f_1[x(\eta),\eta]\sqrt{1+\dot{x}^2(\eta)}-f_2[y(\eta),\eta]\sqrt{1+\dot{y}^2(\eta)},$$

$$\dot{x}=\frac{dx}{d\eta},\qquad \dot{y}=\frac{dy}{d\eta}, \tag{12}$$

$$f_1[x(\eta),\eta]=-\frac{1}{v(x(\eta),\eta)},\qquad f_2[y(\eta),\eta]=\frac{1}{v(y(\eta),\eta)}.$$

For the case of refracted rays the moving point is $(x^*,x^*,z^*)\in G$ and for the case of the rays reflected from Γ, the moving point belongs to the manifold Γ.

If the curve γ is an extremal, i.e., if it realizes the minimum of the functional $J(\gamma)$, the pair of functions $[x(z),y(z)]$ form a seismic ray: either refracted, if $(x^*,x^*,z^*)\in G$, or reflected from Γ,

if $(x^*, x^*, z^*) \in \Gamma$. Here $J(\gamma)$ becomes a function of the coordinates of the finite end-points of the extremal γ: $J(\gamma) = \tau(x, y, z)$; the value of this function coincides with the travel-time from the point (x, z) to the point (y, z) and is sometimes called the geodesic distance or eikonal (Courant 1962). The point (x^*, x^*, z^*) along with manifolds G and Γ will be called initial according to $\tau(x^*, x^*, z^*) = 0$. On the initial manifolds G and Γ there are transversality conditions, which are obtained from stationary state conditions of the functional $J(\gamma)$ with a moving end. In the case of refracted rays we have

$$\dot{x}(z^*) = -\dot{y}(z^*) = \infty \tag{13}$$

and in the case of reflected rays (Elsgolz 1969) one can easily obtain

$$\frac{h' + \dot{x}}{\sqrt{1+\dot{x}^2}} + \frac{h' + \dot{y}}{\sqrt{1+\dot{y}^2}} = 0, \qquad h' = \left.\frac{dh}{d\xi}\right|_{\xi = x^*} \tag{14}$$

as the condition, describing the law of ray reflection for the moving end.

Considering the function $\tau(x, y, z)$, introduced in the way described above, a theory of direct and inverse kinematic problems in an inhomogeneous medium with interfaces can be developed in the most natural form. Here one may use Hamilton formalism, developed in analytical mechanics, in variational calculus, in the general theory of equations of the first order partial derivatives, as well as in other fields of mathematics, for example, in the theory of optimal control. The application of the results of the latter is especially interesting from the point of view of creating algorithms for the solution of kinematic problems.

It is known (see Sect. 1) that $\tau(x, y, z)$ satisfies the Hamilton-Jacobi equation

$$\frac{\partial \tau}{\partial z} + \sqrt{f_1^2(x, z) - \left(\frac{\partial \tau}{\partial x}\right)^2} + \sqrt{f_2^2(y, z) - \left(\frac{\partial \tau}{\partial y}\right)^2} = 0 \tag{15}$$

where $f_1 = -v^{-1}(x, z)$, $f_2 = v^{-1}(y, z)$. If $z = 0$ is the Earth's surface, then in direct seismic problems the problem of determining the function

$$\tau(x, y, 0) = \varphi(x, y), \qquad 0 \leqq x \leqq y \leqq L \tag{16}$$

is posed, with a function of velocity distribution $v(\xi, \eta)$ and the interface Γ given (or functions f_1, f_2 in Hamilton-Jacobi equation). Here the function $\varphi(x, y)$ presents a complete set of travel time curves of refracted or reflected waves.

The inverse problem, of the determination of the velocity function $v(\xi, \eta)$ and the interface Γ, can now be posed, using the observed travel time curves, as an inverse problem for the differential Hamilton-Jacobi equation.

In kinematic problems it is natural to pass on to the system of ordinary differential equations equivalent to the Hamilton-Jacobi equation and to determine characteristic curves (from the Hamilton system we have $dz/ds = 1$, where s is a parameter, therefore the parameter s is identified with z)

$$\frac{dx}{dz} = \frac{\partial H}{\partial p}, \quad \frac{dy}{dz} = \frac{\partial H}{\partial q}, \quad \frac{dp}{dz} = -\frac{\partial H}{\partial x}, \quad \frac{dq}{dz} = -\frac{\partial H}{\partial y}, \tag{17}$$

$$\frac{d\tau}{dz} = p\frac{\partial H}{\partial p} + q\frac{\partial H}{\partial q} - H \tag{18}$$

$$\frac{dr}{dz} = -\frac{\partial H}{\partial z} \tag{19}$$

where r is a dual variable with respect to z. The system (17) is complete and if the solution of the system is known, then $\tau(z)$ can be found by integration from (18). Equation (19) determines the characteristic function $r(z)$, which can help in solving the inverse problem. In (17)–(19) the Hamilton function is expressed with independent variables

$$H = H(x, y, p, q, z) = \sqrt{f_1^2(x, z) - p^2} + \sqrt{f_2^2(y, z) - q^2}. \tag{20}$$

In direct problems the Cauchy problem is considered for the canonical Hamilton system (17) and Eq. (18), with the initial conditions defined at the moving end. For refracted waves these initial conditions are of the form

$$x(z^*) = x^*, \quad y(z^*) = x^*, \quad p(z^*) = f_1(x^*, z^*),$$
$$q(z^*) = f_2(x^*, z^*), \quad \tau(z^*) = 0. \tag{21}$$

For reflected waves the initial conditions are different only for p and q:

$$p(z^*) = \frac{f_1(x^*, z^*)}{\sqrt{1+\dot{x}^2(z^*)}}\,\dot{x}(z^*),$$
$$q(z^*) = \frac{-f_2(x^*, z^*)}{\sqrt{1+\dot{y}^2(z^*)}}\,\dot{y}(z^*), \quad z^* = h(x^*). \tag{22}$$

The values $\dot{x}$, $\dot{y}$ are given in accordance with (14) and determine various values of the reflection angles from the same interface point.

Note that in the case of a three-dimensional medium x, y, p, q are vector-functions $x = (x_1, x_2)$, $y = (y_1, y_2)$, $p = (p_1, p_2)$, $q = (q_1, q_2)$.

In inverse problems unknown functions $f_1(x, z)$, $f_2(y, z)$ [or the function v, see (12)] are included in the right-hand sides of (17) - (19) but the final conditions with $z = 0$ are given:

$$x(0) = x^0, \quad y(0) = y^0, \quad p(0) = \left.\frac{\partial \varphi}{\partial x}\right|_{\substack{x = x^0 \\ y = y^0}},$$
$$q(0) = \left.\frac{\partial \varphi}{\partial y}\right|_{\substack{x = x^0 \\ y = y^0}}, \quad \tau(0) = \varphi(x^0, y^0).$$

Some assumption (see Sect. 1) about the medium and a particular choice of the points (x^0, y^0) on the plane $z = 0$ allow one to determine f_1, f_2 approximately, thus reducing the inverse problem to the Cauchy problem similar to (17), (18), (21), the integration of system (17)–(18) being in reverse order (i.e. from the final to the initial point).

Thus the geodesic distance $\tau(x, y, z)$, introduced in the above manner, allows one to consider kinematic seismic problems in inhomogeneous media from a uniform viewpoint.

The approach described is called a characteristic method since its essence lies in the use of the notion of the characteristic curve, i.e., of the ray $[x(z), y(z)]$, along with travel time $\tau(z)$.

References

Alekseev, A.S., Lavrentiev, M.M., Mukhometov, R.G., Nersesov, I.L., Romanov, V.G.: A numerical method of determination of the structure of the Earth's upper mantle. In: Mathematical problems of geophysics, Vol. 2, M.M. Lavrentiev and A.S. Alekseev, eds.: pp. 143–165. Novosibirsk: Computing Center, Acad. Sci. USSR (Sib. Branch) 1971

Anikonov, Yu.E.: On the uniqueness of the solution of inverse problems. In: Mathematical problems of geophysics, Vol. 1, M.M. Lavrentiev and A.S. Alekseev, eds.: pp. 26–40. Novosibirsk: Computing Center, Acad. Sci. USSR (Sib. Branch) 1969

Anikonov, Yu.E.: On geometric methods of investigation of inverse problems. In: Mathematical problems of geophysics, Vol. 2, M.M. Lavrentiev and A.S. Alekseev, eds.: pp. 7–63. Novosibirsk: Computing Center, Acad. Sci. USSR (Sib. Branch) 1971

Anikonov, Yu.E.: Some partial solutions of the inverse kinematic problem. In: Mathematical problems of geophysics, Vol. 4, M.M. Lavrentiev and A.S. Alekseev, eds.: pp. 30–60. Novosibirsk: Computing Center, Acad. Sci. USSR (Sib. Branch) 1974

Anikonov, Yu.E., Shasheva, N.P.: Formulae for the inverse kinematic seismic problems. In: Mathematical problems of geophysics, Vol. 2, M.M. Lavrentiev and A.S. Alekseev, eds.: pp. 57–68. Novosibirsk: Computing Center, Acad. Sci. USSR (Sib. Branch) 1971

Belonosova, A.V., Alekseev, A.S.: On a statement of the inverse kinematic problem of seismology for the two-dimensional continuously nonhomogeneous medium. In: Some methods and algorithms of geophysical data interpretation, M.M. Lavrentiev, ed.: pp. 137–154. Moscow: Nauka 1967

Courant, R., Hilbert, D.: Methods of Mathematical Physics, Vol. 2. New York: Interscience 1962

Elsgolz, L.E.: Differential equations and calculus of variation. Moscow: Nauka 1969

Jobert, G.: Travel-time equations and energy propagation for an elastic heterogeneous, isotropic medium: correction and approximate inversion. Bull. Seismol. Soc. Am. **63**, 1967–1972, 1973

Lavrentiev, M.M.: Some improperly posed problems of mathematical physics. Springer Tract. Nat. Phil., Vol. 11. Berlin, Heidelberg, New York: Springer 1967

Lavrentiev, M.M., Romanov, V.G.: On the three linearized inverse problems for hyperbolic equations. Dokl. Akad. Nauk USSR **171** (6), 1279–1281, 1966

Lavrentiev, M.M., Romanov, V.G., Vasiliev, V.G.: Multidimensional inverse problems for differential equations. Lecture Notes in Mathematics, Vol. 167. Berlin, Heidelberg, New York: Springer 1970

Mukhometov, R.G.: An inverse kinematic seismic problem on the plane. In: Mathematical problems of geophysics, Vol. 6, Part 2, M.M. Lavrentiev and A.S. Alekseev, eds.: pp. 243–254. Novosibirsk: Computing Center, Acad. Sci. USSR (Sib. Branch) 1975

Romanov, M.E.: A characteristic method of numerical solution of the inverse kinematic problem of seismology. In: Mathematical problems of geophysics, Vol. 3, M.M. Lavrentiev and A.S. Alekseev, eds.: pp. 328–346. Novosibirsk: Computing Center, Acad. Sci. USSR (Sib. Branch) 1972

Romanov, M.E.: On numerical solution of an integral geometry problem. In: Mathematical problems of geophysics, Vol. 6, Part 1, M.M. Lavrentiev and A.S. Alekseev, eds.: pp. 289–297. Novosibirsk: Computing Center, Acad. Sci. USSR (Sib. Branch) 1975

Romanov, V.G.: Integral geometry and inverse problems for hyperbolic equations. Springer Tracts. Nat. Phil., Vol. 26. Berlin, Heidelberg, New York: Springer 1974a

Romanov, V.G.: On a class of uniqueness of the solution of the inverse kinematic problem. In: Mathematical problems of geophysics, Vol. 4, M.M. Lavrentiev and A.S. Alekseev, eds.. pp. 147–164. Novosibirsk: Computing Center, Acad. Sci. USSR (Sib. Branch) 1974b

Received January 15, 1978; Revised Version November 5, 1979
Accepted March 17, 1980

J. Geophys. 48, 181–194, 1980

Journal of Geophysics

Geomagnetic Induction Studies in Scandinavia

I. Determination of the Inductive Response Function from the Magnetometer Array Data

A.G. Jones
Institut für Geophysik, Gievenbecker Weg 61, D-4400 Münster, Federal Republic of Germany

Abstract. Data from the Münster IMS Magnetometer Array (Küppers et al. 1979) have been analysed in the frequency domain to derive the inductive response function, $C(\omega, 0)$, from the ratio of the vertical magnetic field to the spatial gradient of the horizontal magnetic field. The response function was best determined by statistical frequency analysis techniques after the spatial gradients had been derived by least-squares fitting of two-dimensional second-order polynomials to the observations, with the constraint imposed that the solutions be curl-free.

The derived response function was found to obey two different causality requirements, and most of the inequality constraints imposed on it (Weidelt 1972). A preliminary model, in which conductivity is a function of depth only and which explains the major details of the observed response, is presented. It has a highly resistive uppermost layer of some $10^4\,\Omega$m and of the order of 30 km thick, underlain by a layer of about 125 Ωm to a depth of around 140 km, where a transition takes place to a highly conducting (3 Ωm) half space.

The effect of various non-uniform sources on the observations of $C(\omega, k)$ for this 1D model is illustrated.

Key words: Magnetometer arrays – Response function analysis – Geomagnetic induction studies in Scandinavia.

1. Introduction

This is the first in a series of papers treating various aspects of geomagnetic induction in Scandinavia as observed by the Münster IMS Magnetometer Array (Küppers et al. 1979) and by concurrent telluric field recording at some locations. This paper will deal principally with the inductive response function, $C(\omega, k)$, its properties, its determination and its validity, from magnetometer array data. Other papers will be concerned with the more traditional geomagnetic depth sounding (GDS) and magneto-telluric (MT) techniques, modelling studies, and a geophysical interpretation of all the results obtained.

Geomagnetic induction studies using large arrays of magnetometers began with the introduction of a cheap but very fieldworthy instrument by Gough and Reitzel (1967). Since that time, many array studies have been made in various countries (see the map of world activity in Lilley 1975) and reviews by Porath and Dziewonski (1971), Gough (1973a, b), Frazer (1974) and Lilley (1975) summarise the usual methods employed to analyse and present the data. These comprise *qualitative* analysis methods; e.g., mapping Fourier terms, induction vectors, inspection of magnetograms; *quantitative* analysis methods; e.g., separation of internal and external parts, determination of horizontal layering from Z/H ratios, determination of horizontal layering employing the ratio of the vertical field to the spatial gradient of the horizontal field; and *modelling* techniques, from primitive depth-of-line-current determinations to general 2D or 3D numerical or analytical techniques.

Of all these, the method involving the determination of the ratio of the vertical magnetic field to the horizontal spatial gradient, hereafter referred to as the HSG (horizontal spatial gradient) method, has attracted the least interest. Since the relationship between this ratio and the magnetotelluric impedance was shown by Schmucker (1970), Kuckes (1973a, b), and in a global sense by Berdichevsky et al. (1976), this HSG method has only been appied to magnetometer array data by Kuckes (1973a), Lilley and Sloane (1976) and Woods and Lilley (1979). The first two studies cited above derived the spatial gradients from published maps of the Fourier horizontal field components, which did not permit any determination of the coherence between the vertical field and the horizontal spatial gradient field. Woods and Lilley fitted fields at Fourier harmonics to three complex two-dimensional surfaces, one for each component, but restricted themselves to quiet daily variations and the first four harmonics thereof.

