THERMAL AND MECHANICAL EVOLUTION OF CONTINENTAL CONVERGENCE ZONES: ZAGROS AND HIMALAYAS

by

George Peter Bird

A.B., Harvard College (1972)

SUBMITTED IN

PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

August, 1976

Signature of Author...George Peter Bird...........................
Department of Earth and Planetary Sciences
August, 1976

Certified by...M. Nafi Toksöz...................................
Thesis Supervisor

Accepted by...Theodore R. Madden..................................
Chairman, Departmental Committee on Graduate Students

THERMAL AND MECHANICAL EVOLUTION OF CONTINENTAL CONVERGENCE ZONES: ZAGROS AND HIMALAYAS

by

George Peter Bird

Submitted to the Department of
Earth and Planetary Sciences on
August 17, 1976 in partial
fulfillment of the requirements
for the degree of Doctor of Philosophy

ABSTRACT

Plate-tectonic theory explains the disposition of rock types in mountain belts resulting from the convergence of two continents, such as the Zagros, Himalayas, Alps, Urals, and Appalachians. It does not explain the plutonism and metamorphism on the updip side of the former subduction zone, or the crustal thickening on both sides. The young Zagros and the mature Himalayan-Tibetan convergence zones are modelled by finite-difference (thermal) and finite-element (mechanical) techniques to determine the role of various mechanisms which thicken and melt the crust.

Simple thermal models show that radioactive heating is unimportant until the post-orogenic period. Possible sources of crustal melting are shear-strain heating or intrusion from below. Variations of temperature,stress, and width in subduction zones are studied by an approximate analytical technique. The model predicts that continental subduction zones will shear by faulting down to depths of 10-20 km, where deformation by nonlinear dislocation creep will take over at temperatures well below the solidus. For any assumed material the shear zone as a whole is unstable, with stress proportional to the inverse square root of plate velocity.

A finite-element method is developed for the solution of two-dimensional plane-strain mechanical problems. It solves the problem of the instantaneous flow field of an incompressible material with a spatially-varying nonlinear rheology (dislocation creep) and a plasticity condition (earthquake faulting). Solutions include the stream function, velocity, strain rate, deviatoric stress, seismicity, and planes of plastic slip (faults).

Geologic evidence from the Zagros range implies that a continental collision has occurred there within the last 2 m.y. Gravity, heat flow, surface wave spectra and dispersion, seismicity, and fault plane solutions show that present deformation involves basement shortening. Finite element models

of this process show that the Tethyan oceanic slab is weak or detached. Crustal shortening is occurring by uniform horizontal strain and decollement of the crust from the lithosphere. This decollement starts when shear stresses reach about 100 bars at the base of the crust. Earthquakes in the Zagros occur at shear stresses of 100-800 bars. Both results suggest deformation mechanisms that have not yet been observed in the laboratory.

The Himalayan range is in a more mature stage of convergence, and crustal detachment has developed into the overthrusting of one layer of crust upon another. Finite element models require an anomalously weak shear zone, and only the properties of this zone are well resolved. It shears at stress levels of about 200 bars either by creep of water-weakened quartz or by faulting at this low stress, as in the Zagros. Frictional heating in this zone is too weak to have melted the Himalayan granites unless plate motions in the past exceeded velocities of 60 cm/year. The preferred hypothesis is that they formed during the detachment of the oceanic slab and before the overthrusting. The present convergence velocity in the Himalayas is 2 cm/year, so the greatest part of the plate convergence is being absorbed in China.

The adjacent plateau of Tibet is also a product of the collision. Rayleigh wave group velocities indicate that the crustal thickness is between 55 and 70 km if low-velocity and low-density mantle underlies Tibet. Attenuation of Rayleigh waves indicates a partially-molten layer at a depth of 60-80 km with a Q below 10. These high temperatures are attributed to thermal erosion of the lithosphere by induced convection before the collision. The Himalayas did not immediately form after the collision because plate motion was absorbed as horizontal shortening in Tibet. The present elevation of Tibet is in lithostatic equilibrium with the forces causing overthrusting in the Himalayas and strike-slip faulting in China.

Continental collision tectonics is dominated by gravity forces because continental crust can be deformed by shear stress of hundreds of bars. The convergence of the Zagros can be explained as a result of driving forces from the Red Sea alone. Convergence of the Himalayas requires only one-third of the potential driving forces acting on the Indian plate.

The faulting stress of about 200 bars needed to reproduce the seismicity of the Zagros and Himalayas is an order of magnitude less than laboratory experiments predict. The discrepancy is most likely explained by melting on earthquake fault planes. This could reduce friction and allow an earthquake to propagate out of the high-stress source region to relieve strain over a large area.

Thesis supervisor: M. Nafi Toksöz, Professor of Geophysics

ACKNOWLEDGEMENTS

My advisor Nafi Toksöz deserves great credit for selecting the most interesting problem left in the theory of plate tectonics, for keeping me in touch with everyone and everything that could help with its solution, and for patiently analyzing my sometimes half-baked theories. Norman Sleep contributed his versatile version of the finite-difference temperature program. Jerome Connor and Gilbert Strang provided guidance in the finite element formulation of the deformation problem. Christopher Goetze helped with data, references, and discussions on the physical properties of rocks. Claude Allegre suggested a mechanism for rock melting and provided thermodynamic and geologic data to test it. Peter Molnar contributed fault-plane solutions, maps, photographs, and helpful discussions about Tibet.

Most of the gravity data used was sent by Elmer J. Hauer of the DOD Gravity Services Division of the Defense Mapping Agency. Borehole temperatures in Iran were supplied by M. A. Nabegh of the National Iranian Oil Company. The seismic studies were advanced by Howard Patton, who ran eigenfunction calculations, and by Richard Buck and Steve Wilk, who helped to process the data. Mansour Niazi sent a map and useful references on the Zagros. Finally, I would like to thank

Raymon Brown, Albert Hsui, Nezihi Canitez, Kevin Burke and Frank McClintock for their encouraging and helpful suggestions throughout the project.

This work was supported by the Advanced Research Projects Agency and administered by the Air Force Office of Scientific Research under contract F44620-75-C-0064 with the Massachusetts Institute of Technology.

TABLE OF CONTENTS

		PAGE
ABSTRACT		2
ACKNOWLEDGEMENTS		4
CHAPTER 1:	PLATE TECTONIC ORIGIN OF CERTAIN MOUNTAIN BELTS	10
	1.1 Common Features of Continental-Collision Ranges	11
	1.2 Development of the Continental-Collision Hypothesis	13
	1.3 Unsolved Problems and the Work Performed	17
CHAPTER 2:	THERMAL ASPECTS OF CONTINENTAL CONVERGENCE	20
	2.1 Initial Geotherm in Undisturbed Continental Plates	21
	2.2 Regional Thermal Evolution Modelled by Finite Difference Methods	25
	2.2.1 Finite Difference Method	25
	2.2.2 Sources of Heat	26
	2.2.3 Model Results	31
	2.3 Temperature and Stress in Shear Zones	37
	2.3.1 Computational Strategy	37
	2.3.2 Deformation Mechanisms and Shear Zone Equilibria	41
	2.3.3 Implications for Melting, Seismicity, and Stability	52
	Tables	56
	Figures	58

		PAGE
CHAPTER 3:	NUMERICAL MODELLING OF PLATE DEFORMATION WITH FINITE ELEMENTS	67
	3.1 The Need for Understanding of Instantaneous Velocity Fields	68
	3.2 Flow Laws Applied to Continental Convergence Models	71
	3.3 Finite Element Formulation of the Viscous Deformation Problem	77
	3.4 Solution Techniques for Nonlinear Materials	87
	3.4.1 The Variable Viscosity Method	87
	3.4.2 The Initial Stress Method	88
	3.5 Error Estimates and Test Problems	94
	Figures	106
CHAPTER 4:	THE ZAGROS MOUNTAINS: DETACHMENT OF CRUST AND SLAB	119
	4.1 Geologic Outline and Tectonic Theories of the Zagros	120
	4.2 Gravity Data and Crustal Roots	126
	4.3 Seismicity of the Zagros Region	134
	4.4 Heat Flow and Thermal Models	141
	4.5 Force Balance Responsible for Crustal Detachment	147
	4.6 Finite Element Models of Continuing Deformation	153
	4.7 Conclusions	169
	Tables	173
	Figures	179

PAGE

CHAPTER 5: THE HIMALAYAS: ADVANCED CRUSTAL OVERTHRUSTING 212

5.1 Tectonic Outline of the Range 214

5.2 Tectonic History 219

5.3 Gravity Data and Crustal Roots 227

5.4 Thermal Models 234

5.5 Estimates of Uplift and Subsidence Rates 240

5.6 Finite Element Models of Continuing Deformation 246

5.7 The Frictional Theory of Himalayan Magma Genesis 259

5.8 Conclusions 267

Tables 270

Figures 272

CHAPTER 6: THE TIBETAN PLATEAU: TECTONICS BEHIND A CONTINENTAL SUTURE 306

6.1 Geomorphology and Geology of Tibet 308

6.2 Geophysical Evidence 312

6.3 Rayleigh Wave Group Velocities 316

6.4 Rayleigh Wave Attenuation 322

6.5 Tectonic Shortening of Tibet 331

Tables 335

Figures 340

PAGE

CHAPTER 7: SPECULATIONS ABOUT MOUNTAINS, DRIVING FORCES, AND EARTHQUAKES 360

7.1 The Significance of Mountain Topography 361

7.2 The Driving Forces in Continental Convergence 368

7.3 Speculations on the Mechanism of Earthquakes 376

Figures 380

REFERENCES 385

APPENDICES 403

A. Table of Symbols Employed 403

B. Shear-Zone Modelling Technique 408

C. A Plasticity Criterion for Faulting 417

BIOGRAPHICAL NOTE 423

CHAPTER 1: PLATE TECTONIC ORIGIN OF CERTAIN MOUNTAIN BELTS

Among the mountain systems of the world, there are at least five which exhibit such strong similarities that a common mode of formation is implied. These are the Zagros, the Himalayas, the Alps, the Urals, and the southern Appalachians in order of increasing age. The theory that these ranges result from horizontal convergence of continental masses, first articulated in the nineteenth century, has received such strong support from the discoveries of sea-floor spreading and subduction as to be no longer controversial. This thesis will begin with a summary of the evidence for a plate-tectonic origin of these mountains and proceed to a quantitative investigation of the mechanisms which produce heating and deformation. The study is restricted to the two youngest and most active ranges, where geologic information can be supplemented by geophysical observations. However, unless the thickness of the lithosphere or the style of its motions has changed during the Phanerozoic, the results obtained should be equally applicable to those older chains whose internal complexity is now exposed.

1.1 Common Features of Continental-Collision Ranges

The first common feature of these ranges is their great and continuous length, in contrast to a small and constant width. The Urals stretch 3000 km from Novaya Zemlya to the Ural River, the Himalayas 2500 km from Nanga Parbat to Namche Barwa, and the Appalachians at least 1500 km from Atlanta to Vermont, with possible continuations either buried in the Gulf Coast or confused with the earlier Caledonian-Taconic orogeny. Even the Alps extend 800 km in a curve from Monaco to Vienna. Yet in each case the zone of deformation on the updip side of the plate suture is a relatively constant 200-300 km. As the known geology suggests simultaneous rather than migrating deformation in each range, this topographic evidence enables us to rule out such hypothetical deep mantle sources as upwelling juvenile material (Artyushkov, 1973) or thermal plumes (Parmentier and Turcotte, 1975). The constant width of the mountains implies a comparable controlling vertical dimension, perhaps the thickness of the lithosphere.

The fact that each range is presently surrounded by continental lands, yet extends on at least one end to the margin of deep ocean, suggests that the lineation controlling its formation was the margin of a former continent. The discovery of an ophiolite suite of ultramafic rocks and deep sea sediments in four of the ranges (excluding the Appalachians) confirms this theory (Dewey and Bird, 1970). Another common characteristic indicating a former continental margin is the accumulation before deformation of many kilometers of sediments on a slowly subsiding basement. Without continental drift, the juxtaposition of these well-sorted "miogeosynclinal" sediments with coarse volcanic greywackes of the "eugeosyncline" is difficult to explain, except by the convenient appearance and disappearance of barrier ridges (Saxena, 1971).

This thesis concentrates on the deformational phase, in which there is the common feature of monovergent thrusting, with the sense of eugeosyncline overriding miogeosyncline. Except in the Zagros, which are very young, this process leads to the overthrusting of great slabs of metamorphic basement rock over the sediments of the miogeosyncline. This is probably the direct cause of the anomalous crustal thickness or "roots" below each range. Simultaneous crustal thickening may (Tibet) or may not (Alps) occur on the other side of the suture. Melting of crust and the emplacement of granites in the overthrusting basement are common. As the next sections will show, plate tectonics has given a clear explanation of the location of these events, but no quantitative theory of the mechanism.

1.2 Development of the Continental-Collision Hypothesis

The idea that continents "collide" and produce mountains is not an outgrowth of the 1960's concept of plate tectonics, but extensively predates it. While the geologic evidence for horizontal compression is clear, this was at first ascribed to contraction of the Earth, according to the theory of Lord Kelvin. When this mechanism fell into disfavor at the beginning of the twentieth century because of the discovery of radioactivity, another mechanism was needed.

One of the early proponents of continental drift, Alfred Wegener (1915) suggested that the equatorial Alpine-Himalayan system was a result of the _Polflucht_ of continents and their convergence at the equator. His estimate of the total Indian-Eurasian convergence (8,000 km) was essentially the same as that of Dietz and Holden (1970). Developing this idea, Emile Argand (1924) referred to it as the mobilistic concept, and differentiated between the "creative block" (which subducts and deforms) and the "antagonist continent" which overthrusts. He correctly concluded that the creative block in the Alps was Europe, and that in the Himalayas it was India. He also attributed the widespread _plissement du fond_, or Cenozoic basement folding of Asia, to the Himalayan collision, thus anticipating Molnar and Tapponnier (1975).

Little advancement took place for several decades, because it was impossible to say from surface geology when, or how far, or at what depth the continents moved. But the influence of the concept is seen, as when Gansser (1964) in his summary of Himalayan geology emphasizes the fundamental importance of the Indus melange zone (plate suture) and Main Central Thrust, and takes pains to show that the rocks of the Higher Himalayas are of Indian and not Asian origin.

McKenzie (1969) first discussed continental collisions in terms of the subduction of thick lithospheric plates containing continental crust. He showed that steady-state subduction of continents is impossible because of buoyancy constraints, and so explained the general cause of orogenies. Dewey and Bird (1970) showed that the longstanding geologic problem of mio- and eugeosynclines was resolved by continental drift, with the miogeosyncline being the former continental shelf of one continent, and the eugeosyncline the volcanic and tectonic equivalent of the Andes: a volcanic arc above a former oceanic subduction zone. This concept rested upon their work on the ophiolite suites of the suture zones, which they were able to correlate with mid-ocean ridge rocks obtained by deep-sea drilling.

Another geometric possibility was suggested by Oxburgh (1972) under the name of flake tectonics: that the overthrust basement forming the high peaks is transported across the suture from its original position on the subducting plate. According to this theory the African and Asian plates are subducting in the Alps and Himalayas respectively, the reverse of Argand's concept. This theory is clearly not applicable to the Himalayas, Urals, or Appalachians because there the overthrust sheets and andesite volcanic lines are found on opposite sides of the suture, contrary to the prediction of flake tectonics. A variant of flake tectonics, in which the overthrust sheet does not cross the suture, but is still found on the eugeosynclinal side, was proposed by Sorokhtin (1973). The same objection applies.

Simultaneously, Powell and Conaghan (1973) conducted a reinterpretation in terms of plate tectonics of the geologic material of Gansser and others. Like Argand, they concluded that the overthrust crystalline slab in the Himalayas was a fragment of the subducting Indian plate driven over itself. They noted the additional complication that this overthrusting did not occur during the collision itself, but in a later tectonic phase. Mattauer (1975) reached the same conclusion from a study of foliations and lineations in the Himalayas, and argued that the overthrusting of a crustal slab on a horizontal decollement had analogues in the Hercynian Massif Central and the Scandinavian Caledonides.

Recently, interest has also focused on the Zagros Mountains as a possible younger analogue in the first stages of collision. Haynes and McQuillan (1974) showed that the Zagros Crush Zone was a classic continental suture with ophiolites and that there also the deformation was concentrated in the subducting plate. They felt that the continental blocks are not yet in contact and that continental margin sediments only are being deformed. However, Bird, Toksöz, and Sleep (1975) argued from gravity and seismic data that a crustal overthrust of Himalayan proportions is presently developing.

1.3 Unsolved Problems and the Work Performed

Clearly the basic assumption of plate tectonics, that plates overthrust each other without deforming, is too restrictive to apply to continental convergence zones. From this assumption one would predict that the margin of one continent would subduct under the other until buoyancy prevented further motion. Thus all crustal thickening, uplift, metamorphism and intrusion would occur on the down-dip or eugeosynclinal side of the plate suture. The following geologic observations cannot be accommodated by such a theory: i) plate motions do not stop immediately after collisions but may continue for over 40 million years (Powell and Conaghan, 1973; Molnar and Tapponnier, 1975); ii) in some cases (Tibet) a high volcanic plateau extending over 1000 km back from the plate suture is formed; iii) overthrusting and crustal thickening occur within the subducting plate (Gansser, 1964; Mattauer, 1975); and iv) this crustal overthrust is metamorphosed and intruded by granites during the orogeny (Naylor, 1971; LeFort, 1975).

The approach taken in this thesis is that deformation is controlled by temperature, which fixes the boundary between elastic-brittle rock mechanics and thermally activated processes of creep. In Chapter 2 we examine the role of various heating mechanisms during an orogeny by a finite-difference technique and show that shear-strain heating is dominant. This implies that the thermal and mechanical problems are linked. A hybrid finite difference/analytic method is developed for studying this interaction in "one-dimensional" subduction zones, and predicting the variation of stress and temperature with depth, rock type, and plate velocity. Thus we can calculate temperatures from assumed rock parameters.

In Chapter 3 we develop a two-dimensional finite element technique for solving the problem of instantaneous flow of incompressible rocks subject to faulting and nonlinear creep. With this technique we can test whether the rock properties assumed and the inferred temperatures would result in deformations that are geologically plausible and consistent with seismologic data.

In Chapter 4 we apply these techniques to the Zagros Mountains, which are at an early stage in the orogenic process. Crustal detachment is shown to be taking place, and the strength of crustal rocks is found to be well below the upper limits set by rock mechanics experiments.

Chapter 5 examines the formation of the Himalayas by crustal overthrusting. Finite element modelling of the continuing deformation confirms the results of Chapter 4 regarding the low strength of rocks in the shear zone. This implies that different mechanisms operated in the past to form the Himalayan granites. Two possible mechanisms investigated are high plate velocities and disruption of the subcrustal lithosphere.

Chapter 6 continues the study of the mature Himalayan-Tibetan convergence zone with an investigation of the origin of the Tibetan uplift. Rayleigh wave propagation and attenuation show that Tibet has a thick crust and low-velocity mantle, with partial melting near the Moho. These observations combined with gravity and geologic constraints are used to show that Tibet was weakened by convective heating before the collision and subsequently suffered major horizontal compression.

Chapter 7 discusses the implications of the low rock strengths implied by the preceding models. Forces arising from topography in mountain belts are shown to be comparable to those causing deformation. This implies that the known driving mechanisms of plate tectonics are strong enough to explain the formation of mountains. The fact that natural earthquakes require stresses an order of magnitude below laboratory stick-slip behavior is shown to support the McKenzie and Brune (1972) model of fault melting in major earthquakes.

CHAPTER 2: THERMAL ASPECTS OF CONTINENTAL CONVERGENCE

In order to understand the effects of subduction on the temperatures in a continental plate, I begin with purely thermal models in which the motions of the rock are prescribed. An initial geotherm for the continental margins is developed and used as an input to various two-dimensional finite-difference thermal models. These models illustrate the effect of different subduction geometries and rates. Because shear-strain heating on faults appears to be the only possible source of melting, a technique is introduced for modelling the mechanics of shear zones in detail. By finding the equilibrium temperature where heat production and consumption are balanced, the extent of seismicity, shear, and melting can be determined for a wide variety of materials and conditions.

2.1 Initial Geotherm in Undisturbed Continental Plates

Most published geotherms for continents are unsatisfactory because they do not take account of convective heat transport in the asthenosphere. Broad scale "passive" convection at least is required by the motion of the plates, and more rapid "active" convection is predicted if the Rayleigh number is calculated using Kohlstedt and Goetze's (1974) olivine flow law. Ignoring convection, many authors have overestimated mantle conductivity in order to obtain reasonable temperatures at the phase changes. For example, Clark and Ringwood (1964) used the conductivity equation of MacDonald (1959) with an opacity of 5 cm^{-1}. However, Schatz (1971) has shown that olivine opacities under mantle conditions are considerably higher and that radiative heat transfer constitutes less than half of the total to temperatures of 1500°K.

Therefore, the geotherm used in this work has been based on the empirical conductivity formula of Schatz and Simmons (1972) and has not been extended beyond the lithosphere. Since the average composition of continental crust is not well known, a temperature-dependent conductivity is not used in this region. Instead an average value of 2.5×10^5 ergs/cm-sec-°C is taken from values on Rockport granite at 300°C given by Clark (1966).

The two remaining parameters of importance are the mantle heat flux and the crustal radioactivity contribution. The sum of these is the total heat flux, which is taken to be 1.3×10^{-6} cal/cm^2-sec (54 mW/m^2), a typical value for non-orogenic continental areas with sedimentary cover (Clark and Ringwood, 1964). It is also in the range (1.2-1.6 HFU) reported by Diment et al. (1970) for the Atlantic coastal plain of the U.S., a region which serves as an informal model of a pre-collision continental margin.

The geotherm is fixed at its base by the requirement of partial melting in the asthenosphere. This is the mechanism most likely to explain the sudden drop of seismic shear velocities (Press, 1970) and increase in attenuation at the base of the lithosphere. Thus for thermal studies the base of the lithosphere is defined as the top of the low-velocity zone. Various studies of surface-wave dispersion across non-tectonic continental shields have shown that this depth is about 120 km (Brune and Dorman, 1963; Toksőz et al., 1967; Knopoff and Fouda, 1975). The temperature which produces partial melting at this pressure (37 kb) is harder to obtain because it is dependent on the concentration of water and CO_2 in the mantle. Kay et al. (1970) estimate it is 1350°C, based on a peridotite mantle with a bulk H_2O content of 0.2% and partial melting of 1.4%. This figure is 300° below the dry solidus at that pressure, yet high enough to be consistent with observed eruption temperatures at island arcs of 1200°C (Osborn, 1969). Note that a 100°C error in this figure would affect crustal temperatures by a maximum of 20°C because of the fixed heat flux.

Using these constraints, a geotherm was calculated using an iterative one-dimensional finite-difference technique programmed by Sleep (1973). Different values of mantle heat flux were tried, with the residual heat production being placed in the crust, concentrated in the upper half relative to the lower half by a ratio of 3:1. The successful geotherm giving a temperature of 1350°C at 120 km depth is given in Table 2.1. It has a mantle heat flux of 0.725 HFU (30.3 mW/m^2), which is within the range 0.8 ± 0.1 HFU obtained by Roy _et_ _al_. (1968) in the eastern United States by the technique of plotting heat flow against heat production. It requires a total crustal heat production of 0.56 HFU. If 75% of this heat production were concentrated in the upper half of the crust, it would imply a value of 2.65×10^{-13} cal/cm^3-sec, not far from the value (2.2×10^{-13}) used by Clark and Ringwood (1964). The crust in all models is assumed to be 33 km thick, a figure obtained from a worldwide compilation of seismic refraction results in continental regions of zero elevation and/or zero Bouguer gravity anomaly (Woollard, 1959, 1968). This is also the thickness in the prototype Atlantic coast area of the U.S. (Pakiser and Steinhardt, 1964).

It should be noted that there are departures in this thesis from the use of this initial geotherm. In section 2.2, some models calculated at an early stage use the MacDonald (1959) conductivity in the mantle. However, as mantle flux and crustal heat production were very similar the calculated temperatures above the Moho are not affected by the lower mantle temperatures. In Chapter 4, a slightly lower crustal radioactivity is needed to match the observed surface heat flow. Fortunately, as shown in the following section, the thermal history of convergence is dominated by shear-strain heating and thus relatively insensitive to the initial geotherm, which must always remain conjectural.

2.2 Regional Thermal Evolution Modelled by Finite Difference Methods

Subducting continental crust may be heated simultaneously by many mechanisms, including the conductive, radioactive, adiabatic, frictional, chemical, and intrusive contributions. All of these are a function of the geometry and velocity of subduction, which may be complex. Analytic methods are incapable of dealing with such a variety of spatially-varying or nonlinear terms, and so it is necessary to employ an approximate finite-difference method.

2.2.1 Finite Difference Method

In these models we solve the heat equation

$$\rho C_p \frac{\partial T}{\partial t} = \nabla \cdot (K \vec{\nabla} T) + H \tag{2.1}$$

where ρ = density, gm/cm^3

C_p = specific heat, ergs/gm-°C

T = temperature, °C

K = conductivity, ergs/cm-sec-°C

H = heat production, ergs/cm^3-sec

in two dimensions by finite differences. The algorithm is an implicit alternating-direction technique of Peaceman and Rachford (1955). References to the programming, accuracy, and stability of this method include Minear and Toksöz (1970a,b 1971), Toksöz et al. (1971, 1973), Sleep (1973), Bird et al.

(1975), and Toksöz and Bird (1976). If one observes the stability conditions for time step size

$$\kappa \Delta t/\Delta z^2 < .25 \tag{2.2}$$

(where κ = thermal diffusivity, cm^2/sec and

Δt = time step in one iteration, sec

Δz = shortest grid interval, cm)

and conductivity contrasts:

$$\frac{2\kappa}{\Delta x} > \frac{d\kappa}{dx} \; ; \quad \frac{2\kappa}{\Delta z} > \frac{d\kappa}{dz} \tag{2.3}$$

then the technique is accurate to 1% in the treatment of an infinite heat source of two grid points dimension (Toksöz *et al.*, 1971). When translation of temperatures is added the errors increase, but remain less than 10% of the temperature anomaly for a 5 km grid.

This was the vertical grid spacing in the models whose parameters are to be found in Table 2.2. The horizontal spacing was set to accommodate the dip of the slabs. A 41x41 point grid was used in all cases, with boundary conditions of zero temperature at the surface, zero heat flux through the sides, and fixed mantle flux into the base. The heat production term H included the adiabatic, radioactive, and frictional sources described in the next section.

2.2.2 Sources of Heat

Adiabatic heating affects all rocks which move downward and experience increasing pressure, according to the equation (Toksöz *et al.*, 1971)

$$\frac{dQ}{dt} = \rho \alpha g \, TV \sin(\theta) \tag{2.4}$$

where Q = internal heat energy, erg/cm^3

ρ = rock density, gm/cm^3

g = gravitational acceleration, cm/sec^2

α = thermal expansion coefficient, $°C^{-1}$

T = absolute temperature, °K

V = plate velocity, cm/sec

θ = dip of subduction zone.

Thus the potential temperature increase due to this effect is

$$\Delta T_A \cong \alpha g \, \overline{T} \, \Delta Z / Cp \tag{2.5}$$

where $\overline{T}$ = average temperature during subduction, °K

ΔZ = distance subducted vertically, cm

Cp = specific heat, erg/gm-°C.

Taking upper limits of $\overline{T}$ = 500°C, ΔZ = 40 km, and values $\alpha = 2.4 \times 10^{-5} °C^{-1}$ (Skinner, 1966) for granite, and Cp = 7.95×10^6 erg/gm-°C (Ingersoll *et al.*, 1954) we obtain an upper limit of 9.2° for ΔT_A. Clearly this will never be an important effect.

The importance of radioactive heating has been stressed by Dewey and Burke (1973), who argue that Tibetan and Himalayan granites are caused by radioactive self-heating of overthickened crust. I will only point out here that the maximum heating rate due to this effect is

$$\frac{d\Delta T_R}{dt} = H/\rho Cp \tag{2.6}$$

where H = radioactive heat production, erg/cm^3-sec.

Taking the upper crustal H from section 2.1 we obtain an increase of only 16.4°C/million years. This value of H is not likely to be incorrect by more than a factor of two on a regional basis. Localized regions of larger H are possible, but will not retain their radioactive heat unless their dimension is several tens of kilometers, and larger than the scale distance (defined below). The importance of radioactive heat is really determined by the amount of time available for accumulation and the extent to which it is prevented from conducting away. It will not be important during active subduction at typical plate velocities.

A negative heat source, or sink, is provided by various metamorphic reactions which involve the expulsion of mineralogic water. This effect would be principally important in the top sedimentary layer, because the majority of crustal rocks are products of previous metamorphic events. It is not known to what extent these sediments are subducted, as opposed to being scraped and eroded off at the surface. In addition, the kinetics of the reactions are determined by whether the water is free to migrate out of the system. Because of these uncertainties no metamorphic heat terms have been included in these calculations. This includes the basalt-eclogite or plagioclase-garnet-peridotite phase change, which releases unimportant amounts of heat (Sleep, 1973).

Another mechanism hard to evaluate is intrusion. The asthenosphere is very mobile and deformable and if the lithosphere were sufficiently disrupted during convergence, the asthenosphere material might come into contact with the base of the crust. In that case the temperature of the contact would instantaneously be the average temperature (500°C + 1350°C)/2 = 925°C, which is sufficient for melting of granite. This is an entirely *ad hoc* mechanism, difficult to prove and impossible to quantify, and thus not included in thermal models. It is, however, discussed further in Chapters 4 and 5.

Heating by conduction can be very accurately treated. Being diffusive, it affects a thickness of the lithosphere proportional to the square root of time:

$$S = (\kappa t)^{1/2} \qquad (2.7)$$

where S = scale distance, cm

κ = thermal diffusivity, cm^2/sec

For a value of κ of .0113 cm^2/sec derived from the above figures, and a time of 4 m.y. (obtained by dividing the thickness of a typical range by a typical plate velocity of 5 cm/year), S = 11.9 km. Thus heating of the crust through the lithosphere from below will be negligible, but heat conducted in through the upper contact is important. When overthrusting occurs within one plate the conductive heating can be calculated using the geotherm of Table 2.1 as an initial condition in the upper plate. However, the initial temperature in an Andean-type overthrusting plate margin is entirely unknown. Our only constraint is that magmas are generated at at least 1200°C at a depth of about 100 km in an oceanic subduction zone such as exists before the collision. Since this constraint is satisfied by the geotherm just mentioned, it is used as an approximation in the calculations of this section.

Shear-strain heating from permanent strains is

$$\frac{dQ}{dt} = \sum \sigma_i \, \dot{e}_i \qquad (2.8)$$

where σ_i are stress deviators in dynes/cm^2

$\dot{e}_i$ are non-compressional strain rates in sec^{-1}.

The plate tectonics approximation is that all strain is two-dimensional and occurs in a narrow interplate shear zone. In this case we can integrate across the shear zone to obtain:

$$A = \int \tau \dot{e} dz = \tau V \qquad (2.9)$$

where A = heat production in ergs/cm^2-sec

τ = shear stress on the plane of shearing, dynes/cm^2

V = relative plate velocity, cm/sec

Thus this term is independent of shear zone width, which is unknown. It does require an estimate of stress, which may or may not depend on shear zone width. Detailed estimates of τ will be obtained in section 2.3, but in this section we merely assume that τ is approximately equal to the shear stress required to support the Earth's largest topographic features (1 kb). Given a plate velocity of 5 cm/year, A is thus on the order of 3.8 HFU (158 mW/m^2). It is easy to see that this will be the dominant heat source during continental collisions.

2.2.3 Model Results

As described by previous authors, geologic movements are represented by transferring temperatures an integral number of grid points in each time step. In models I and II this motion is the initial collision of continents, with transient subduction of the continental margin. Temperatures are translated diagonally across the grid from the slab hinge point. In model III, which represents the second phase of overthrusting within the continent, there is a point at which the slab flattens out and translates horizontally. This motion obviously does not conserve the lithospheric material ahead of the slab. However, a more elaborate translation routine that subducted this lithosphere out of the way would have made no difference in the calculated temperatures, as all motion would have been parallel to isotherms.

These three models are presented to show the effects of velocity or quiescence and of the amount of frictional heating on the thermal regime. They embody three different subduction geometries, illustrating that the effect of this geometry is small. Details of the parameters of each model are given in Table 2.2. Heat flow profiles presented with each model are obtained from the temperatures one grid point away from the boundary. The Bouguer gravity anomalies calculated for models I and II are based on an infinite two-dimensional structure with crust/mantle density contrast of -0.43 g/cc and thermal expansions of 3×10^{-5} (upper crust), 2.5×10^{-5} (lower crust), and 3.2×10^{-5} (mantle) from Skinner (1966).

Figure 2.1 represents the initial subduction of a continental margin at the low velocity of 1 cm/year, taken to maximize the effect of radioactivity. Enough of the attached oceanic lithosphere is included to cover the edge of the continent, but we do not claim to have produced steady-state oceanic-subduction temperatures above the fault. Most of the continental crust and lithosphere are unaffected by the event. Frictional and conductive heating create minor warming, to 400°C, in the upper crust. But as the depressed heat flow shows, the influx of cold material dominates over the frictional heating everywhere because that heating is distributed over a band of width 2S = 33 km (eq. 2.7). The gravity field is also controlled by compositional effects for the first 180 km behind the trench.

In model II (Figure 2.2) a number of changes enhance the effects of friction. Velocity is increased to 6 cm/year, giving less time for dissipation. Dip is decreased to 15° to increase the area of the fault. And shear stress is raised to a high of 1.5 kb (Table 2.2) at shallow depths where the rocks are coldest and presumably strongest. The net result is that the portion of the slab affected by the event is thinner (15 km) but hotter. Frictional heating actually raises temperatures to the water-saturated granite solidus of Tuttle and Bowen (1958) below depths of 30 km. This could create a small volume of granite (limited by available water) which would be located 100-200 km behind the trench if it were able to rise diapirically. However, it is doubtful that such high values of shear stress are possible.

To balance forces on the overthrusting wedge, an average horizontal stress of 4.64 kb is required from 0-50 km depth at the point x = 230 km in Figure 2.2. Comparable values are required everywhere in the wedge, even in its shallow portions. Even if the plates can support such stress, and if unknown mechanisms exist to create it, it is questionable whether the rocks of the subduction zone would not creep rapidly at lower values. Thus another mechanism must be sought to create the granites found in the Himalayas (Gansser, 1964), the Appalachians (King, 1950), and the Urals (Carte Tectonique Internationale De L'Europe, 1962).

Because all of these ranges are over 40 m.y. old, radioactivity may be an important heat source. To test this, we allowed the calculation to proceed for 30 m.y. more without any motion (Figure 2.3A). Although a great quantity of heat is produced, most of it goes to restore the region to its original geotherm and heat flow. The highest crustal temperature reached is 700°C at the lower tip, and this will be insufficient to produce melting of the anhydrous migmatitic or gabbroic rocks most likely found in the lower crust. Furthermore, uplift and erosion, which are not included here, will tend to destroy the crustal bulge and lower its temperature. For example, Alpine uplift rates of over 1 mm/year (Pavoni, 1975) have reduced the crustal thickness in that range to a maximum of 50 km (Rybach, 1975). The temperatures calculated here must be regarded as upper limits for a crust of reasonable radioactivity.

To see how much time is required for radioactive melting we continue the calculation out to 100 m.y. in Figure 2.3B. Here at last the probable solidus is exceeded in a small region and pluton formation could occur. However, the time required is excessive, for granites are known to have formed within 30 m.y. after the collision in the Appalachians, Alps, and Himalayas (Naylor, 1971; Allegre _et al._, 1974; Hamet and Allegre, 1976).

Finally, in model III (Figure 2.4) we leave these extreme cases and consider the most probable evolution of a typical collision belt. In this model we recognize that the major crustal underthrusting is probably within the subducting plate rather than behind the plate suture (Powell and Conaghan, 1973; Mattauer, 1975). We also employ the more accurate Schatz and Simmons (1972) mantle conductivity in the lithosphere, and a higher conductivity in the asthenosphere to simulate convection. The velocity of 5 cm/year is close to the present convergence rates of the Zagros and Himalayas (Bird et al., 1975; Molnar and Tapponier, 1975). Most importantly, frictional stress is reduced to values generally less than 1 kb (Table 2.2) that are more consistent with known rock mechanics. In the shallow part of each fault zone a frictional mechanism is assumed, with the stress in the upper 10 km being limited to 60% of the confining pressure. At greater depths, the stress is adjusted so as not to exceed the temperature-dependent creep strength of natural polycrystalline quartzites (Tullis, 1971), given in detail in section 2.3. The quiescent part of the calculation (Figure 2.4C) is not extended past 40 m.y. because a detailed treatment of postorogenic uplift is beyond the scope of this thesis.

In this model no melting is predicted. Maximum temperatures on the faults are approximately 500°C during subduction, and maximum lower crustal temperatures after the 35 m.y. quiet period are again about 700°C. The different geometries and mantle thermal parameters are seen to have very little influence on the result. The production of granite melts in a short orogenic collision event apparently requires an extraordinary mechanism.

2.3 Temperature and Stress in Shear Zones

The detailed structure of the shear zone between a subducting and an overthrusting plate is of interest for several reasons. First, the frictional heat from this zone appears to be a possible source of crustal melting. Second, stresses acting between the plates must be transmitted across the zone, which is likely to be anomalously weak. Finally, a knowledge of the width and maximum metamorphic grade of these zones would assist the study of old pre-Cambrian collision belts (Dewey and Burke, 1973).

2.3.1 Computational Strategy

The two-dimensional finite-difference program is not a suitable tool for such studies. To find the structure of a zone several points of the grid must lie across it. A minimum of two is required so that half of the frictional heat may be placed in the subducting and half in the stationary plate. Thus a very fine grid is desirable. However, because material moves through a subduction zone the whole zone must be modelled at one time, implying that the number of grid points goes as the inverse square of the grid interval. For a constant amount of time modelled, equation (2.2) dictates that the number of iterations rises by the same factor. Thus the expense of the computation rises as the inverse fourth power of the grid interval. Most models in this thesis were performed with a 5 km interval; 2 km would be the minimum financially feasible interval. A shear zone composed of two such points would have the thermal inertia of a 4 km layer and would tend to underestimate temperature maxima produced by frictional heating.

Another problem is that the programs available evaluate the contribution of heat sources by explicit means, based on the last temperature solution. If the major source of heat is from frictional work according to a temperature-dependent flow law, such a technique will yield violent artificial oscillations of stress with each iteration. These oscillations may also excite undesirable temperature oscillations in the spatial dimensions. On the other hand, analytical solutions have difficulty incorporating the subduction geometry, initial geotherm, and variable radioactive heating which are easily handled by finite-difference methods.

The method used here is to separate the effects of frictional heating from all others and to treat them approximately with analytic formulas. The first step is to use the finite-difference techniques of the last section to solve for the artificial case of subduction without frictional heating. This problem requires an initial condition and a time interval. The time interval is that period over which a subduction zone changes its equilibrium from oceanic to continental, and this process is presumed to begin with the subduction of the voluminous silicic continental rise sediments before the collision. Emery et al. (1970) give profiles along the Atlantic coast of the United States from which a typical width of continental rise deposits of 400 km was selected. So the finite-difference calculation without friction is carried out for 400 km of subduction, at the end of which a radioactive continental crust is introduced as in Figure 2.5.

The ideal initial condition would be the equilibrium temperature field of an Andean-type continental margin. As this is not known, the standard quiet continental margin geotherm of Table 2.1 is used as an estimate. Because of the large amount of subduction in the transition period, any error introduced in this way is strongly damped. The temperatures obtained are very low (200°C at 45 km depth) and relatively insensitive to plate velocity (Bird et al., 1975; Figure 5). They are probably accurate to ±20° numerically and ±50°C considering the uncertain initial condition.

The complete solution is obtained by adding to this solution the thermal effect of a thin layer of frictional heat sources. This addition can be performed in regions where the heat equation (eq. 2.1) is linear in temperature. Temperature effects on density and specific heat have been ignored in all calculations, so the time-variation term is linear. We have assumed a temperature-invariant crustal conductivity (section 2.1), so that the addition may be performed as long as frictional heating does not affect the mantle. This is not a serious constraint, because the models of the previous section showed the mantle to be unaffected by subduction of the continental margin to 50 km depth at velocities as low as 1 cm/year.

The greatest non-linearity comes from the adiabatic and frictional heating terms. Because a non-physical temperature is used in the finite-difference scheme to evaluate adiabatic heating, an error of less than 4°C is introduced (eq. 2.5). This is ignored. The zone of frictional heating itself must be modelled in detail because in the case of rock creep the contribution to the heat term H will be an exponential function of temperature in this region. Turcotte and Oxburgh (1969) considered this problem for linear viscous thermally-activated creep in a one-dimensional shear zone that varied only across its width. Their solution required the parameterization of distance and the evaluation of double integrals of singular functions. In the case of nonlinear material these integrals are intractable, and the theory does not allow for non-symmetric heat flow away from the zone or variations along its length (as a function of depth).

Since variation with depth is great in the upper 50 km, this method concentrates on these variations instead of variations through the thickness. The shear zone is considered to be thin and well-mixed, so that constant temperature and strain-rate through the thickness are reasonable approximations. The thickness W (see Figure 2.5) is left as a free parameter, to be determined within geologic limits by the physics of the subduction zone. At each point in the subduction zone, it is necessary to conserve energy. Also, the width must vary in such a way that stress and dissipation are minimized. Finally, the stress must obey the faulting or the creep relations of the particular rock unless it should melt. The numerical means by which these constraints are expressed and satisfied have been set aside as Appendix B to maintain the continuity of this chapter.

2.3.2 Deformation Mechanisms and Shear Zone Equilibria

This section attempts to illustrate the range of possible temperatures and stresses in subduction shear zones by means of the approximate calculations outlined in Appendix B. Final answers are not possible because of our ignorance of trench tectonics, and the fraction of sediment which is actually subducted. Also, creep data is only available for a limited number of minerals (olivine, quartz, calcite, halite, ice) in sufficient quantity to justify extrapolation to geologic conditions. Fortunately, no harm is done in neglecting minerals that are harder than these, for the levels of stress and thus the temperature will be determined by the weakest component of the system. Also, because of the nonlinear character of rock creep, the estimated fraction of each weak mineral enters the equations with a fractional exponent, so that errors of quantity have little effect.

All calculations were performed for slabs dipping at 30°. This number is rather arbitrarily chosen. However, the dip angle θ enters the equations only in (B.5) where it has a small effect, and in (B.10). If stress and temperature are considered functions of s, the distance down-dip (rather than depth Z), then similar results will be obtained at all dips. Smaller dips will give slightly lower temperatures and higher stresses, and vice versa.

There will always be a zone of seismic or stable sliding according to the frictional law (B.5) in the shallowest part of the shear zone. The rocks enter the zone at near 0°C, and only rock salt or ice would creep appreciably at that temperature. In equating the shear stress of (B.5) to (B.1), we assume that the energy radiated seismically is a small fraction of that which goes into heating. Since seismic stress drops in subduction zones are generally 1-200 bars (Wyss, 1970), this is a good approximation in models having frictional stresses of over one kilobar. The value of the coefficient of friction used in all models is $\mu = 0.60$. This value is representative of a wide variety of rock types (Stesky _et_ _al_., 1974) at low temperatures, up to stresses approaching the ultimate strength of the rock (G/10). Byerly

(1966) found that this value of μ was approached after a few millimeters of sliding in granite samples having a wide variety in initial surface roughness.

The Mohr envelope is assumed to converge at the origin with no initial cohesion, because the problem involves sliding between two plates and not within an unbroken rock.

The greatest uncertainty concerning the faulting region is over the effect of pore pressure exerted by water. High pore pressures act to reduce the effective normal stress on the sliding surface and thus the shear stress also. In the unlikely event that there were an oversupply of interstitial water in a closed system, the effective strength of the rock could be reduced almost to zero (Hubbert and Rubey, 1959). Here we make the relatively neutral assumption that the system is open, and that water is available. Thus the value of ρ appearing in equation (B.5) is the corrected density (2.67-1.03 = 1.64 g/cc) of crustal rocks.

If no creeping minerals are present, temperatures will rapidly rise above 600°C. As shown in Figures 2.6 and 2.7, this temperature is reached at 14 km depth with a velocity of 5 cm/year, and at 28 km if the velocity is 1 cm/year. Soon after this the faults reach the water-saturated granite solidus of Tuttle and Bowen (1958). This solidus should be appropriate for granitic upper crustal rocks, as well as for sediments derived from them and having a similar bulk chemical composition. However, the formation of the first melts may consume all the available pore water, and after that the melting will be controlled by the water released in dehydration of micas. For this reason, we also compute a model using the higher granite solidus of Brown and Fyfe (1970) which assumes the presence of muscovite but no free water.

With either solidus, melting first appears at a depth of 30 (15)* km at a plate velocity of 1(5) cm/year. The highest shear stress in the subduction zone is reached just below the melting temperature: 4.0 (2.4) kilobars. After this point the behavior of the melted zone is unknown. It may form itself into plutonic accumulations and rise from the fault surface. Because the flow law of the melt is unknown, the zone of melting is restricted to 1 cm in these models. In this way we obtain a minimum stress estimate because negligible work is required to provide the latent heat of fusion.

However, intracrystalline dislocation creep is likely to be activated at temperatures well below the solidus. On observational grounds, this deformation mechanism obeys the law

$$\dot{e}_{zz} = D\ (\sigma_{zz}-P)^n \quad \exp\left(\frac{-Q_a\ -PV_a}{kT}\right) \tag{2.10}$$

when samples are deformed uniaxially,

---

*Numbers in parentheses result from models at 5 cm/year.

where P = pressure, dynes/cm^2

Q_a = activation energy, ergs/mole

V_a = activation volume, cm^3/mole

k = Boltzmann constant = 8.314×10^7 erg/°K-mole

(Goetze and Brace, 1972; Weertman and Weertman, 1975; Carter, 1975). Because we require the flow law in a plane-strain form such as

$$\left[\dot{e}_{xz}^2 + \left(\dot{e}_{xx}-\dot{e}_{zz}\right)^2\right]^{1/2} = E\left[\tau_{xz}^2 + \frac{(\sigma_{xx}-\sigma_{zz})^2}{4}\right]^{n/2} \bullet$$

$$\exp\left(\frac{-Q_a \quad -PV_a}{kT}\right) \tag{2.11}$$

it will be necessary to make some assumptions about what controls the "viscosity". Two plausible possibilities are that viscosity is controlled by the maximum shear stress, which in the uniaxial experiment is $1/2\ (\sigma_{zz} - P)$; or that it is controlled by the second invariant of the stress tensor $J_2 = 1/3\ (\sigma_{zz} - P)^2$. In the former case we derive $E = 3(2)^{n-1}D$ and in the latter $E = \sqrt{(3)^{n+1}}D$. For a value of $n = 3$, this amounts to a difference of 4 versus 3 units of predicted strain rate. Out of ignorance, the two estimates were simply averaged. A further simplification was the neglect of the pressure term. This is necessary because no determinations of V_a have been made and reasonable because laboratory experiments fall in the same pressure range as these models. In terms of equation (B.3), we now have

$$A = E^{-1/n} = \left[D/2\left(3(2)^{n-1} + (3)\frac{n+1}{2}\right)\right]^{-1/n} \qquad (2.12)$$

and

$$B = Q_a/nk \qquad (2.13)$$

and it remains to select the best values of n, Q_a, and D from the literature.

On the assumption that externally applied stress is balanced by internal stress created by dislocations, it is possible to predict an equilibrium dislocation density proportional to the square of shear stress. This relationship has been experimentally confirmed in the case of olivine (Goetze and Kohlstedt, 1973; Kohlstedt and Goetze, 1974). The velocity of each dislocation is believed to be limited by a variety of stress-linear diffusional processes including dislocation climb required to accommodate an arbitrary strain and diffusion of pinning impurities.

Thus theoretically

$$\dot{e} = cb\rho_d V_d = c(\beta\sigma/Gb)^2\,\gamma\,(G/bkT)\cdot(\sigma/G) \qquad (2.14)$$

where c, β, γ are constants,

b = Burger's dislocation vector, cm

ρ_d = dislocation density, cm^{-2}

V_d = dislocation velocity , cm/sec

G = shear modulus, $dynes/cm^2$

(Weertman and Weertman, 1975). The cubic relationship appears to break down at high stresses (over one kilobar) as shown by Post (1973) because of some unknown mechanism allowing higher dislocation velocities. Fortunately, this breakdown usually occurs at higher than geologic stresses. Post and Griggs (1973) analyzed the Fennoscandian uplift data and concluded that $n = 3.2 \pm 0.3$ in the upper mantle. However, experiments in the high-stress range are apt to give $n > 3$, which cannot properly be extrapolated to geologic conditions.

The mineral quartz is one of the most important mechanically in continental crust, and it appears to share this behavior. Heard and Carter (1968) deformed natural quartzite at 2-23 kb and obtained $n = 6.5$. However, later work by Tullis (1971) provided sufficient data at stresses below 10 kb to determine a value of $n = 3.24$ in this range. Parrish et al. (1976) found $n = 2.6$, and very similar values of D and Q_a. Therefore a value of $n = 3.0$ was chosen as most likely representative of the behavior of quartz in the one-decade extrapolation to geologic stresses. Values of $D = 7.60 \times 10^{18}$ $(\text{dynes/cm}^2)^{-3}$ /sec and $Q_a = 4.43$ kcal/mole were selected from the data of Tullis.

The actual desired flow law is not that of polycrystalline quartz, but that of a granite or mixed sediment containing about 30% quartz. To estimate this we make the simple assumption

$$\dot{e}_{rock} = c\,\dot{e}_{mineral} \tag{2.15}$$

where c is the mineral concentration by volume, and obtain $A = 3.44 \times 10^{8}\ (\mathrm{sec})^{1/3}$ - dynes/cm^2 for granite. As the value of c enters into A through a cube-root, it is not critical. By assuming (2.15) we obtain an upper limit on stress and temperature, assuming that other minerals do not creep at the same conditions. This seems to be true of feldspars, from limited experimental data (Carter, 1971) and from the frequent occurrence of "augen gneisses" in which the quartz has obviously flowed while the feldspar has not. The behavior of the micas is more problematical, as they are not stable at high enough temperatures to permit laboratory testing.

When this granite flow law is input to the model, melting does not occur. Instead, appreciable creep begins at 20 (13) km depth with a velocity of 1 (5) cm/year (Figures 2.6, 2.7). The temperature stabilizes at 340-400 (400-480)°C, and shear stress drops off rapidly with depth. The shear zone expands rapidly (from an initial 1 cm) and in the 5 cm/year case it reaches 5.9 km thickness at 25 km depth and 16 km thickness at 50 km. The 1 cm/year case reached a thickness of 17 km at 35 km depth and was restricted from further expansion. This was because seismic refraction results commonly show a "Conrad discontinuity" in mid-crust, with higher velocities and presumably non-granitic rocks beneath.

This expansion of the shear zone unfortunately reduces the accuracy of the computational method. For example, at 5 cm/year and 25 km depth, the shear zone is producing a heat flux of 127 mW/m^2. If this heat production were uniformly distributed, the excess frictional temperature in the shear zone would be parabolic, with a maximum contrast of 29°C between the center and the margins. From equation (2.11) it is clear that such a temperature contrast should be accompanied by a stress contrast of about 0.66; or more physically, a strain-rate contrast of 3.4 times at constant stress. Thus the heat production will not in fact be uniformly distributed, and the model contains a contradiction in the assumption of uniform strain and temperature. This effect is less important when there are weaker rocks in the shear zone, because less flux implies a lower thermal gradient.

One such rock might be limestone, which in equatorial latitudes would be a prominent continental shelf deposit. Data on calcite creep comes mainly from studies of Yule marble by Heard and Raleigh (1972) and of Solnhofen limestone by Rutter and Schmid (1975) and Schmid (1976). As with olivine and quartz, early experiments were performed at a high stress level (1 to 4 kb), leading to high and variable values of n (7-8), However, later work by Schmid found n = 2.05, Q_a = 50.4 kcal/mole, and D = 7.08 x 10^{13} (dynes/cm^2)$^{-2.05}$/sec. His values are accepted, as they are obtained from the finer-grained rock type of interest, in the appropriate stress range of 100-1000 bars. No concentration correction is required for limestone, but the shear zone is limited to 5 km thickness, the largest amount that could plausibly be deposited.

Creep of limestone begins at 15 (10) km depth, and stabilizes at temperatures of 220-300 (270-320)°C. Stress drops off very rapidly, reaching essentially zero at a depth of 50 km. Very little creeping of the granite basement will occur if it is overlain by limestone.

The final material considered may not in fact be a geologic material. This is "wet" quartz, artificially grown under non-equilibrium conditions and containing H_2O concentrations as high as 8000 ppm (Heard and Carter, 1968). This water creates a substantial reduction in the creep strength, which was shown by Balderman (1974) to be continuous at all temperatures. He obtained a flow law with parameters $n = 3.64$, $Q_a = 31.6$ kcal/mole, and $D = 2.96 \times 10^{35}$ $(\text{dynes/cm}^2)^{-3.64}$/sec. This agrees within experimental scatter with Heard and Carter and is weaker by 40% than results of Hobbs et al. (1972). It enters these calculations in the same way as "dry" quartz, with a correction for an assumed volume concentration of 30%.

Blacic (1975) argues that water-weakening is a geologic effect because naturally-deformed quartzite is usually deformed by basal slip rather than prismatic. This agrees better with results on wet quartz in the laboratory. However, attempts to induce water-weakening in the lab have not been successful; Parrish et al. (1976) deformed natural quartz packed in talc (a source of water) and obtained results essentially the same as Tullis (1971), as much as forty times as strong as Balderman's artificial specimens. There is a further difficulty that all specimens to date have been deformed in the most favorable (0^+) orientation for basal slip, whereas an arbitrary natural strain requires five independent slip systems, any one of which may be rate-limiting. For all of these reasons, the results shown in Figures 2.6 and 2.7 probably represent the most extreme effects of water-weakening possible. They are comparable to the results on limestone, but with equilibrium temperatures some 40°C lower and stresses lower by 20-50%.

Two points about the accuracy of the computations can be shown from the results. First, the value of shear stress in the zone depends rather weakly on creep strength. At the same temperature and strain rate, "dry" granite should be stronger than "wet" granite by a factor of about 48 (T = 400°C, $\dot{e} = 10^{-13}$/sec). But the calculated shear stresses differ by only a factor of five when integrated from the surface to 50 km. This is because the work done is determined by the creep temperatures of the rocks, which differ by much less. This means that if the stress law contains errors of 5% introduced in equation (2.12) and 26% from equation (2.15) (caused by a factor of two error in concentration), the final calculated shear zone stress will still be accurate within 12%. An error of a decade in strain rate in extrapolating laboratory results would imply a stress law error factor of $(10)^{1/n}$, which implies a final solution error of 38% when n=3.

Second, the temperature stabilizes as soon as creep sets in. After that point it increases at about the same rate as the T_0 temperatures in the no-friction case. This makes the internal heating term evaluated in equation (B.12) a small fraction of the total heat production. Therefore, the calculation of this term by a backward difference is not an important source of error in the creep calculations.

2.3.3 Implications for Melting, Seismicity, and Stability

The models calculated in the last section have the property of being mutually exclusive. That is, if two minerals subject to creep are present in the subduction zone simultaneously, the stress and temperature will be controlled by the weaker. The same principle determines that no melting will occur if the same motion can be accommodated by creep at a lower temperature. Since quartz (whether "wet" or "dry") is a major component of continental crust and sediments, no melting can be expected to occur at the plate velocities modelled, up to 5 cm/year.

The approximate velocity that would be required (assuming "dry" quartz creep and no calcite, halite, or "wet" quartz present) can be estimated in the following way: The total frictional dissipation per unit fault area is equal to shear stress times slip distance. This distance is constrained by geologic observations and bouyancy forces and is not a function of plate velocity. At a plate velocity $V > 5$ cm/year, the creep strength of the rock is raised by the factor $(V/5)^{1/3}$. However, in warming from the calculated 460°C equilibrium to the 643°C melting point, a factor of 7.6 reduction in the creep strength occurs because of the thermal activation term. So the total work available is decreased by the factor $13.0\ V^{-1/3}$. Still, this reduced work must produce a maximum temperature increase that is 1.7 times greater than at 5 cm/year, to reach melting. This is only possible if the heat is spread over a narrower zone because of higher velocities. From equation (2.7) it is clear that the subduction time must be reduced by the ratio

$$\frac{t}{t_{(5)}} = \frac{5}{V} = .00205\, V^{2/3} \tag{2.16}$$

A velocity of roughly 108 cm/year would be required. This is an order of magnitude faster than any relative plate velocities yet determined from marine magnetic anomalies (LePichon, 1968). Whether such a velocity might have occurred briefly and gone unnoticed is a question discussed in Chapter 5. If calcite is present, if free water can diffuse into the quartz of the granite, or if the creep of micaceous minerals is significant, even higher and less acceptable plate velocities are required.

An interesting prediction of the model is that seismicity should not occur more than 50 km down-dip in the subduction zone (at 1 cm/year) or 25 km downdip at 5 cm/year. Unfortunately, it is not possible to test this prediction at present with the earthquake locations of the International Seismological Center, because the accuracy of depth location of shallow events is too poor to distinguish between shear-zone and mid-plate earthquakes. A local array capable of determining depths to within 5 km would be required. In the Zagros, Nowroozi (1971) has relocated earthquakes with an estimated error in horizontal position of 10-20 km, and most of his cross-sections show a minimum of seismicity at shallow depths just northeast of the original subduction zone. However, this may only indicate that this subduction zone is no longer active. If more precise locations were available, it would be possible by such modelling to determine the effective coefficient of friction at depth in subduction zones, and thus assess the effects of fluid pore pressure and any deviations from the frictional law observed in the laboratory. This subject is discussed further in sections 5.6 and 7.3.