In this work, the spatial gradients are determined by least-squares fitting of 2D second-order surfaces to the observed fields, with the curl-free constraint imposed on the solutions, and the inductive response function is derived by statistical frequency analysis methods, with corresponding coherences and confidence interval estimations. The response function can be determined down to short periods (100 s) due to the large gradients observed at these periods in the auroral zone. The derived function is tested for physical realisability, i.e., causality and validity, and a preliminary one-dimensional model that explains the major details is given. The effects of non-uniform source fields, in the form of non-zero values for Price's (1962) wavenumber, and of Gaussian electrojets on the theoretical response function calculated from the 1D model is illustrated.

2. Theory

In this section the inductive response function is defined and its properties stated. For a complete exposition, the reader is referred to Schmucker (1970), Weidelt (1972, 1978), Kuckes (1973a, b) and Lilley (1975).

0340-062X/80/0181/$02.80

2.1. The Inductive Response Function $C(\omega, k)$

Defining the inductive response function, $C(\omega, k)$, for a single wave vector field $\mathbf{k}=(k_x, k_y)$, by the relationship

$$C(\omega,k)=\frac{H_z(\omega)}{\frac{\partial}{\partial x}H_x(\omega)+\frac{\partial}{\partial y}H_y(\omega)}, \tag{1}$$

where $k=(k_x^2+k_y^2)^{\frac{1}{2}}$, the absolute value of the tangential wave number $(k=2\pi/\lambda)$,

and $\mathbf{H}(\omega)=(H_x(\omega), H_y(\omega), H_z(\omega))$, the three orthogonal components of the magnetic field at frequency ω,

it is simple to derive that, *over a* 1 D *earth* (i.e., conductivity a function of depth only, $\sigma(z)$), this function is related to the magneto-telluric impedance function $Z_{xy}(\omega, k)$, by

$$C(\omega,k)=\frac{1}{i\omega\mu_0}\cdot Z_{xy}(\omega,k) \tag{2}$$

(Schmucker 1970; Kuckes 1973a, b; Schmucker and Weidelt 1975). This is a direct consequence of the fact that, for a 1 D earth, the magneto-telluric impedance tensor, $\mathbf{Z}$, assumes the well-known Cagniard type form, viz.

$$\mathbf{Z}=\begin{bmatrix} 0 & Z_{xy} \\ -Z_{xy} & 0 \end{bmatrix} \tag{3}$$

(Cagniard 1953). Following the notation of Weidelt (1972), the real and imaginary parts of $C(\omega, k)$ shall be denoted by $g(\omega)$ and $-h(\omega)$ respectively, from

$$C(\omega,k)=g(\omega,k)-ih(\omega,k). \tag{4}$$

2.2. General Properties of a Response Function

Consider a simple time-independent linear system as illustrated in Fig. 1. The input, $i(t)$, is related to the output $o(t)$, by the convolution integral

$$o(t)=\int_{-\infty}^{\infty} h(\tau)\, i(t-\tau)\, d\tau$$

where $h(\tau)$ describes the dynamic characteristics of the system, and is commonly termed the impulse response function. The function $h(\tau)$ must display certain characteristics if the system is to be *physically realisable;* it must not respond before receiving any input. This causality condition requires that

$$h(\tau)=0 \qquad \tau<0 \tag{5}$$

hence the integral should only be determined in the range $(0, \infty)$, i.e., for all positive lags.

Fourier transformation of the above convolution integral yields

$$O(\omega)=H(\omega)\cdot I(\omega) \tag{6}$$

where $H(\omega)$ is termed the transfer function and is related to $h(\tau)$ by

$$H(\omega)=\int_0^{\infty} h(\tau)\, e^{-j\omega\tau}\, d\tau.$$

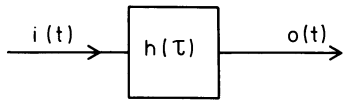


Fig. 1. Single input - single output linear system with the input, $i(t)$, related to the output, $o(t)$, by the dynamic characteristics of the system, $h(\tau)$

As shown by Solodovnikov (1952), (repeated in Kanasewich 1973), the real and imaginary parts of a causal transfer function describing a linear system are *not* independent, but are related by the Hilbert transform, viz.

$$H_R(\omega)=\mathscr{H}\{H_I(\omega)\}=\frac{1}{\pi}P\int_{-\infty}^{\infty}\frac{H_I(w)}{\omega-w}\,dw \tag{7a}$$

$$H_I(\omega)=-\mathscr{H}\{H_R(\omega)\}=-\frac{1}{\pi}P\int_{-\infty}^{\infty}\frac{H_R(w)}{\omega-w}\,dw \tag{7b}$$

where $H_R(\omega)$ and $H_I(\omega)$ are the real and imaginary parts of $H(\omega)$ respectively, and P denotes the Cauchy principal value of the integral. For minimum phase systems, the relationships may be rewritten to show that if $H(\omega)=|H(\omega)|\, e^{i\phi(\omega)}$, then $\log|H(\omega)|$ is related to $\phi(\omega)$ by the Hilbert transform. Relationships (7a) and (7b) are termed the Kramers-Kronig relations in atomic scattering theory and Bode's relations in servo-mechanics.

2.3. Properties of the Inductive Response Function

The inductive response function, $C(\omega, k)$, as defined by (1), possesses certain properties due to the fact that the earth has a finite, positive conductivity, i.e., $0<\sigma<\infty$. Recalling that $C=g-ih$ (dependence on frequency and wavenumber assumed), and defining the operator D to denote

$$Df=\frac{\omega\, df}{d\omega}=\frac{df}{d\log\omega}=-\frac{df}{d\log T},$$

the following inequalities apply (Weidelt 1972, Eqs. 2.30–2.34)

$$g\geqq 0 \qquad h\geqq 0, \tag{8a, b}$$

$$Dg\leqq 0, \tag{9}$$

$$0\leqq -D|C|\leqq|C|, \tag{9a, b}$$

$$|DC|\leqq h, \qquad |C+DC|\leqq g, \tag{11a, b}$$

$$|D^2C|\leqq h, \qquad |C+2DC+D^2C|\leqq g. \tag{12a, b}$$

The limiting values of the terms g and h at high and low frequencies are

$$g=\begin{cases}\dfrac{1}{k}\tanh(kd) & \text{for } \omega\to 0 \qquad (13a)\\[2ex] \dfrac{1}{(2|\omega|\mu_0\sigma_1)^{\frac{1}{2}}} & \text{for } \omega\to\pm\infty, \qquad (13b)\end{cases}$$

$$h=\begin{cases}0 & \text{for } \omega\to 0 \qquad (13c)\\[2ex] \dfrac{\pm 1}{(2|\omega|\mu_0\sigma_1)^{\frac{1}{2}}} & \text{for } \omega\to\pm\infty \qquad (13d)\end{cases}$$

where k is as given in Eq. (1)
d is the depth of the perfect conductor
σ_1 is the conductivity of the top layer

and μ_0 is the permeability of free space,
and the function C satisfies

$$C(-\omega)=C^*(\omega) \tag{14}$$

such that g is symmetric and h anti-symmetric about the $\omega=0$ axis.

As shown by Weidelt [1972, Eq. (2.19)], the inductive response function admits the representation

$$C(\omega)=\int_0^\infty \frac{a(\lambda)}{\lambda+j\omega}\,d\lambda$$

where $a(\lambda)$ is a generalised function given by

$$a(\lambda)=-\lim_{\varepsilon\to+0}\frac{1}{\pi}\operatorname{Im}(C(j\lambda+\varepsilon))\geqq 0$$

(Schmucker and Weidelt 1975, the non-negativity of which ensures that C must be a smooth function of frequency. Using the result

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{j\omega t}}{\lambda+j\omega}\,d\omega=\frac{1}{2\pi j}\int_{-\infty}^{\infty}\frac{e^{j\omega t}}{\omega-j\lambda}\,d\omega$$
$$=\begin{cases}0 & t<0\\ e^{-\lambda t} & t>0\end{cases}$$

($\lambda>0$) obtained by closing the contour for $t<0$ in the lower ω-halfplane, and for $t>0$ in the upper, it follows that

$$c(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}C(\omega)e^{j\omega t}\,d\omega$$
$$=\begin{cases}0 & t<0\\ \int_0^\infty a(\lambda)e^{-\lambda t}\,d\lambda & t>0\end{cases}$$

which gives directly the consequence that

$c(t)>0 \quad$ for $0<t<\infty$.

This constraint that the function be a positive real function has applications in many branches of physics, for example, that the Laplace transform of a linear network driving-point-impedance be a positive real function is one of the necessary conditions in order for it to be realized as a Brune or Bott-Duffin network (Ferris 1962).

Further restrictions on the necessary form of $c(t)$ may be derived by differentiating $c(t)$ with respect to time. This yields that in the range $0<t<\infty$,

$$c'(t)<0, \quad c''(t)>0, \quad c'''(t)<0$$

etc.

These conditions on the sign, slope and curvature result in a very strong constraint on the permitted form of the earth's inductive impulse response function, $c(t)$.

3. Determination of $C(\omega, k)$

To determine $C(\omega, k)$ from Eq. (1), it is necessary to derive the spatial gradients of the horizontal field component at each frequency. This was accomplished in this work by fitting a second order 2D polynomial of the form

$$H(x,y)=h_0+h_1x+h_2y+h_3x^2+h_4y^2+h_5xy+\delta H_x \tag{15}$$

(similarly for D and Z), to the real and imaginary parts of each of the 3 components at each Fourier harmonic. (Note: to ease notation, the forms $H_x\equiv H$, $H_y\equiv D$ and $H_z\equiv Z$ have been used.) This is exactly the same procedure as that adopted in a similar analysis by Woods and Lilley (1979).

However, in order to determine surfaces which were as physically meaningful as possible, the solutions were constrained to be curl-free, that is

$$\frac{\partial H_x}{\partial y}=\frac{\partial H_y}{\partial x}$$

thus permitting no vertical current flow, $j_z=0$.

This causes a coupling of the surfaces for (H, D) because

$$h_2=d_1,$$
$$2h_4=d_5,$$
$$h_5=2d_3,$$

must be upheld for a curl-free solution. Hence, the two independent equations for H and D of 6 complex unknowns each, are reduced to a combined form of 9 complex unknowns. This necessitates information from a minimum of 5 stations, each station contributing 4 degrees of freedom to the solution, the locally observed real and imaginary parts of the two horizontal components.

The spatial gradient terms at the origin are hence

$$\frac{\partial H_x}{\partial x}\triangleq h_1$$
$$\frac{\partial H_y}{\partial y}\equiv d_2$$

and the denominator in Eq. (1) is (h_1+d_2).

For the value of $H_z(\omega)$, 2D surfaces of the form of (15) were also fitted to the observed vertical fields. This was because what is being attempted by this method is the derivation of a parameter that permits interpretation in terms of a *regional* 1D conductivity-depth distribution, and not one that is grossly perturbed by *local* induction anomalies. In this respect, the vertical magnetic field component observed at any one location may not truly reflect the regional field, but may be - and often is - highly indicative of more local structure. Hence, to try to remove this local effect at the central point of the investigation area, the observed H_z variations were also fitted by the polynomial of form (15). Thus, the numerator in Eq. (1) is (z_0).

One very important consequence of representing the field by a polynomial of second order is that the relationship (1) reduces to

$$C(\omega,0)=\frac{\tilde{H}_z}{\frac{\partial}{\partial x}\tilde{H}_x+\frac{\partial}{\partial y}\tilde{H}_y}$$

(Weidelt 1978, 142-143), where $\tilde{H}(x,y,0)$ is the magnetic field determined from fitting equations of the form (15). Thus, the derived inductive response function can be interpreted without regard to non-uniform source field configuration, and is henceforth written as $C(\omega)$.

Three techniques were compared to determine estimates of $C(\omega)$ from the derived horizontal spatial gradient and smoothed vertical field terms.

Method 1. At each Fourier harmonic, j, the raw inductive response function $C_j(\omega)$ was derived from

$$\hat{C}_j = \frac{(z_0)}{(h_1 + d_2)} \tag{16a}$$

(dependence on frequency assumed). These $\hat{C}_j$'s were then band-averaged to give a smoothed $\bar{C}_1$ with the acceptance criteria that, for each estimate, inequalities (8a) and (8b) be upheld, i.e., Real $(\hat{C}_j) \geqq 0$ and Imaginary $(\hat{C}_j) \leqq 0$. Hence, estimate (1) is given by

$$\bar{C}_1(\bar{\omega}) = \langle C_j(\omega) \rangle. \tag{16b}$$

Method 2. The raw estimates of numerator and denominator of Eq. (1) were frequency band-averaged, and the smoothed $\bar{C}_2$ was given by

$$\bar{C}_2(\bar{\omega}) = \frac{\langle (z_0) \rangle}{\langle (h_1 + d_2) \rangle} \cdot \tag{17}$$

Method 3. Equation (1) describes a linear system with an input related to an output by an impulse response function, or, in the frequency domain, a frequency response function. This is illustrated in Fig. 1. In this formalism, Eq. (1) can be considered exactly as Eq. (6) with

$$I(\omega) \equiv \frac{\partial H_x(\omega)}{\partial x} + \frac{\partial H_y(\omega)}{\partial y}, \tag{18a}$$

$$O(\omega) \equiv H_z(\omega), \quad \text{and} \tag{18b}$$

$$H(\omega) \equiv C(\omega) \tag{18c}$$

Eq. (5) may be solved by method of cross-spectral analysis invoking the Wiener-Hopf integral equation and the Wiener-Khintchine theorem, to derive the least-squares estimate of $H(\omega)$. However, as shown in Jones (1980), there exist two possible forms of the least-squares estimate. Form (1) is

$$\bar{C}_L(\bar{\omega}) = \frac{S_{io}(\bar{\omega})}{S_{ii}(\bar{\omega})} \tag{19}$$

where $S_{ab}(\bar{\omega}) = \langle A^*(\omega) B(\omega) \rangle$, i.e., the cross-spectra between $A(\omega)$ and $B(\omega)$, and is *downward-biased* for noise occurring on the input, i.e., the spatial gradient term (18a). Form (2) is

$$\bar{C}_U(\omega) = \frac{S_{oo}(\bar{\omega})}{S_{oi}(\bar{\omega})} \tag{20}$$

and is *upward-biased* for noise occurring on the output, i.e., the smoothed vertical magnetic field term (18b). The smoothed $\bar{C}_3$ estimate was given by

$$\bar{C}_3(\bar{\omega}) = \frac{\bar{C}_L(\bar{\omega}) + \bar{C}_U(\bar{\omega})}{2}. \tag{21}$$

Estimate $\bar{C}_3(\bar{\omega})$ is itself only unbiased by noise contributions if the input signal-to-noise ratio exactly equals the output signal-to-noise ratio (Jones 1980). The true unbiased response function lies in the interval (C_L, C_U).