One final conclusion from these models that is independent of the deformation mechanism is that lower velocities imply higher stresses. Even with a nonlinear rheology, this is contrary to intuition. It occurs because both the creep and melting mechanisms require a relatively constant temperature regardless of velocity. At a lower velocity the region warmed by the fault is thicker and more heat is required. The thickness of the region is proportional to $V^{-1/2}$ (2.7) and because the slip distance is unchanged, the stress must vary by the same exponent. This law does not apply exactly to the models calculated because of the influence of the shallow faulting region where stress is independent of velocith. It is more appropriate as an integral approximation:

$$\int_0^\infty \tau(s)\,ds \approx cV^{-\frac{1}{2}} \tag{2.17}$$

The importance of this effect is that it makes any given shear zone unstable with respect to long-term velocity perturbations. The short-term velocity variations imposed on a shear zone by the periodicity of large earthquakes (Smith, 1974) is not important because there is not time for the temperature of the zone to change significantly. However if a constant force, such as the bouyant resistance to continental subduction, reduces the plate velocity for a million years or more, then the resisting stress in the subduction zone will rise. This in turn will decrease the velocity even more. If at this stage a competitive shear zone should form to take up a portion of the plate velocity, its resistance will drop as the old zone stiffens. This is probably an important factor in the detachment of large slabs of crust from the subducting plate, as described by Mattauer (1975).

Table 2.1 Initial Continental Geotherm

Depth, km	Temperature, $^{\circ}C$
0	0
5	109
10	206
15	291
20	364
25	433
30	498
35	558
40	615
45	667
50	719
55	770
60	820
65	870
70	918
75	966
80	1013
85	1060
90	1105
95	1149
100	1192
105	1233
110	1274
115	1313
120	1352

TABLE 2.2

Parameters of Finite-difference Thermal Models

Model	Figure	Elapsed time (m.y.)	Velocity cm/year	Dip °	Mantle conductivity	Crustal H erg/g-sec	Fault Stress bars
I	2.1	8.	1.	30	McDonald (1959) 50/cm opacity	3.9×10^{-6} U 1.1×10^{-6} L	1,000 (5-80 km depth)
II	2.2	2.8	6.	15	McDonald (1959) 50/cm opacity	3.9×10^{-6} U 1.1×10^{-6} L	1,500 (10-20 km)
	2.3A	32.8	0.	-			1,200 (25-35 km)
	2.3B	102.8	0.	-			1,000 (40-80 km)
III	2.4A	1.3	5.	30	Lithosphere: Schatz and Simmons (1972)	3.25×10^{-6} U	(0.6) ρgZ (0-10 km)
	2.4B	5.4	5.	15/0		1.08×10^{-6} L	945-542 (15-25 km)
	2.4C	40.4	0.	-	Asthenosphere: 3.6×10^{6} erg/cm-sec°C		460-189 (30-60 km)

U signifies upper half and

L signifies the lower half of the 33 km - thick continental crust

Chapter 2 - Figure Captions

Figure 2.1 Results of thermal model I. Temperature field, heat flow, and Bouguer gravity anomaly during subduction of a continental margin at 1 cm/year. Continental crust is left white and lithosphere below the Moho is shaded. Temperatures in °C, heat flow in $\mu cal/cm^2$-sec, gravity in milligals. For complete parameters see Table 2.2.

Figure 2.2 Results of thermal model II. Temperature field, heat flow, and Bouguer gravity anomaly during subduction of a continental margin at 6 cm/year. See Figure 2.1 legend.

Figure 2.3 Temperature field and heat flow evolved from Figure 2.2 after 30 m.y. (A) and 100 m.y. (B) without motion of the slab. Dashed lines show former slab outlines. Hatched area in (B) is partially melted.

Figure 2.4 Results of thermal model III. Thermal history of a representative continental convergence zone. Continental crust is dotted; heat flow in mW/m^2; horizontal tic marks are 50 km apart. In (A) the continental margin is subducted at 5 cm/year. In (B) the plate becomes jammed and crustal overthrusting occurs. In

(C) the region is shown after 35 m.y. without additional motion. Complete parameters are in Table 2.2.

Figure 2.5 Schematic diagram of geometry for section 2.3.

Figure 2.6 Shear stress (top) and temperatures (bottom) in a subduction zone with a dip of 30° and a plate velocity of 1 cm/year. From top to bottom, curves represent (1) melting of granite containing muscovite; (2) melting of granite in the presence of water; (3) creep of granite containing 30% "dry" quartz' (4) creep in a 5 km layer of limestone; (5) creep of granite containing 30% "wet" quartz.

Figure 2.7 Shear stress (top) and temperatures (bottom) in a subduction zone with a dip of 30° and a plate velocity of 5 cm/year. From top to bottom, curves represent (1) melting of granite containing muscovite; (2) melting of granite in the presence of water; (3) creep of granite containing 30% "dry" quartz; (4) creep in a 5 km layer of limestone; (5) creep of granite containing 30% "wet" quartz.

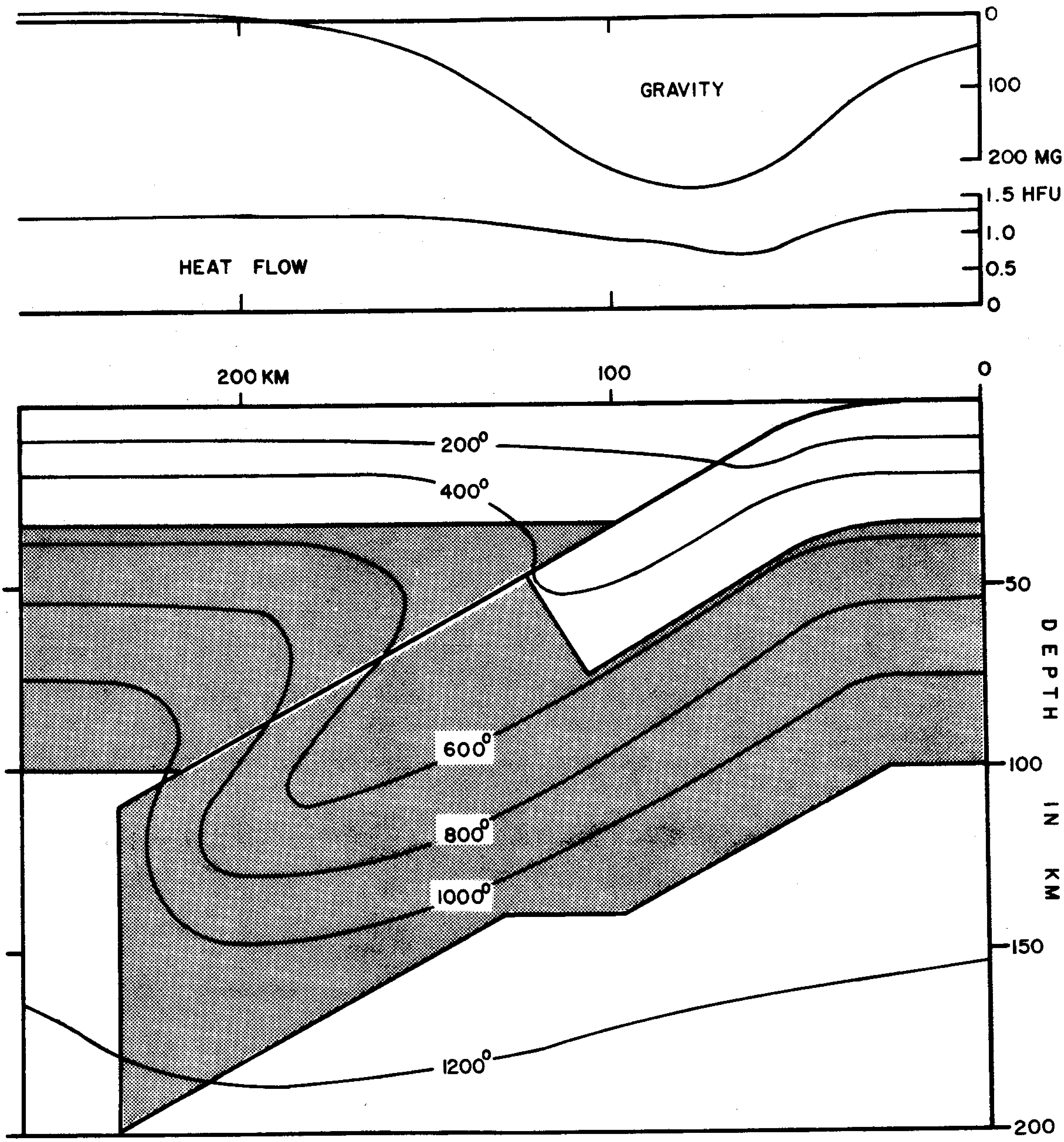
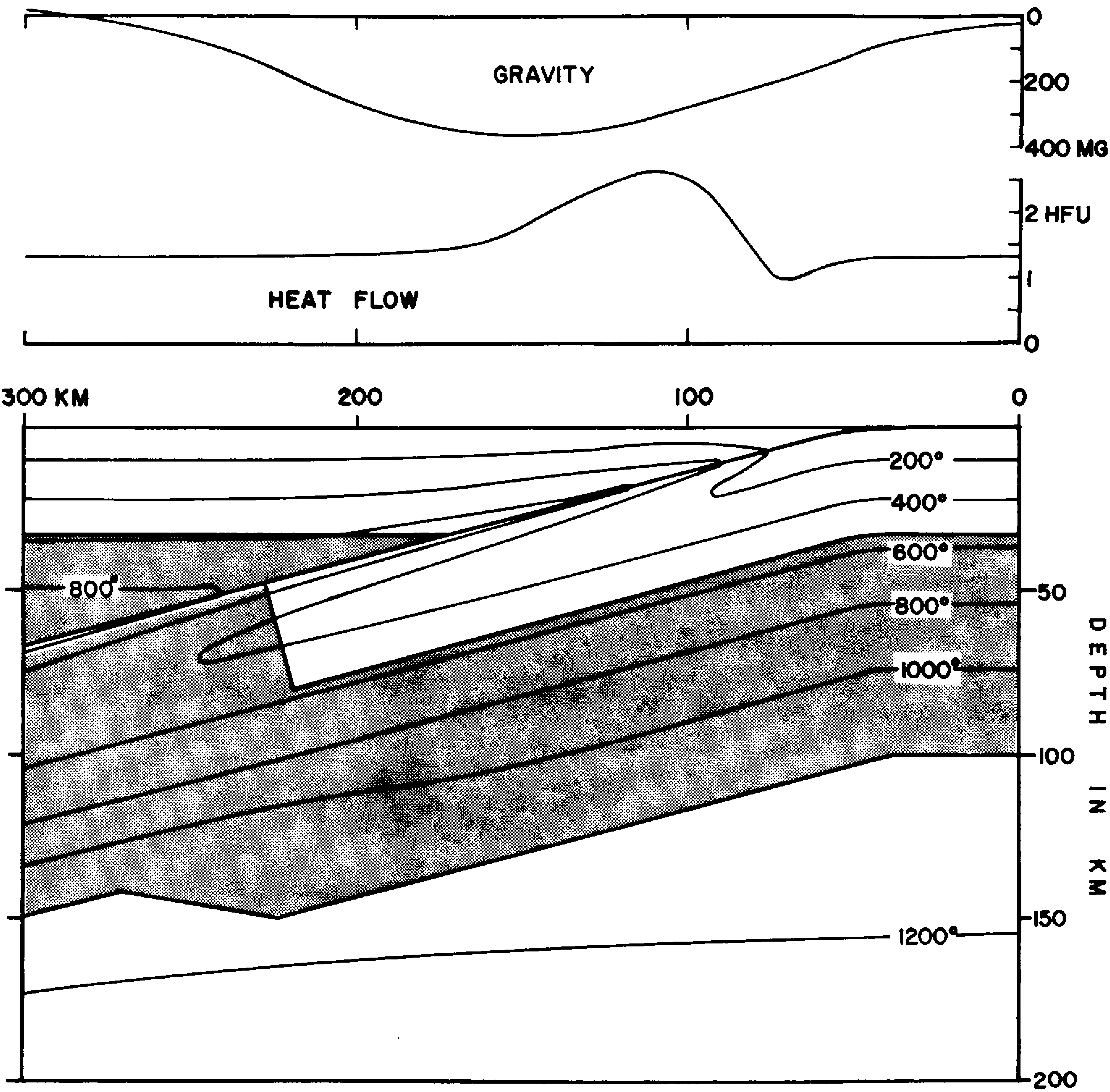

GRAVITY
0
100
200 MG
HEAT FLOW
1.5 HFU
1.0
0.5
0
200 KM
100
0
200°
400°
600°
800°
1000°
1200°
50
100
150
200
DEPTH IN KM

Fig. 2.1



GRAVITY
HEAT FLOW
0
200
400 MG
2 HFU
1
0
300 KM
200
100
0
200°
400°
600°
800°
800°
1000°
1200°
50
100
150
200
DEPTH IN KM

Fig. 2.2

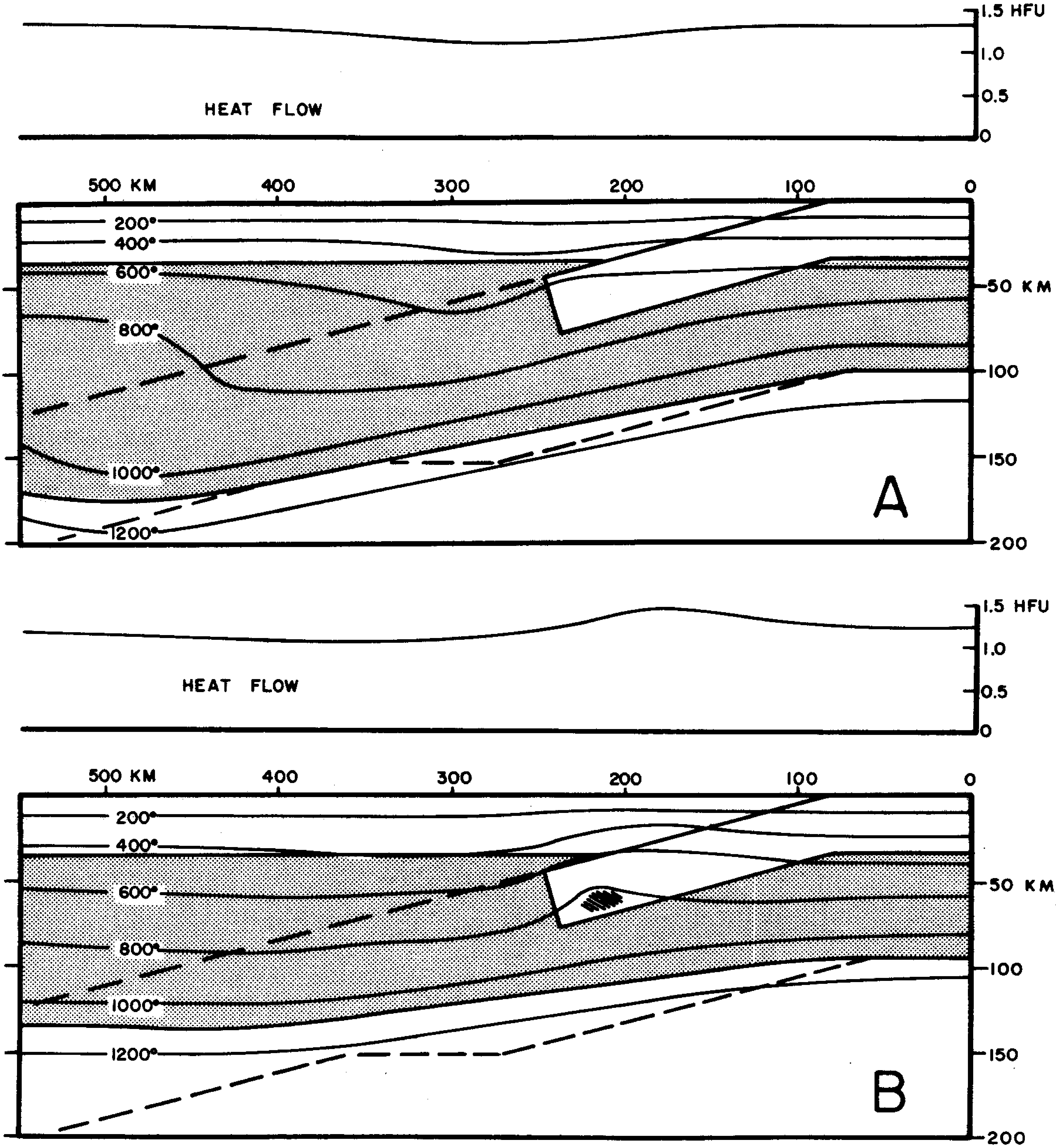

1.5 HFU
1.0
0.5
0
HEAT FLOW
500 KM
400
300
200
100
0
200°
400°
600°
800°
1000°
1200°
50 KM
100
150
200
A
1.5 HFU
1.0
0.5
0
HEAT FLOW
500 KM
400
300
200
100
0
200°
400°
600°
800°
1000°
1200°
50 KM
100
150
200
B

Fig. 2.3


80
40
0
50
100
150
KM
200
400
600
800
1000
1200
1400
A 1.3 M.Y.
B 5.4 M.Y.
C 40.4 M.Y.

Fig. 2.4

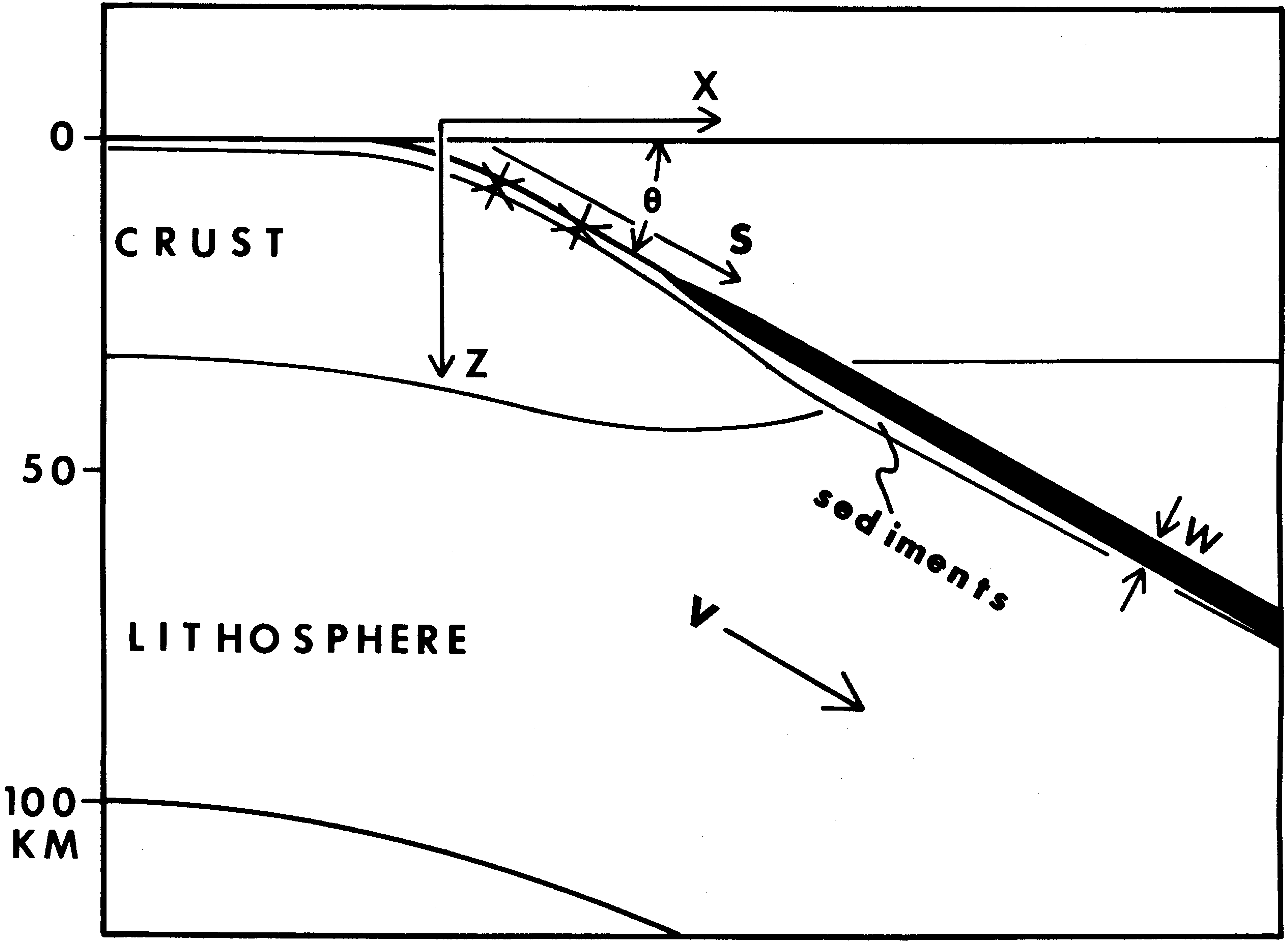


Fig. 2.5


V = 1 CM/YR
SHEAR STRESS IN KB
4
3
2
1
0
1
2
3
4
5
0
10
20
30
40
50
DEPTH IN KM (θ = 30°)
800
600
400
200
0
TEMPERATURE IN °C
1
2
3
4
5
0
10
20
30
40
50

Fig. 2.6


V = 5 CM/YR
SHEAR STRESS IN KB
DEPTH IN KM (θ = 30°)
TEMPERATURE IN °C
0
1
2
0
200
400
600
800
0
10
20
30
40
50
1
2
3
4
5

Fig. 2.7

CHAPTER 3: NUMERICAL MODELLING OF PLATE DEFORMATION WITH FINITE ELEMENTS

It would be ideal, for the evaluation of geologic hypotheses, if we had the numerical capacity to solve three dimensional, nonlinear, time-varying, linked thermal and mechanical problems. Since this is presently uneconomical, I have concentrated instead on the present-day patterns of deformation, for which the greatest number of seismological and other constraints are available. When the problem is posed in this way, the temperature field is an input parameter but not an unknown.

In this chapter the need for such modelling is reviewed, as well as the nonlinearity of rock mechanical properties which makes it difficult. A finite element technique, comparable to those used in engineering studies of elastic plate bending, is developed for two dimensional incompressible viscous flow. This is then generalized by either of two iteration techniques to accommodate an arbitrary nonlinear material. The resulting method is extremely flexible and comparatively economical. Accuracy is found to be good except around singularities, and in all cases the numerical errors are much less than the uncertainty in input parameters.

3.1 The need for understanding of instantaneous velocity fields.

It is part of the work of geologists to observe field evidence of the displacement of rock masses and to form hypotheses about the motions which emplaced them and the forces responsible. Unfortunately, after all the exposed rocks have been examined there may still be insufficient information to evaluate a hypothesis. Erosion or tectonic overprinting may destroy evidence, and even in the best circumstances it is impossible to know the topography, gravity anomalies, seismicity, fault-plane solutions, or heat flow of the past. For these reasons the study of modern tectonic analogues where the quantitative constraints of geophysics can be applied is essential.

In the deep ocean basins it is possible to measure recent motions by means of paleomagnitism, but no comparable tool exists for the study of intracontinental tectonics. Direct velocity measurements by surveying commonly require at least a decade before the accumulated displacement exceeds the measurement error. Even then the short term velocity may not be representative because of the effect of earthquakes. For example, the axis of the Siwilik Hills in the vicinity of Hardwar, India was uplifted 13 cm in the 1905 Kangra earthquake. But in the subsequent period 1908-1947 it rose only 2.5 cm more (Chugh, 1974). In such a region measurements over a century may be required to obtain an average velocity. And no geophysical technique to date has the precision to record the movements of rocks at depth.

A straightforward calculation of these velocities would require a knowledge of both driving forces and the flow law of rocks. Of the two, it is the driving force that is relatively well known. If the value of the gravitational constant and the radius of the Earth are not changing with time, then all long-wavelength deviatoric stresses in the Earth must result from horizontal density contrasts embodying gravitational potential energy. The density contrasts within 10 km of sea level are known from the Earth's topography; those at greater depths can be inferred from gravity anomalies with seismological constraints to reduce non-uniqueness. A very powerful constraint is the knowledge that the asthenosphere cannot support deviatoric stresses above a few bars (Kohlstedt and Goetze, 1974; Cathles, 1975), so that the stress state of the lithosphere is virtually independent of the deep mantle except in the vicinity of downgoing slabs.

Thus the problem is one of finding flow laws which when combined with approximately known driving forces will properly predict the secondary fields of topography, seismicity, and fault-plane solutions. Whether this is posed as a forward or an inverse problem, the flow-law information obtained is not the essential result. Rather, it is the velocities obtained in the solution with the correct flow law which allow extrapolations to the geologic past.

3.2 Flow Laws Applied to Continental Convergence Models

Fortunately, our knowledge of the Earth's flow law is improving rapidly through experimental studies. Each documented deformation mechanism which is intelligently extrapolated to geologic conditions gives an upper limit on the strength of rock, because the additional mechanisms yet unknown can only act to further weaken it. In the models to follow, flow laws weaker than experiments have demonstrated are invoked only when changes in boundary conditions or temperature cannot produce a successful model.

One well-documented deformation mechanism that makes only a negligable contribution to the steady-state flow field is elasticity. Rocks constantly undergo strain buildup and release cycles because of their elastic response to the periodic stress changes caused by earthquakes. But if a rock is stationary the long-term average strain from the elastic mechanism is zero. It can only undergo a permanent elastic strain if it is in motion through the stress field. In that case the average elastic strain rate is related to average stress by:

$$\overline{\dot{e}_{xz}} \approx \frac{\vec{V} \cdot \vec{\nabla}\, \overline{\tau_{xz}}}{G} \tag{3.1}$$

where τ = shear stress, dynes/cm^2

V = particle velocity, cm/sec

G = Young's modulus, dynes/cm^2

For a typical continental collision problem, V=5 cm/year, $|\nabla \tau|$ is roughly 1 kilobar/100 km; and G=340 kilobars. This gives $\dot{e}_{ELASTIC} = 5 \times 10^{-17}$ sec^{-1}, equivalent to a strain of 1% in 6.3 million years. Strains of this magnitude can be safely neglected if other deformation mechanisms give apparent viscosities of less than 2×10^{25} poise. As will be seen, this approximation simplifies the mathematics of the problem by making the flow incompressible. It is also consistent with the goal of obtaining an upper limit to strength in the initial models.

The upper mantle is commonly believed to be composed mainly of olivine and pyroxene, with varying amounts of garnet, feldspar, and spinels. Of these, olivine is probably most susceptible to creep, and has been the subject of the most intensive study. Carter (1971) has shown that much of the apparent plasticity of orthopyroxene in the laboratory is the result of transformations of clinopyroxene that would not occur under mantle conditions. Ross and Nielsen (1975) have found from laboratory experiments that pyroxenes undergo strain hardening while olivine recrystallizes and continues to shear. Thus the creep contribution of pyroxene is probably negligible. For the creep law of olivine we use the data of Kohlstedt and Goetze (1974) derived from low-stress experiments on dry single crystals. For reasons given in the last chapter, this low-stress data is most likely to extrapolate correctly to mantle conditions. From the data we obtain constants $n=3.0$, $A=6.82\times10^4$ dynes/cm^2, and $B=21000°K$ for the equation: $\tau^* =$

$$\left[\frac{(\sigma_{xx}-\sigma_{zz})^2}{4}+\tau_{xz}^2\right]^{1/2} = A\cdot\exp\left(\frac{B}{T}\right)\left[(\dot{e}_{xx}-\dot{e}_{zz})^2+\dot{e}_{xz}^2\right]^{1/2n} \tag{3.2}$$

This flow law falls in the range determined by Post (1973) for creep of polycrystalline Mt. Burnett Dunite, although Post's stresses may be too high because of oxidation of his samples. Because of our ignorance of exact composition, no correction is made for the effect of the stiffer minerals. Since these are likely to make up less than 25% of the upper mantle (Clark and Ringwood, 1964) the correction would be less than 9% of A (equations 2.15 and 3.2).

This flow law may be too stiff if there is significant H_2O in the mantle. Many authors have confirmed the result of Carter and Ave'Lallemant (1974) that a small amount of water reduces the dislocation creep activation energy by as much as 30 to 40 kcal/mole, with a resultant strength decrease of up to two orders of magnitude. However, the Fennoscandian uplift data indicates an asthenosphere viscosity of 1.3×10^{20} poise at a strain rate of 10^{-15}/sec (Van Bemmelan and Berlage, 1935, reported in Cathles, 1975). This neatly matches the Kohlstedt and Goetze dry flow law if the asthenosphere is at 1450°C, but it can only be matched to a "wet" flow law with temperatures around 1070°C, which is unacceptably low (see section 2.1). It appears that dislocation creep is sufficient to explain the low strength of the asthenosphere without invoking volatiles or partial melting. In any case, the question of partial melting as a deformation mechanism is sidestepped in this thesis because models are extended no further down than the 1100°C isotherm where strength becomes negligible in comparison with the shallow lithosphere.

In the parts of the mantle just below the Moho where temperatures are as low as 550°C, any geologic deformation rate requires a stress of tens of kilobars according to the n=3 flow law just discussed. Experimental data show a gradual transition in the 1.5-10. kb shear stress range through higher and higher n values to perfect plasticity (Kohlstedt and Goetze, 1974). The mechanism is poorly understood, but limited data (Post, 1973) shows that such stresses produce rapid creep events in polycrystalline samples which may be related to mantle earthquakes. In the models, a simple cut-off plastic limit at a differential stress of 3 kilobars was substituted for the gradual transition. Details of high-stress rheology are lost in any case because of the need to limit viscosities to 10^{25} poise.

Above the Moho in the continental crust we obtain an upper strength bound by assuming that only quartz creeps and that it is not hydrolitically weakened. The assumption seems to be justified by experiments of Tullis _et al_. (1976), who deformed Westerly granite and found that quartz alone recrystallizes. They also found a significantly greater strength in Hale albitite with a lower fraction of quartz. Using this assumption, the flow law was derived in section 2.3.2: $n=3.0$, $A=3.4 \times 10^8$ dynes/cm^2, and $B=7445°K$. That section also contains a discussion of the frictional faulting law of the upper crust.

An accurate representation of the lower continental crust is hampered by our lack of knowledge of its constitution. Seismologists have frequently assumed on the basis of velocities that below the Conrad discontinuity it consists of pure gabbro. This simplistic hypothesis is suspect because such rocks are not found in old orogenic belts that have been subjected to deep erosion. Instead, inspection of these zones reveals a "sequence of basal granulites overlain in turn by anorthosites, amphibolite facies gneisses, and migmatites intruded by granites, greenschist facies phyllites and finally supracrustal sediments. . . . The continental crust is thus characterized by complex lateral and vertical compositional variations overprinted by pronounced metamorphic layering." (Salisbury and Fountain, 1976) In an earlier paper Christensen and Fountain (1975) showed that a granulite model was also consistent with seismic velocity data, and deduced a quartz concentration of 6-23% (in a bronzite-quartz-An_{29} assemblage) for the lower crust of the Canadian shield. In the initial models the lower crust is represented as creeping by the deformation of its quartz constituent of

14%, leading to the value $A=4.44x10^8$ dynes/cm^2 (which is 28% stronger than the upper crust).

The complete strength profile of this initial model is shown in Figure 3.1, which utilizes the undisturbed geotherm of Table 2.1. At a typical tectonic strain rate, it shows that permanent deformation will take place by faulting down to 20 km, creep in the lower crust, plasticity in the uppermost mantle, and creep in the remainder of the lithosphere. It also shows that the mechanical lithosphere is 20 to 30 km thinner than the thermal lithosphere defined as the region below the solidus. But the most striking feature is the very weak (100-200 bar) zone that is predicted to occur at the base of the continental crust.

The significance of this zone is that it effectively decouples the strong upper crust and lithosphere. In response to large tectonic stresses during continental collisions, it is likely that the two may behave as separate plates. Faults beginning in the brittle upper crust should flatten out with depth and pass into this horizontal zone of <u>decollement</u>, rather than penetrating the entire lithosphere. This explains one of the fundamental geologic observations about continental zones: that large sheets of crust become detached from the lithosphere and overthrust adjacent crust. It explains the occurence of thick crustal roots with twice normal thickness, yet no evidence for a mantle layer in the middle. And the large shear strains absorbed in this zone are the cause of the regional horizontal foliations discussed by Mattauer (1976).

3.3 Finite Element Formulation of the Viscous Deformation Problem

The formulation presented here is for two-dimensional plane-strain flow. Because of the great length of continental collision ranges, this is thought to be a good approximation in a central vertical cross-section. For a slow creeping flow the static form of the stress equilibrium equation is appropriate. In two dimensions

$$\frac{\partial p_{xx}}{\partial x} + \frac{\partial p_{xz}}{\partial z} = 0$$

$$\frac{\partial p_{zz}}{\partial z} + \frac{\partial p_{xz}}{\partial x} + \rho g = 0 \qquad (3.3\ a,b)$$

applies in the interior of the domain, where

x, z = horizontal and vertical coordinates in a left-handed system, cm

p_{xx} = horizontal normal stress, dynes/cm^2 (extension positive)

p_{zz} = vertical normal stress, dynes/cm^2 (extension positive)

p_{xz} = shear stress, dynes/cm^2

For stress continuity on the boundary of the domain we require

$$p_{xx}\cos(\vec{\nu},\vec{x}) + p_{xz}\cos(\vec{\nu},\vec{z}) = h_x$$

$$p_{zz}\cos(\vec{\nu},\vec{z}) + p_{xz}\cos(\vec{\nu},\vec{x}) = h_z \qquad (3.4\ a,b)$$

where $\vec{\nu}$ = outward normal vector of surface,

h = applied external forces, dynes/cm^2.

However, the gravitational forces in (3.3b) are very large and the variations of interest are small. To promote numerical accuracy, we define deviatoric stresses relative to the normal lithostatic pressure: $\tau_{xz} \equiv p_{xz}$;

$$\sigma_{xx} \equiv p_{xx} + \int_0^z g(r)\rho_0(r)\,dr \;; \quad \sigma_{zz} \equiv p_{zz} + \int_0^z g(r)\rho_0(r)\,dr$$

(3.5 a,b,c)

In these terms (3.3 a,b) becomes:

$$\frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \tau_{xz}}{\partial z} = 0$$

$$\frac{\partial \sigma_{zz}}{\partial z} + \frac{\partial \tau_{xz}}{\partial x} + \Delta\rho\, g = 0$$

(3.6 a,b)

The corresponding deviatoric boundary forces and conditions are:

$$f_x \equiv h_x + \int_0^z g(r)\rho_0(r)\,dr = \sigma_{xx}\cos(\vec{\nu},\vec{x}) + \tau_{xz}\cos(\vec{\nu},\vec{z})$$

$$f_z \equiv h_z + \int_0^z g(r)\rho_0(r)\,dr = \sigma_{zz}\cos(\vec{\nu},\vec{z}) + \tau_{xz}\cos(\vec{\nu},\vec{x})$$

(3.7 a,b)

In order to proceed further we require strain-rate definitions and a constitutive law. For reasons to become apparent later, it is necessary to develop a linear viscous methodology before handing the nonlinear materials described in the last section. Define

$$\dot{e}_{xx} = \frac{\partial u}{\partial x} \; ; \quad \dot{e}_{zz} = \frac{\partial w}{\partial z} \; ; \quad \dot{e}_{xz} = \frac{\partial u}{\partial z} + \frac{\partial w}{\partial x} \qquad (3.8\ a,b,c)$$

where $\dot{e}$ = strain rate in sec^{-1}

u,w = horizontal and vertical velocities in cm/sec.

Note that we use the "~~engineering~~" rather than the "tensor" definition of shear strain.

We now define a linear viscosity η in poises as a function of space only by the relations

$$(\sigma_{xx} - \sigma_{zz}) = 2\eta(\dot{e}_{xx} - \dot{e}_{zz}) \; ; \quad \tau_{xz} = \eta \dot{e}_{xz} \qquad (3.9\ a,b)$$

These definitions are appropriate for an incompressible material such that

$$\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0 \qquad (3.10)$$

In this case the pressure P (dynes/cm^2) is not related to the strain rates but is given by

$$P = -\frac{1}{2}(\sigma_{xx} + \sigma_{zz}) \qquad (3.11)$$

Stress equilibrium is now related to velocity and pressure by

$$-\frac{\partial P}{\partial x} + 2\frac{\partial}{\partial x}\left(\eta \frac{\partial u}{\partial x}\right) + \frac{\partial}{\partial z}\left(\eta\left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}\right)\right) = 0$$

$$g\Delta\rho - \frac{\partial P}{\partial z} + 2\frac{\partial}{\partial z}\left(\eta \frac{\partial w}{\partial z}\right) + \frac{\partial}{\partial x}\left(\eta\left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}\right)\right) = 0 \qquad (3.12\ a,b)$$

in the interior, and, on the boundary

$$\left(2\eta\frac{\partial u}{\partial x}-P\right)\cos(\vec{\nu},\vec{x})+\eta\left(\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x}\right)\cos(\vec{\nu},\vec{z})-f_x=0$$

$$\left(2\eta\frac{\partial w}{\partial z}-P\right)\cos(\vec{\nu},\vec{z})+\eta\left(\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x}\right)\cos(\vec{\nu},\vec{x})-f_z=0$$

(3.13 a,b)

Since the four left-hand sides of (3.12) and (3.13) are zero everywhere on interior and boundary, we can multiply the equations by any arbitrary finite weighting functions and still have zero everywhere. We choose as weighting functions perturbations $\delta u(x,z)$ and $\delta w(x,z)$ of the velocity field. These are a pair of scalar functions of position. If (3.12) and (3.13) hold, then their weighted integral is also zero:

$$\iint_A \left\{\left[\left(-\frac{\partial P}{\partial x}+2\frac{\partial}{\partial x}\left(\eta\frac{\partial u}{\partial x}\right)+\frac{\partial}{\partial z}\left(\eta\left(\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x}\right)\right)\right]\delta u+\right.\right.$$
$$\left.\left[\Delta\rho g-\frac{\partial P}{\partial z}+2\frac{\partial}{\partial z}\left(\eta\frac{\partial w}{\partial z}\right)+\frac{\partial}{\partial x}\left(\eta\left(\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x}\right)\right)\right]\delta w\right\}dA$$
$$-\oint_S\left\{\left[\left(2\eta\frac{\partial u}{\partial x}-P\right)\cos(\vec{\nu},\vec{x})+\eta\left(\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x}\right)\cos(\vec{\nu},\vec{z})-f_x\right]\delta u\right.$$
$$\left.+\left[\left(2\eta\frac{\partial w}{\partial z}-P\right)\cos(\vec{\nu},\vec{z})+\eta\left(\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x}\right)\cos(\vec{\nu},\vec{x})-f_z\right]\delta w\right\}dS$$
$$=0$$

(3.14)

If we require (3.14) to hold for an infinite number of orthogonal variations of velocity (forming a basis which spans the infinite-dimensional space of possible velocity fields) then the mathematical step from (3.12) and (3.13) to the weak form (3.14) can be reversed. Since δ u and δ w need only span the space of u and w, we can restrict them to be both continuous and incompressible (equation 3.10). Because of their continuity we can integrate (3.14) by Green's theorem. After the cancellation of numerous boundary integrals (because of the minus sign chosen in (3.14)), we obtain

$$\iint_A \Big[\Delta\rho g\,\delta w + P\frac{\partial}{\partial x}(\delta u) - 2\eta\frac{\partial u}{\partial x}\frac{\partial}{\partial x}(\delta u) - \eta\frac{\partial u}{\partial z}\frac{\partial}{\partial z}(\delta u) - \eta\frac{\partial w}{\partial x}\frac{\partial}{\partial z}(\delta u) + P\frac{\partial}{\partial z}(\delta w) - 2\eta\frac{\partial w}{\partial z}\frac{\partial}{\partial z}(\delta w) - \eta\frac{\partial u}{\partial z}\frac{\partial}{\partial x}(\delta w) - \eta\frac{\partial w}{\partial x}\frac{\partial}{\partial x}(\delta w)\Big]dA + \oint_S \big[f_x\delta u + f_z\delta w\big]dS = 0 \tag{3.15}$$

Because the variations are incompressible, the terms involving pressure now drop out and we have the stress equilibrium equation in a weak form involving only velocities and viscosities:

$$\iint_A \Big[-\Delta\rho g\,\delta w + 2\eta\frac{\partial u}{\partial x}\frac{\partial}{\partial x}(\delta u) + \eta\left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}\right)\frac{\partial}{\partial z}(\delta u) + 2\eta\frac{\partial w}{\partial z}\frac{\partial}{\partial z}(\delta w) + \eta\left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}\right)\frac{\partial}{\partial x}(\delta w)\Big]dA - \oint_S \big[f_x\delta u + f_z\delta w\big]dS = 0 \tag{3.16}$$

In order to incorporate the incompressibility (3.10) directly into this equation, one can define a stream function $\emptyset$ in units of cm^2/sec such that

$$u = \frac{\partial \phi}{\partial z} \quad ; \quad w = -\frac{\partial \phi}{\partial x} \tag{3.17}$$

Equation (3.16) becomes

$$\iint_A \left[\Delta\rho g \frac{\partial}{\partial x}(\delta\phi) + 4\eta \frac{\partial^2 \phi}{\partial x \partial z} \frac{\partial^2}{\partial x \partial z}(\delta\phi) + \eta\left(\frac{\partial^2 \phi}{\partial z^2} - \frac{\partial^2 \phi}{\partial x^2}\right) \cdot \left(\frac{\partial^2}{\partial z^2}(\delta\phi) - \frac{\partial^2}{\partial x^2}(\delta\phi)\right)\right] dA - \oint_S \left[f_x \frac{\partial}{\partial z}(\delta\phi) - f_z \frac{\partial}{\partial x}(\delta\phi)\right] dS = 0 \tag{3.18}$$

The essential approximation of the method is to replace the infinite-dimensional space of possible stream functions by an N-dimensional subspace spanned by orthogonal component stream functions:

$$\phi(x,z) = \sum_{i=1}^{N} c_i \phi_i(x,z) \tag{3.19}$$

In this case the velocity field is described by the N constants c_i, and the space of possible velocity variations is spanned by the derivitives of the N functions ϕ_i. Thus the approximate form of stress equilibrium reduces to a set of N equations in N unknowns:

$$\iint_A \left[\Delta\rho g \frac{\partial}{\partial x}\phi_i + 4\eta \frac{\partial^2 \phi_i}{\partial x \partial z} \sum_{j=1}^{N} c_j \frac{\partial^2 \phi_j}{\partial x \partial z} + \eta\left(\frac{\partial^2 \phi_i}{\partial z^2} - \frac{\partial^2 \phi_i}{\partial x^2}\right) \cdot \left(\sum_{j=1}^{N} c_j \left(\frac{\partial^2 \phi_j}{\partial z^2} - \frac{\partial^2 \phi_j}{\partial x^2}\right)\right)\right] dA - \oint_S \left[f_x \frac{\partial \phi_i}{\partial z} - f_z \frac{\partial \phi_i}{\partial x}\right] dS = 0 \; ; \; i = 1, 2, 3 \ldots N \tag{3.20}$$

This can be expressed as a set of simultaneous linear equations

$$\underset{\sim}{K}_{ij}\,\vec{c}_i = \vec{F}_i \tag{3.21}$$

if

$$\underset{\sim}{K}_{ij} = \iint_A \left[4\eta \left(\frac{\partial^2 \phi_i}{\partial x \partial z}\right)\cdot\left(\frac{\partial^2 \phi_j}{\partial x \partial z}\right) + \eta\left(\frac{\partial^2 \phi_i}{\partial z^2} - \frac{\partial^2 \phi_i}{\partial x^2}\right)\cdot \left(\frac{\partial^2 \phi_j}{\partial z^2} - \frac{\partial^2 \phi_j}{\partial x^2}\right)\right] dA \tag{3.22}$$

and

$$\vec{F}_i = \iint -\Delta\rho g \frac{\partial \phi_i}{\partial x}\, dA + \oint_S \left(f_x \frac{\partial \phi_i}{\partial z} - f_z \frac{\partial \phi_i}{\partial x}\right) dS \tag{3.23}$$

Note that it was not necessary to assume a constant viscosity in this derivation. A spatially-varying viscosity was assumed in (3.12) but when the problem is cast into a weak form the derivitives of viscosity drop out. Thus the viscous problem can be solved by finite elements in a single step, without the iteration that is required in finite-difference methods for variable viscosity (Andrews, 1972).

Further efficiency is obtained by designing the ϕ_i functions to be non-zero in only a limited part of the domain. Then many of the constants in $\underset{\sim}{K}_{ij}$ are zero and the matrix can be arranged as a banded symmetric positive-definite one. This is accomplished by dividing the domain into (triangular) subregions called elements, with each triangular vertex referred to as a node. The summation factors c_i are then identified with values of $\emptyset$ or its derivitives at a particular node and $\emptyset_i$ is represented within adjacent elements by polynomials.

The type of element used in this thesis was developed by Bazeley et al. (1965) for problems of thin-plate bending. In that application, the normal plate displacement corresponds to our stream function. The element has three unknowns per node, which are identified with $\emptyset$, W, and u. These nine values from the three nodes of each element are used to determine the coefficients of a two-dimensional cubic polynomial representation of $\emptyset$ in the element interior. Thus the representation of the stream function is cubic, the velocity fields are quadratic, and the variation of strain-rate is linear in each element domain. The possible deformations of such an element are shown in Figure 3.2. Actually, the cubic polynomial has ten constants and one must be fixed arbitrarily. This is chosen to suppress the non-physical degree of freedom in which the interior of the element rotates without any motion of (or net flow between) the nodes.

The one failing of this type of element is that it is "incompatible." That is, the velocity fields within adjacent elements cannot be exactly matched at element boundaries.

The velocity normal to each boundary is determined by the nodal values at the ends of the boundary and therefore is the same on both sides. However, the parallel component of velocity includes a small contribution from nodal values at the third vertex and thus may differ on opposite sides of the boundary. Another way of expressing this is that the stream function is continuous at element boundaries but its normal derivitive may not be. Full continuity would require the use of a quintic interpolation polynomial with 21 unknowns per element.

Theoretically, it has been shown that such nonconforming elements converge to exact solutions in the limit of infinitesimally small elements (Bazeley *et al.*, 1965). However, this is of little practical interest. Because of the discontinuities, it is no longer possible to prove that the finite element grid is "stiffer" (requires more energy to deform) than the real structure it represents; it may err in either direction. However, this may be a virtue. According to Zienkiewicz (1971), "for practical engineering purposes in most cases the accuracy obtained by the non-conforming triangle is adequate. Indeed it gives, at most practical subdivisions, results superior to those attainable with equivalent conforming triangles. This may well be due to the fact that the solution now does not follow the energy bounds...". The fully-conforming element is very rarely used because it requires 27 times as much computational effort.

The integral in equation (3.22) is evaluated numerically using 7 integration points in each element. With the proper positions and weighting factors (Zienkiewicz, 1971), such integration is exact for polynomial integrands up to fifth order. Each second derivitive of $\emptyset_i$ occuring in the equation is linear in space, so a cubic variation of viscosity is permissable. This viscosity is determined from its values at integration points, allowing the effects of temperature gradients to be approximately represented. This variation of viscosity applies only within each element, and there is no restriction on viscosity differences between adjacent elements. A geologic contact between different rock types is easily represented by defining its shape with nodes and inputting different constitutive laws for the elements on opposite sides.

One advantage of the finite element method is the great ease with which velocity boundary conditions can be included in the problem. The velocity at a node is fixed by replacing one of the N equations (3.20) with an equation of the nodal variable to the desired value. If no velocity or force conditions are specified on the domain boundary, it will automatically be a free or no-work boundary.

3.4 Solution Techniques for Nonlinear Materials

When the material is nonlinear in strain-rate it is not possible to follow the development of the preceeding section and obtain simultaneous linear equations. This is because the total stress at a point cannot be expressed as the sum of stresses caused by the strains associated with the component ϕ_i functions. Instead, it is necessary to solve the linear problem iteratively using artificial viscosities until a solution is found which also obeys the nonlinear constitutive law within an acceptable error. This can be done by either the "variable viscosity" or the "initial stress" techniques (Zienkiewicz, 1971). The "initial strain" technique is not applicable to problems with perfect plasticity because strain-rate is not uniquely defined as a function of stress.

3.4.1 The Variable Viscosity Method

The variable viscosity technique is simpler, although it is generally considered to be more expensive and slower to converge. One starts with an initial guess of the average strain-rate of the problem, evaluates the corresponding stresses from the material constitutive law, and estimates a viscosity from the ratio in each element. A linear viscous solution is then performed, leading to an improved estimate of local strain-rate. After each step, the new viscosity is given by

$$\eta_{i+1}(x,z) = \tau_i^*(x,z) \Big/ \dot{e}_i^*(x,z) \tag{3.24}$$

where
$$\tau^* = \sqrt{\tfrac{1}{4}(\sigma_{xx} - \sigma_{zz})^2 + \tau_{xz}^2} \tag{3.25}$$

and
$$\dot{e}^* = \sqrt{(\dot{e}_{xx} - \dot{e}_{zz})^2 + 4\dot{e}_{xz}^2} \tag{3.26}$$

at each of the seven integration points per element. The stresses are evaluated from the calculated strains with a constitutive law of arbitrary form. In the case of plasticity by faulting, the stress-strain relation will be anisotropic as shown in Appendix C.

This technique converges slowly but reliably as shown in Figure 3.3. It has the advantage that the results of the previous iteration are discarded and no numerical errors in stress equilibrium can accumulate. Its disadvantage is that the stiffness matrix $\underset{\sim}{K}$ must be recomputed at every step and re-factored to obtain a solution. If the number of unknowns is N and the matrix bandwidth is M then the time required for each iteration will vary as NM^2. This tends to preclude models with very fine grids.

3.4.2 The Initial Stress Method

The initial stress technique is more complex and consists of summing solutions (which all satisfy equilibrium and incompressibility) to find the solution that satisfies the nonlinear flow law. After each solution the accumulated calculated stress is

$$(\sigma_{xx}-\sigma_{zz}) = \sum_{i=1}^{n}(\Delta\sigma_{xx}-\Delta\sigma_{zz})_i = \sum_{i=1}^{n} 2\eta_i(\Delta\dot{e}_{xx}-\Delta\dot{e}_{zz}) ;$$

$$\tau_{xz} = \sum_{i=1}^{n}(\Delta\tau_{xz})_i = \sum_{i=1}^{n} 2\eta_i(\Delta\dot{e}_{xz})_i \qquad (3.27\ a,b)$$

This is separated into the correct stress from the flow law and an "unbalanced stress:"

$$(\sigma_{xx}-\sigma_{zz}) = m(\dot{e}_{xx},\dot{e}_{zz},\dot{e}_{xz}) + (\overline{\sigma}_{xx}-\overline{\sigma}_{zz}) ;$$

$$\tau_{xz} = r(\dot{e}_{xx},\dot{e}_{zz},\dot{e}_{xz}) + \overline{\tau}_{xz} \qquad (3.28\ a,b)$$

The accumulated stress is set to its proper values (m or r) and in the next iteration the unbalanced stresses are treated as an "initial stress" or pre-stress in the viscous problem. In these subsequent iterations no external forces, body forces, or non-zero boundary conditions are imposed, because they entered the cumulative stress and velocity fields with the first iteration. In subsequent iterations the stress equilibrium equation is corrected for pre-stress. With no body forces it becomes:

$$\frac{\partial}{\partial x}(\Delta\sigma_{xx}-\overline{\sigma}_{xx}) + \frac{\partial}{\partial z}(\Delta\tau_{xz}-\overline{\tau}_{xz}) = 0$$

$$\frac{\partial}{\partial z}(\Delta\sigma_{zz}-\overline{\sigma}_{zz}) + \frac{\partial}{\partial x}(\Delta\tau_{xz}-\overline{\tau}_{xz}) = 0 \qquad (3.29\ a,b)$$

In following the steps of the last section the same stiffness matrix $\underset{\sim}{K}$ is obtained, but there is a different equation for the load vector involving only unbalanced stresses:

$$\vec{F}_i = \iint_A \left[\left(\overline{\sigma_{zz}} - \overline{\sigma_{xx}} \right) \frac{\partial^2 \phi_i}{\partial x \partial z} + \overline{\tau}_{xz} \left(\frac{\partial^2 \phi_i}{\partial x^2} - \frac{\partial^2 \phi_i}{\partial z^2} \right) \right] dA \tag{3.30}$$

This is also numerically integrated in each element. The effect of solving the problem with the new load vector is to distribute the unbalanced stresses in a way that satisfies stress equilibrium. Since the new load creates new displacements, the procedure must be iterated. This iteration process is illustrated in Figure 3.4.

A pure initial stress method is very slow to converge in the case of severe nonlinearity. A common compromise is to apply the load and boundary conditions in increments, re-evaluate viscosities at the beginning of each step, and converge in that step by initial-stress iterations. This incremental strategy is described by Sarne (1974), and a flow chart is shown in Figure 3.5. It is particulary advantageous to design load steps so that the plastic limit is just barely exceeded at a few integration points per step. If β_i is the coefficient of the load at load step i, the program should survey non-plastic regions to estimate

$$\beta_{i+1} = (1+\epsilon)\beta_i \cdot \text{minimum}\left(\frac{\tau_p}{\tau^*}\right) \tag{3.31}$$

where τ_p is the local plastic limit shear stress (Marcal and King, 1967). Then after an integration point has reached the plastic limit, convergence is assisted by formulating a new, incremental, constitutive law which allows further deformation along the same axis without further stress. This is accomplished by expressing the incremental stress-strain law in matrix form.

$$\begin{Bmatrix} \Delta\sigma_{xx} - \Delta\sigma_{zz} \\ \Delta\tau_{xz} \end{Bmatrix} = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} \begin{Bmatrix} \Delta\dot{e}_{xx} - \Delta\dot{e}_{zz} \\ \Delta\dot{e}_{xz} \end{Bmatrix} \tag{3.32}$$

and requiring that the dot product of the incremental stress vector with the vector of stresses at the onset of plasticity $\{\sigma_0, \tau_0\}$ be zero. This determines the elements of $\underset{\sim}{D}$:

$$\underset{\sim}{D}_{11} = 2\eta\left(1 - \frac{\sigma_0^2}{\sigma_0^2 + \tau_0^2}\right); \quad \underset{\sim}{D}_{22} = \eta\left(1 - \frac{\tau_0^2}{\sigma_0^2 + \tau_0^2}\right);$$

$$\underset{\sim}{D}_{12} = 2\,\underset{\sim}{D}_{21} = \frac{-\eta\,\sigma_0\tau_0}{\sigma_0^2 + \tau_0^2} \tag{3.33 a,b,c}$$

where $\sigma_0 = \frac{1}{2}(\sigma_{xx} - \sigma_{zz})$ at the onset of plasticity (Sarne, 1974). This new constitutive law is incorporated into the stiffness matrix for the next load step according to the equation:

$$\underset{\sim}{K}_{ij} = \iint_A \Big[2\,D_{11}\left(\frac{\partial^2\phi_i}{\partial x\,\partial z}\right)\left(\frac{\partial^2\phi_j}{\partial x\,\partial z}\right) +$$

$$2 D_{21} \left(\frac{\partial^2 \phi_i}{\partial x \partial z}\right)\left(\frac{\partial^2 \phi_j}{\partial z^2} - \frac{\partial^2 \phi_j}{\partial x^2}\right) +$$

$$D_{12} \left(\frac{\partial^2 \phi_i}{\partial z^2} - \frac{\partial^2 \phi_i}{\partial x^2}\right)\left(\frac{\partial^2 \phi_j}{\partial x \partial z}\right) +$$

$$D_{22} \left(\frac{\partial^2 \phi_i}{\partial z^2} - \frac{\partial^2 \phi_i}{\partial x^2}\right)\left(\frac{\partial^2 \phi_j}{\partial z^2} - \frac{\partial^2 \phi_j}{\partial x^2}\right)\Bigg] dA$$

(3.34)

A pitfall of the initial stress technique is the possibility of non-converging oscillations caused by excessively low viscosities. As shown in Figure 3.5, the sense of strain may reverse at each iteration and cause a large unbalanced stress. This is prevented by monitoring each element for change of shearing direction and increasing its viscosity to the value of equation (3.24) if necessary. The necessity for these corrections undermines the main advantage of this technique, which is that the stiffness matrix can be used over and over without recomputing it. If there are no oscillations, then the stiffness matrix $\underset{\sim}{K}$ is factored into upper and lower triangular band matrices which are transposes of each other by the technique of Rutischauser (1966). Thereafter, each solution requires only NM operations and is faster by the factor of the bandwidth. On the other hand, a large number of load steps may be required if the extent of faulting plasticity is great.

Both techniques have been employed for calculations presented in this thesis. The initial stress method was programmed first. It was found to require as much as 440 iterations divided among 18 load steps (4.8 minutes of IBM 370/165 computer time) to solve a representative problem with stress errors of less than 10% at 80-95% of the test points. Much of the time is spent in recalculating the stiffness matrix to avoid oscillations (Figure 3.6), and the occurrence of plasticity in the higher load levels greatly slows convergence. Surprisingly, the variable viscosity method was found to be much more efficient. On the same problem, it required only 14 iterations (26 seconds computer time) to obtain an answer with a root mean square stress error of 2% and a maximum error of 15%. Thus this method is strongly preferred, and was used to calculate most of the Zagros models in Chapter 4 and all of the Himalayan models in Chapter 5. As discussed in the next section, it is also more accurate than the initial stress method.

3.5 Error Estimates and Test Problems

A successful solution is one which satisfies boundary conditions, incompressibility, stress equilibrium, and the flow law at every point. With finite elements the boundary conditions present no problem. Force boundary conditions are exactly represented and velocity boundary conditions that are piecewise-quadratic or lower in order can be input with proper values of the velocity and stream functions at nodes. The only fundamental difficulty is in the treatment of curved boundaries, where the accuracy is limited by the number of nodes per unit length one can afford to use in representing them. By a slight modification of the element discussed here it could be made "isoparametric" (Zienkiewicz, 1971) and used to represent boundaries of up to cubic shape. Fortunately, the boundaries of geologic problems are usually arbitrary and can be taken as flat. Likewise, there is little justification for a more sophisticated treatment of internal geologic contacts given our present state of knowledge.

Incompressibility is exactly satisfied at every point through the use of a stream function. The incompatability of the elements is in the shear mode, so incompressibility is not violated at element boundaries.