The estimate of the coherence between the input and output of a single input - single output linear system, in this case between the horizontal spatial gradients and the vertical magnetic field, is given by

$$\hat{\gamma}^2_{io}(\bar{\omega}) = \frac{|S_{io}(\bar{\omega})|^2}{S_{ii}(\bar{\omega}) S_{oo}(\bar{\omega})} \cdot \tag{22}$$

However, this estimate is relatively well known to be a biased function. For example, although the *true* coherence between two random data sets is zero, the *estimate* of the coherence is non-zero due to the bias associated with the cross-spectral estimate, $S_{io}(\bar{\omega})$. The expectation value of the estimate of the ordinary coherence function between two random data sets is

$$E[\hat{\gamma}^2_{r_1 r_2}] = \frac{2}{n}$$

where n is the number of degrees of freedom associated with the smoothed spectral value (each Fourier harmonic has two degrees of freedom, one for the real and one for the imaginary part). Jones (1977, 1979) suggested a method of correcting for this bias by normalising the coherence value after transforming the function into a domain where it exhibits a normal distribution. This normalised transformed ordinary coherency function, N_{xy}, given by

$$\hat{N}_{xy} = \frac{\tanh^{-1}(\hat{\gamma}_{xy})}{\tanh^{-1}((2/n)^{1/2})} \tag{23a}$$

has a variance of

$$\sigma^2_{N_{xy}} = \frac{1}{n \operatorname{arctanh}^2((2/n)^{1/2})}, \tag{23b}$$

where $\hat{\gamma}_{xy}$ is the estimated ordinary coherency [the positive square root of the estimated ordinary coherence given by (22)], and n is the number of degrees of freedom associated with the estimate,

and there is an expectation value of

$$E[\hat{N}_{xy}] = 1,$$

for random data sets.

The derived inductive response function from (21), $\bar{C}_3(\bar{\omega})$, is only an estimate of the true response function, $C(\omega)$, and it can be shown that, if the total noise terms constitute a white series, then a confidence region for $C(\omega)$ can be specified at a confidence level of $100(1-\alpha)\%$ from

$$|\bar{C}_3(\bar{\omega}) - C(\omega)|^2 \leqq \frac{2}{n-2} F_{2:n-2;\,1-\alpha}(1 - \hat{\gamma}^2_{io}(\bar{\omega})) \frac{S_{oo}(\bar{\omega})}{S_{ii}(\bar{\omega})} \tag{24}$$

(see, for example, Bendat and Piersol 1971, 199–203), where $F_{2:n-2;\,1-\alpha}$ is the Fisher likelihood function with a $(1-\alpha)$ confidence coefficient,

$S_{oo}(\bar{\omega})$ is the smoothed estimate of the autospectra of the output, defined by (18b),

$S_{ii}(\bar{\omega})$ is the smoothed estimate of the autospectra of the input, defined by (18a),

and $\hat{\gamma}^2_{io}(\bar{\omega})$ is the estimate of the coherence function between the input and the output, given by (22).

4. Application

The application of this method is only worthwhile if interpretable response functions are determined. In this respect, the *major* assumption made is that the observed magnetic fields result from induction by a non-uniform source over an earth which

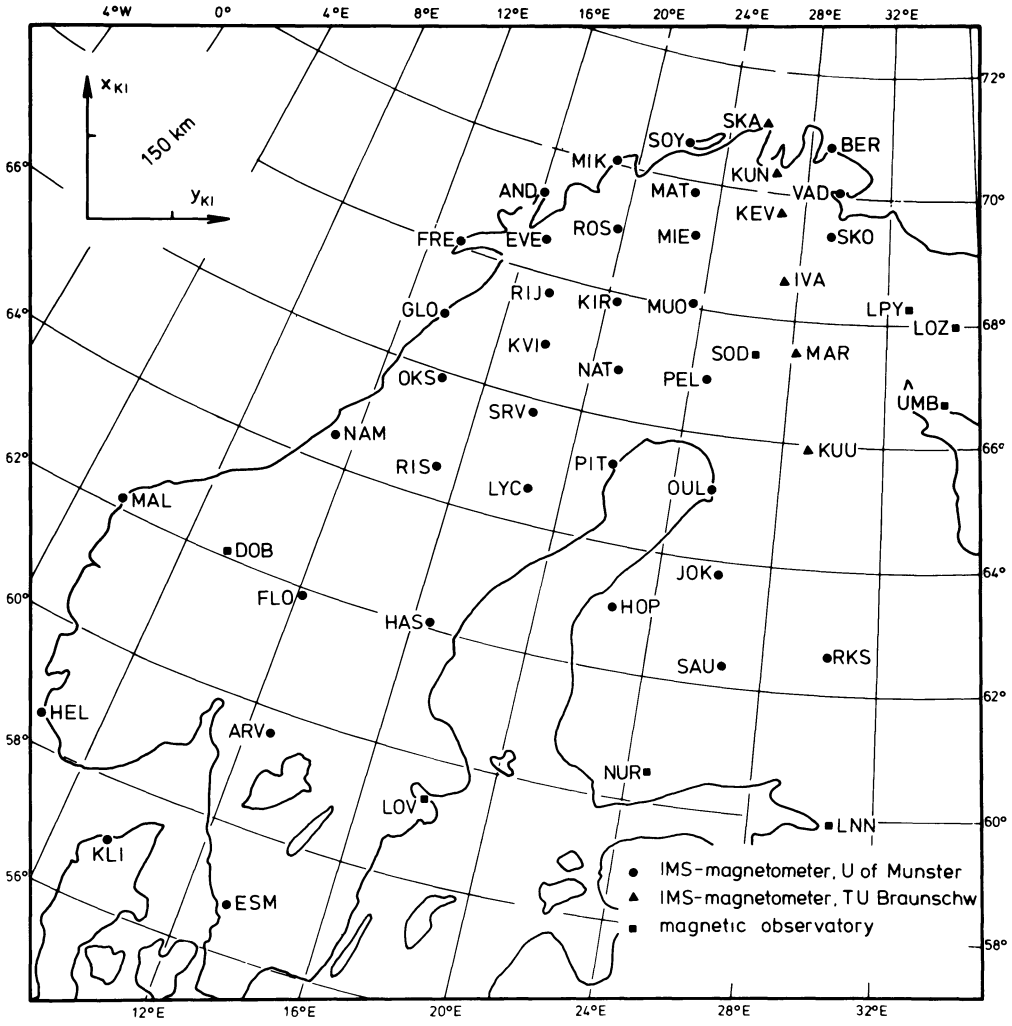


Fig. 2. Station map in geographic coordinates. Also indicated are the axes of the Kiruna cartesian coordinate system (see Küppers et al. 1979)

has a 1D conductivity distribution, i.e., that conductivity is a function of depth only, $\sigma(z)$.

In following papers, it will be illustrated that the region around Kiruna (KIR) (see Fig. 2) displays small anomalous induction effects, and that magneto-telluric data recorded at Nattavaara (NAT) exhibits, to a first approximation, one-dimensionally. Hence data from the zone centered on Kiruna should be interpretable in terms of a 1D structure. The nearest obvious 2D or 3D structure is the coast, which is more than 300 km from Kiruna.

The principal reasons for the operation of the array were coordinated observations with other groups of the ionospheric current systems for the International Magnetospheric Study (Küppers et al. 1979). Hence many events of profoundly different current configurations were digitised for analysis (see for example Baumjohann et al. 1978, 1980; Mersmann et al. 1979; Untiedt et al. 1978; and Küppers et al. 1979). Some of these have proved suitable for determination of $C(\omega)$ due to the strong spatial gradients observed.

As an example, consider the magnetogram illustrated in Fig. 3. This is a record of the magnetic field variations recorded by the Kiruna (KIR, Fig. 2) magnetometer between 15:00 UT and 23:00 UT on 2 December 1977. This event is of special world-wide interest as it occurred during a very favourable configuration of three different satellites. Study of the event has inferred extremely complex auroral zone current structure over Scandinavia during the interval.

The magnetic data from the stations EVE, RIJ, KVI, SRV, ROS, KIR, NAT, MIE, MUO, and PEL (Fig. 2) were digitised, conditioned and fast Fourier transformed to give the raw Fourier harmonics at each station. Each time series was of 2,880 points with a sampling interval of 10 seconds which was extended by zeroes to 4,096 giving 2,049 harmonics, each harmonic having 1.4 degrees of freedom (0.7 each for the real and imaginary parts respectively).

Using the procedures outlined in Sect. 3, the fields were fitted by polynomials at each Fourier harmonic. The fit was by least-squares methods which minimised the Euclidean norm between the solutions of the overdetermined system (Golub 1965), i.e., minimised $|\delta H_x + \delta H_y|^2$ and $|H_z^2|$ for the horizontal and vertical field solutions respectively. The 10 stations employed give 40 degrees of freedom for the horizontal field and 20 for the vertical field. The polynomials require 18 and 12 degrees of freedom for the horizontal and vertical solution respectively. Hence the combined horizontal field solution has 22 degrees of freedom whilst that of the vertical field has 8 degrees of freedom.

The estimates of C for the three different methods described in Sect. 3 were frequency band averaged with a constant Q filter

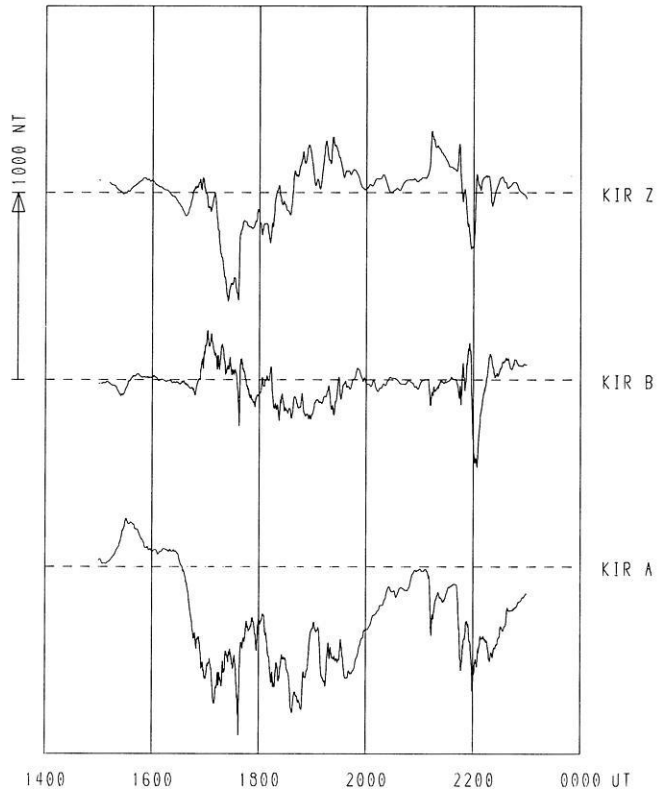


Fig. 3. Magnetic data recorded at Kiruna in the interval 15:00–23:00 UT on 2nd December 1977. *A*, component toward north; *B*, toward east, *Z*, vertically downwards

of box-car form with $Q=0.5$. These solutions for g and h (where $C=g-ih$) are illustrated in Fig. 4a and 4b respectively. It is apparent from these figures that method 3, employing statistical frequency analysis techniques, produces far smoother estimates than either of the other two methods.

In Fig. 5 are displayed the downward-biased and upward-biased estimates, from Eqs. (19) and (20) respectively, of $C_3(\omega)$, together with the 95% confidence intervals, calculated from (24). It is apparent from the figure that there is little bias due to noise in the estimates, but, at long periods, the confidence intervals are large. In Fig. 6 is displayed the coherence function between the input and output of the system, and also the normalised transformed ordinary coherency, as defined by (23a), N_{io}. The two series appear well correlated, even at short periods of the order of hundreds of seconds.

A total of 8 events were analysed in the manner described above, and the individual estimates for g and h were averaged in a weighted manner by the algorithm

$$\bar{g}(\bar{\omega}_j)=\frac{\sum_{i=1}^{8} N_i(\bar{\omega}_j)\cdot g_i(\bar{\omega}_j)}{\sum_{i=1}^{8} N_i(\bar{\omega}_j)} \tag{25}$$

similarly for $\bar{h}(\bar{\omega}_j)$. The sample weighted variances of the estimates were determined using the algorithm which gives an efficient and consistent unbiased estimate

$$s_g^2=\frac{n}{n-1}\,\frac{\sum N_i(g_i-\bar{g})^2}{\sum N_i} \tag{26}$$

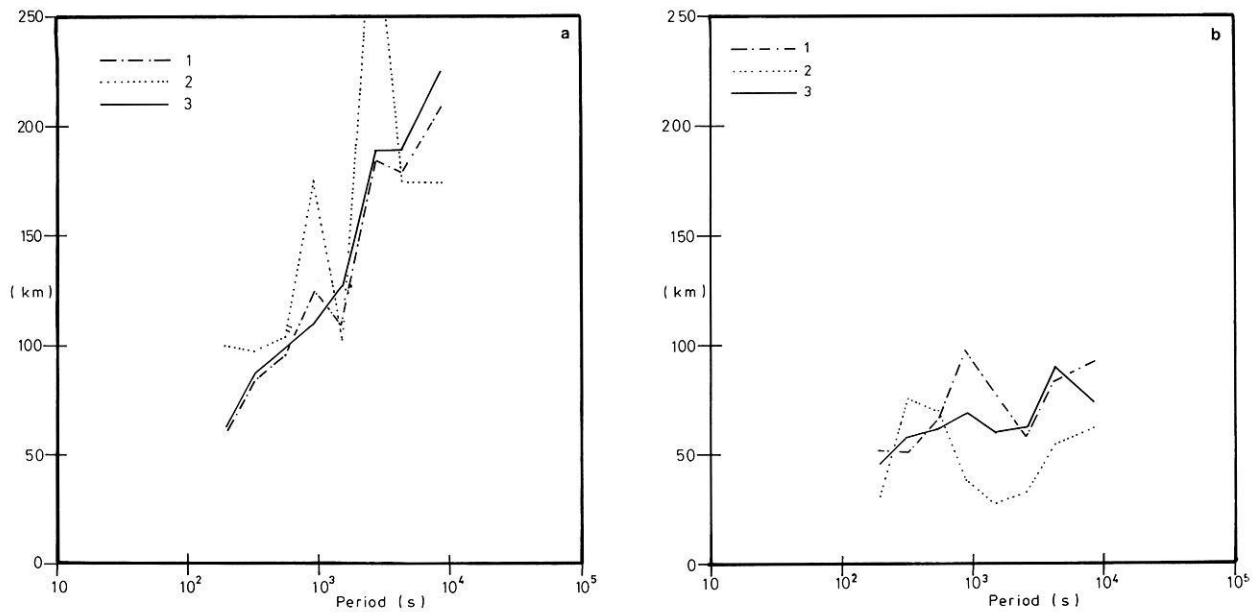


Fig. 4. a Determinations of g from the three methods outlined in Sect. 3; (-·-·-) method 1, (·····) method 2, (———) method 3. **b** As for (**a**) but for h

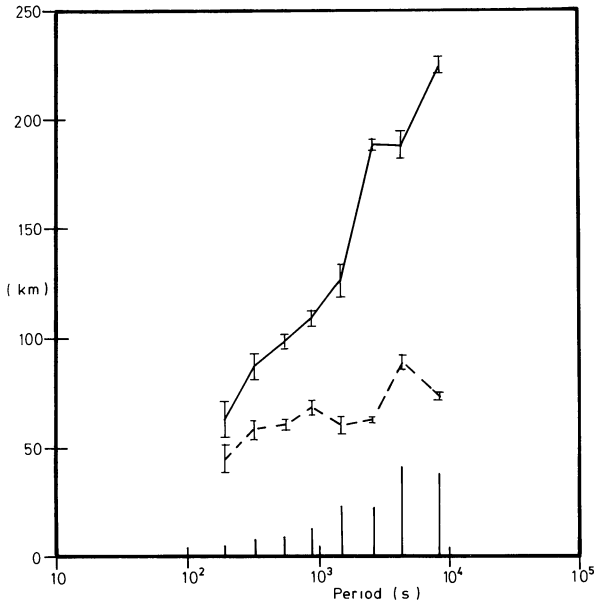


Fig. 5. The estimates from method 3 for g (*full line*) and h (*dashed line*) showing the downward- and upward-biased estimates as error bar ranges. The 95 % confidence intervals are given by the magnitude of the *solid lines* at the base of the figure

(dependence on frequency assumed)

where s_g^2 - is the sample weighted variance of the estimate $\bar{g}$
n - number of estimates
N_i - the function, as defined by (23a) for the i'th estimate
g_i - the i'th estimate of g
$\bar{g}$ - the weighted average, given by (25)

similarly for s_h^2.