The question of how well the technique satisfies the stress equilibrium equation (3.6) has two parts. The first is, to what extent can a finite number of integral tests represent the original differential equation? Corresponding to every degree of freedom and unknown in the problem there is an equation (3.20) which can be interpreted as requiring minimization of the dissipation rate with respect to that degree of freedom (R.B. Bird, 1960). These constraints are designed to be non-redundant and to eliminate long wavelength errors in stress equilibrium, but an infinite-dimensional family of short wavelength errors is still possible. If a compatible linear-strain element had been used, it would be possible to prove that the dissipation of the error (which is equal to the error in the dissipation) is reduced at a rate proportional to the fourth power of the grid interval in problems with no stress singularities (Strang and Fix, 1972). However, with an incompatible element no such general theory is available and it will be necessary to judge performance from test problems.

Available analytic solutions for stress-cubic incompressible materials seem to be restricted to the cases of constant strain, radially-varying strain, and strain around a crack-tip (all in a homogeneous medium). Inhomogeneous material is only solvable in the trivial case of one-dimensional property variations and constant stress. The first three cases are considered here, and the effect of inhomogeneity is discussed below.

Constant strain and constant stress are the appropriate solutions in the problem of compression of a uniform rectangular body on a frictionless foundation by a constant velocity or pressure boundary condition on the side opposite the foundation. The nature of the solution is unaffected by all material properties except the Poisson's ratio, which in this case is one-half, implying horizontal extension equal to the vertical shortening. A square grid of 25 nodes and 32 elements was employed: the lack of a characteristic dimension in the problem makes fine subdivisions meaningless. A boundary condition of constant pressure on one side was applied for both linear viscous and nonlinear (n=3) rheologies.

In the viscous case, the stream function error was no more than 1.7%. This agrees with results of Bazeley _et al_. (1965), who found 1.5% error in displacement for plate bending problems using this element. The derivitive of the stream function (velocity) contains larger errors up to 4.0%. Stress intensity in the sense of element averages had a root-mean-square error of 2.8% and a maximum error of 4.6%. Stress direction was accurate to a RMS error of 1.8° and a maximum of 3.8°. In the nonlinear case stress intensity was actually more uniform, with 2.5% RMS error and 3.5% maximum. Stress directions were still good: 2.2° RMS error and 4.3° maximum. A direct consequence of the nonlinearity is that with this comparable stress field the errors in stream function (to 5.3%) and velocity (to 7.6%) were amplified.

The same problem solved with conforming elements would have given the exact solution within the limits of computer representation. Thus this case illustrates the effect of the non-conforming element on accuracy. It is believed that the larger errors in problems to follow are not the result of inter-element discontinuities, but of the use of a finite number of linear-strain elements to represent strain fields of large curvatures. In such a situation the discontinuities may even allow a better fit by giving more freedom to the grid.

The second problem considered was the plane-strain compression of a hollow cylinder consisting of incompressible material. This is another case in which velocities are independent of material properties. Radial symmetry requires that $V_\theta \equiv 0$ and incompressibility implies

$$\frac{\partial V_r}{\partial r} + \frac{V_r}{r} = 0 \tag{3.35}$$

leading directly to

$$V_r = \frac{c}{r} \tag{3.36}$$

and

$$(\dot{e}_{\theta\theta} - \dot{e}_{rr}) = \frac{2c}{r^2} \tag{3.37}$$

Deviatoric stress should vary as r^{-2} in the viscous case and $r^{-2/3}$ in the case of n=3 creep. A radius ratio of 6:1 was chosen to allow a large stress variation. Because the stream function does not allow for internal compressible voids it is necessary to make a cut in the cylinder to let the central vacancy "escape." Having done this, we might as well restrict the model to one quadrant because of symmetry. The grid used is shown in Figure 3.7. The problem was solved with displacement boundary conditions instead of pressures this time. This is forced upon us by a peculiarity of the element used: in order to represent an impermeable boundary the stream function value must be fixed at the nodes as well as the velocities. In order to render both radial cut sides of the grid in Figure 3.7 impermeable we are forced to fix the net flux between the sides.

The problem was solved for three different rheologies: viscous, viscous-plastic, and cubic creep (n=3). In each case the velocities of all nodes were accurate within 1.6%, so they have not been plotted. Instead we show in Figure 3.8 the stress results from central elements along the line x=z. Each plotted point represents a single element average:

$$\langle \sigma_{xx} - \sigma_{zz} \rangle = \frac{\iint_{A_{el}} (\sigma_{xx} - \sigma_{zz})\, dA}{A_{el}} \; ; \; \langle \tau_{xz} \rangle = \frac{\iint_{A_{el}} \tau_{xz}\, dA}{A_{el}} \qquad (3.38\ \mathrm{a,b})$$

and is plotted at the radius corresponding to the element centroid. A good fit is obtained in each case, and the largest errors (up to 17% in the viscous case and 10% in the nonlinear case) occur as a result of tremendous stress variations within single elements. An oscillation is seen around the correct stress values. This is common with finite elements, and results from the fact that the elements alternate between inward and outward-pointing orientations in the grid. If one performs a further smoothing operation by averaging the radius and stress values between adjacent elements, the errors are reduced to 5% in the viscous case and 3% in the nonlinear.

One of the few known solutions in which velocities do depend on rheology is that of the extreme near-field around the tip of a static plane-strain tension crack in an incompressible medium. Stress solutions for both the linear and cube-root strain hardening laws are given by McClintock (1971), but unfortunately the displacement field has not been obtained. By substituting strain rate for strain we can use the stress solutions for the similar problem of slow crack opening in a creeping material.

The solution was performed using the grid in Figure 3.9. Because the displacements are known only at infinite distance we must restrict the model to a small dimensionless region around the crack tip and use force boundary conditions obtained from the analytic stress solutions. By looking at stresses at a smaller radius closer to the singularity we can test whether the applied stresses have been properly amplified and distributed. The results from elements in the third ring are shown in Figure 3.10.

Again there is a fair solution containing oscillations around the correct value in each case. Only the stress component $\tau_{r\theta}$ in the viscous case shows a consistent error, by underestimating this component some 20%. The fact that tension along the walls of the crack does not vanish in the non-linear case is correctly represented by the solution. As in the cylinder problem, there is no noticable loss of accuracy in going from the linear to nonlinear flow law. This seems to confirm that the largest part of the errors is due to the limited flexibility of the element in a rapidly-varying strain field.

In the geologic problems of interest there are no such singularities of stress. Therefore in the limit of very fine grids the errors of stress will be reduced to those caused by element incompatibility, some 4%. In realistic applications with coarse grids oscillations can be expected and reduced by averaging across element pairs. The maximum error contained in the oscillations will be about 20%, but the general stress level will be correctly represented within 10%.

The second requirement for maintaining stress equilibrium is that numerical errors in the solution of the approximate form (3.21) be kept small. These errors are of the order of 10^{-5} in the test problems just presented, but they are aggravated in geologic problems by the large stiffness contrasts between different parts of the model. In addition, these numerical errors may be accumulated when using the initial stress technique because incremental stress solutions are summed.

It is possible to obtain an estimate of these errors by using the calculated solution stresses to evaluate

$$\epsilon_i = \iint_A \left[(\sigma_{xx} - \sigma_{zz}) \frac{\partial^2 \phi_i}{\partial x \partial z} + \tau_{xz} \left(\frac{\partial^2 \phi_i}{\partial z^2} - \frac{\partial^2 \phi_i}{\partial x^2} \right) \right] dA - \vec{F}_i \tag{3.39}$$

for each of the degrees of freedom and its corresponding stream function component $\phi_i\ (x,z)$. According to equation (3.22) the quantities ϵ_i should all be zero, and the actual values can be interpreted as averages over several adjacent elements of the stress errors introduced by numerical round-off errors. The most convenient values for this purpose are the ϵ_i corresponding to stream function variables. The stream function component ϕ_i associated with a stream function unknown has a value of 1 cm^2/sec at that node and zero at others (for ϕ_i associated with velocity unknowns the amplitude depends on grid dimensions). Therefore the area integral of its second derivitive is also of order one independent of grid dimension and allows us to interpret the resulting ϵ_i (actually in units of dynes/sec) as if they were in stress units of dynes/cm^2.

Since the prevailing stress level, grid dimension, and flow amplitude were approximately equal in all geologic calculations performed, these quantities ϵ_i were found to be most closely related to the magnitude of the largest viscosity allowed in any element. This is plausible because in the numerical representation of an infinitely stiff element some small artificial strain would creep into the flow solution. The size of the resulting stress error would depend on the viscosity by which such small strain errors were multiplied. Using the variable viscosity method to solve a typical geologic problem, a maximum viscosity of 1.3×10^{24} poise resulted in a maximum stress error of 5 bars (0.5%). A higher limit of 4.3×10^{25} poise gave a maximum stress error of 23 bars (2.9%). Regardless of how cold the rocks may be, larger viscosities are not physically realistic because we have neglected elastic effects that might give an effective viscosity limit of this order (section 3.2). The numerical stress errors in the variable viscosity method are thus easily kept within the accuracy limitations set by element incompatability.

Obtaining this degree of numerical accuracy requires some care. Both programs had a feature which checked for unacceptably high viscosities in each step and reduced them by increasing rock temperature by an appropriate amount (in that element only). All stiffness matrices were assembled and solved in double-precision arithmetic (16 decimal digits). And finally, the stream function variables were rescaled to keep them the same order of magnitude as the velocity variables. If this is not done, adjacent diagonal elements of the stiffness matrix will differ by a factor of element area in square centimeters (up to 10^{13}) causing severe problems.

Unfortunately, the error test just described was not programmed into the early version of the program used for initial-stress method calculations. Thus the effect of iteration on error accumulation is not known. An upper limit can be estimated by assuming that the stress field increment in each iteration had the same magnitude as the final solution. In that case a random-walk model would predict that errors would be amplified by the square root of the number of iterations, or about 20 times. This would imply stress errors of as much as 460 bars. However, this assumption is not true; stresses have low amplitudes in the first load steps, and they have decaying amplitudes within each load step as convergence is approached. Realistically, we estimate that net emplification is by ten or less. Because this would still imply substantial error, most of the calculations that were performed by the initial-stress method have been repeated using variable viscosity.

Convergence toward the flow law is never exact, but can be pursued to arbitrary precision if there is sufficient computer time. Furthermore, the extent of the error can be exactly determined at any point desired. With the initial stress method, computing was usually halted when the stress error fell below 10% at 80-95% of the test points (the seven integration points in each element). With the variable viscosity method, it is economical to force convergence to a maximum error of 20% and a root-mean-square error of 2-5%. Most of the large percent errors occur at points of low ambient stress, while a majority of test points have errors of less than 3%. This is shown by the fact that a typical model (variable viscosity) had an RMS stress residual of 18 bars and a maximum of 160 bars when the prevailing stress was as much as 1,500 bars. Thus convergence at the test points can easily be made to fall within the experimental uncertainties of the flow law.

A more fundamental problem is the impossibility of obtaining convergence at more than 7 points in each element, particularly when there are temperature gradients across them. Although the exponential variation of viscosity with temperature could be incorporated into the stiffness matrix by a more sophisticated integration scheme, no two-dimensional element in the literature is able to reproduce the exponential variation of strain through space (at constant stress) that actually occurs. The element used here, with its linear variation of strain, is sometimes stiffer than the actual rocks in the presence of temperature variations. A test was made of a granite shear zone one element thick, in which the plane of maximum shear was parallel to the isotherms. When the temperature contrast was 100°C (range 300-400°), implying a rock stiffness range of a factor of 6.9, the elements were found to be 33% stiffer than the analytic solution. When the temperature contrast was doubled (300-500°C), implying a stiffness range of 28.8 times, the element was found to be 2.2 times as strong as the actual aggregate. This is not a fundamental error because it can be removed by using a finer grid for the problem. But it is presently the most important factor limiting our ability to represent the real Earth in economical models. It dictates that in problems involving the entire lithosphere it will not be feasible to determine the strength of the rocks to better than a factor of two.

Chapter 3 - Figure Captions

Figure 3.1 Profile of strength of continental lithosphere with depth at two orogenic strain rates equal to 5 cm/year velocity change over 100 km and 10 km respectively. Down to 20 km depth the stress limit is set by faulting with a hydrostatic pore-pressure correction. Creep of quartz occurs in lower crust down to Moho at 33 km. Mantle olivine creeps plasticly or according to a temperature-dependent nonlinear law. Geotherm of Table 2.1 was used in calculation of strength.

Figure 3.2 Schematic diagram showing the degrees of freedom of the linear strain triangular finite element used in all calculations. Solid dots are nodes. Open circles show the seven numerical integration points for the initial stress method; triangles are integration points in the variable viscosity method.

Figure 3.3 Convergence of variable viscosity iteration scheme. Curved line represents a cubic-creep flow law for simple uniaxial extension, and solid dot is the analytic solution with a stress boundary condition. Straight lines are successive linear viscous approximations with different viscosities.

Figure 3.4 Convergence of initial stress iteration scheme within one load step. Curved line is the cubic-creep flow law for simple uniaxial extension. Solid lines are increments of stress and strain rate in each iteration. Dotted lines show stress corrections which create unbalanced stresses to drive the next step.

Figure 3.5 Schematic flow chart of computations with the multiple-load-step, initial stress method. Diamond-shaped boxes represent decision points. Program begins with stresses scaled down by 10^{-7} (10^{-21} for velocity) to avoid pasticity in the first step.

Figure 3.6 An example of a non-converging oscillation possible with the initial-stress technique. Program is attempting to get from point A to B on the flow law curve, but selects too low a viscosity. Oscillation will stop growing when viscosity exceeds the slope of the flow law at each extreme.

Figure 3.7 Grid of finite elements representing a hollow incompressible cylinder in cross-section. Because of symmetry only one quadrant is modelled.

Figure 3.8 Comparison of analytic and finite-element stress solutions for the hollow incompressible cylinder with three different rheologies. Analytic solutions shown by curves; other symbols represent averaged stresses within elements at different radii.

Figure 3.9 Finite element grid used for problem of a plane strain tension crack in an incompressible material. Because of symmetry only one side of one end of the crack is modelled.

Figure 3.10 Comparison of analytic and finite element stress solutions for the plane strain tension crack tip with viscous and cubic (n=3) creep rheologies. Curves are analytic solutions from McClintock (1971) for different values of θ at constant radius from the cracktip. Symbols are average element stresses from elements 19-31 in the middle ring of Figure 3.9.


SHEAR STRENGTH, IN BARS
DEPTH, IN KM
0
20
40
60
80
100
10^1
10^2
10^3
10^4
$\dot{e} = 1.6 \times 10^{-13}$
$\dot{e} = 1.6 \times 10^{-14}$

Fig. 3.1

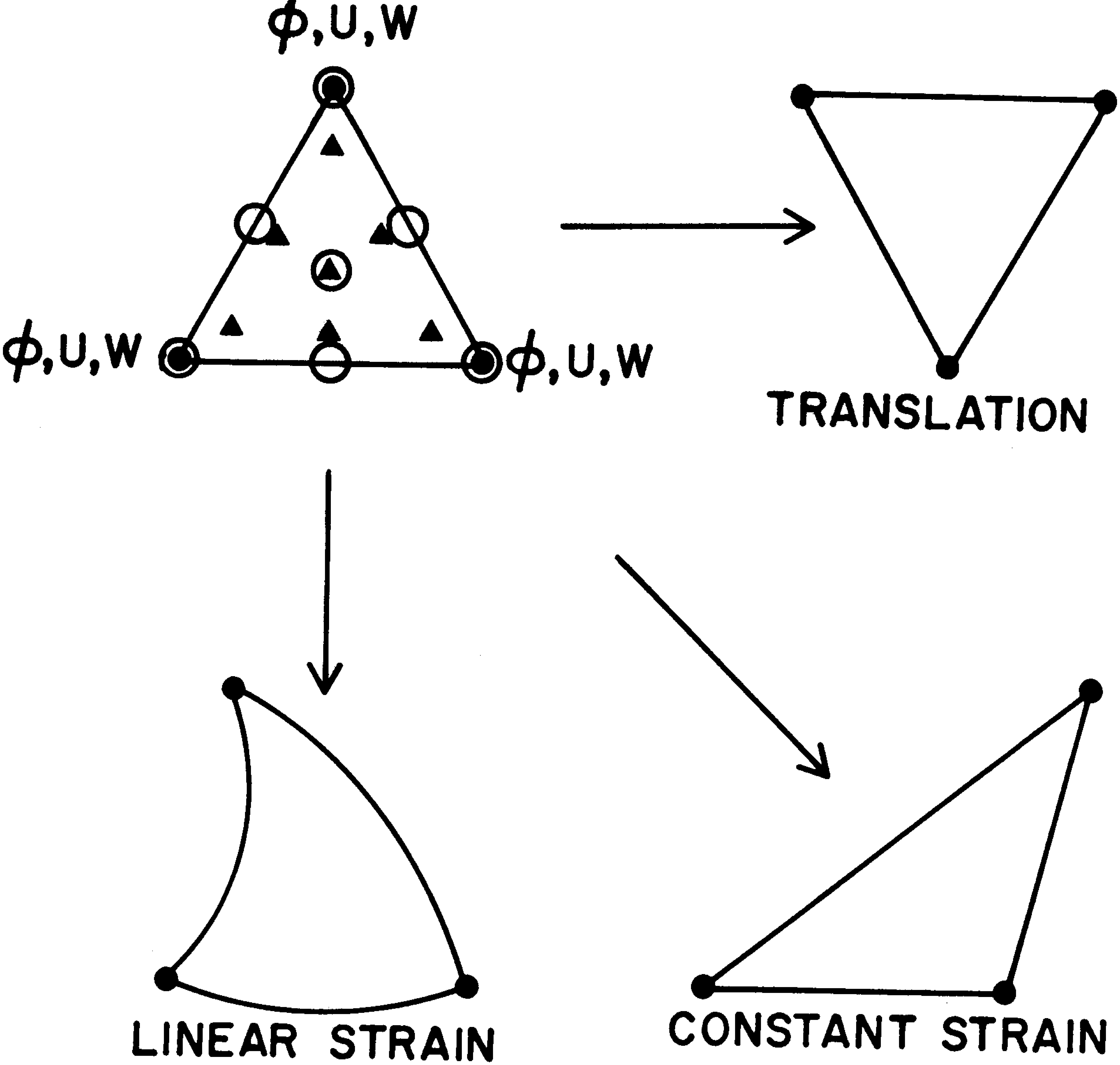


Fig. 3.2

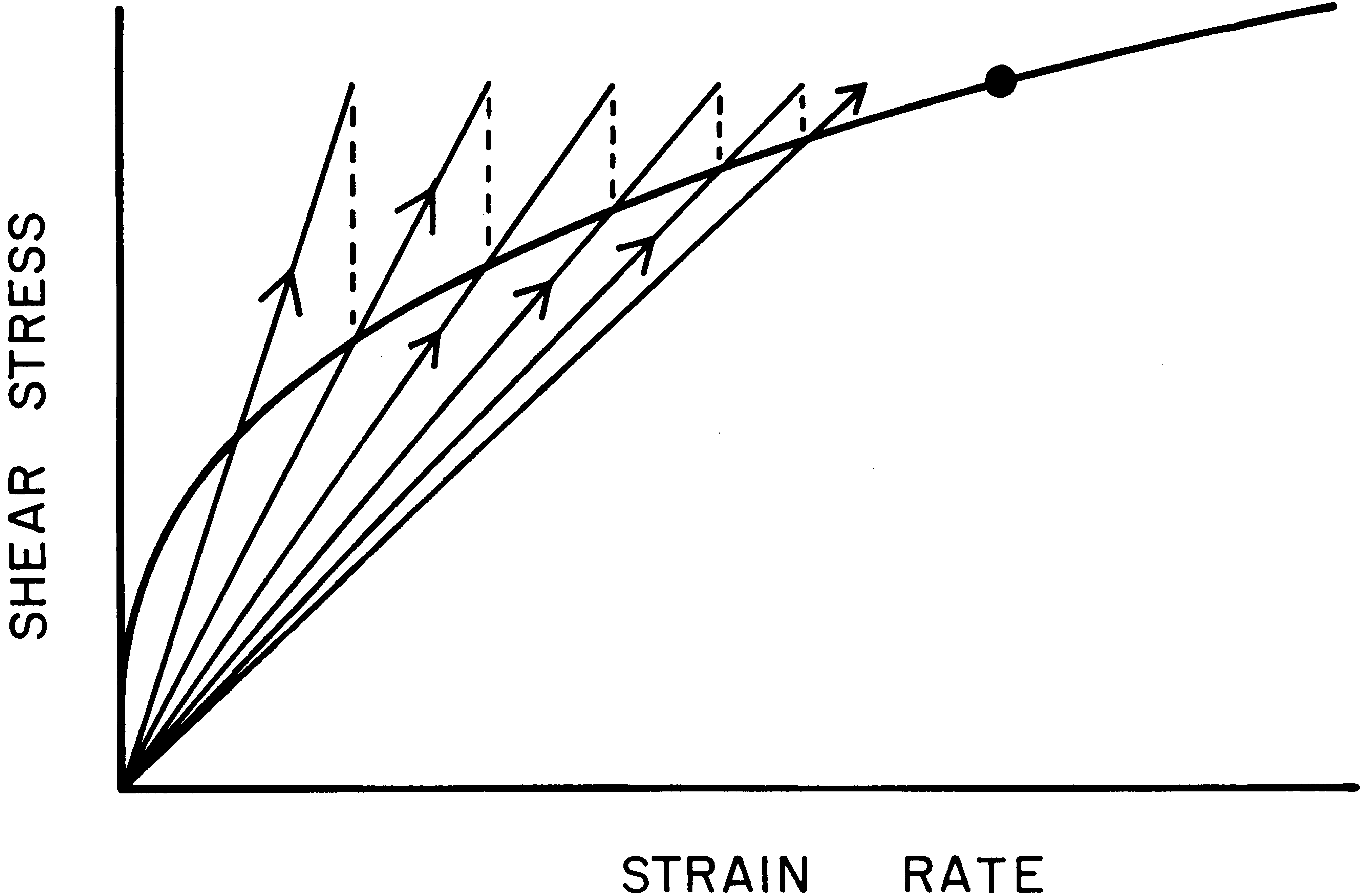


Fig. 3.3

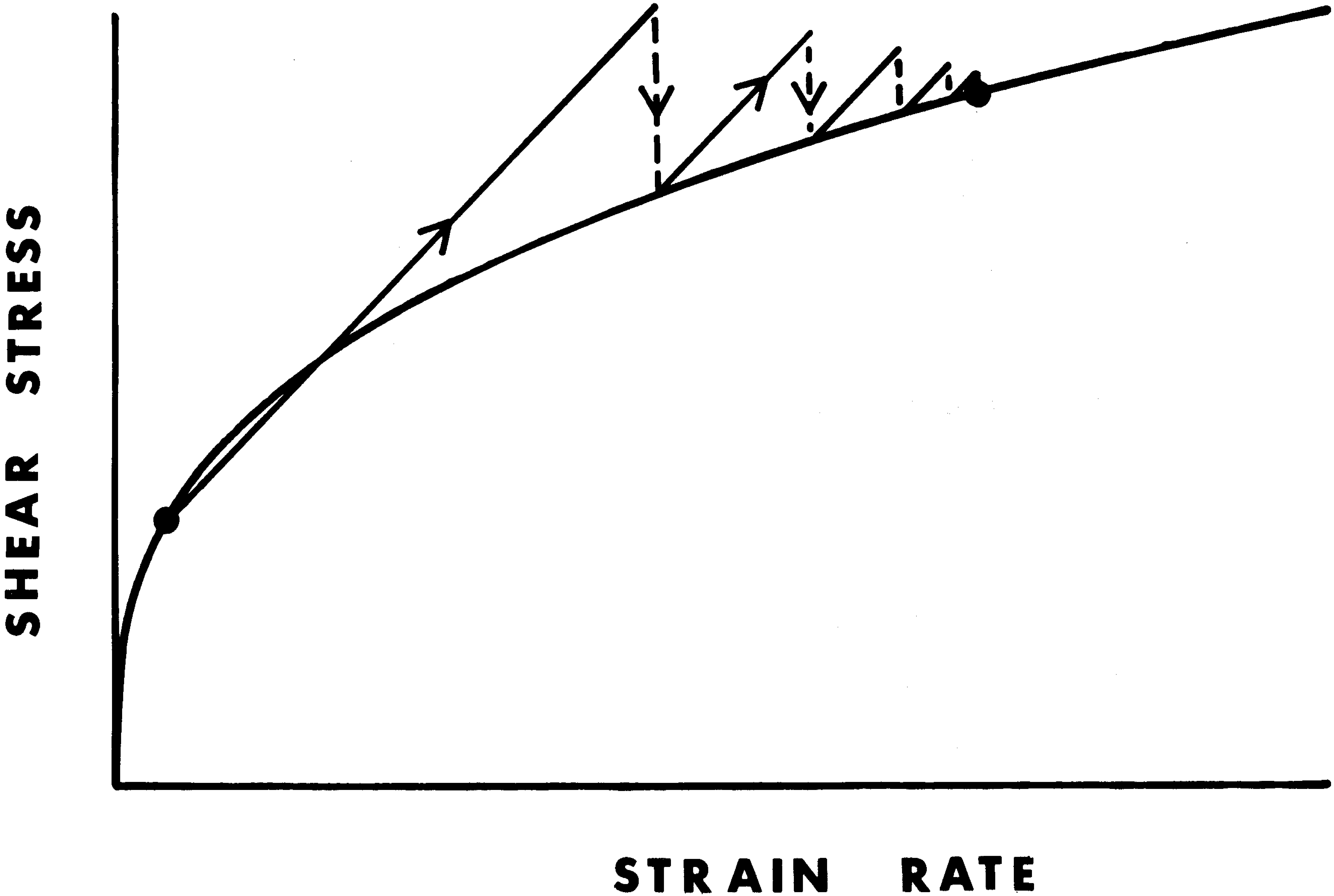


Fig. 3.4

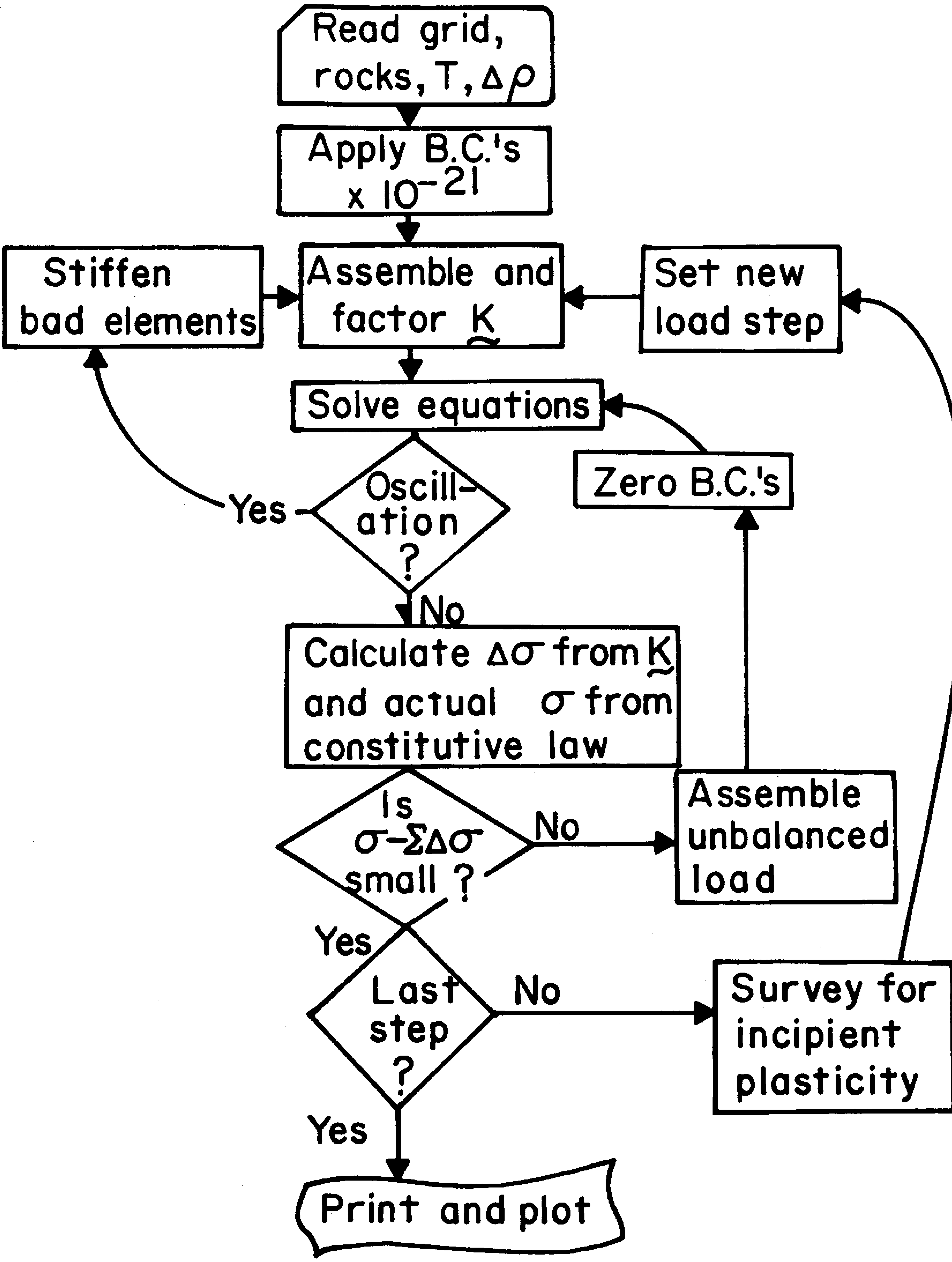

Read grid, rocks, T, $\Delta\rho$
Apply B.C.'s x 10^{-21}
Stiffen bad elements
Assemble and factor $\underset{\sim}{K}$
Set new load step
Solve equations
Zero B.C.'s
Yes
Oscill-ation ?
No
Calculate $\Delta\sigma$ from $\underset{\sim}{K}$ and actual σ from constitutive law
Is $\sigma - \Sigma\Delta\sigma$ small ?
No
Assemble unbalanced load
Yes
Last step ?
No
Survey for incipient plasticity
Yes
Print and plot