Variances, and their square root, the standard deviation, are in themselves very weak statistical parameters. They only describe the dispersion, or scatter, of the individual estimates which are employed to determine the mean value. A far more powerful statistical description is given by the confidence intervals, at a certain probability level, of the determined mean values. Because the true weighted variance, σ^2, is unknown and is only estimated by the sample weighted variance, given by (26), statements regarding possible future values of the sample weighted mean, from (25), must be made by employing the Student-t, rather than the normal, distribution. The probability that any mean, $\bar{x}$, will exceed the true mean, μ_x, plus an associated confidence interval, is given by

$$\text{Prob}\left[\bar{x} > \mu_x + \frac{s t_{n-1:\alpha}}{n^{1/2}}\right] = \alpha$$

where μ_x is the true mean
$\bar{x}$ is the sample (weighted) mean
s is the sample (weighted) standard deviation
$t_{n-1:\alpha}$ is the Student-t distribution at the α confidence level for $(n-1)$ degrees of freedom
α is the probability level
and n is the total number of degrees of freedom.

Accordingly, the $100(1-2\alpha)\%$ confidence intervals for $\bar{x}$ are given by

$$\left(\bar{x} - \frac{s t_{n-1:\alpha}}{n^{1/2}}\right) \leqq \mu_x \leqq \left(\bar{x} + \frac{s t_{n-1:\alpha}}{n^{1/2}}\right). \tag{27}$$

Figure 7a and b display the individual estimates of $g_i(\bar{\omega})$ and $h_i(\bar{\omega})$ from the 8 events, with their relative weighting being crudely indicated by the type of symbol used.

The weighted means, from (25), are joined by the full lines, and the relative 95 % confidence intervals for these means from (24) and (25) are also shown. These mean values have been further smoothed by a Hanning window, $(\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$, to give the dashed lines indicated in Fig. 7a and b and given in Table 1 together with the confidence intervals.

The statistical method described above for determining the means, and the confidence limits of those means, of the parameters $\bar{g}$ and $\bar{h}$ only necessitates assuming that for these data the Central Limit Theory is valid, i.e., that the probability density functions of the estimates of g and h tend very quickly to the Normal form as the number of estimates is increased. This is a very weak statistical assumption to make when compared to that made for determining confidence intervals from equations

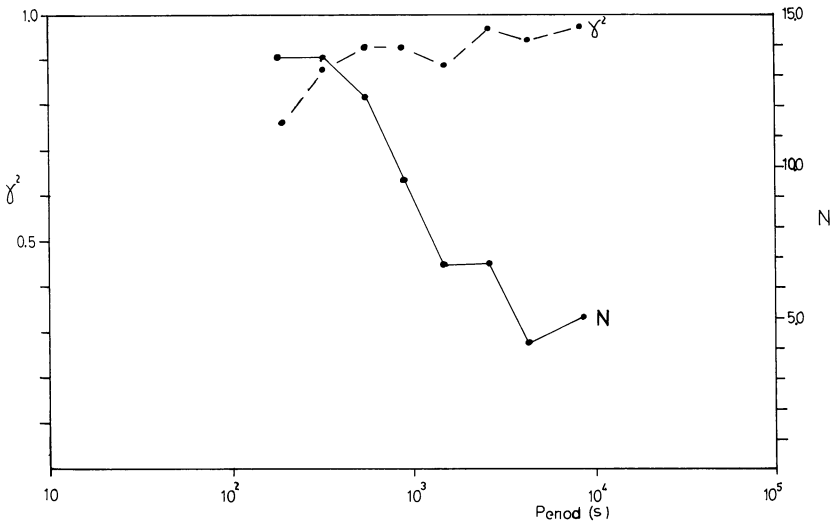


Fig. 6. Coherence functions between the derived vertical field and spatial gradient of the horizontal field; (- - -) γ^2, usual ordinary coherence function; (——) N, normalised transformed ordinary coherency function

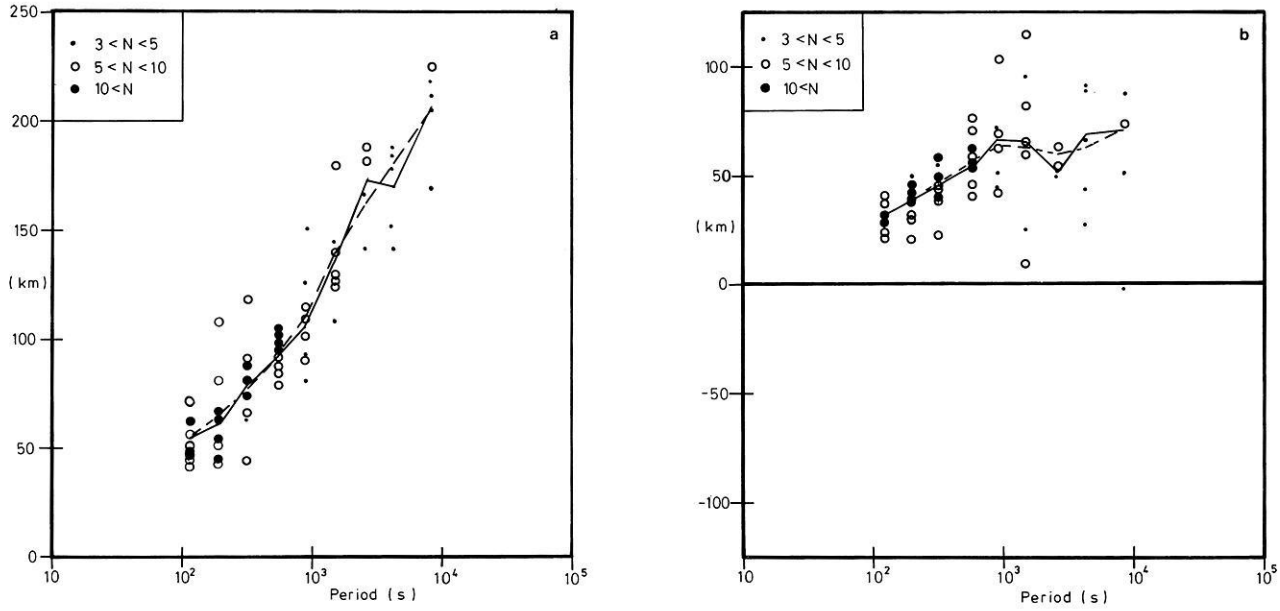


Fig. 7. a Individual event determinations of g with the corresponding value of the normalised transformed ordinary coherence being indicated by symbol type: $\bullet \equiv 3 < N < 5$; $\circ \equiv 5 < N < 10$; $\bullet \equiv 10 < N$. The *full line* indicates the weighted ensemble average, and the *dashed line* is the Hanning-smoothed average. **b** As (**a**) but for h

Table 1. Determined response function and confidence intervals

Period(s)	$\bar{g}$ (km)	$\bar{h}$ (km)	r_g (km)	r_h (km)	ρ_a (Ωm)	ϕ_a (degrees)
115	55.2	31.5	9.2	4.9	277	60
190	64.2	38.6	15.6	6.7	231	60
320	77.7	47.6	15.0	7.8	204	58
535	92.4	57.5	6.2	8.7	175	58
890	110.6	63.6	16.0	17.0	144	61
1,490	138.0	62.7	19.6	41.5	122	66
2,480	163,0	60.0	32.7	9.5	96	70
4,090	179.7	63.8	25.6	35.0	70	70
8,200	206,0	71.5	37.9	42.0	46	71

$\bar{g}$ and $\bar{h}$ are the smoothed real and imaginary parts of the response function, averaged over eight independent events

r_g and r_h are the 95 % confidence intervals of the parameters $\bar{g}$ and $\bar{h}$ respectively

ρ_a and ϕ_a are the apparent resistivity and phase of the derived inductive response function

of the form (24). This method also has another advantage over the more usual statistical technique, of, for example, averaging the ensemble estimates of (H_z) and $(\partial H_x/\partial x + \partial H_y/\partial y)$, in that higher weighting can be given to those estimates of $C_i(\omega)$ which are considered more superior (Eq. (25)).

5. Validity

Having obtained the smoothed values of $\bar{C}(\bar{\omega})$, as displayed in Fig. 7a and b and given in Table 1, there are certain questions that can be posed regarding the validity of the inductive response function so determined.

5.1. *Does the Function Describe a Causal System?*

As detailed in Sect. 2.2, there are two tests which may be applied to determine whether or not the estimated response function describes a physically realisable, i.e., causal, linear system. It is known that at the periods of interest, 10^2–10^5 s, the Earth responds inductively in a causal manner. Hence any response function that describes the induction process must comply to conditions (5), (7a), and (7b).

5.1.1. Is $c(t)=0$ for $t<0$? Defining the forward Fourier transform operator by $\mathcal{F}$, given by

$$X(\omega) = \mathcal{F}\{x(t)\} = \int_{-\infty}^{\infty} x(t)\, e^{-j\omega t}\, dt \tag{28a}$$

and the inverse Fourier transform operator by $\mathcal{F}^{-1}$,

$$x(t) = \mathcal{F}^{-1}\{X(\omega)\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega)\, e^{j\omega t}\, d\omega \tag{28b}$$

it is easy to prove that

$$x(-t) = \mathcal{F}\{X(\omega)\} = \frac{1}{2\pi}\, \mathcal{F}\mathcal{F}\{x(t)\}$$

(Champeney 1973). If $x(t)$ is a purely real series, then from symmetry relations

$$x(-t) = \mathcal{F}^{-1}\{X^*(\omega)\} \tag{29}$$

where $X^*(\omega)$ is the complex conjugate of (28a).

Hence, for a response function, the inverse Fourier transform gives the impulse response function for all positive lags, and the inverse Fourier transform of the complex conjugate of the

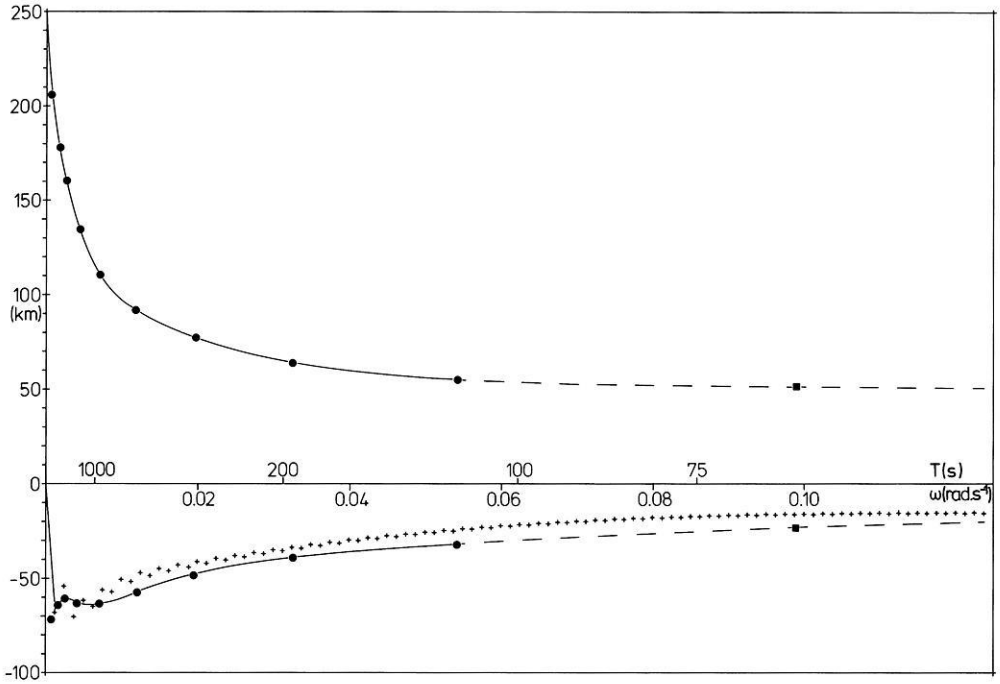


Fig. 8. Interpolation of g (positive) and $-h$ (negative) in the observation range given by *solid lines*. *Dashed lines* are the extrapolations of these functions to higher frequencies using the parameters from a two-layer model as aids (*squares*). *Crosses* are the function $g_H = -\mathscr{H}\{g\}$, the Hilbert transform of g

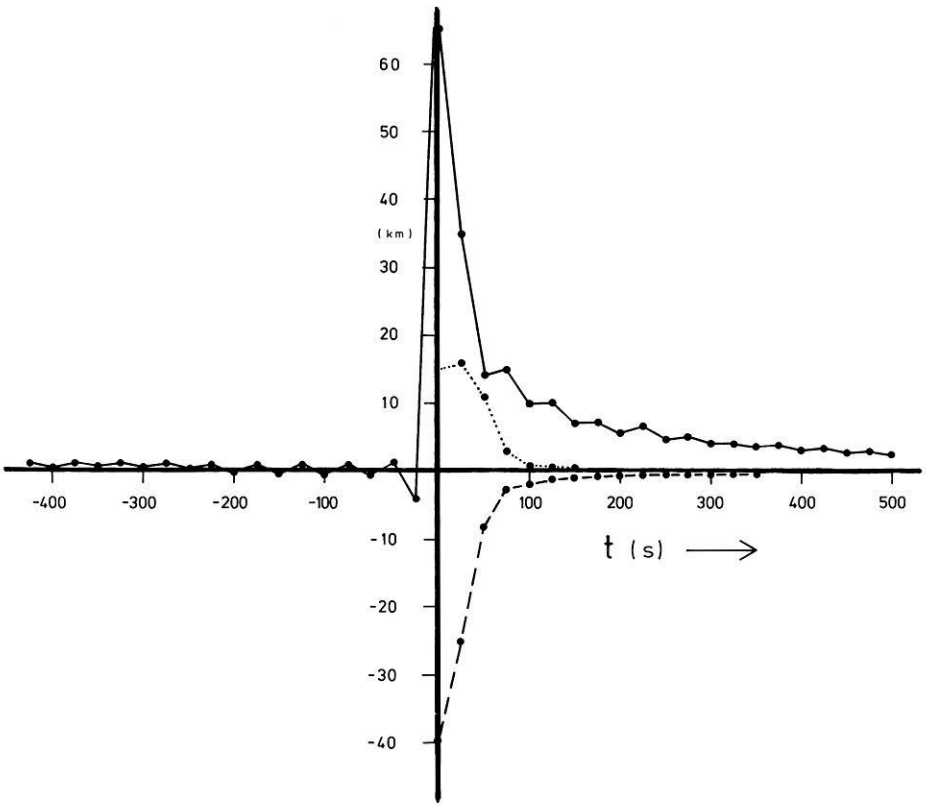


Fig. 9. Inductive impulse response function, $c(t)$ (————), with its first differential, $c'(t)$ (— — — —), and second differential, $c''(t)$ (·····), with respect to time

response function gives the impluse response function for all negative lags, i.e.,

$$c(\tau)=\begin{cases}\mathscr{F}^{-1}\{C(\omega)\} & \tau \geqq 0 \quad (30\,a)\\ \mathscr{F}^{-1}\{C^*(\omega)\} & \tau \leqq 0. \quad (30\,b)\end{cases}$$

The inductive response function given in Table 1 was interpolated by fitting cubic spline functions to give estimates which were equally spaced in the frequency domain. This fit is illustrated in Fig. 8. These values were then employed to derive the impulse response function $c(\tau)$, for both positive and negative

lags, by utilising expressions (30a) and (30b). The resulting impulse response function is illustrated in Fig. 9, from which it can be seen that the function does indeed comply with the requirement of a causal system that $c(\tau)=0$ for $\tau<0$.

5.1.2. Are the real and imaginary parts of $C(\bar{\omega})$ related by the Hilbert transform? There are two techniques for deriving the Hilbert transform of a function. The first is to evaluate the integral form directly, as expressed by Eq. (7a) and (7b), the second employs Fourier transformation. Both, however, require that the function be well described over a sufficiently long interval that the basic assumptions that have to be made regarding the function outside the interval be valid.

For $\omega=0$, then from (13)

$$h=0$$

and $$\lim_{\omega\to 0}(g)=\frac{1}{k}\tanh(kd).$$

Assuming kd are such that $\tanh(kd)\simeq kd$, then $g\simeq d$. For this, $kd\lesssim 0.5$ such that the wavelength, λ, of the field must be

$$\lambda>\pi\cdot d.$$

Hence reasonable values of d ensure $g\simeq d$ for the wavelengths of the fields analysed. A value of $d=250$ km was adopted.

At the very high positive frequencies, g and h tend towards expressions (13b) and (13d), i.e., the same values. Assuming that Scandinavia is overlain by a resistive top layer of about $10^4\,\Omega$m, which is inferred by the work of Westerlund (1972), of thickness at least 10 km, then the period at which this is the skin depth is given by

$$\delta=0.5(\rho T)^{1/2}$$

where δ is the skin depth in km
ρ is the layer resistivity in Ωm
and T is the period in seconds
and for $\delta=10$ km with $\rho=10^4\,\Omega$m gives $T=0.04$ s.
Hence, for $\omega>160$ rads^{-1} ($T<0.04$ s)

$$g(\omega)=h(\omega)=\frac{63}{\omega^{1/2}}\text{ km}$$

or

$$g(T)=h(T)=25\,T^{\frac{1}{2}}\text{ km}.$$

At the intermediate periods, $100<T<0.04$, it is necessary to assume a conductivity distribution in order to aid the interpolation of the data in this range. Using an algorithm given by Fischer and Schnegg (in press 1980), a two-layer model was derived which was compatible with the three shortest period $|\bar{C}(\bar{\omega})|$ values. This model consisted of a top layer of $10^4\,\Omega$m and thickness 30 km, underlain by a more conducting layer of 125 Ωm. These values are indicated in Fig. 8 where the solid lines are the interpolation using the derived values of $\bar{C}(\bar{\omega})$ (Table 1) and the dashed lines are the extrapolation constrained

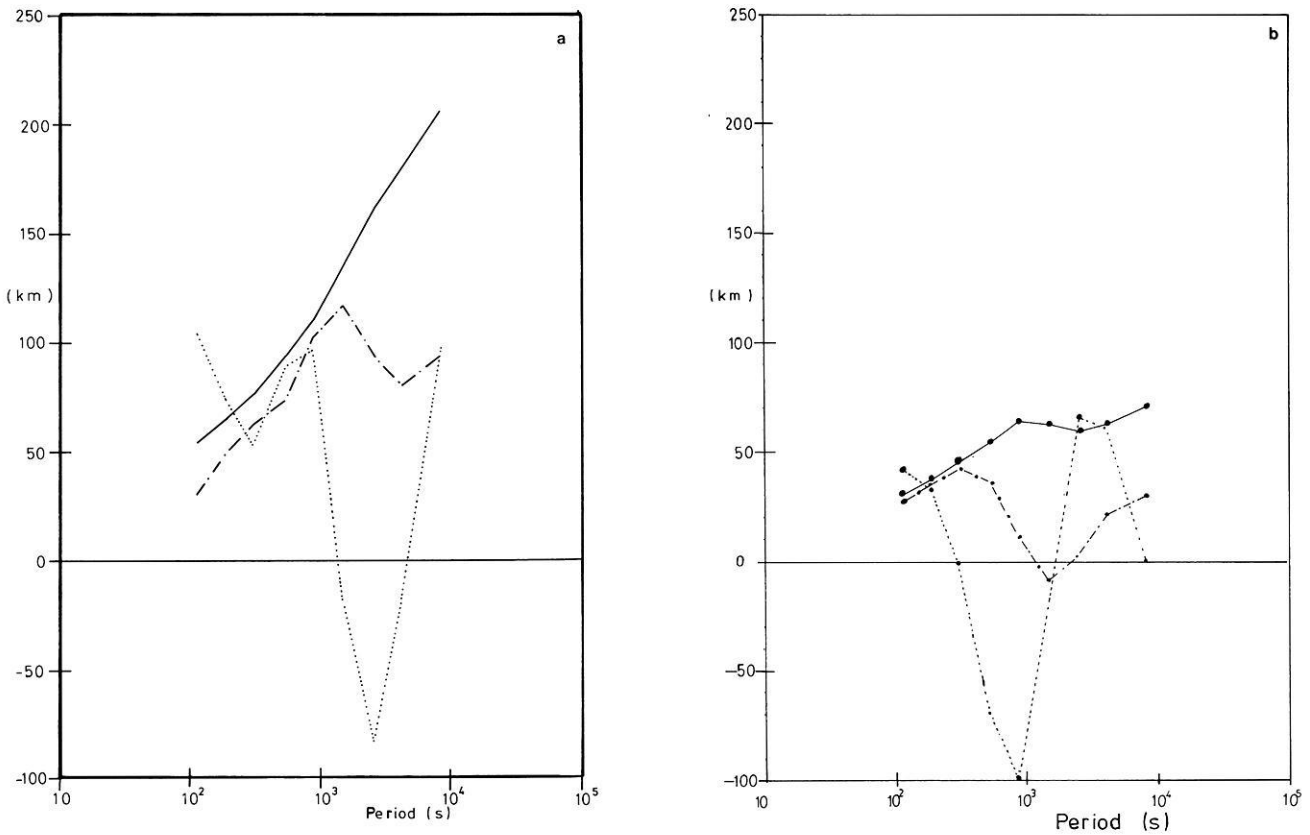


Fig. 10. a The function $g(\omega)$ (———) with its first derivative (- - -) and second derivative (·····) with respect to log T. **b** As 10a but for the function $h(\omega)$

Table 2. Tests of Weidelt's inequality constraints

Period	8a	8b	9	10a	10b	11a	11b	12a	12b
115	+	+	+	+	+	−	+	−	−
190	+	+	+	+	+	−	+	−	+
320	+	+	+	+	+	−	+	−	+
535	+	+	+	+	+	−	+	−	+
890	+	+	+	+	+	−	+	−	+
1,490	+	+	+	+	+	−	+	+	+
2,480	+	+	+	+	+	−	+	−	+
4,090	+	+	+	+	+	−	+	−	+
8,200	+	+	+	+	+	−	+	−	+

\+ the inequality was upheld by the response function
− the inequality was not upheld by the response function

by the 2-layer 1 D model described above. Although these somewhat severe assumptions about the form of $\bar{C}(\bar{\omega})$ outside the observation interval have been made, it must be remembered that the kernel of the Hilbert transform integral possesses a simple pole at ω_0. Hence, any reasonable extrapolation of the function outside the interval will not seriously affect the transformation for values of ω_0 which are in the observation interval and not too close to the high and low frequency bounds.

Using the procedure for evaluating Hilbert transforms given by Čížek (1970), the real part of the inductive response function g, was Hilbert transformed in the range $0<\omega<0.628\,\mathrm{rad\,s^{-1}}$, where the necessary values for extrapolation were taken as described above. The derived function $g_H(\omega)=-\mathscr{H}(g(\omega))$ is also displayed in Fig. 8 where a very good correlation between it and $-h(\omega)$, i.e., the imaginary part of $\bar{C}(\bar{\omega})$, in the range $0.015<\omega<0.040\,\mathrm{rad\,s^{-1}}$ can be seen. The underestimate of the function $g_H(\omega)$ is possibly an artifact of the enforced cut-off of the slowly decaying function $g(\omega)$ at high frequencies.

This comparison also gives much weight to the argument that the values determined for the function $C(\omega)$ describe a *physically realisable* system.

5.2.1. Do the estimates of $C(\omega)$ comply with the inequality constraints? Weidelt (1972) lists 22 inequality constraints imposed on $C(\omega)$, 9 of which are given here by Eqs. (8)–(12). In order to discover if the estimates of $C(\omega)$ satisfy these 9 constraints, the two curves for $g(T)$ and $h(T)$ were differentiated by determining the derivative of the Lagrangian interpolation polynomial of degree 2 relevant to the three neighbouring values of g_i and h_i on a logarithmic period scale (Hildebrand 1956).

The derivatives $\partial g/\partial(\log T)$ and $\partial h/\partial(\log T)$ are illustrated, together with $g(T)$ and $h(T)$, in Figs. 10a and b respectively. The second derivatives $\partial^2 g(T)/\partial(\log T)^2$ and $\partial^2 h(T)/\partial(\log T)^2$ were derived by differentiating, in a similar manner, the first derivatives. These are also illustrated in Figs. 10a and b. The constraints given in Eqs. (8)–(12) were tested for, and the results are compiled in Table 2. It is apparent from Table 2 that inequalities (11a) and (12a), viz.

$$|DC|\leqq h, \quad |D^2 C|\leqq h$$

are not upheld at any period, with the one exception of (12a) at 1490 s. This was also found to be fairly consistently true of all the individual event estimates of $C(\omega)$, whereas the other constraints were, in general, upheld. This is taken as an indication that the data are not totally consistent with a 1 D Earth conductivity-depth distribution model, but arise from effects due to 2- or 3-dimensionality. However, it is believed that a 1 D model is a good approximation.

5.2.2. Does $c(t)$ Display the Required Form? As described in Sect. 2.3 (see acknowledgements for reference), the inductive impulse response function, $c(t)$, given by the inverse Fourier transform of $C(\omega)$, i.e., $c(t)=\mathscr{F}^{-1}\{C(\omega)\}$, must display the form

$$c(t)>0,\ c'(t)<0,\ c''(t)>0 \quad \text{for } 0<t<\infty.$$

The estimated function $c(t)$, illustrated in Fig. 9, is certainly a positive function in the range $t>0$, and differentiation of this function, undertaken in the same manner as described in 5.2.1, showed that the above conditions on the gradient and curvature of $c(t)$ are also fully upheld (Fig. 9). These three requirements on the form of $c(t)$ are a very strong constraint imposed on it, and that they are all displayed by the estimate determined is a powerful indication that $\bar{c}(t)$, and hence $\bar{C}(\bar{\omega})$, is interpretable in terms of the conductivity structure under northern Sweden.

6. Implications

A full interpretation of the geomagnetic and magneto-telluric results will be presented in a later paper in this series. However, it is considered worthwhile to examine the implications of the determined $\bar{C}(\bar{\omega})$ function in a preliminary manner here.

The ρ^*-z^* inversion of the $\bar{C}(\bar{\omega})$ data listed in Table 1 is shown, with the 95 % confidence limits, in Fig. 11. The data show a clear trend towards a more conducting layer at depth. Also included in the figure is the two-layer model determined from the first three points, using the algorithm of Fischer and Schnegg (in press 1980), and the preliminary results of a Monte-Carlo inversion, of the type proposed in Jones and Hutton (1979), of the magneto-telluric data recorded at Nattavaara (NAT, see Fig. 2) in northern Sweden, some 120 km south of Kiruna (Jones and Olafsdottir 1979).

The apparent resistivity and phase data, derived from

$$\bar{\rho}_a(\bar{\omega})=\omega\mu_0|\bar{C}(\bar{\omega})|^2,$$

$$\bar{\phi}_a(\bar{\omega})=-\tan^{-1}\left(\frac{\mathrm{Re}(\bar{C}(\bar{\omega}))}{\mathrm{Im}(\bar{C}(\bar{\omega}))}\right)$$

is shown in Fig. 12. A 3-layer 1 D model of parameters $\rho_1=10^4\,\Omega$m; $\rho_2=125\,\Omega$m; $\rho_3=3.5\,\Omega$m; $d_1=30$ km; $d_2=140$ km, is also illustrated. Although the correspondence between the data and the theoretically calculated values is not good, the essential features of a resistive upper layer of about $10^4\,\Omega$m underlain by a layer of some 100 Ωm, with a highly conducting basement at a depth of around 140 km are well described.