Fig. 3.5

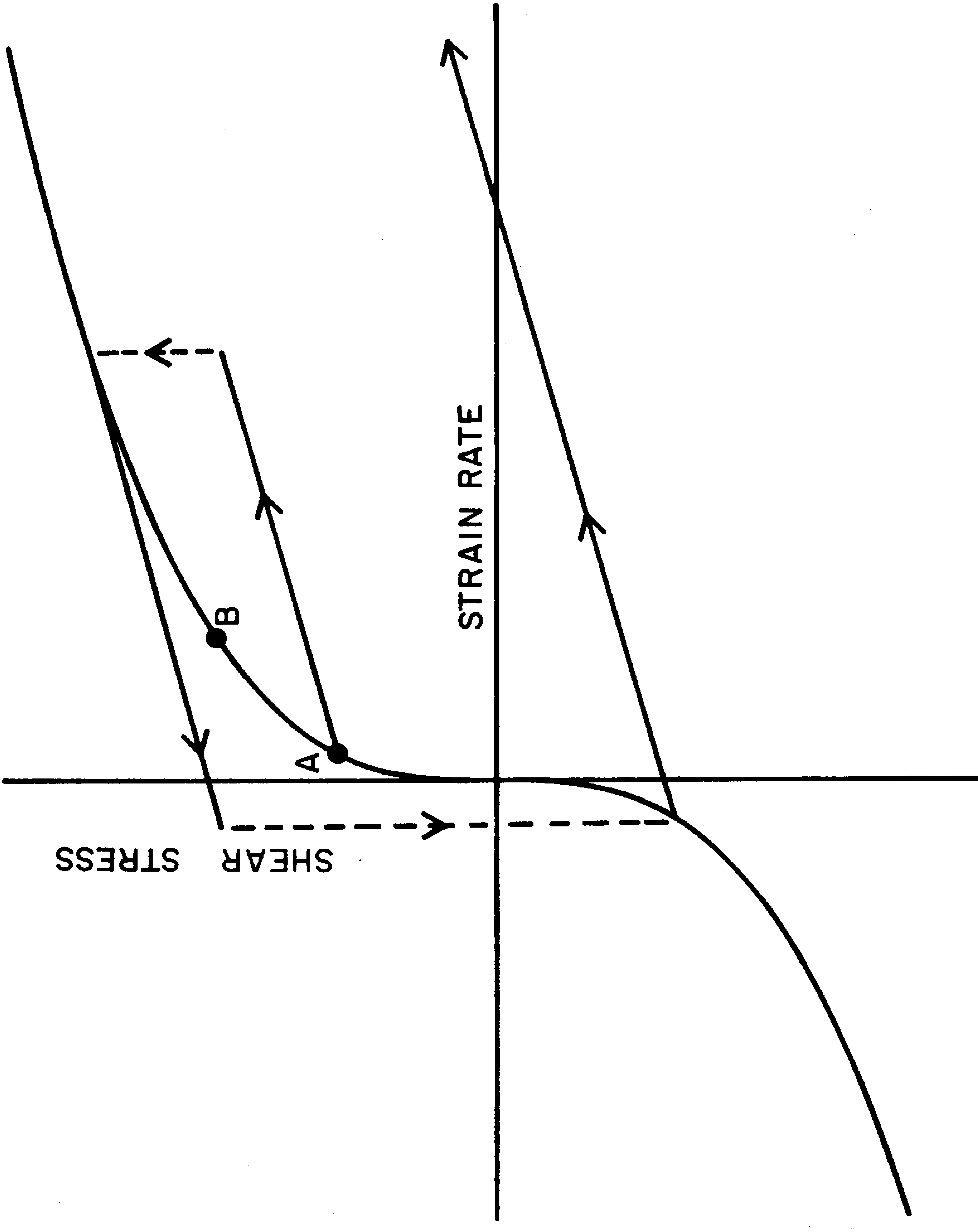


Fig. 3.6

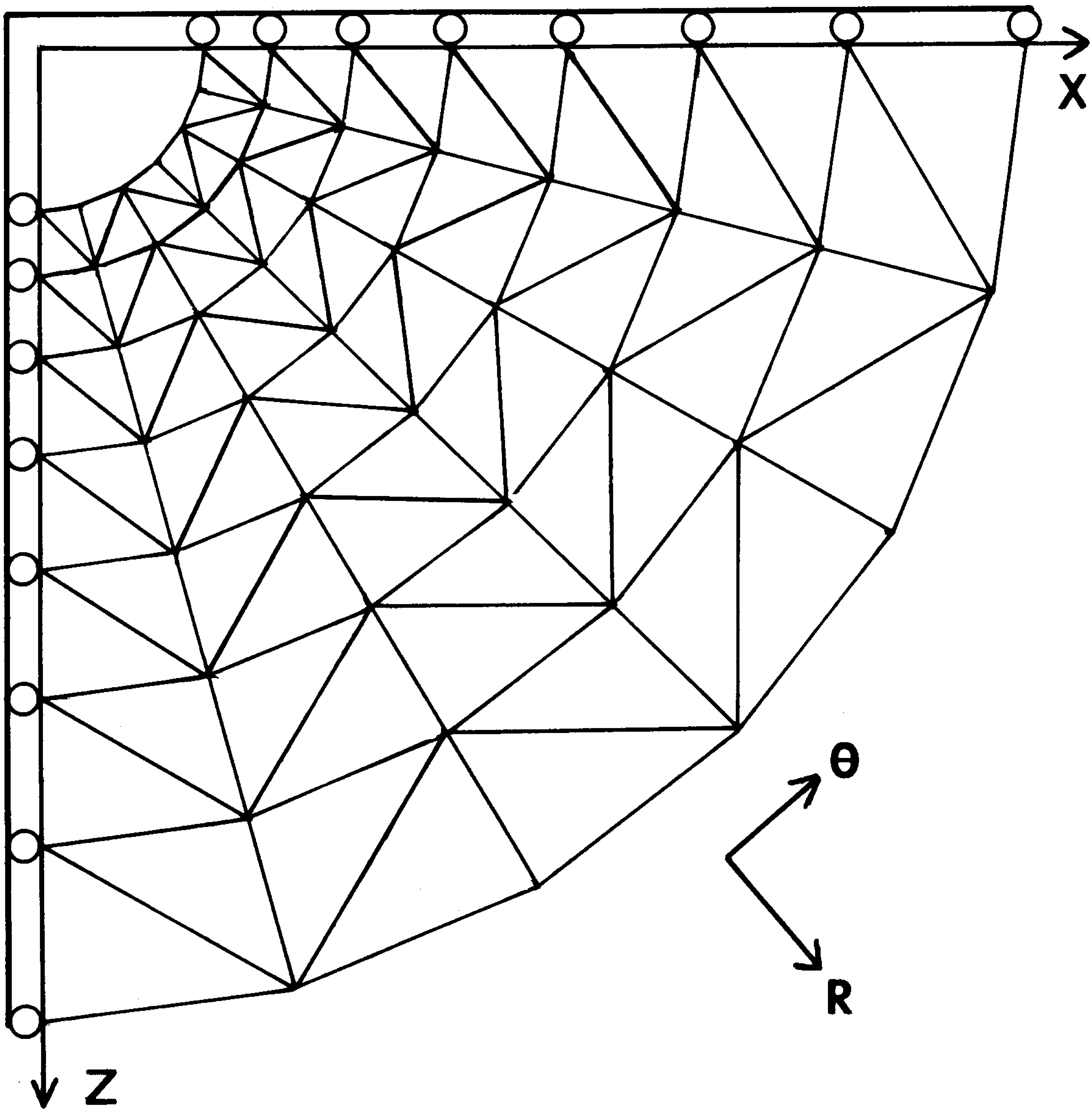


Fig. 3.7

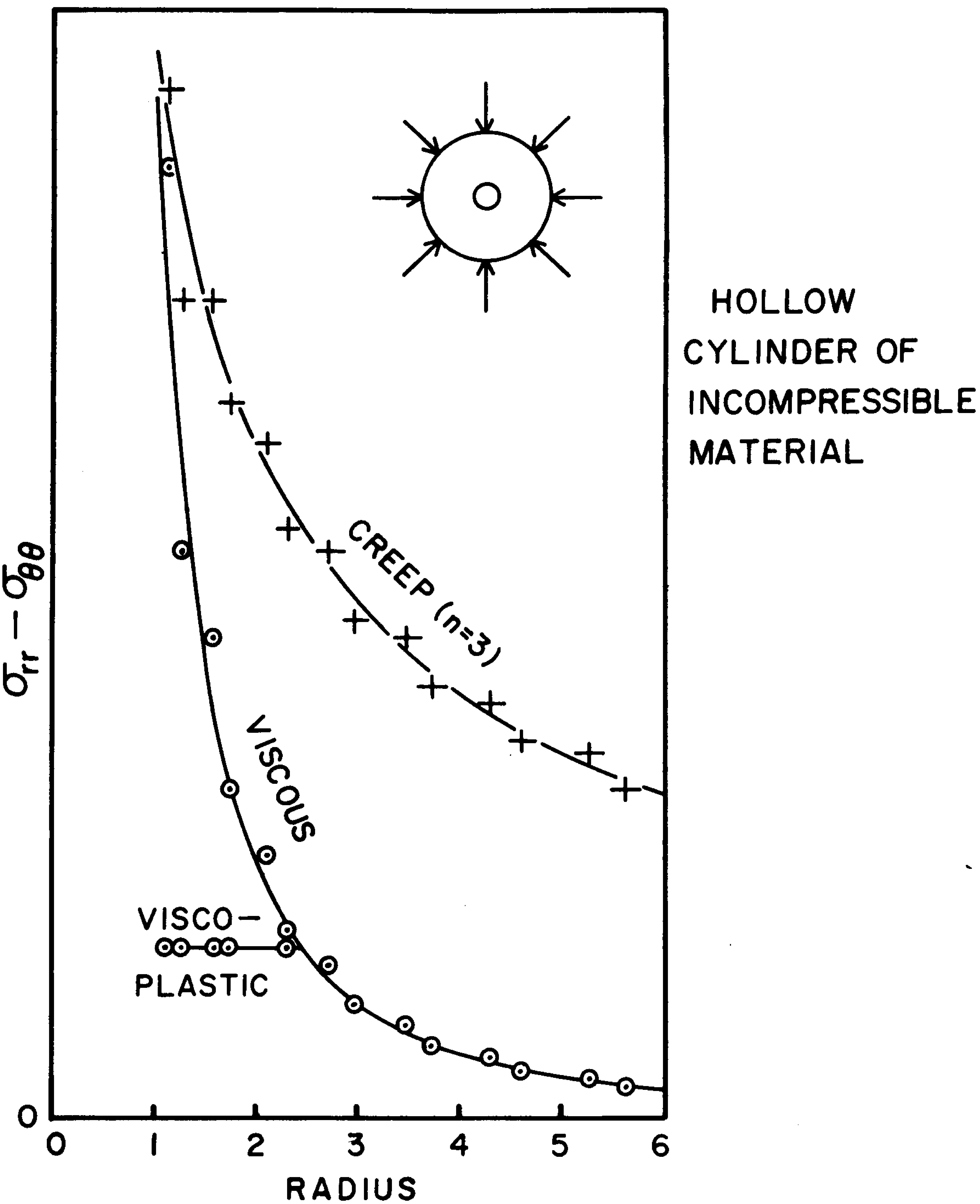

HOLLOW
CYLINDER OF
INCOMPRESSIBLE
MATERIAL
CREEP (n=3)
VISCOUS
VISCO-
PLASTIC
$\sigma_{rr} - \sigma_{\theta\theta}$
RADIUS
0
0
1
2
3
4
5
6

Fig. 3.8

Fig. 3.9

PLANE STRAIN TENSION CRACK TIP

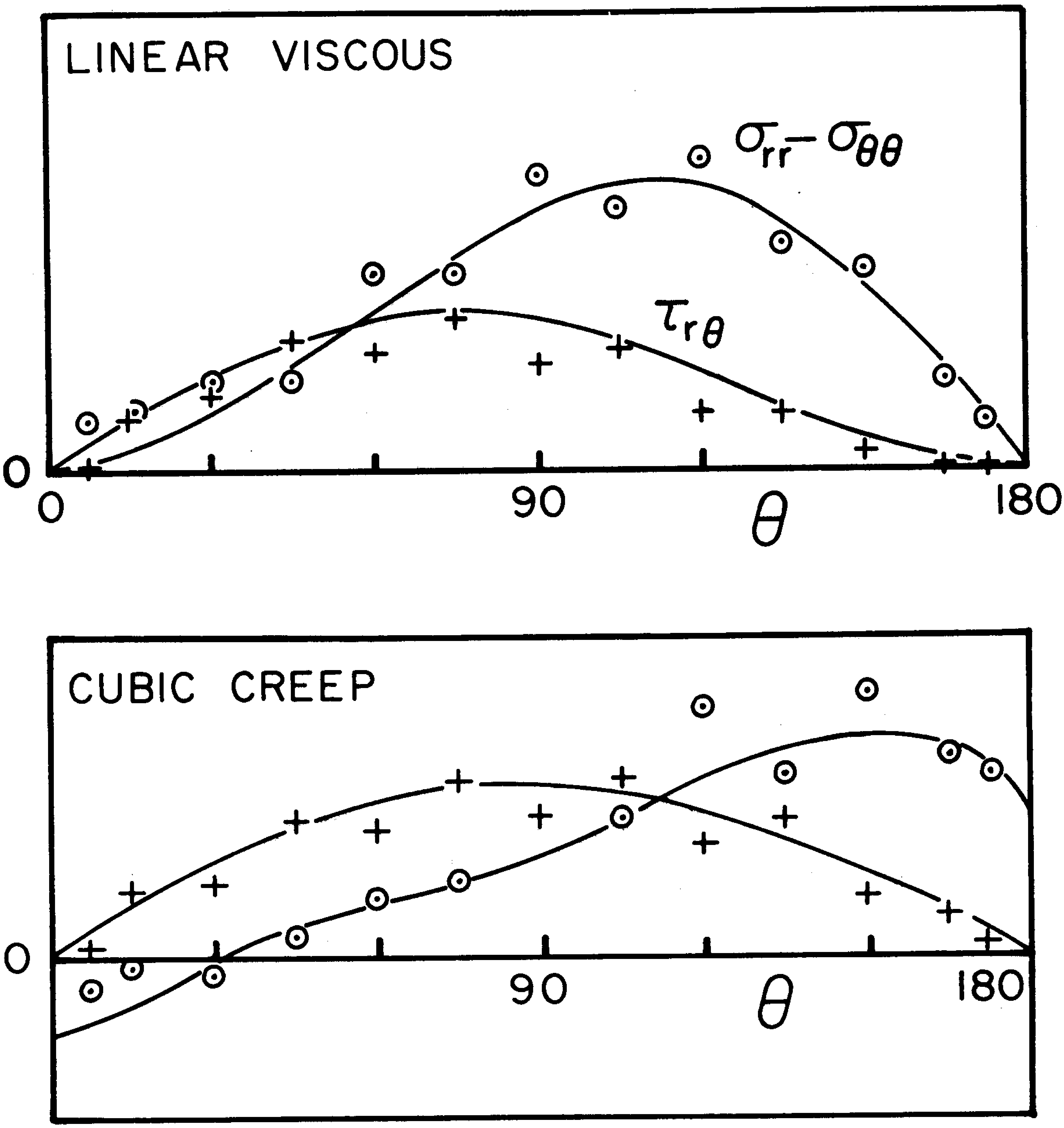


Fig. 3.10

CHAPTER 4: THE ZAGROS MOUNTAINS: DETACHMENT OF CRUST AND SLAB

The Zagros Mts. are forming at a plate boundary between Africa and Eurasia. Convergence and subduction of the Tethys at this line have gone on since the late Jurassic (Dewey *et al.*, 1973). During the last 9 m.y. these plates have converged at about 2.7 cm/yr about a finite rotation pole at 31°N, 35°W. Since the Arabian shield split away from Africa the rate has been increased. Interpreting the results of Phillips (1970) in terms of 2 cm/yr opening of the Red Sea in the last 1.5 m.y., a recent Zagros convergence rate of 4.7 cm/yr can be obtained. The pole for this motion at 25°N, 37°W is well aligned with the strike of the range.

This is an ideal place to study the early stages of continental convergence, because we have simple plain-strain deformation of horizontally-layered rocks at near-equilibrium temperatures. In the first part of the chapter the geology is summarized and the collision is dated. Then gravity, heat flow, and seismic evidence are used to show that crustal shortening and thickening is underway. This process is modelled with finite elements, which reveal that only models in which the oceanic slab is assumed to be mechanically unimportant (*i.e.*, weak or detached) are successful. The models also give upper limits on the stresses at which rocks creep and fault in the continental crust- rather lower limits than laboratory experiments predict.

4.1 Geologic Outline and Tectonic Theories of the Zagros

The first evidence of Tertiary orogenic events is found not in the range itself but to the northeast in central Iran. This is the occurrence over a wide area of voluminous andesitic lavas with minor associated rhyolites, dacites, and basalts. According to Azizbekov (1971) this volcanism was most active in the Eocene, when it cut a wide swath across northwest Iran, dividing into a southern Kāshān-Muriān and a northern Elburz branch. Dewey _et al_. (1973) state that these volcanics had Cretaceous antecedents which Azizbekov does not mention. In any case, volcanism has continued in these locations to the Quaternary, because there are large volcanic cones near Tabriz, Tazd, Tehran, and Khāsh.

The long-term stability, composition, and linear pattern of this volcanism suggested an analogy with volcanic island arcs. Crawford (1972) was perhaps the first to propose that these eruptions were the result of subduction of unspecified material from the southwest beneath central Iran at the line of the Zagros Thrust (Figure 4.1). Dewey _et al_. (1973) agreed, and used the relative positions of Africa and Europe (as constrained by Atlantic paleomagnetic anomalies) to suggest that this underthrusting of the Tethys Ocean has continued since late Jurassic and consumed some 3,000 km of oceanic lithosphere.

The existence of a plate suture at the Zagros Crush Zone was earlier suggested by Dewey and Bird (1970) on the basis of its ultramafic rocks associated with deep-water sediments in a tectonic melange. With the exception of some olivine basalts erupted near Zanjan these are the only basic rocks of the Zagros region. In a comprehensive article Haynes and McQuillan (1974) confirmed that the Crush Zone is a former trench, containing serpentenized harzburgites, pyroxenites, and dunites associated with chromite in a sheared matrix of colored marls, conglomerates, exotic limestone blocks, and radiolarian cherts. Hallam (1974) reported that the direct contacts of the ultramafic bodies with very silicious cherts (containing no turbidites) are non-intrusive. This is consistent with the theory of ophiolite formation during sea-floor spreading in the deep ocean with later tectonic emplacement of the ultramafics among turbidite trench sediments (Dewey and Bird, 1971).

This Crush Zone marks the northeastern limit of the Zagros range, and a complete discontinuity of geology. To the southwest there is a wide 200 km belt of thick shallow-marine sediments, predominantly limestones and marls with occasional sandstone and evaporite. This conformable sequence was deposited on a subsiding basement in an area that has been free of tectonism and intrusion at least since the Cambrian (Stöcklin, 1968). Wells drilled to depths of 4.8 km produce petroleum from Mesozoic formations, indicating a tremendous thickness of sediments and a quiescent Tertiary history (Stöcklin and Nabavi, 1973). Estimates of the total thickness are commonly 7 km (Lees, 1952) to 12 km (Falcon, 1969). This "geosynclinal" region of thick shallow water sediments was very recently the continental shelf of the northeastern margin of the Arabian-African plate (Dewey and Bird, 1970; Haynes and McQuillan, 1974), facing an open arm of the Tethys Ocean.

The continental collision between Arabia and Iran which began the deformation of this geosyncline was most likely a late Pliocene event. The geomorphology of the Zagros is of the youthful first-order type in which anti-clinal ridges have been only slightly affected by erosion. Deformed marine sediments of Pliocene age are found in the Imbricate Zone (Wells, 1969; Haynes and McQuillan, 1974) giving an upper limit of 5 m.y. for the collision (Berggren, 1972).

The character of the deformation of the continental shelf is uniform from the straits of Hormoz to the syntaxial bend in southeastern Turkey. Its principal component is a NE-SW horizontal compression directed perpendicular to the Crush Zone. The intensity of the resulting deformation de-creases with distance from the suture, beginning as intense folding with steeply (70° NE) dipping thrust faults and passing into a sequence of about a dozen open cylindrical folds which decrease in amplitude and die out below alluvial cover. This alluvial cover of Pliocene-Quaternary age is derived from the rising folds and deposited in the young downwarping axis of the Persian Gulf-Mesopotamian Trough. From the plate suture to the last detectable deformation the range is 160-300 km in width (Figure 4.1).

This regional deformation is modified by local secondary elements. Wrench faulting is common, particularly in the vicinity opposite the Oman Peninsula, where stiff Pre-Cambrian structural elements within the Arabian shield are apparently causing unequal strains. Salt domes have risen diapirically through the section from their original position in the basal Cambrian or Pre-Cambrian. In the region 48°-52° East the wavelength of folding is markedly decreased, which may be the result of evaporite layers at shallower depths in the section, (Paul Tapponnier, personal communication, 1975).

Because of the well-known weakness and mobilism of halite (Heard, 1972) there is likely to be some degree of *decollement* between the sediments and the basement, as well as between different competent layers of sediment. Starting from this premise, geologists have reached very different conclusions about the origin of deformation. One school of thought is that the deformation is entirely restricted to the sediments, which slide and fold on top of a passive basement in response to forces from the northeast. The opposite view is that surface folding reflects local basement shortening, which may however have a different structural style, such as thrust faulting. This debate is reminiscent of the controversy over the formation of the Jura Mts. of Switzerland, which have a similar open folded geomorphology. These type-section Jurassic limestones overlying Triassic evaporites were once thought to have been folded by lateral pressure from the Alps, but are now regarded as the result of steep thrust faults in the underlying basement (Holmes, 1965).

In the case of the Zagros, the former position is taken by Dewey and Bird (1970), Haynes and McQuillan (1974) and Chapple (1975). Dewey and (John) Bird stated that the salt horizons acted as a lubricant permitting independent folding of the overlying rocks. Haynes and McQuillan suggest that the salt might take an active as well as a passive role in folding because of its buoyant upwelling in the cores of anticlines. And Chapple showed with a simple stress model that simultaneous plasticity of a wide thin sheet does not violate stress equilibrium if there is a weak foundation and a topographic gradient in the direction of shortening.

Basement involvement in the Zagros folding is put forward by Lees (1952), Falcon (1969), and (Peter) Bird et al. (1975). Lees argues that analagous structures in Venezuela, Sumatra, and Nigeria show flexural folding of the basement in the cores of anticlines. Falcon describes a number of cases from economic geology of inaccurate extrapolations to depth, and concludes cautiously: "the probability is that the basement, being of variable composition, fractured irregularly, not on a simple pattern, and that the disharmony between it and the overlying sediments is variable and complex. Surface geology can tell us nothing definite about these things."

The purpose of this chapter is to go beyond geologic evidence by considering the implications of gravity, seismicity, and heat flow for the problem of present-day deformations in the Zagros. The understanding of this belt, which is at such an early stage in the orogenic process, is essential to the unravelling of the history of its older and more complex analogues.

4.2 Gravity Data and Crustal Roots

In the absence of published seismic refraction data, the most powerful tool for determining the distribution of continental crust beneath the Zagros is the interpretation of gravity. Even an approximate determination of the crustal thickness is helpful in selecting between the models of Haynes and McQuillan (1974) and Dewey and Bird (1970). According to Haynes and McQuillan the Zagros have been formed by thickening of sediments originally deposited on oceanic crust; while Dewey and Bird interpreted the geosyncline as a former continental shelf underlain by continental crust.

A compilation of all open-literature gravity data in Eurasia has been made by the Gravity Services Division of the Defense Mapping Agency as reported by Wilcox et al. (1972). This data includes over 1,460 readings in the area between 26-38° North by 44-57° East. Most surveys were based at one of the thirteen gravity bases established by the Institute of Geophysics of Tehran University and connected with the International Calibration Line through Tehran. A smoothed, summary diagram of this data is presented in Figure 4.2. It shows a broad linear Bouguer anomaly of over -200 mGal intensity centered around and parallel to the plate suture at the Crush Zone. The linearity of this pattern is maintained between 46° and 56° East (from the Caspian Sea to the straits of Hormoz). To reduce scatter, all the data in this region has been assembled into 1° by 1° average Bouguer anomalies and projected onto a single cross-section of the range. The number of values in each average varied from 1 to 111, with an average number of 17. For purposes of projection the plate suture was represented as a straight line passing through 32.2° N, 50.8° E with an azimuth of 131.8°. The averaged points are plotted in Figure 4.3 according to the position of the midpoint of the 1° rectangle.

This data still shows considerable scatter in the folded belt. While some amplitude reductions might be caused by end effects or the averaging performed, the greatest part is probably due to unequal degrees of deformation and tectonic development along the strike of the range (Figure 4.2). This could be interpreted as showing that the central portion (50°-54°E) of the Arabian continental shelf collided with Iran first. A model curve has been fitted to the average anomaly amplitude of the entire region.

For purposes of interpretation it was necessary to make some assumptions about the depth of the anomalous masses. A large part of the sedimentary sequence is limestone, and this commonly has a density range of 2.4-2.8 g/cc (Birch, 1942) comparable to that of upper crustal metamorphic basement. The thick sediments are not likely to be responsible for more than about 35 mGal of the anomaly (2π G (0.1 g/cc) (8km)). Unfortunately the wavelength of the anomaly is much too large for any possible lithospheric sources to be eliminated on that basis.

However, very limited seismological data tends to rule out low-density sources in the mantle. In a Rayleigh wave phase velocity study of paths from Shiraz to Jerusalem, Helwan, and Addis Ababa, which cross a part of the Zagros and the adjacent Arabian shield, Knopoff and Fouda (1975) found high velocities comparable to the central United States. They interpreted these as resulting from a thick 100-140 km lithosphere with shear velocity 4.55 km/sec. If normal, thick lithosphere of this type extended under the continental shelf of Arabia at the time of the collision, and the collision was no earlier than Pliocene, then the results of Section 2.2.3 show that it should still be cold and dense.

A further check on the integrity of the mantle beneath the Zagros was made by searching for lithospheric shear wave arrivals (S_n) at the WWSSN station at Shiraz (SHI) from events along the strike of the Zagros. As shown by Molnar and Oliver (1969), high frequency and amplitude of S_n arrivals are good indicators of lithospheric thickness and low temperature. A clear arrival of greater than 1Hz frequency was recorded from the July 1, 1971 event at 36.4°N, 43.4°E at a distance of 10.3°. This record is shown in Figure 4.4. Four other, smaller S_n arrivals (Molnar and Oliver quality factor 2) were obtained at distances up to 15° from Shiraz. Two other small arrivals were observed by Molnar and Oliver from the same directions. This seems to indicate that no major thermal or tectonic events have disrupted the lithosphere beneath the folded belt.

With these possibilities tentatively eliminated, the most likely remaining source of Bouguer anomalies is a changing depth to Moho, either because the crust has been thickened or because it has been depressed beneath thickened sediments. A density contrast of 0.43 g/cc between crust and mantle was assumed, and the reference crustal thickness was taken as 33 km (Woollard, 1959) to match the models of Chapter 2. This is consistent with the value 35±8 km obtained by Knopoff and Fouda.

To the north the anomaly approaches a constant amplitude of -130 mGal, which is characteristic of much of central Iran. This implies either a crustal thickness of 40 km or anomalous high mantle temperatures. We assume that immediately behind the suture the crust is responsible because low heat flow is a general feature of arc-trench gaps (Toksöz et al., 1971). To calculate the edge effect, the Iranian crust is assumed to be terminated at the upper surface of the Arabian lithosphere, dipping 30° NE from the Crush Zone to reach a depth of 100 km beneath the volcanic arc (Figure 4.3). Also subtracted from the anomaly at the southwestern end is the gravitational effect of 2 kilometers of alluvial fill (using the curve of Skeels, 1940) at the NE edge of the Mesopotamian Trough. This must be regarded as schematic, as the depth of the fill has not been published.