7. Source Field Effects

In this work, the inductive response function $C(\omega,0)$ has been explicitly derived in a strict manner due to fitting the observed fields to a second order polynomial. Thus, it is not necessary to consider the effects of non-uniform source fields on the derived function when interpreting the data. Other methods may derive response functions which are however affected by the non-uniformity of the source field, for example, standard magneto-telluric analysis ($Z(\omega,k)$), the internal-external field ratios ($S(\omega,k)$), or functional representations of the field other than a

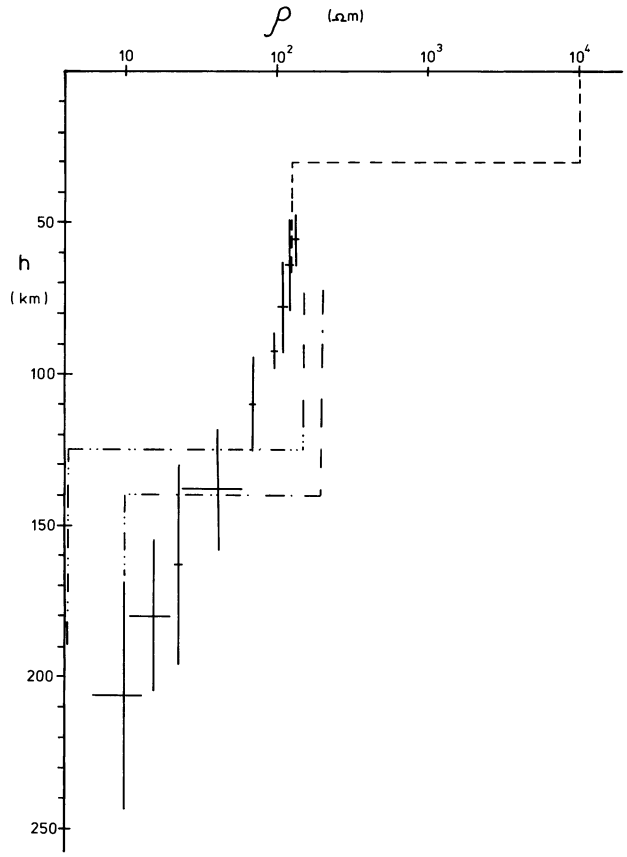


Fig. 11. The Schmucker $\rho^* - z^*$ inversion of the derived function with corresponding 95 % confidence intervals. Also shown in the Fischer-Schnegg two-layer inversion (- - -) of the first three data points and the bounds of the Monte-Carlo inversion (- · · - · · -) of the long period magneto-telluric data recorded at NAT (see Fig. 2)

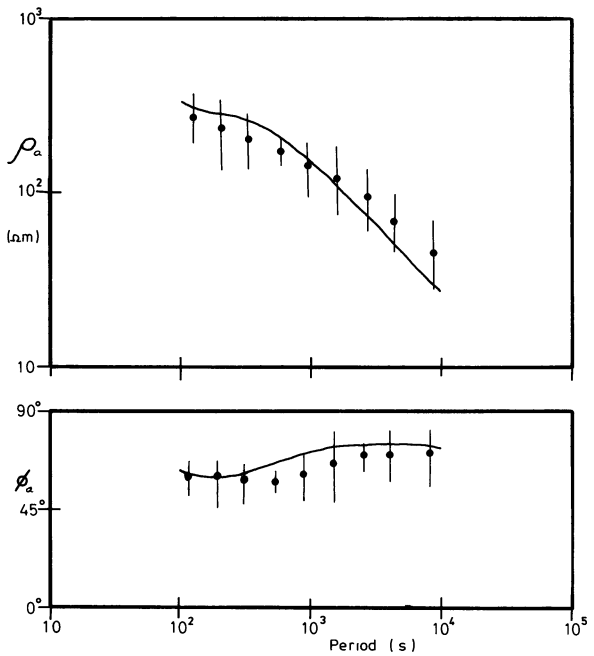


Fig. 12. The fit of a theoretical model of parameters: $\rho_1 = 10^4\,\Omega$m; $\rho_2 = 125\,\Omega$m; $\rho_3 = 3.5\,\Omega$m; $d_1 = 30$ km; $d_2 = 140$ km; to the derived inductive response function, expressed in terms of apparent resistivity and phase (*dots*), both with associated 95 % confidence limits shown

second-order polynomials in order to derive $C(\omega, k)$ from the HSG method. Hence, it is considered a worthwhile exercise to investigate the effects on $C(\omega, 0)$ of non-zero values of wavenumber for the derived one-dimensional conductivity model, consistent with the magnetic data observed in northern Sweden.

As was shown by Price (1962) and Srivastava (1965), the effect of a non-uniform source field is to change the layer propagation constant from k_i, where k_i is given by

$$k_i = k_i(\omega, 0) = \left(\frac{\omega \mu_i \sigma_i}{2}\right)^{1/2} + j\left(\frac{\omega \mu_i \sigma_i}{2}\right)^{1/2}$$

$k(\omega, 0)$-uniform field propagation constant in the i'th layer μ_i-magnetic permeability of the i'th layer, σ_i-conductivity of the i'th layer

to a value of

$$k_i(\omega, k) = \frac{1}{2^{\frac{1}{2}}}\left(([|k_i(\omega,0)|^4 + k^4]^{1/2} + k^2)^{1/2} + j([|k_i(\omega,0)|^4 + k^4]^{1/2} - k^2)^{1/2}\right)$$

where $k = 2\pi/\lambda$ and is the Price wavenumber.

The correction only becomes important for k of the order of $|k_i(\omega, 0)|$. For the three layers of the model proposed in Sect. 6, and for the longest period of interest [i.e., the smallest possible value of $k_i(\omega, 0)$], the following hold:

layer 1 $k \sim |k_1(\omega, 0)|$ when $\lambda \sim 10^4$ km,
layer 2 $k \sim |k_2(\omega, 0)|$ when $\lambda \sim 10^3$ km,
layer 3 $k \sim |k_3(\omega, 0)|$ when $\lambda \sim 200$ km,

hence the uppermost highly resistive layer is very sensitive to non-uniform fields. Figure 13 illustrates the effects of finite wavelengths on the parameters g and h. As can be seen from the figure, wavelengths of the order of 2×10^3 km or greater do not seriously affect g and h for periods less than 10^4 s.

Representation of the source configuration by a wavenumber, k, implicitly assumes, analogously to Fourier transformation, that the source is repeated in form every λ (wavelength $\lambda = 2\pi/k$). This approximation is valid directly underneath an electrojet maximum, where the effective wavelength, λ_{eff}, is equal to four times the electrojet half-width, but becomes increasingly worse with distance away from the maximum. Using the programme of Hibbs and Jones (1976a, b), the variation of $C(\omega)$ with distance away from the centre of a Gaussian-shaped electrojet of half-width 500 km over the Earth model proposed in Sect. 6 was calculated and is also illustrated in Fig. 13. As can be seen by the spread of the determined g and h functions, under the centre of the electrojet $\lambda_{\text{eff}} = 2{,}000$ km $= 4 \times$ half-width. Moving away from the centre, $\lambda_{\text{eff}} = \infty$ at the point directly underneath the half-width (i.e. ± 250 km). Moving yet further away gives values *greater than* those for $k = 0$, this arises because the effective wavenumber is complex.

The sources in Scandinavia are, as a rule, moving rather than stationary. As shown by Hermance (1978), this feature aids the magneto-telluric relationship, and hence correspondingly the inductive response function, in that a moving source has a larger

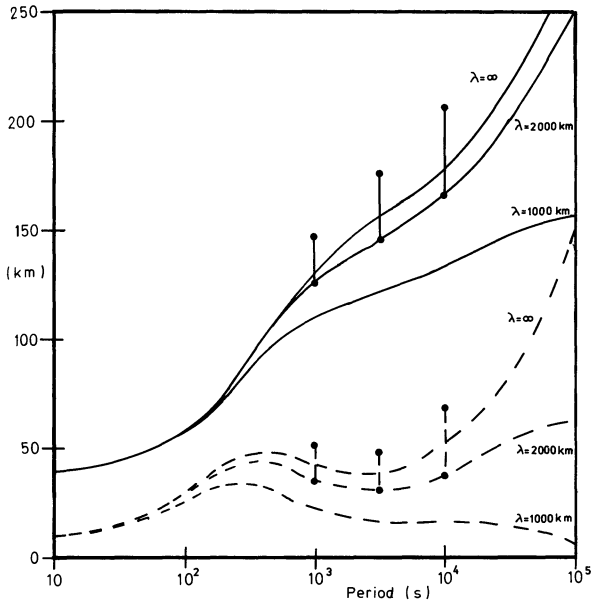


Fig. 13. Source effect on the parameters g (——) and h (— — —) for wavelengths of ∞, 2,000 km and 1,000 km. Also shown are the bounds of the parameters derived for a Gaussian electrojet source of half-width 500 km. The lower points correspond to measurements directly underneath the electrojet maximum, whilst the upper ones are for a position ± 700 km from this maximum

"effective width" through averaging induction effects over a larger volume of the Earth during its lateral movement. If a Gaussian electrojet of half-width 500 km (as in the previous discussion) moves with a velocity of $500\,\mathrm{m\,s^{-1}}$, then $C(\omega,k) \sim C(\omega,0)$ for periods up to 5 h (2×10^4 s).

8. Conclusions

That magnetometer array studies are able not only to map the conductivity anomalies but also to determine the conductivity-depth distribution, where valid, has been known for a decade (Schmucker 1970). However, apart from three other studies, two of which derived the spatial gradient from Fourier maps (Kuckes 1973a; Lilley and Sloane 1976) and the third restricted to long periods (Woods and Lilley, 1979), the technique has not been much employed.

In this paper, the spatial gradient was determined by least-squares methods and the inductive response function by statistical frequency analysis techniques. The derived function was shown to be causal by two methods, *i*) the impulse response function $\bar{c}(\tau)$ did not react before receiving any input, and *ii*) the Hilbert Transform of the real part of $\bar{C}(\bar{\omega})$ closely resembles the imaginary part of $-\bar{C}((\bar{\omega})$. It was also shown to pass 7 of Weidelt's (1972) 9 inequality constraints imposed on it, and to be fully consistent in the time domain.

Because of the rather special property of the inductive response function, $C(\omega,k)$, that it is a function of k^2 and not of k, representation of the observed fields in terms of second-order polynomials results in a determination of $C(\omega,0)$, i.e., the uniform field response. Hence, this permits the data to be interpreted without regard to source-field configuration. A very preliminary conductivity-depth profile that explains the major details of the derived response function consists of three layers with the parameters ; $\rho_1 = 10^4\,\Omega$m; $\rho_2 = 125\,\Omega$m; $\rho_3 = 3.5\,\Omega$m; $d_1 = 30$ km; $d_2 = 140$ km. Using this model as a base, the effects of source contribution were shown to be important only for stationary sources of very small wavelength ($\lambda < 1{,}000$ km) and at a point directly underneath the source maximum. Although the parameter k, wavenumber ($k = 2\pi/\lambda$), is independent of the model parameters (Edwards et al. 1980), its effect closely resembles that of moving the conductive basement nearer to the surface. Hence it could be expected that other induction studies in the region, whose derived responses are affected by source-structure, would indicate a less deep highly conducting basement. Where magnetic variations exhibit a wavenumber which is a frequency constant, i.e., $k \neq fn(\omega)$, the derived conductivity profile from such studies may be transformed, by the algorithm given in Weidelt (1972), to yield the true conductivity-depth profile.

Acknowledgements. The author wishes to thank all those acknowledged in the paper by Küppers et al. (1979), and in addition J. Untiedt for continual support and guidance, F. Küppers for many reasons, and W. Baumjohann for discussion concerning source structure. The author is also very grateful to P. Weidelt for detailing the conditions placed on the inductive impulse response function, $c(t)$, as described in Sect. 2.3. All three referees are thanked for their constructive comments on an earlier version of this manuscript.

This work was supported in full by grants from the Deutsche Forschungsgemeinschaft, which are gratefully acknowledged.

References

Baumjohann, W., Greenwald, R.A., Küppers, F.: Joint magnetometer array and radar backscatter observations of auroral currents in northern Scandinavia. J. Geophys. **44**, 373–383, 1978

Baumjohann, W., Untiedt, J., Greenwald, R.A.: Joint two-dimensional observations of ground magnetic and ionospheric electric fields associated with auroral zone currents. 1. Three-dimensional current flows associated with a substorm-intensified eastward electrojet. J. Geophys. Res. **85**, 1963–1978, 1980

Bendat, J.S., Piersol, A.G.: Random data: analysis and measurement procedures. New York: Wiley-Interscience 1971

Berdichevsky, M.N., Fainberg, E.B., Rotanova, N.M., Smirnov, J.B., Vanyan, L.L.: Deep electromagnetic investigations. Ann. Géophys. **32**, 143–155, 1976

Cagniard, I.: Basic theory of the magnetotelluric method of geophysical prospecting. Geophysics **18**, 605–635, 1953

Champeney, D.C.: Fourier transforms and their physical application. London: Academic Press 1973

Čižek, V.: Discrete Hilbert Transform. IEEE Trans. Audio Electroacoust. **AU-18,** 340–343, 1970

Edwards, R.N., Bailey, R.C., Garland, G.D.: Crustal and upper mantle electrical conductivity studies with natural and artificial sources. Geophysics, submitted 1980

Ferris, C.D.: Linear network theory. Columbus, Ohio: Charles E. Merrill Books Inc. 1962

Fischer, G., Schnegg, P.-A.: The dispersion relations of the magnetotelluric response and their incidence on the inversion problem. Geophys. J.R. Astron. Soc. in press. 1980

Frazer, M.C.: Geomagnetic sounding with arrays of magnetometers. Rev. Geophys. Space Phys. **12**, 401–420, 1974

Golub, G.: Numerical methods for solving linear least-squares problems. Numerische Math. **7**, 206–216, 1965

Gough, D.I.: The interpretation of magnetometer array studies. Geophys. J.R. Astron. Soc. **35**, 85–98, 1973a

Gough, D.I.: The geophysical significance of geomagnetic variation anomalies. Phys. Earth Planet. Inter **7**, 379–388, 1973b

Gough, D.I., Reitzel, J.S.: A portable three component magnetic variometer. J. Geomagn. Geoelectr. **19**, 203–215, 1967

Hermance, J.F.: Electromagnetic induction in the Earth by moving ionospheric current systems. Geophys. J.R. Astron. Soc. **55**, 557–576, 1978

Hibbs, R.D., Jones, F.W.: The calculation of the electromagnetic fields of a sheet current source with arbitrary spatial intensity distribution over a layered half-space-I. The general method and results. Geophys. J.R. Astron. Soc. **46**, 433–452, 1976a

Hibbs, R.D., Jones, F.W.: The calculation of the electromagnetic fields of a sheet current source with arbitrary spatial intensity distribution over a layered half space – II. The computer program and its application. Geophys. J.R. Astron. Soc. **46**, 453–465, 1976b

Hildebrand, F.B.: Introduction to numerical analysis. New York: McGraw-Hill 1956

Jones, A.G.: Geomagnetic induction studies in southern Scotland. Ph.D. Thesis, University of Edinburgh 1977

Jones, A.G.: New coherence functions useful for determining the signal to random noise ratio in multivariate studies. IEEE Trans. Acoust., Speech, Signal Processing, submitted 1979

Jones, A.G.: On the reduction of bias in response function estimation for hand-digitised data. Geophys. J.R. Astron. Soc. submitted 1980

Jones, A.G., Hutton, R.: A multi-station magnetotelluric study in southern Scotland – II. Monte-Carlo inversion of the data and its geophysical and tectonic implications. Geophys. J.R. Astron. Soc. **56**, 351–368, 1979

Jones, A.G., Olafsdottir, B.: Geomagnetic induction studies in northern Scandinavia. Geophys. J.R. Astron. Soc. **57**, 265, 1979

Kanasewich, E.R.: Time sequence analysis in geophysics. Edmonton, Canada: University of Alberta Press 1973

Kuckes, A.F.: Relations between electrical conductivity of a mantle and fluctuating magnetic fields. Geophys. J.R. Astron. Soc. **32**, 119–131, 1973a

Kuckes, A.F.: Correspondence between the magnetotelluric and field penetration depth analysis for measuring electrical conductivity. Geophys. J.R. Astron. Soc. **32**, 381–385, 1973b

Küppers, F., Untiedt, J., Baumjohann, W., Lange, K., Jones, A.G.: A two-dimensional magnetometer array for ground-based observations of auroral zone electric currents during the International Magnetospheric Study. J. Geophys. **46**, 429–450, 1979

Lilley, F.E.M.: Magnetometer array studies: A review of the interpretation of observed fields. Phys. Earth Planet. Inter. **10**, 231–240, 1975

Lilley, F.E.M., Sloane, M.N.: On estimating electrical conductivity using gradient data from magnetometer arrays. J. Geomagn. Geoelectr. **28**, 321–328, 1976

Mersmann, U., Baumjohann, W., Küppers, F., Lange, K.: Analysis of an eastward electrojet by means of upward continuation of ground-based magnetometer data. J. Geophys. **45**, 281–298, 1979

Price, A.T.: The theory of magnetotelluric fields when the source field is considered. J. Geophys. Res. **67**, 1907–1918, 1962

Porath, H., Dziewonski, A.: Crustal resistivity anomalies from geomagnetic deep sounding studies. Rev. Geophys. Space Phys. **9**, 891–915, 1971

Schmucker, U.: Anomalies of Geomagnetic Variations in the Southwestern United States. Bull. Scripps Inst. Oceanogr., Univ. Calif. Press **13**, 1970

Schmucker, U., Weidelt, P.: Electromagnetic Induction in the Earth. Lect. Notes, Aarhus University, 1975

Solodovnikov, V.V.: Introduction to the statistical dynamics of automatic control systems. Chapter 1. (Translation edited by J.B. Thomas and L.A. Zadeh, Dover Publications Inc., New York, 1960), 1952

Srivastava, S.P.: Method of interpretation of magneto-telluric data when source field is considered. J. Geophys. Res. **70**, 945–954, 1965

Untiedt, J., Pellinen, R., Küppers, F., Opgenoorth, H.J., Pelster, W.D., Baumjohann, W., Ranta, H., Kangas, J., Czechowsky, P., Heikkila, W.J.: Observations of the initial development of an auroral and magnetic substorm at magnetic midnight. J. Geophys. **45**, 41–65, 1978

Weidelt, P.: The inverse problem of geomagnetic induction. J. Geophys. **38**, 257–289, 1972

Weidelt, P.: Entwicklung und Erdprobung eines Verfahrens zur Inversion zweidimensionaler Leitfähigkeitsstrukturen in E-Polarisation. Habilitationsschrift Math.-Naturwiss. Fak. Univ. Göttingen, 1978

Westerlund, S.: Magnetotelluric experiments in the range 0.01 Hz to 10 kHz. KGO Report No. 72:10, Kiruna Geophysical Observatory, 1972

Woods, D.V., Lilley, F.E.M.: Geomagnetic induction in central Australia. J. Geomagn. Geoelectr. **31**, 449–458, 1979

Received April 11, 1980; Revised Version May 12, 1980
Accepted May 23, 1980

J. Geophys. 48, 195–198, 1980

Journal of Geophysics

Comments on Polarization and Coherence

J. C. Samson
Institute of Earth and Planetary Physics, Department of Physics, University of Alberta, Edmonton, Alberta, Canada T6G2J1

Abstract. Measures of the frequency-dependent coherence. and the frequency-dependent degree of polarization have been used for some time in describing the statistical characteristics of multichannel geophysical data. The degree of polarization is rotationally invariant, suggesting that this measure is more useful than the coherence in describing the statistical characteristics of the spatial components of waves. In the present paper generalizations of the bivariate coherence and the degree of polarization for data with an arbitrary number of channels are developed, and the relationship of the coherence to the degree of polarization is described. Some suggestions for applications of the measure of the degree of polarization in the selection and filtering of geophysical data are given.

Key words: Coherence – Polarization – Polarization filters.

Introduction

Measures of the coherence and degree of polarization of vector time-series have been used for some time in evaluating the statistical characteristics of geophysical processes. Jones (1979) has indicated that there does not yet seem to be much appreciation of the close connections between coherence and polarization, at least in the geophysical literature. Jones has limited his discussion and conclusions to two-dimensional processes, but has stated that 'they are equally valid however in three dimensions'. This statement might be somewhat misleading, and consequently I would like to present a more comprehensive evaluation of the measures of coherence and polarization in processes of arbitrary dimension n.

There are numerous reasons for extending these concepts to vector processes of arbitrary dimension, rather than restricting the discussion to three-dimensional spaces. Some examples will serve to illustrate this point. In magnetotellurics, four time series are measured, and the vector process is $\mathbf{x}^T(t)=[e_1(t), e_2(t), b_1(t), b_2(t)]$ where e_j and b_j $(j=1,2)$ are the electric and magnetic field components respectively. The descriptions of VLF waves in the magnetosphere require a 6-dimensional vector, with 3 magnetic and 3 electric field components (Storey and Lefeuvre 1979). Finally, any array of instruments can be considered to be an $m \times p$ dimensional process, where m is the number of instruments and p is the number of spatial components measured by each instrument.

The Spectral Matrix, Coherence, and Polarization

The information in the vector process $\mathbf{x}^T(t)=[x_1(t), x_2(t) \ldots x_n(t)]$ can be conveniently represented in the frequency domain by using the spectral matrix $\mathbf{S}$ where

$$\mathbf{S}(f,\delta)=\int_{f-\delta}^{f+\delta}\int_{-\infty}^{\infty}\mathbf{C}(\tau)e^{-2\pi i g\tau}d\tau\,dg, \tag{1}$$

$$\mathbf{C}(\tau)=\langle\mathbf{x}(t)\mathbf{x}^T(t+\tau)\rangle, \tag{2}$$

and $\langle\ \rangle$ denotes the expectation.

The coherence, γ^2, in a coordinate system with bases (j,k) is defined by

$$\gamma_{jk}^2=S_{jk}S_{kj}/(S_{jj}S_{kk}), \tag{3}$$

with $0\leqq\gamma_{jk}^2\leqq 1$. For arbitrary directions $\mathbf{r}_1$ and $\mathbf{r}_2$ ($\mathbf{r}_1$ and $\mathbf{r}_2$ are real and orthogonal, i.e., $\mathbf{r}_1^T\mathbf{r}_2=0$), Eq. (3) becomes

$$\gamma^2(\mathbf{r}_1,\mathbf{r}_2)=(\mathbf{r}_1^T\mathbf{S}\mathbf{r}_2\ \mathbf{r}_2^T\mathbf{S}\mathbf{r}_1)/(\mathbf{r}_1^T\mathbf{S}\mathbf{r}_1\,\mathbf{r}_2^T\mathbf{S}\mathbf{r}_2). \tag{4}$$

Equation (4) can be written in a more standard operator format by noting that

$$\mathbf{r}_1^T\mathbf{S}\mathbf{r}_2=\mathrm{Tr}((\mathbf{r}_2\mathbf{r}_1^T)\mathbf{S}), \tag{5}$$

where Tr denotes the sum of the diagonal elements. Then,

$$\gamma^2(\mathbf{r}_1,\mathbf{r}_2)=\mathrm{Tr}(\mathbf{R}_{12}\mathbf{S})\,\mathrm{Tr}(\mathbf{R}_{21}\mathbf{S})/(\mathrm{Tr}(\mathbf{R}_{11}\mathbf{S})\,\mathrm{Tr}(\mathbf{R}_{22}\mathbf{S})) \tag{6}$$

where $\mathbf{R}_{jk}=\mathbf{r}_j\mathbf{r}_K^T$.

For reasons to be discussed later, it is necessary to further generalize the definition of the coherence by defining this parameter for vector $\mathbf{u}_j$ in a unitary space. A linear vector space will be called unitary if the components of the vector are from the field of complex numbers, and the inner product has the following properties

$$\mathbf{u}_j^\dagger\mathbf{u}_k=(\mathbf{u}_k^\dagger\mathbf{u}_j)^*, \tag{7a}$$

$$(a\mathbf{u}_j)^\dagger\mathbf{u}_k=a\mathbf{u}_j^\dagger\mathbf{u}_k, \tag{7b}$$

$$(\mathbf{u}_j+\mathbf{u}_k)^\dagger\mathbf{u}_l=\mathbf{u}_j^\dagger\mathbf{u}_l+\mathbf{u}_k^\dagger\mathbf{u}_l, \quad \text{and} \tag{7c}$$

$$\mathbf{u}_j^\dagger\mathbf{u}_j=b. \tag{7d}$$

0340-062X/80/0195/$01.00

The symbol † denotes the hermitean adjoint or complex conjugate of the transpose, a and b are real numbers, and b is non-negative.

For vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ in a unitary space ($\mathbf{u}_1^\dagger \mathbf{u}_2 = 0$), the coherence is

$$\gamma^2(\mathbf{u}_1, \mathbf{u}_2) = \mathrm{Tr}(\mathbf{u}_{12}\mathbf{S})\mathrm{Tr}(\mathbf{u}_{21}\mathbf{S})/(\mathrm{Tr}(\mathbf{u}_{11}\mathbf{S})\mathrm{Tr}(\mathbf{u}_{22}\mathbf{S})) \tag{8}$$

where $\mathbf{U}_{jk} = \mathbf{u}_j \mathbf{u}_k^\dagger$. Henceforth, the vectors $\mathbf{r}_j$ will always be in a real space, and the vectors $\mathbf{u}_j$ will be in a unitary space.

In defining the degree of polarization for $n=2$, Jones has used the expansion (Born and Wolf 1964)

$$\mathbf{S} = \mathbf{P} + \mathbf{N} \tag{9}$$

where $\mathbf{P}$ is totally polarized, $\det[\mathbf{P}] = 0$, and $\mathbf{N}$ is unpolarized, $\mathbf{N} = \begin{bmatrix} N & 0 \\ 0 & N \end{bmatrix}$. The degree of polarization R is given by the ratio of the polarized power to the total power, or

$$R = \mathrm{Tr}\,\mathbf{P}/\mathrm{Tr}\,\mathbf{S}. \tag{10}$$

The expansion in Eq. (8) is not possible for $n>2$ (Samson 1973), and consequently an alternative method must be found to define the degree of polarization.

We first expand $\mathbf{S}$ in the form

$$\mathbf{S} = \sum_{j=1}^{n} \varepsilon_j \mathbf{u}_j \mathbf{u}_j^\dagger \tag{11}$$

where the ε_j and $\mathbf{u}_j$ are respectively the eigenvalues and eigenvectors of $\mathbf{S}$. If $\mathbf{S}$ is purely polarized or is a *pure state*, then $\mathbf{S}$ has only one nonzero eigenvalue and

$$\mathbf{S} = \varepsilon_1 \mathbf{u}_1 \mathbf{u}_1^\dagger. \tag{12}$$

In this case, the vector process can be considered to be 'polarized' in the direction of the unitary vector $\mathbf{u}_1$. Since the information on the polarization is in the eigenvalues of $\mathbf{S}$, we must consider the characteristic equation for the eigenvalues, which is given by

$$\sum_{l=0}^{n} \alpha_l \varepsilon^l = 0. \tag{13}$$

If $\alpha_{n-2} = 0$, then $\mathbf{S}$ is a pure state.

The first three coefficients of the characteristic equation can be written in terms of the invariants $\mathrm{Tr}\,\mathbf{S}$ and $\mathrm{Tr}\,\mathbf{S}^2$ (Samson and Olson 1980) giving

$$\begin{aligned} \alpha_n &= 1.0 \\ \alpha_{n-1} &= -\mathrm{Tr}\,\mathbf{S} \\ \alpha_{n-2} &= \tfrac{1}{2}((\mathrm{Tr}\,\mathbf{S})^2 - \mathrm{Tr}\,\mathbf{S}^2). \end{aligned} \tag{14}$$

Since

$$(\mathrm{Tr}\,\mathbf{S})^2 - (\mathrm{Tr}\,\mathbf{S}^2) \leqq \frac{n-1}{n}(\mathrm{Tr}\,\mathbf{S})^2, \tag{15}$$

(Samson and Olson 1980) we have

$$0 \leqq n(2\alpha_{n-2})/(n-1)(\mathrm{Tr}\,\mathbf{S})^2 \leqq 1, \tag{16}$$

and we can choose our degree of polarization R^2 to be

$$R^2 = (n\,\mathrm{Tr}\,(\mathbf{S}^2) - (\mathrm{Tr}\,\mathbf{S})^2)/((n-1)(\mathrm{Tr}\,\mathbf{S})^2). \tag{17}$$

If $n=2$, Eq. (16) becomes

$$R^2 = (2(\mathrm{Tr}\,\mathbf{S}^2) - (\mathrm{Tr}\,\mathbf{S})^2)/(\mathrm{Tr}\,\mathbf{S})^2 = 1 - \frac{4\det[\mathbf{S}]}{(\mathrm{Tr}\,\mathbf{S})^2} \tag{18}$$

which is the parameter derived by Born and Wolf (1964).

A Comparison of Coherence and the Degree of Polarization

Having derived the generalizations for the coherence [Eq. (8)], and the degree of polarization [Eq. (17)], we are now in a position to compare the two parameters. Jones (1979) pointed out that in two dimensions, the maximum value of $\gamma^2(\mathbf{r}_1, \mathbf{r}_2)$, under a real rotation, obeys the equality

$$\max(\gamma^2(\mathbf{r}_1, \mathbf{r}_2)) = R^2 \qquad (n=2) \tag{19}$$

Relation (19) was derived earlier by Wolf (1959) and by Parrent and Roman (1960) in the study of partially polarized light.

In generalizing the case to $n>2$, I shall show that Eq. (19) is *not true* in general, and that the maximum in the coherence is found by a unitary, not orthonormal (real) transformation. Maximization in a unitary space leads to a maximum value for the coherence which is, for all n, similar to the value for $n=2$. Maximization in a real space leads to quite different results.

The maximization of $\gamma^2(\mathbf{u}_1, \mathbf{u}_2)$ is an extremum problem in a unitary (complex) vector space, and we must determine stationary values of $\gamma^2(\mathbf{u}_1, \mathbf{u}_2)$ where $\mathbf{u}_1$ and $\mathbf{u}_2$ are orthogonal vectors in a unitary space. Thus we must find the stationary points $\mathbf{u}_1$ and $\mathbf{u}_2$ where

$$\begin{aligned} &\gamma^2(\mathbf{u}_1 + \delta\mathbf{u}_1, \mathbf{u}_2) - \gamma^2(\mathbf{u}_1, \mathbf{u}_2) = 0 \quad \text{and} \\ &\gamma^2(\mathbf{u}_1, \mathbf{u}_2 + \delta\mathbf{u}_2) - \gamma^2(\mathbf{u}_1, \mathbf{u}_2) = 0 \end{aligned} \tag{20}$$

with $\delta\mathbf{u}_1$ and $\delta\mathbf{u}_2$ arbitrarily small changes, and subject to the constraint $\mathbf{u}_1^\dagger \mathbf{u}_2 = 0$ (orthogonality).