The remainder of the anomaly is probably due to crustal thickening in the Zagros. A good fit is obtained if the Moho is assumed to dip only about 3° from a slightly depressed depth of 35 km under the Mesopotamian Trough to a maximum of 49 km below the Crush Zone. This represents an anomalous thickening of 2-17 km. A figure of 14 km anomalous thickness was independently estimated by Akasheh (1975) on the grounds that the average travel-time residual for P-waves from 120 events was +0.5 sec at Shiraz, in the center of the range. Such thicknesses are inconsistent with the model of Haynes and McQuillan. Even with an initial sediment thickness of 15 km, shortening by a factor of 2.3-3.3 would be required, and this is inconsistent with the observed gentle folding (Stöcklin, 1968). Haynes and McQuillan reached their conclusion that oceanic crust underlies the sediments largely on the basis that basaltic fragments are contained in pre-Cambrian salt upwelling as domes. It seems more likely that both salt and basalt were deposited in rifted grabens on continental crust during the pre-Cambrian breakup of Arabia-Africa from an unknown proto-continent. Such an association is found today in the Afar triangle (Tazieff, 1972).

If, instead, the present thickness results from 2-17 km thickening of normal continental crust, less dramatic shortening is required. It is still not possible to explain this much thickening by sediment shortening alone, however. If a 12 km layer originally 370 km wide had been shortened to a present wedge shape 14 to 28 km thick by 210 km wide, then 160 kilometers of the continental crust on which these shallow-water sediments were deposited would have had to disappear. Though it is not impossible for such an amount of bouyant material to be pulled down into the mantle by a long oceanic slab of lithosphere, it is hard to imagine this happening without the tension producing deep earthquakes.

It seems more likely that both sediments and crust have been shortened by equal amounts, about 60 km out of an original width of 270. Not only does this imply a collision date (1.3 m.y.) consistent with geology, but it gives an average strain (22%) that can be reconciled with observed folding. In the case of cylindrical circular folds, this amount of shortening would imply maximum limb dips of θ=69° according to the formula

$$\frac{\theta}{\sin(\theta)} = \frac{\ell_1}{\ell_2} \qquad (4.1)$$

where ℓ_1 and ℓ_2 are the original and final horizontal widths of the folded beds. This overall shortening figure of 22% agrees very well with the average 21% shortening of the competent Asmari limestone shown in three sections by Lees. These sections, reproduced by Abdalian (1963), are from the southwestern half of the range.

The model proposed here for the Zagros is one in which a 35-49 km crust including about 9 km of deformed sediments overlies a thick, high-velocity lithosphere. This model was tested by comparing its predicted Rayleigh wave group velocities with observed seismograms at Shiraz from nearby events sending waves along the strike of the range. The parameters of the seven events used are listed in Table 4.1. The small region to be studied limits the distance range to about 16°, so the long-period waves are insufficiently dispersed to allow detailed interpretation.

The long-period vertical records were digitized and subjected to the multiple-filter alogrithm of Dziewonski et al. (1969) with a Gaussian filter of constant relative bandwidth:

$$L(\omega) = \exp\left(-30\left(\frac{\omega-\omega_0}{\omega_0}\right)^2\right) \tag{4.2}$$

Resulting spectral amplitudes were summed in frequency-group velocity space, so that the noise contained in each record makes an equal contribution to the composite plot. This plot is shown in Figure 4.5, along with the group velocity curve of a successful model as calculated with the program by Harkrider based on the matrix method of Haskell (1951). The parameters of the model are in Table 4.2.

The fit of this simple model is excellent, though of course not unique. The usual trade-off between crustal thickness and mantle velocity cannot occur because the upper mantle shear velocity is fixed from S_n observations (4.65 km/sec). However, different crustal thicknesses could be fit by varying the lower crustal velocity. In a similar way, thicker sedimentary layers could be used if the velocity was raised. However, a high shear velocity of 3.00 km/sec appropriate for dense limestone (Press, 1966) is already employed in this model with 9 km of sediments. If one attempted to fit the data with a 14-28 km sediment layer as discussed above, higher and less plausible velocities would be required. It appears that the combination of gravity and group velocity data require some amount of basement shortening in the Zagros.

The mechanism by which basement shortening of this magnitude might occur is investigated in the following sections with heat-flow and seismic evidence compared to the predictions of finite-difference and finite-element models.

4.3 Seismicity of the Zagros Region

The Zagros continental convergence zone is as seismically active as either the Chilean or the Himalayan subduction zones. This statement is based on the number of earthquakes with magnitudes of 5.0-5.9 occurring during 3/63-12/70 according to the Coast and Geodetic Survey. (There are not enough larger events to give valid comparisons, and the detection of magnitude 4 earthquakes is biased by station densities.) In each case, an area of 1500 by 1100 km containing the subduction zone was sampled to a depth of 150 km. The Zagros had 76 events in that period, slightly more than the Himalayas (60) and slightly less than Chile (91). From the surface-wave moments of large earthquakes in Asia, Chen and Molnar (1976) have estimated that seismic slips account for 1.0-2.7 cm/year convergence in that area. The convergence rate in the Zagros is probably no less.

As in the Himalayas, this seismicity is concentrated at shallow depths. Akasheh (1972) has estimated that over 90% of the seismic energy release in the period 1903-1970 occurred at depths of less than 40 km. Nowroozi (1971) relocated some 200 earthquakes in the Zagros region during 1950-65, and only 14 of these were placed at over 100 km depth. These few deep events have no pattern, and may merely be mislocated. In short, there is no evidence from seismicity that the downgoing slab which created the Neogene volcanism in Iran is still present.

There are only two pieces of evidence which suggest such a slab. First, Akasheh (1975) observed that P-wave travel-time anomalies at Shiraz (in the center of the Zagros) were more positive (slow) from the southwest than from the northeast. This he interpreted in terms of a high-velocity slab dipping northeast from Shiraz. However, the anomalies in the northeast direction are still positive, and can be simply explained as a result of a small amount of crustal thickening. Furthermore, early arrivals through a slab should be restricted to a narrow range of azimuths and dip angles (Sleep, 1973), whereas in this case it is the slow southwestern arrivals which occupy a narrow range of azimuths.

The existence of descending slabs was originally established in part by the observation of low attenuation in narrow dipping regions (Oliver and Isacks, 1976). Thus when Molnar and Oliver (1969) observed high-frequency S waves at Shiraz from deep (200 km) earthquakes in the Hindu Kush, they also suggested the presence of a dormant or inactive slab. Perhaps there is a former slab in the area, too warm to produce earthquakes, yet still a diffuse zone of low temperature and attenuation. The difficulty with this theory is one of timing. Around the margins of the Pacific, the depth of seismic zones can be approximately predicted from subduction rates by assuming that the slab requires 10 m.y. to become warm and stop producing earthquakes (Isacks _et al._, 1968). As discussed in the last sections, geology restricts the age of any Zagros slab to a maximum of 5 m.y., with a most probable age of 2 m.y. These figures could be reconciled if the slab had subducted at unusually low rates in the recent past, reaching thermal equilibrium at shallow depths (as in the Aegean Sea). Since the relative positions of Africa and Europe are only well fixed at anomaly 5 (10 m.y.) time, this is a possibility. Alternatively, some aspect of the continental collision may have caused the slab to detach and sink independently into the mantle. According to this hypothesis, the lack of deep earthquakes is caused by low stress in the slab rather than high temperature.

The present seismicity does not indicate any continuing subduction at the Zagros Thrust. As shown in Figure 4.1, it falls in a wide band between the Crush Zone and the Mesopotamian Trough, which corresponds closely to the region of Recent folding and negative Bouguer gravity anomalies. When the locations of these earthquakes, determined by Nowroozi (1971), are projected onto a section perpendicular to the range, no further pattern appears (Figure 4.6). If the given depths are accurate, the seismicity suggests a general compressive strain throughout the lithosphere of the Zagros.

The sense of deformation is confirmed by fault-plane solutions of eight events published by Nowroozi (1972), McKenzie (1972), and Akasheh (1973). These solutions are projected onto the section of Figure 4.6 by rotating the intermediate-stress axis (nodal line) to a horizontal position parallel to the strike of the range. In six cases this rotation was a small one about the vertical axis; in the other two cases the axis of rotation was within 20° of vertical. All eight solutions show thrust faulting with NE-SW shortening, in agreement with the stress pattern implied by folding. Because only a few solutions are constrained by dilatational arrivals, the small variations in the dip of the compressive axis may not be significant. Two of these events (9/18/66 and 9/14/68) were analyzed simultaneously by Nowroozi and McKenzie: in the first case the interpretations were similar and in the second the solution of McKenzie has been used. Also consistent with this pattern (but not included in the figure) are eight fault-plane solutions from reported first arrivals by Canitez (1969). These were predominantly thrust faults like the others; in the cases involving a strike-slip component the compressional axis was still horizontal and NE-SW.

It is interesting that even at the Crush Zone itself there is thrust faulting, rather than the normal faulting commonly associated with oceanic trenches (Isacks _et al._, 1968). Apparently the Arabian plate is free from any large bending torques along its northeastern margin. This is weak evidence for the absence of a downgoing slab in the Zagros.

If the depths of these thrusting events could be well determined, the question of whether the present orogeny includes crust as well as sediments could be answered. Locations based on teleseismic P-wave arrivals alone lack the accuracy to resolve depth within the uppermost 30 km. However, by the technique of Tsai and Aki (1970) it may be possible to extract this information from amplitude spectra of Rayleigh waves excited by the events. This method requires that the fault-plane solution of the event and the layered Earth structure both be known.

Of the eight events with published fault-plane solutions, four gave clean Rayleigh wave records at one or more of the nearby WWSSN stations at Jerusalem, Addis Ababa, Lahore, or New Delhi. Although all these paths are short and continental, none is ideal. The path to Jerusalem crosses the Dead Sea rift, a probable strike-slip plate boundary (McKenzie, 1972). The path to Addis Ababa crosses the narrow mouth of the spreading Red Sea, and the paths to India go through the folded belt of Afghanistan.

Using the equation of Tsai and Aki (1970) in conjunction with the Rayleigh wave eigenfunctions of the Zagros earth model developed in the last section, theoretical spectra were predicted for each event at each station as a function of source depth. The predicted curves are almost identical to those derived with the Gutenberg continental model (Dorman *et al.*, 1960), so this is not a major source of error. A sample set of curves is shown in Figure 4.7; unfortunately, for the event and station geometries available no sharp spectral "holes" are predicted and the spectra for events of different depths may be quite similar.

The calculated model spectra were compared with spectra obtained from the digitized long-period records of each station, including the group velocity window 5.0-2.5 km/sec. With one exception, no higher-mode arrivals were visible in the records. All spectra were corrected for instrument amplitude response. Results for the May 30, 1968 event were inconclusive; a good record was only available at AAE and its spectrum did not match any of the model curves. Likewise the Sept. 18, 1966 event could not be placed because the AAE record indicated a depth of 9-12 km, while the JER record suggested it was either deeper (25 km) or more shallow (3 km). However, the two remaining events appear to be definitely below the sedimentary level. Waveforms and their spectra at AAE and JER from the June 23, 1968 event are shown in Figure 4.8 and fitted to a focal depth near 50 km. The poorer fit of the short periods at JER is believed to result from some aliasing of the strong 5-second higher mode into the 10-50 second period range of the model curves. Also, for the Sept. 14, 1968 event the LAH record indicates a focal depth near 35 km, as shown in Figure 4.9. Less confidence is placed in this location because the same event recorded at NDI gave an ambiguous spectrum.

Although this experiment was not generally a success, it did indicate that one and possibly two of the four events studied are too deep to have originated in the sedimentary layer. This supports the theory of basement involvement in Zagros tectonics. With our present knowledge of rock mechanics, it is hard to see how sediments which fold pliably at zero confining pressure could also generate magnitude 7 earthquakes (Akasheh, 1973). Horizontal shortening of the underlying crust is a more plausible source for these destructive events.

4.4 Heat Flow and Thermal Models

There is only one published measurement of heat-flow in the Zagros range, performed by Coster (1947). However, the National Iranian Oil Company has recorded bottom-hole temperatures during the drilling of a number of deep 3-5 km wells, and has kindly made them available. In this section those temperatures are used to make heat-flow estimates for five additional fields, and the possible tectonic implications are discussed. In addition, a group of small granitic intrusions which may have formed during the present orogeny are assigned a tentative source.

Coster's work was at the Masjid-i-Sulaiman oil field at 31°59'N by 49°18'E, in the center of the folded belt. After making conductivity measurements for the Asmari limestone and overlying Fars and Bakhtiari conglomerate formations, he determined heat flow from nineteen wells in a small region. These determinations had a range of 0.53-1.22 HFU (22-51 mW/m^2) and an average of 0.88 HFU (36 mW/m^2). This low value is more typical of continental shields than shelf or orogenic areas, and tends to support the low age assigned to the Zagros orogeny.

All the available bottom-hole temperatures in the Zagros are shown in Figure 4.10. As a group they imply a regional heat flow around 1.0 HFU (42 mW/m^2) if we assume a conductivity value of 5.2×10^{-3} cal/cm-sec/°C for Asmari limestone is representative (Clark, 1966). They also suggest a high temperature intercept of about 29°C (85°F), warmer than the regional annual temperature average of about 23°C (73°F) (Grosvenor _et al_., 1966). Since most of the data is from deeper than one kilometer, this is not likely to be the result of late Pleistocene climatic changes. It could be the result of uplift and erosion of 350 meters in very recent times.

In an attempt to find heat-flow constraints on the recent tectonics, data from individual wells was separately analyzed. Much of the data was rejected because of apparent temperature discontinuities or reversals, probably caused by groundwater movement or drilling effects. Wells with less than six measurements were also rejected because the continuity could not be evaluated. This left six wells, with the data shown in Figure 4.11. This data was fitted to theoretical curves calculated with Asmari limestone parameters: $K_0=5.2\times10^{-3}$ cal/cm-sec-°C and dK/dt- -9.5×10^{-6} cal/cm-sec-°C^2. The conductivity increase of 20% due to wetting and compression is assumed to occur linearly in the first 2 km. Temperature curves were integrated from the two initial values mentioned above (23° and 29°C) at heat-flow intervals of .05 HFU. The best-fitting curve was selected by a least-squares criterion with temperature the dependent variable.

As one of the six wells is in Coster's field the method can be calibrated. The data gave a heat flow of 0.90 HFU, the same as Coster's average value. There seems to be a trend toward higher heat flow in wells closer to the southwest margin of the mountains. The highest value of 1.15 HFU occurs at the Ahwaz field, which is located over a thrust fault in the last anticline before the Mesopotamian Trough. While a heat flow difference of this magnitude (.25 HFU) would not normally be considered significant, there are several reasons why it may be in this case. First, the Ahwaz data is very linear, while the Masjid-i-Sulaiman value is supported by Coster's result. Second, the wells are so deep that the temperature contrasts we are trying to detect amount to about 30°C. Finally, all the wells are in a comparable structural position on the axis of an anticline in the same formation.

Two approaches can be taken in the interpretation of this constrast, depending on which value is assumed to be normal. In an earlier paper, Bird et al., (1975) assumed the lower value was normal and ascribed the higher ones to tectonic frictional heating. The thermal model incorporating this hypothesis is shown as Figure 4.12; it assumes that the crustal thickening in the Zagros has been caused in part by 27 km of slip on a new thrust fault dipping 15° NE from the margin of the Mesopotamian Trough. If the shear stress on the fault is assumed to follow the pore-pressure-corrected frictional law

$$\tau_p = \mu \bar{\rho} g z = (0.6)(2.67 - 1.03)(980.)\,z \tag{4.3}$$

then a good match to the data is obtained, as shown in Figure 4.13. Although shear stresses are allowed to rise as high as 4 kb in this model, the highest stress level actually required by the surface data would be about 1.4 kb beneath Agha Jari.

While this model is consistent with all the constraints discussed in previous sections, it has the disadvantage of requiring an unusually low initial heat flow for the continental shelf. Also, there is no direct evidence for the existence of this thrust fault except for the short 50 km segment mapped by Stöcklin and Nabavi (1973) and the fact that the mountains are rising while the adjacent plains subside. A somewhat simpler explanation is possible if we consider the low heat flow in the interior to be anomalous.

This low heat flow could be a result of crustal shortening and thickening (with or without major thrust faults) if the deformational stress were low. For example, a strain of 20% caused by a stress of 1 kb results in a dissipative heating of only 9°C. Thus the geothermal gradient could actually be depressed at an early stage in the orogeny. This possibility is sketched (not calculated) in Figure 4.14, which was derived from the following assumptions, i) the initial heat flow of the Arabian continental shield was 1.13 HFU, derived from a geotherm like that of section 2.1 but with crustal radioactivity reduced 30%; ii) the initial crustal thickness was 33 km, and the anomalous thickness shown in Figure 4.3 gives a measure of the local strain; iii) shear-strain heating was negligible; and iv) the thickening has occurred too recently for the geotherm to re-equilibrate. As shown, this model also gives a good match of the heat flow gradient.

A choice between these two models can only be made by numerically modelling the deformation to test which is self-consistent. This will be done in section 4.6. Of course, it may develop that the apparent difference in heat flows is fortuitous.

The only other evidence for thermal anomalies in the Zagros is geologic and is given by Wells (1969):

> ...partly kaolinized but very young acidic intrusives are found in four separate outcrops (probably connected at depth) in the Crush Zone between Bandar Abbas and Sirjan, some 190 km south-east of Neyriz. They are true granites, which intrude and thermally metamorphose both earlier basic intrusives and crushed sediments, and they are totally unfoliated even though they outcrop within the margin of the Crush Zone.

Unlike the basic intrusives of the Crush Zone, which can be interpreted as products of ancient sea-floor spreading, these granites present a real problem.

As shown in Chapter 2, no melting of crust is expected this early in a continental collision. Even if the shear zone models of that chapter are wrong and melting does occur, the granites should be found in central Iran and not in the top of the Crush Zone. The only remaining source of heat would seem to be some kind of diapiric upwelling of the mantle, whereby the hot, mobile asthenosphere comes into contact with continental crust and melts it. If the oceanic slab of lithosphere once attached to Arabia has in fact detached, this would be the expected result (Figure 4.15). Mechanical evidence tending to support this detachment is given later in the chapter. If on the other hand the Tethyan slab is still attached, then these granites must have risen from great depths, ascending at a shallow angle through the subducting material of the Crush Zone. Neither hypothesis is presently susceptible to proof.

4.5 Force Balance Responsible for Crustal Detachment

The results of the previous section indicate crustal shortening of some type is presently active in the Zagros. The details of the mechanism will be discussed in the next section. Here we attempt to find a balance between known forces acting on subducting continental crust and those needed to detach it from the lithosphere. The assumption that detachment occurs is based on the observation in section 3.2 that the lithosphere is much stronger than the lower crust. It is also required by the observed concentration of deformation at the edge of the Arabian plate; if the lithosphere were not a strong stabilizing influence this shortening would spread over the whole plate, as in eastern Asia.

The assumed geometry is shown in Figure 4.15. We take the block of detached crust to be 200 km wide to match the observed width of mountain ranges, and assume that the oceanic slab has dropped off. (Otherwise it would exert a "suction" on the crust, making detachment much harder to explain.) Detachment is encouraged by friction against the overriding plate and by the buoyancy of the low-density crust relative to the mantle. It is opposed by creep strength of the lower crust and by friction at the "toe" of the overthrust trailing edge of the block. I will quantify each of these forces, considering only the components in the direction of detachment motion. By finding the force balance at the initiation of detachment we avoid the geometric complications of plate deformation which develop later.

The primary variable is the vertical displacement of the leading edge or tip of the crust, called Δz_t. Since the shear stress $\tau(s)$ is known in the original subduction zone from models in Chapter 2, the total frictional force on the leading end is just:

$$f_{sz}(V,\theta,A,B,n,\mu,\Delta z_t) = \int_0^{\frac{\Delta z_t}{\sin\theta}} \tau(V,\theta,A,B,n,\mu,s)\,ds \tag{4.4}$$

and depends on plate velocity, dip, and the rock strength in the subduction zone. As noted earlier, this term is only weakly dependent on θ; I take $\theta = 30°$.

The buoyancy term f_b is calculated as the difference between the downdip component of gravity forces acting on the crust and the updip component of the lithostatic pressure of the asthenosphere. If ρ_c is crustal density and ρ_m is mantle density and l is crustal thickness, then

$$f_b = g\sin(\theta)\int_{\max(l,\Delta z_t)}^{l+\Delta z_t} \left(l\rho_c + (z-l)\rho_m\right)dz -$$

$$g\cos(\theta)\, l\rho_c \Delta z_t =$$

$$g\sin(\theta)\Delta\rho \begin{cases} \Delta z_t^2/(2\tan\theta) & ; \Delta z_t \le l \\ l(\Delta z_t - l/2)/\tan\theta & ; \Delta z_t > l \end{cases} \tag{4.5}$$

This could also be calculated as the derivative of the total gravitational potential energy with respect to downdip displacement. I take $\Delta\rho = 0.43$ g/cc.

To find the creep resistance to detachment we first define the detachment velocity ΔV_d by

$$\Delta V_d = \int_0^{\ell} \dot{e}_{xz}(\tau, T(z), A, B, n)\, dz \qquad (4.6)$$

in the flat part of the plate outside the subduction zone. Using the initial constitutive model for the lower crust developed in section 3.2, we evaluate $\Delta V_d = 1.55\times10^{-26}$ cm/year for a unit shear stress of $\tau = 1$ dyne/cm^2. Then using the cubic (n=3.0) flow law, the total resistance to detachment from creep is

$$f_c = \int_0^{200\,\mathrm{km}} \sqrt[3]{\frac{\Delta V_d}{1.55\times10^{-26}}}\, dx = 5.84\times10^{15}\sqrt[3]{\Delta V_d} \qquad (4.7)$$

where f_c is in dynes/cm and ΔV_d in cm/year. Some models are also considered where this resistance is arbitrarily reduced by a factor of ten.

The final term is f_f, the frictional faulting resistance to the new overthrust that develops.

$$f_f = \int_0^{a} \frac{\mu \rho_c g z}{\sin\beta}\, dz = \frac{a^2 \mu \rho_c g}{2\sin\beta} \qquad (4.8)$$

if "a" is the depth at which the creep stress becomes less than the faulting stress. This depends on ΔV_d, A, μ, and the width over which the creep occurs (W). For $\mu = .6$, W = 5 km, and V_d = 5 cm/year, "a" is 23 km in the initial strength model, and 17 km in the reduced strength model. Zagros fault-plane solutions suggest $\beta = 45°$, so $f_f = 5.87\times10^{15}$ dyne/cm in the initial model and 3.21×10^{15} dyne/cm in the reduced-strength model. This is comparable in size to the creep resistance f_c.

To find the detachment point we set

$$f_{sz} + f_b = f_c + f_f \tag{4.9}$$

with the condition $\Delta V_d = V/2$. This represents the point at which the crust is only subducting half as fast as the lithosphere beneath it. At this point two shear zones are competing to absorb the relative slip between plates, one above the crustal block and one at its base. As shown in Chapter 2, both zones are unstable and become stronger at lower velocities. This instability will cause a very rapid transfer of slip, as the old subduction zone cools off and becomes stronger and the new shear zone begins to warm up from shear-strain heating.

These instability points, as a function of subduction zone and lower crustal strength, are shown in Table 4.3. For comparison, the ΔZ_t value for the Zagros is now around 16 km.

This disagrees strongly with the predictions of the initial strength model, especially as limestone is likely to be present at depth in the Zagros Crush Zone. In that case, this analysis predicts that the crust would not detach at all, but be completely subducted.

To get even an approximate match, the strength of the lower crust must be reduced from the value based on creep of dry quartz only. When it is reduced ten times, the predicted amounts of subduction are 25 and 31 km for limestone and "wet"-quartz-granite respectively in the Crush Zone. These are still rather too large, and since the main opposition to detachment in these cases comes from the overthrusting "toe", the shear stress on this toe was arbitrarily reduced to 200 bars. This has less effect than one might suppose, because the same low stress limit must also be applied at the subducting end of the slab for consistency. The effect of the change is to reduce the predicted ΔZ_t to 22 km.

Of course, the shape of the Moho in the Zagros is different from the simple shape assumed. Furthermore, complete detachment may not have occurred yet. These effects are investigated with finite elements in the next section. What this section clearly shows is that some reduction of strength from the initial model will be essential to a successful model of Zagros tectonics. This reduction of strength below that of dry quartz could have several causes.

The geotherm of Chapter 2 could be in error. But, since a Moho temperature of 750°C would be required to weaken dry quartz to this extent, and this is equal to the dry melting point of granite (Brown and Fyfe, 1970), this explanation is not attractive. Quartz could also be reduced to the necessary strength by a crystal water content still substantially less than in the artificial specimens of Hobbs *et al*. (1972) and Balderman (1974). Or (most likely) substantial creep could be taking place in some other mineral. This problem points out the need for further laboratory work on crustal minerals.

It is worth noting that we can only explain the recent tectonics of the Zagros by scaling down the strength of the crust until deformational forces are comparable to buoyancy forces. The buoyancy force in turn is dependent on the crust being in contact with a soft mobile material (asthenosphere) which does not participate in the general subduction. If an oceanic slab were still attached to the leading edge of the continental crust, it would be pulled down to much greater depths. This effect is apparent in the more sophisticated models of the next section.

4.6 Finite Element Models of Continuing Deformation

The Zagros range is 1800 km long and less than 300 km wide. On this basis, and because most fold axes parallel the range, it appears that plane strain gives a good approximation to the present deformations everywhere except near the mouth of the Persian Gulf. This allows us to solve the flow problem in two dimensions instead of three, and to concentrate seismic data from a limited period onto a single cross-section. The grid of elements used to represent this cross-section is shown in Figure 4.16. It has 63 modes, 96 elements, and 188 degrees of freedom before the imposition of boundary conditions.

The figure also shows the five important rock types in the Zagros which the various elements represent: mantle, crust, subducted continental rise sediments in the Crush Zone, Cambrian salt, and the thick sedimentary section. It is not possible on this scale to include the various sedimentary formations, and we allow only one row of elements to the whole section. The salt is represented by an underlying row of elements 1 km thick, invisible in the figure. The mechanical properties initially assumed for these rocks are given in Table 4.4.

The mantle creep parameters are derived, as described in the last chapter, from Kohlstedt and Goetze (1974). A pressure correction of +16°K/km was incorporated into the constant B, making the assumption that creep can be predicted as a function of the fraction of the absolute melting temperature (Weertman and Weertman, 1975) and using the melting curve of Davis and England (1964). The constitutive law of the crust was also described in Chapters 2 and 3. Subducted sediments in the Crush Zone probably include a high percentage of limestone, so the Solnhofen limestone parameters of Schmid (1976) are utilized.

To obtain a creep parameter for the evaporite, the halite results of Heard (1972) were scaled up in strength by a factor of 2.15. This causes the 1-km zone in the model to deform with the strength of a 100 meter layer of salt, a more plausible thickness. With or without this correction the salt represents an almost stress-free horizon. The proper constitutive law for the sediments is really completely unknown. We can only observe that they fold at zero confining pressure without faulting; hence they must be very weak. Since the aggregate is much weaker than the creep laws of the minerals predict, some other non-elastic deformation process is probably occurring, such as dilatational strain through cracking. Although this process would be nonlinear in stress, there is no particular justification for the exponent (n=3) employed other than consistency. The sediment creep parameters were arbitrarily set to allow strong folding at a low stress of 250 bars, and held constant in all models. The importance of this sedimentary layer in the model is not its strength but its mobility. By flowing around rising and sinking areas of basement, it allows crustal deformations that would otherwise be resisted by gravity (or in the model by boundary conditions)

The temperature model used in all calculations was the conservative horizontal-shortening model of Figure 4.14. Besid incorporating a more "normal" initial geotherm, it has no local anomalies to bias the flow solution. It will be shown that the results favor this model with low shear-strain heating. The temperatures in the Crush Zone were calculated by the method of Appendix B using the same creep parameters, dimensions, and velocity as in the mechanical model. When creep parameters in the Crush Zone were changed, new compatible shear-zone temperatures were also calculated. Boundary conditions were imposed on the left-hand side of the grid, at the top, and sometimes on the right. Because differential densities are used and the asthenosphere is very weak, the base of the model is left free of deviatoric stress. The horizontal component of velocity only is set to zero at X=0 to provide a reference point for the velocity solution. The vertical component of velocity was initially set to zero at Z=0 for the same reason. Actual rates of erosion and deposition would be preferable as vertical velocity boundary conditions (if they were known), but they are typically small with respect to relative plate velocities. The top boundary condition allows the vertical stress ρ_{zz} to be non-zero at sea level, thus simulating the topographic load effect. Crustal elements extending below the 33 km reference depth are assigned a density anomaly of -0.43 g/cc in equation (3.23) to represent the uplifting force of these roots. Plate convergence at a rate of 4.7 cm/year was enforced either at all nodes on the right boundary ("ridge-push" models) or on the three lithospheric nodes of the lower left boundary ("slab-pull" models).

The successful model which is sought is defined as one which passes the following tests:

i) Predicted earthquakes should be evenly distributed between the Crush Zone and the edge of the Persian Gulf-Mesopotamian Trough, with none outside this area. This corresponds to the range X = 150-400 km in the models. Predicted earthquakes are not in general required to match the relocated hypocenters in depth, because they are probably inaccurate, but some events must be predicted below the sedimentary layer.

ii) All predicted faults must be of the thrusting type, with an average dip close to 45°.

iii) The surface sediments must be undergoing horizontal compression in the SW half of the range.

iv) Some amount of crustal thickening must be continuing in the area where thickened crust is now observed.