Using a Lagrangian multiplier formalism, we look for the vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ which give stationary values for the function A where

$$\begin{aligned} A = &(\mathbf{u}_1^\dagger \mathbf{S}\mathbf{u}_2)(\mathbf{u}_2^\dagger \mathbf{S}\mathbf{u}_1) + \lambda_0(1 - \mathbf{u}_1^\dagger \mathbf{S}\mathbf{u}_1 \mathbf{u}_2^\dagger \mathbf{S}\mathbf{u}_2) \\ &+ \lambda_1 \mathbf{u}_1^\dagger \mathbf{u}_2 + \lambda_2 \mathbf{u}_2^\dagger \mathbf{u}_1. \end{aligned} \tag{21}$$

The denominator in Eq. (8) has been eliminated by using the constraint with multiplier λ_0. The derivatives $\partial/\partial\lambda_0$, $\partial/\partial\lambda_1$, $\partial/\partial\lambda_2$ give the constraints in finding the stationary values. Then we also have the two equations

$$\begin{aligned} &A(\mathbf{u}_j + \delta\mathbf{u}_j, \mathbf{u}_k) - A(\mathbf{u}_j, \mathbf{u}_k) = 0 \\ &(j=1,\ k=2;\ j=2,\ k=1). \end{aligned} \tag{22}$$

Expanding (22) and neglecting terms second order in $\mathbf{u}_1$ and $\mathbf{u}_2$ we get the two vector-equations

$$\begin{aligned} &\delta\mathbf{u}_j^\dagger(\mathbf{u}_k^\dagger \mathbf{S}\mathbf{u}_j \mathbf{S}\mathbf{u}_k - \lambda_0 \mathbf{u}_k^\dagger \mathbf{S}\mathbf{u}_k \mathbf{S}\mathbf{u}_j + \lambda_j \mathbf{u}_k) \\ &+ (\mathbf{u}_j^\dagger \mathbf{S}\mathbf{u}_k \mathbf{u}_k^\dagger \mathbf{S} - \lambda_0 \mathbf{u}_k^\dagger \mathbf{S}\mathbf{u}_k \mathbf{u}_j^\dagger \mathbf{S} + \lambda_k \mathbf{u}_k^\dagger)\,\delta\mathbf{u}_j = 0 \\ &(j=1,\ k=2;\ j=2, k=1). \end{aligned} \tag{23}$$

Thus we now have two equations of the form

$$\delta\mathbf{u}_1^\dagger \mathbf{a}_1 + \mathbf{a}_2^\dagger \delta\mathbf{u}_1 = 0 \tag{24a}$$

$$\delta\mathbf{u}_2^\dagger \mathbf{b}_1 + \mathbf{b}_2^\dagger \delta\mathbf{u}_2 = 0. \tag{24b}$$

If we consider the variation in the direction $i\delta\mathbf{u}_1$ and $i\delta\mathbf{u}_2$, then Eqs. (24a) and (24b) become

$$\delta\mathbf{u}_1^\dagger\mathbf{a}_1-\mathbf{a}_2^\dagger\delta\mathbf{u}_1=0 \tag{25a}$$

$$\delta\mathbf{u}_2^\dagger\mathbf{b}_1-\mathbf{b}_2^\dagger\delta\mathbf{u}_2=0. \tag{25b}$$

It then follows that

$$\mathbf{a}_1=\mathbf{a}_2=\mathbf{b}_1=\mathbf{b}_2=0 \tag{26}$$

giving four vector-equations.

To determine λ_0 we use the equation

$$\mathbf{u}_1^\dagger\mathbf{a}_1=0 \tag{27}$$

giving $\lambda_0=\gamma^2(\mathbf{u}_1,\mathbf{u}_2)$. We also note that $\mathbf{b}_1-\mathbf{b}_2=\mathbf{0}$, and consequently $\lambda_1=\lambda_2^*$. Finally we solve for λ_1 using

$$\mathbf{u}_2^\dagger\mathbf{a}_1=0 \quad \text{and} \tag{28a}$$

$$\mathbf{b}_2^\dagger\mathbf{u}_1=0, \tag{28b}$$

giving the two equations

$$\lambda_1=\mathbf{u}_2^\dagger\mathbf{S}\mathbf{u}_1\frac{\mathbf{u}_1^\dagger\mathbf{S}\mathbf{u}}{\mathbf{u}_1^\dagger\mathbf{u}_1}(1-\gamma^2(\mathbf{u}_1,\mathbf{u}_2)), \quad \text{and} \tag{29a}$$

$$\lambda_1=\mathbf{u}_2^\dagger\mathbf{S}\mathbf{u}_1\frac{\mathbf{u}_2^\dagger\mathbf{S}\mathbf{u}_2}{\mathbf{u}_2^\dagger\mathbf{u}_2}(1-\gamma^2(\mathbf{u}_1,\mathbf{u}_2)). \tag{29b}$$

The multiplier λ_1 is nonzero only if

$$\frac{\mathbf{u}_1^\dagger\mathbf{S}\mathbf{u}_1}{\mathbf{u}_1^\dagger\mathbf{u}_1}=\frac{\mathbf{u}_2^\dagger\mathbf{S}\mathbf{u}_2}{\mathbf{u}_2^\dagger\mathbf{u}_2}. \tag{30}$$

Since the value of $\gamma^2(\mathbf{u}_1,\mathbf{u}_2)$ is independent of the magnitudes of $\mathbf{u}_1$ and $\mathbf{u}_2$, we can choose $\mathbf{u}_1^\dagger\mathbf{u}_1=\mathbf{u}_2^\dagger\mathbf{u}_2$, and consequently the condition (30) becomes $\mathbf{u}_1^\dagger\mathbf{S}\mathbf{u}_1=\mathbf{u}_2^\dagger\mathbf{S}\mathbf{u}_2$.

If either $\mathbf{u}_2^\dagger\mathbf{S}\mathbf{u}_1=0$ or $\gamma^2(\mathbf{u}_1,\mathbf{u}_2)=1$, then we also have a solution for the stationary values. The former case obviously gives an absolute minimum since $\gamma^2=0$. The latter case gives an absolute maximum, and in this case $\mathbf{u}_1$ and $\mathbf{u}_2$ are not uniquely defined, but need only lie in the plane for which $\gamma^2=1$.

The equations $\mathbf{a}_1=\mathbf{0}$ and $\mathbf{b}_2=\mathbf{0}$ can now be written

$$\mathbf{S}\mathbf{u}_2-\alpha\mathbf{S}\mathbf{u}_1+\beta\mathbf{u}_2=0 \tag{31a}$$

$$\mathbf{S}\mathbf{u}_1-\alpha^*\mathbf{S}\mathbf{u}_2+\beta\mathbf{u}_1=0 \tag{31b}$$

where

$$\alpha=\frac{\gamma^2\mathbf{u}_1\mathbf{S}\mathbf{u}_1}{\mathbf{u}_2^\dagger\mathbf{S}\mathbf{u}_1}=\frac{\gamma^2\mathbf{u}_2\mathbf{S}\mathbf{u}_2}{\mathbf{u}_2^\dagger\mathbf{S}\mathbf{u}_1}, \quad \text{and}$$

$$\beta=\frac{\mathbf{u}_1^\dagger\mathbf{S}\mathbf{u}_1}{\mathbf{u}_1^\dagger\mathbf{u}_1}(1-\gamma^2(\mathbf{u}_1,\mathbf{u}_2)).$$

Solving (31a) and (31b) for $\mathbf{u}_1$ we find

$$\mathbf{u}_1=(\alpha\beta)^{-1}(\alpha^2-1)\mathbf{S}\mathbf{u}_2-\alpha^{-1}\mathbf{u}_2. \tag{32}$$

Substituting this value for $\mathbf{u}_1$ into Eq. (31a) we obtain the equation

$$\mathbf{A}\mathbf{u}_2+\beta\mathbf{u}_2=0 \tag{33}$$

where

$$\mathbf{A}=\left(\frac{1-\alpha^2}{\beta}\right)\mathbf{S}^2+2\mathbf{S}.$$

Equation (33) is an eigenvector equation, with eigenvalue β. Since $\mathbf{A}$ is a polynomial of the matrix $\mathbf{S}$, the n eigenvalues of $\mathbf{A}$ are $\left(\frac{1-\alpha^2}{\beta}\right)\varepsilon_j^2+2\varepsilon_j$ $(j=1,n)$, and we have the n equations

$$\beta^2=(1-\alpha^2)\varepsilon_j^2+2\beta\varepsilon_j \qquad (j=1,n). \tag{34}$$

Choosing any two equations and solving for γ^2 gives

$$(\gamma^2(\mathbf{u}_1,\mathbf{u}_2))=\frac{(\varepsilon_j-\varepsilon_k)^2}{(\varepsilon_j+\varepsilon_k)^2}. \tag{35}$$

An absolute maximum occurs if we choose ε_j to be the maximum eigenvalue of $\mathbf{S}$, and ε_k to be the minimum eigenvalue. Consequently

$$\max(\gamma^2(\mathbf{u}_1,\mathbf{u}_2))=\frac{(\varepsilon_1-\varepsilon_n)^2}{(\varepsilon_1+\varepsilon_n)^2}, \qquad \varepsilon_1\geqq\varepsilon_2\ldots\geqq\varepsilon_n. \tag{36}$$

In general, since the space of real vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ is a subspace of the n-dimensional unitary space, we have

$$\max(\gamma^2(\mathbf{r}_1,\mathbf{r}_2))\leqq\max(\gamma^2(\mathbf{u}_1,\mathbf{u}_2)). \tag{37}$$

The equality applies only when $\mathbf{A}$ in Eq. (33) is a real matrix. For $n=2$ $\mathbf{A}$ can always be written as a real matrix, with a suitable choice for $(1-\alpha^2)/\beta$, and thus the equality is true giving

$$\max(\gamma^2(\mathbf{r}_1,\mathbf{r}_2))=\frac{(\varepsilon_1-\varepsilon_2)^2}{(\varepsilon_1+\varepsilon_2)^2} \qquad (n=2). \tag{38}$$

To complete the discussion, we must consider the relative values of $\max(\gamma^2(\mathbf{u}_1,\mathbf{u}_2))$ and R^2 for arbitrary n. We first rewrite R^2 in the form

$$R^2=\frac{n}{n-1}\frac{\mathrm{Tr}(\mathbf{S}^2)}{(\mathrm{Tr}\,\mathbf{S})^2}-\frac{1}{n-1}. \tag{39}$$

Now we wish to maximize R^2 by varying the eigenvalues $\varepsilon_j(j=2, n-1)$, and thus we must maximize

$$\mathrm{Tr}(\mathbf{S}^2)/(\mathrm{Tr}\,\mathbf{S})^2=\sum_{j=1}^{n}\varepsilon_j^2/\left(\sum_{j=1}^{n}\varepsilon_j\right)^2 \tag{40}$$

subject to the constraint $\varepsilon_1\geqq\varepsilon_2\ldots\geqq\varepsilon_n\geqq0$.

Each derivative $\frac{\partial}{\partial\varepsilon_j}(\mathrm{Tr}(\mathbf{S}^2)/(\mathrm{Tr}\,\mathbf{S})^2)$ has at most one stationary point in the region allowed by the constraints, and this point is a local minimum. Thus the maximum must be found at the end points $\varepsilon_j=\varepsilon_1$ $(j=2, n-1)$ or $\varepsilon_j=\varepsilon_n$. Substitution of these two possible sets into (40) shows directly that $\varepsilon_j=\varepsilon_n$, and thus

$$\max(R^2)=\frac{(\varepsilon_1-\varepsilon_n)^2}{(\varepsilon_1+(n-1)\varepsilon_n)^2}. \tag{41}$$

Comparison of (35) and (41) then shows that

$$\max(\gamma^2(\mathbf{u}_1,\mathbf{u}_2))\geqq R^2. \tag{42}$$

Discussion

It is clear from the above discussion that the comparison of the coherence values and the degree of polarization is more complicated for $n>2$, and no simple relationship exists to compare

$\max(\gamma^2(\mathbf{u}_1, \mathbf{u}_2))$ or $\max(\gamma^2(\mathbf{r}_1, \mathbf{r}_2))$ with the degree of polarization R^2. For $n=2$, the comparison is far simpler with $\max(\gamma^2(\mathbf{r}_1, \mathbf{r}_2)) = \max(\gamma^2(\mathbf{u}_1, \mathbf{u}_2)) = R^2$ $(n=2)$.

Measures of the degree of polarization of a vector process are in many cases a more objective measure of the statistical characteristics of multichannel data than are the individual coherences. This is particularly true for the analysis of the spatial components of waves. Since the degree of polarization [Eq. (17)] is constructed from the scalar invariants $\mathrm{Tr}\mathbf{S}$ and $\mathrm{Tr}\mathbf{S}^2$, the value of this measure does not depend on the choice of the coordinate system for orienting the instrument. The value of the coherence depends on the choice of the coordinate-system. Thus the degree of polarization might be considered to be a more 'intrinsic' quantity of the waves.

Measures of the degree of polarization can have many practical applications in the analysis of geophysical data. In most cases, geophysicists are interested in extracting the more polarized waveforms from multidimensional data, and reducing the random noise component. To facilitate these studies, data-adaptive filters which use measures of the degree of polarization can be used to selectively enhance the pure states or polarized waveforms. For example, one can filter the n-dimensional data $\mathbf{x}^T(t) = [x_1(t), x_2(t) \ldots x_n(t)]$ by modulating the Fourier transform of $\mathbf{x}^T(t)$ with the frequency-dependent measure of the degree of polarization. The filtered data $\mathbf{y}(t)$ are then given by

$$\mathbf{y}(t) = \frac{1}{2\pi T} \int_{-\infty}^{\infty} R^2(f)\mathbf{s}(f)e^{2\pi i f t}\,df, \quad \text{where} \tag{43}$$

$$\mathbf{s}(f) = \int_{-T/2}^{T/2} \mathbf{x}(t)e^{-2\pi i f t}\,dt, \tag{44}$$

and T is the temporal length of the data. We have used filters of this type with considerable success in extracting waveforms from multichannel ULF magnetometer data and ULF riometer data.

For a second, and final, example of a practical application of measures of polarization consider the problems inherent in the selection of magnetotelluric data. The magnetotelluric theory assumes that $\mathbf{b} = \mathbf{Ze}$ where $\mathbf{b}$ is the horizontal magnetic field vector, $\mathbf{e}$ is the two component electric field vector, and $\mathbf{Z}$ is a complex impedance tensor. When the magnetotelluric vector $\mathbf{x}^T(t) = [b_1(t), b_2(t), e_1(t), e_2(t)]$ has a spectral matrix that is completely polarized $(R^2 = 1)$ at some particular frequency, then the data satisfy the impedance-tensor relation. It is possible that selection of magnetotelluric data by using criteria based on measures of the polarization might be more suitable than the coherence-based methods now being used (Goubau et al. 1978, and references therein).

Acknowledgements. The research was supported by the Natural Sciences and Engineering Research Council of Canada.

References

Born, J., Wolf, E.: Principles of optics. NewYork: Macmillan 1964

Goubau, W.M., Gamble, T.D., Clarke, J.: Magnetotelluric data analysis: removal of bias. Geophysics **43**, 1157–1166, 1978

Jones, A.G.: On the difference between polarisation and coherence. J. Geophys. **45**, 223–229, 1979

Parrent, G.B., Roman, Jr. P.: On the matrix formulation of the theory of partial polarization in terms of observables. Nuovo Cimento **15**, 370–388, 1960

Samson, J.C.: Descriptions of the polarization states of vector processes: applications to ULF magnetic fields. Geophys. J.R. Astron. Soc. **34**, 403–419, 1973

Samson, J.C., Olson, J.V.: Some comments on the descriptions of the polarization states of waves. Geophys. J.R. Astron. Soc. **61**, 115–130, 1980

Storey, L.R.O., Lefeuvre, F.: The analysis of a random electromagnetic wave field in a magnetoplasma – I. The direct problem. Geophys. J.R. Astron. Soc. **56**, 255–270, 1979

Wolf, E.: Coherence properties of partially polarized electromagnetic radiation. Nuovo Cimento **13**, 1165–1181, 1959

Received February 21, 1980; Revised Version May 5, 1980; Accepted May 10, 1980