In the attempt to find such a model, four parameters have been varied; the boundary conditions, the fracture strength of the crust, the creep strength of the crust, and the creep strength of the mantle. All the variations away from the initial strength model have been in the direction of greater weakness, because unknown deformation mechanisms cannot strengthen the rocks. Still, it was not possible to investigate all possible combinations. Instead, preliminary models (most of them not presented here) were calculated with the initial-stress method and used to rank the importance of these variables. (Boundary conditions are the most critical and mantle strength is the least.) Then, with the more accurate variable-viscosity method, various combinations of parameters were systematically eliminated. The models fall into three groups, depending on boundary conditions, and within each group crust and mantle strengths were varied. It will not be possible to describe or reproduce each model in detail, but a complete list of parameters and major results is given in Table 4.5. The greatest number of models were unsuccessful in very similar ways.

The first set were "slab-pull" models, in which the lithosphere was required to descend at 30° dip on the NE margin of the model. The SW end is then left free, but it is necessary in some way to simulate the effect of the Arabian plate not included in the model. This was done by artificially stiffening the elements of the rightmost column, to prevent any differential horizontal slip of the rock layers at the right boundary. Because of this stiffening the stresses in the region (X=390-450 km) may be non-physical.

With the initial rock parameters described above and in Table 4.4, the result was pure subduction of the Arabian plate. This model (Z15) is shown in Figure 4.17. All of its predicted seismicity falls in the Crush Zone, which is incorrect. No crustal thickening occurs, and even the surface sediments are in horizontal tension in the southern half of the Zagros. Deviatoric tension of up to 3.0 kb is required in the subducting slab. This model is not acceptable according to all four criteria.

Compared to the large faulting stresses required in the cold upper crust of this model, the forces produced by crustal buoyancy are small. As a result, the natural solution is hardly different from the simple subduction of oceanic lithosphere. In order to investigate the possibility that a smaller faulting stress was the cause of crustal deformation, an upper limit was placed on the plastic faulting stress for all crustal rocks, including the sediments and Crush Zone material. This limit was successively reduced to 1000 bars, 500 bars, and 250 bars in models Z16, Z17, and Z18 respectively. In each case, the new rheology of the Crush Zone required a recomputation of temperatures in that region, using the method of Appendix B.

The results in each case were very similar. As detailed in Table 4.5, the crust simply subducted in each model, and the seismicity remained unacceptably far NE. The only change was a slight rotation of the fault planes to steeper dips, but this was a trivial result of the change in strength. When the faulting-stress equation of Appendix C is replaced, it is no longer consistent to use the same equation for the angle between σ_1 and the fault planes. Because the lowered faulting criterion assumes no dependence on normal stress, this angle becomes 45° in all cases.

In model Z19 the creep strength of the mantle was reduced to that of water-saturated olivine by reducing the activation energy to 94 kcal/mole (Post, 1973). The low crustal faulting stress of 250 bars was retained. Still, the stress field was dominated by deviatoric tension radiating from the downgoing slab, the crust simply subducted, and the seismicity pattern was unchanged. The lower crust was in a state of vertical compression-horizontal extension, neither promoting nor hindering the detachment of the crust. Therefore no models with reduced lower crustal strength were computed. The defect of these models seems to be the boundary condition, since the attempts to scale down the rock strength and so balance it against buoyancy had no effect.

The second set were "ridge-push" models, with horizontal velocity specified on the SW end and the slab end left free.

The artificial stiffening of the SW end was no longer necessary and was removed. Models Z04, Z05, and Z06 were computed with initial parameters, with lower crustal creep strength reduced by 20 times, and with both faulting and creep strengths reduced (respectively). None of these combinations produced any change from the pattern of simple subduction with seismicity in the Crush Zone, too far NE. The only change from "slab-pull" models was that the deviatoric stress in the crust was now uniformly horizontal compression instead of vertical. However, this produced no shortening except by less than 1 cm/year in the sedimentary layer. These models still fail tests (i) and (iv) for seismicity and crustal thickening.

Keeping the crustal faulting stress at a low 250 bars, the mantle creep activation energy was also reduced to 110 kcal/mole (model Z22) without any effect. When it is further reduced to 100 kcal/mole (Z23) a change in seismicity occurs, as shown in Figure 4.18. Receiving less support from the weak mantle, the crust shortens slightly at the SW end under the Mesopotamian Trough. By reducing Q_a further to the water-saturated value of 94 kcal/mole, a dramatic change can be induced (Z21, Figure 4.19). No longer able to support the stress, the crust and mantle both are shortened together in the Mesopotamian Trough area, with considerable seismicity. However, this seismicity falls in an area which is supposed to be inactive, while little or none is predicted in the Zagros range. This model also fails (i). The shortening at the SW end occurs where it does because the crust there has no buoyant roots to resist compression. This shows the futility of attempting to find a good model by adjusting the mantle strength parameters. If the mantle were very weak, shortening would occur in Arabia where the crust is thinnest. If the mantle is made strong enough to stabilize the crust, it also forces it to subduct (regardless of crustal parameters) by a kind of "suction". This "suction" is caused by the Arabian lithosphere being in contact with Iranian lithosphere across the Crush Zone. Here, part of the imposed horizontal velocity is translated into downward velocity which is transmitted to the crust. This appears to be the basic defect of all models described so far.

The problem will not arise if the Tethyan oceanic slab has actually detached at the edge of the continental crust. This would create a sub-crustal break in the lithosphere underneath the Crush Zone which could explain the young granites as well as the lack of deep earthquakes previously discussed. To model this detachment, I simply removed four lithospheric elements from the Arabian plate, as shown in Figure 4.16. The area which was deleted (replaced by weak asthenosphere) is in fact completely aseismic, as Figure 4.6 depicts.

This new grid was used in a third class of models. Since it was shown that a weak lithosphere encourages instability and unwanted seismicity in the SW, the activation energy was set at an average value of 110 kcal/mole for all these models. This gives a lithosphere sufficiently strong to resist shortening, so this value cannot be distinguished from higher ("drier") values. Since the slab is assumed to have dropped off there is no longer uncertainty as to the source of the driving force; it must be a horizontal compression transmitted through Arabia. The two remaining variables are lower crustal creep strength and upper crustal faulting strength, and there are theoretical results suggesting that these can be independently determined.

Recently both Chapple (1975) and Elliott (1976) have produced analyses of thin-skinned overthrust belts deforming above a passive basement. While they were discussing sedimentary sections, the same geometry applies to the crust of the Zagros. It is thin-skinned (33-49 km thick vs. 300 km wide) and (as model Z21 showed) must be deforming independent of the more rigid sub-crustal lithosphere. Assuming plane strain and homogeneous material with a rigid-plastic constitution, both authors found that the shear stress τ_b on the base of the overthrust belt is independent of the internal plastic limit. Instead, in the case of simultaneous plasticity (faulting) of the whole "thin skin" the derivation of τ_b proceeds like this:

$$\tau_b = \int_0^{\ell} \frac{\partial \tau_{xz}}{\partial z} dz = \int_0^{\ell} \frac{-\partial P_{xx}}{\partial x} dz$$

$$\approx \int_0^{\ell} -\frac{\partial P_{zz}}{\partial x} dz = -\ell \frac{\partial P_{zz}}{\partial x} = \ell \rho g \frac{\partial h(x)}{\partial x} \tag{4.10}$$

where h(x) is the topography above sea level and ℓ is the layer thickness. In the case of the Zagros this implies $\tau_b \approx$ 75-100 bars, increasing from the Persian Gulf toward the Crush Zone. This result implies a strength lower by a factor of five than that of the initial strength model that assumed the only creep was in a fractional content of dry quartz (see equations 4.6, 4.7). The same result was suggested by a different kind of approximate analysis in the last section.

The results of the finite-element models were as predicted. With a low faulting stress of 250 bars (model Z24) but no reduction of the creep strength, subduction without shortening is produced as before. The only novelty was that central Iran on the NE side of the plate suture was horizontally shortened and seismically active. To suppress this unwanted effect, the Iranian plate was given a higher plastic limit in later models. The physics behind this strength contrast are not known. Another refinement introduced at this point was the imposition of a 1.1 millimeter/year uplift in the Zagros and a 7 mm/year subsidence in the Persian Gulf

(Figure 4.16). This rate of uplift would be required to create the existing mountains if the uplift began 2. m.y. ago at the plate margin and gradually migrated inward. Isolated uplifts of anticlinal axes may have been faster, such as the rate of 20m in 1700 years determined by Lees and Falcon (1952). The rate of subsidence is somewhat arbitrary.

When the creep stress of the crust was reduced by a factor of five in model Z26, the results began to improve. As shown in Figure 4.20, crustal shortening results but is not evenly distributed. Seismicity falls in the prescribed range in X but is strongly concentrated to the NE. Since the proper parameters for the sediments are unknown, the seismicity in that shallow layer may be artificial. Neglecting it, the activity is too strongly concentrated near the plate suture. This indicates that the creep strength is still slightly too high.

When the creep parameter A was reduced by a full factor of ten, to 4×10^7 dynes/cm^2, the first truly successful model was obtained. This is model Z25, which is shown in Figures 4.21 and 4.22. Its major features are relatively uniform shortening of the crust between the Crush Zone and the Mesopotamian Trough, independent shortening of the sedimentary layer at higher rates in the SW Zagros, and uniformly distributed crustal seismicity on predominantly 45° thrust planes. This model satisfies all four tests (i-iv). It disagrees with the data only in requiring all seismicity to occur at 60 km depth or less. Some events have been located deeper by Nowroozi (1971), but these may be in error. No attempt was made to reproduce them by weakening the lithosphere, because a strong lithosphere is required to stabilize and localize the deformation.

It only remains to determine the faulting stress of the upper crust. This parameter is independent of topography and crustal roots and cannot be directly calculated. However, limits can be obtained. If the shear strength of the crust were less than the value of τ_b calculated, then detachment would occur within the crust and not at its base (at a lesser value of ℓ). Since seismicity seems to extend at least to the base of the crust we can probably conclude that the faulting shear stress is above 100 bars.

Since we are discussing faulting stresses an order of magnitude below those determined from laboratory rock mechanics, the upper limit is of greater interest. This limit is obtained from the fact that the old subduction zone (the Crush Zone) must be no weaker than the surrounding crust. If it were, subduction would proceed as in model Z15 and there would be no earthquakes in the southwest. It is assumed that the same mechanism produces earthquakes at the same stress level in both places.

Figure 4.25 shows the profiles of shear stress and temperature in the Crush Zone when limestone is present and the faulting stress is 250 bars. At a depth of 42 km (at 4.7 cm/year) the material is warmed to 290°C and creep takes over, rapidly reducing the stress. If the faulting stress were raised, this temperature would be reached proportionately sooner (see other curve) and only a fraction of the upper 40 km would be as strong as the adjacent crust. Thus, if there is limestone in the Crush Zone, 250 bars is close to the upper limit on the faulting stress in the adjacent crust.

To obtain a firmer, higher limit it can be assumed that there are no limestones, but only crustal granites in the Crush Zone. We can also arbitrarily assume that the low creep constant just obtained applies only to lower continental crust and not to the upper crust subducted into the Crush Zone. These would be the most favorable conditions for obtaining a high faulting stress. This limit was found by trial and error using the technique of Appendix B. As shown in Figure 4.25, the faulting-creep transition occurs at 40 km depth when the faulting shear stress is raised to 765 bars.

To confirm that this is also an acceptable solution models Z28 and Z29 were computed. In both the Crush Zone elements were given the parameters of granite, and the only difference was that creep strength in the lower crust was reduced ten and five times respectively from Table 4.4 values. Z28 gave a more even distribution of seismicity in the appropriate band, but predicted a small amount in the Mesopotamian Trough which should be inactive. Z29 did not have this flaw, and was therefore more acceptable, although it predicts more of a concentration of seismicity in the NE Zagros than is observed. This model is shown in Figures 4.23 and 4.24, as an alternative to model Z25. The crustal shortening, sediment detachment, fault plane solutions, and stress patterns are almost identical. The most significant difference is that in model Z29 with the higher faulting stress, twice as much driving force is required to cause the convergence of the plates.

In an attempt to further discriminate between the two successful models, the energy dissipation in zones of plasticity was compared with the actual profile of energy radiation from earthquakes. This energy release was calculated from the empirical formula of Gutenberg (1951) for all the events in the period 1963-1973. While the profiles of energy release were similar in shape, the model figures were a thousand times greater than the energy released as seismic waves. Thus, the seismicity level does not favor either model. Apparently either the frictional dissipation in earthquakes far outweighs the seismic radiation, or the greatest fraction of deformation takes place in an aseismic, stable sliding mode. The alternative explanation would be that rock strengths are a thousand times less than in these models (0.2-0.8 bars), but the existence of mountainous topography at the surface makes this unacceptable.

4.7 Conclusions

The Zagros range is a very young continental convergence zone, at most 5 m.y. and more probably 1-2 m.y. old. It was formed by the collision of the stable continental shelf of the Arabian plate with the Andean-type overthrusting plate margin of central Iran.

According to gravity and surface wave evidence, deformation since the collision has involved shortening and thickening of the crust to create crustal roots of up to 16 km amplitude. Seismicity and fault plane solutions indicate that this process is continuing.

Finite element models show that the Zagros convergence is not presently being driven by the pull of a downgoing oceanic slab. The driving force must therefore be transmitted across Arabia as a deviatoric stress in the lithosphere. The slab has either detached and sunk, or been so warmed during slow convergence that its strength is negligible.

The sub-crustal portions of the Arabian and Eurasian lithospheric plates are not (everywhere) in contact across the plate suture. This implies that (in some places) the continental crust is in contact with the asthenosphere under the Crush Zone. The heat transferred through this contact is the most likely source of the young granites in the Crush Zone reported by Wells (1969). Below the central Zagros, S_n

propagation indicates an intact, shield-type lithosphere.

The shear stress causing detachment at the base of the crust is comparable to the gradient of normal stress produced by the gradient of isostatically-supported topography. Simple theoretical models imply a lower crustal creep strength 5 to 10 times lower than that predicted from creep of dry quartz alone. Finite element models confirm this range, with better results obtained near the lower limit. However, values obtained from finite elements need to be corrected upward some 50% because of the artificial stiffness of the elements spanning a 150°C temperature contrast. The most likely value of the lower continental crust creep parameter A (Table 4.4) is 40 to 80 bars. This implies either water-weakening of the crustal quartz or significant creep of other minerals not yet studied.

The crust also faults and produces earthquakes at stresses lower than laboratory results predict. If there is limestone at depth in the Zagros Crush Zone, the crustal earthquakes are produced by shear stresses of 100-300 bars. If only crustal granite is in the Crush Zone the upper limit is raised to 800 bars. In either case the implication is that some aspect of earthquake generation is not understood.

The lithosphere beneath the Zagros localizes the crustal deformation at the plate margin by resisting shear stresses of up to 1.4 kb without significant deformation. Thus the earthquake activation stress seems to be higher in the mantle than in the crust. In terms of mantle creep, this long-term strength implies an olivine creep activation energy of at least 100 kcal/mole. Some degree of water-weakening of the olivine cannot be ruled out, but it is not saturated as in experimental runs on specimens packed in talc.

The single new thrust fault in the Zagros proposed by Bird *et al.* (1975) is not observed in successful finite element models. Instead, crustal shortening occurs by more evenly distributed strain (on a number of minor faults). However, if actual rocks are in some way strain-weakening, the concentration of slip on a single surface could occur temporarily. On a longer time scale, the deformation is distributed over a belt whose width increases as the deformation front moves away from the suture. This implies that the dip of the Moho and the slope of the topography remain constant, as illustrated in Figure 4.26. The value of these dips is regulated by the creep strength of the lower crust.

The folding of the sedimentary layer is to some extent independent of basement shortening. At present the sediments are inactive in the NE half of the Zagros folded belt, and are shortening rapidly in the SW half. This implies that the sediments are in motion in a southwesterly direction with respect to the rest of the crust.

Models predict that all Zagros earthquakes should occur at depths of 60 km or less. To date, depths of Zagros events have not been determined with sufficient accuracy to test this.

Assuming the upper-limit faulting stress and negligible seismic radiation of energy, the maximum amount of frictional heating of any part of the Zagros is 20°C (except in the Crush Zone). The shear-strain heating at the point where Coster (1947) determined heat flow has been less than 10°C. This implies that crustal thickening reduces the geothermal gradient more than shear-strain can raise it, and that the original equilibrium heat flow of the region was closer to the 1.15 HFU determined at the Ahwaz oilfield. In the early stages of orogeny, temperature at a given depth is decreased, although the temperature of an element of rock will generally increase as it moves down.

The integral through the thickness of the Arabian lithosphere of the deviatoric horizontal stress may be termed the driving force of the orogeny. The present value is between 2.8×10^{15} and 5.5×10^{15} dynes/cm. The lower and upper limits correspond to models with faulting stresses of 250 and 750 bars shear respectively. This driving force is equivalent to a horizontal compression of 560 to 1100 bars distributed over a thickness of 50 km.

Table 4.1

Events used to determine Zagros Rayleigh wave group velocities

Date	Time	Latt.	Lon.	Magnitude	Δ to Shiraz
9-04-64	3:39:36.7	39.8	40.3	5.0	14.4°
8-31-65	7:29:45.8	39.3	40.8	5.1	13.7°
9-13-66	20:23:50.4	39.1	40.7	4.5	13.7°
10-20-67	6:47:38.0	37.9	37.7	4.8	14.9°
10-30-68	16:51:33.5	37.9	38.6	4.9	14.3°
7-02-70	2:24:35.7	38.8	36.7	4.8	16.1°
9-03-70	5:32:09.7	39.6	38.7	5.1	15.2°

Table 4.2

Seismic velocity model for Rayleigh wave group dispersion in the Zagros Mts.

Depth (km)	V_p (km/sec)	Vs	Density (g/cc)
0 - 9	5.20	3.00	2.60
9. - 27.3	6.20	3.58	2.70
27.3 - 45.6	6.40	3.67	2.87
45.6 -125.	8.17	4.65	3.40
125. -325.	8.17	4.45	3.40
325. - ∞	8.80	4.60	3.65

Note: These layers were subdivided in the calculations for group velocities and eigen-functions.

Table 4.3

Maximum Possible Subduction of Continental Crust

(Table of vertical displacement (in km) of the leading edge of continental crust subducting at 30° and 5 cm/year when detachment occurs.)

Mechanism limiting stress in shear zone	Nature of Lower Continental Crust	
	14% Quartz Granulite ($A = 4.44 \times 10^8$)	Weaker Rock ($A = 4.44 \times 10^7$)
Dry Granite melting	53 km	15km
Wet Granite melting	66	15
Granite creep	76	20*
Limestone creep	>100	25
Wet Granite creep	>100	31
200 bar limit	80†	22†

*Strength of granite on top of slab also reduced 90% for consistency.

†Faulting stress at trailing edge of crustal block also set to 200 bars for consistency.

Table 4.4

Mechanical parameters of initial Zagros finite element model

Flow law: $\tau^* = A\,(\dot{e}^*)^{1/n}\ \exp\left(\frac{B}{T}\right)$

$\tau^* \leq \tau_{max}$

where: $\tau^* \equiv \sqrt{1/4\,(\sigma_{xx} - \sigma_{zz})^2 + \tau^2_{xz}}$

$\dot{e}^* \equiv \sqrt{(\dot{e}_{xx} - \dot{e}_{zz})^2 + \dot{e}^2_{xz}}$

Rock	Density g/cc	Elements	A dyne/cm^2	B, °K	n	τ max
Peridotite	3.27	67-96	6.84x10^4	21710-22668	3.	1.5kb
Lower crust	2.84	30,31,33 36,41,42, 47,48,53, 54,59,60, 65,66	4.44x10^8	7445	3.	Appendix C
Middle crust	2.75	35,39,40, 45,46,51, 52,57,58, 63,64	4.00x10^8	7445	3	Appendix C
Upper crust	2.67	28,29,32, 34,37,38, 43,44,49, 50,55,56, 61,62	3.48x10^8	7445.	3.	Appendix C
Crush Zone Limestone	2.67	19-27	867.	12391.	3.	Appendix C
Evaporites	2.16	1-9	2.27x10^{11}	0	3.	-
Continental Shelf Sediments	2.67	10-18	1.0x10^{13}	0	3.	Appendix C

Table 4.5

Parameters and results of Zagros finite element models

Name	Figure	Boundary Conditions	Solution Technique	Lower Crustal A $\frac{\text{dyne}}{\text{cm}^2}$	Crustal τ max, bars	Mantle Qa kcal/mole	RMS un-converged stress bars	Velocity Solution	Maximum seismicity region x, km	Fault plane sense and percentage
Z04	-	Ridge-push	IS	4.4×10^8	$\mu\rho gz$	125	18%*	Pure subduction	120-210	90% thrust 10% normal
Z05	-	Ridge-push	IS	1.9×10^7	$\mu\rho gz$	125	7%*	Pure subduction	140-210	70% indeterminate 30% thrust
Z06	-	Ridge-push	IS	1.9×10^7	200	125	31%*	Pure subduction	150-210	70% indet. 30% thrust
Z11	-	R-P new grid	IS	1.9×10^7	200	125	35%*	Shortens in Persian Gulf	0-240	80% thrust 20% indet.
Z15	4.17	Slab-pull	VV	4.4×10^8	$\mu\rho gz$	125	33.	Pure subduction	110-210	100% thrust (shallow dips)
Z16	-	Slab-pull	VV	4.4×10^8	1000	125	18.	Pure subduction	110-210	100% thrust 30° NE dips
Z17	-	Slab-pull	VV	4.4×10^8	500	125	29.	Pure subduction	80-210	90% thrust 10% normal (in SE)
Z18	-	Slab-pull	VV	4.4×10^8	250	125	48.	Pure subduction	80-210	100% thrust 30% NE dips
Z19	-	Slab-pull	VV	4.4×10^8	250	94	17.	Pure subduction	90-200	80% thrust 10% indet. 10% normal
Z21	4.19	Ridge-push	VV	4.4×10^8	250	94	16.	Shortening in Persian G. minor subduction	350-450 (80-200)	100% thrust 45° dip in SW 30° dip in NE

* For initial-stress models, percentage of points with errors over 10% of stress is given.

Table 4.5, continued

Name	Figure	Boundary Conditions	Solution Technique	Lower Crustal A $\frac{dyne}{cm^2}$	Crustal τ_{max} bars	Mantle Qa kcal/mole	RMS unconverged stress, bars	Velocity Solution	Maximum seismicity region x, km	Fault plane sense and percentage
Z22	-	Ridge-push	VV	4.4×10^8	250	110	11.	Pure subduction	80-200	85% thrust 10% indet. 5% normal
Z23	4.18	Ridge-push	VV	4.4×10^8	250	100	19.	Pure subduction	80-200 (380-450)	85% thrust 10% indet. 5% normal
Z24	-	R-P new grid	VV	4.4×10^8	250	110	13	50% shortening in NE Iran; 50% subduction	0-260	95% thrust 5% indet.
Z25	4.21 4.22	R-P new grid †	VV	4.0×10^7	250	110	18	80% shortening in Zagros; 20% subduction	100-370	100% thrust 45° NE dips
Z26	4.20	R-P new grid †	VV	8.0×10^7	250	110	27.	70% subduction, 30% shortening in Zagros	100-250 (200-380 in sediments)	100% thrust 35° NE dips
Z27	-	R-P new grid †	VV	1.5×10^8	250	110	25	70% subduction, 35% shortening in Zagros	100-260 (200-380 in sediments	100% thrust 40° NE dips
Z28	-	R-P new grid †	VV	4.0×10^7	750	110	58.	90% shortening in Zagros, 10% in Persian Gulf.	150-450	100% thrust 45° NE dips
Z29	4.23 4.24	R-P new grid †	VV	8.0×10^7	750	110	38	80% shortening in Zagros, 20% subduction	130-370	100% thrust

† 0.11 cm/year uplift and 0.67 cm/year subsidence added to boundary conditions.

Chapter 4 - Figure Captions

Figure 4.1 Outline map of the Zagros after Stöcklin and Nabavi (1973): 1, Zagros Thrust and northeast limit of deformation; 2, southwest limit of deformation; 3, recent volcanoes; 4, earthquake epicenters 1950-1965 after Nowroozi (1971); 5, Tertiary volcanic rocks, mainly andesite; 6, Persian Gulf and various seas. Inset shows position of illustration.

Figure 4.2 Smoothed Bouguer gravity anomaly map of Iran after the U.S. Air Force Bouguer Gravity Anomaly Map of Asia. Contour interval 50 mGal. Large low is centered over the plate suture and the Zagros Crush Zone.

Figure 4.3 Gravity data from the Zagros and a proposed model. At bottom, dots show 1° by 1° averages of the Bouguer gravity anomalies from the region 46°-56° E projected onto a cross-section of the range. Model curve is calculated for a two-dimensional structure assuming the shape of the Moho shown at top and a crust-mantle density contrast of 0.43 g/cc. Area marked "Persian Gulf" is effect of low-density alluvium only.

Figure 4.4 Portion of the short period North-South record from the WWSSN station at Shiraz, Iran. Shows

P and high-frequency S_n arrivals from event of July 1, 1971 at 36.4°N, 43.4°E at a distance of 10.3° along the strike of the Zagros.

Figure 4.5 Rayleigh wave group velocities in the Zagros Mts. Contours in 5 decibel intervals of the summed spectral amplitudes of seven LPZ records at Shiraz. Event parameters are in Table 4.1. Model curve corresponds to earth structure of Table 4.2.

Figure 4.6 Earthquake epicenters (dots) and fault-plane solutions (circles) in the Zagros region projected onto a cross-section running NE-SW. Earthquake epicenters are from sections B_1 and B_2 of Nowroozi (1971). Some fault-plane solutions have been slightly rotated to make the nodal line perpendicular to the figure. Letters attached to solutions indicate the source: A for Akasheh (1973), M for McKenzie (1972) and N for Nowroozi (1972). The heavy outline shows the region modelled with finite elements. The dotted line shows the aseismic region deleted from later finite element models.

Figure 4.7 Predicted Rayleigh wave amplitude spectra at the station JER from the event of June 23, 1968. Fault-plane solution of Nowroozi

(1971) and Earth model of Table 4.2 are assumed. Different curves are for different possible source depths, labelled in kilometers. Absolute amplitudes of the curves have been adjusted for legibility.

Figure 4.8 Rayleigh waves and spectra of the June 23, 1968 Zagros earthquake at Addis Ababa and Jerusalem. Bar adjacent to waveform indicates the length of one minute. Smooth model curves compared to actual spectra in each box are adjusted to the same arbitrary amplitude.

Figure 4.9 Rayleigh wave and spectrum from the Sept. 14, 1968 Zagros earthquake at Lahore. Bar indicates length of one minute. Smooth curve in box is the calculated model spectrum, adjusted to the same arbitrary amplitude as the data.

Figure 4.10 Bottom-hole temperatures from 35 deep oilwells in the Zagros. Curve is not statistically determined, but shown as a reference. It assumes a conductivity of 5.2×10^{-3} cal/cm-sec-°C.

Figure 4.11 Bottom-hole temperatures in six Zagros wells and curves fitted by least-squares, subject to surface-intercept constraints. Parameters desribed in the text.

Figure 4.12 A tectonic and thermal model which could explain crustal thickening and heat flow anomalies in the Zagros. The thermal calculation assumes that the subduction zone on the left became inactive one million years ago, and that a fault formed at the right and has slipped 27 km. Frictional heating described in the text. For complete geotherm parameters see Bird et al. (1975).

Figure 4.13 Curve of predicted heat flow of the model in Figure 4.12 compared with heat flow estimates from Zagros wells (small dots). Large dot is average of heat flow measurements by Coster (1947).

Figure 4.14 Schematic thermal regime of the Zagros according to the hypothesis of horizontal crustal shortening without major thrusts or frictional heating. Heat flow data at top are the same as in Figure 4.13. Dotted line is the Moho, which merges into the 500°C isotherm. Thermal structure of the subduction zone is from the 5 cm/year limestone model of Figure 2.7.

Figure 4.15 Schematic diagram of subduction of a continental margin after the detachment of the oceanic slab. Various forces acting on a block of continental crust in different directions are identified. Angles and dimensions used in section 4.5 are defined.

Figure 4.16 Schematic geology of the Zagros in a NE-SW cross-section and the grid of plane strain finite elements used to represent it. Engineering symbols around grid show boundary conditions of the successful model. The dotted portion of the grid was dropped after it was determined that there was no strength in this area. There is no vertical exaggeration.

Figure 4.17 Output of the unsuccessful model Z15. At top, velocity vectors of nodes are shown, with the tail of the vector located at the node. Motionless nodes are shown by crosses. In center, the outline represents the area of the finite element model. Line segments show the direction of principal compressive stress (σ_1). Length of line shows maximum shear stress ($\sigma_1 - \sigma_3$)/2. At bottom, energy dissipation in plastic regions is illustrated, with the size of the symbol proportional to shear stress and strain rate. The two possible fault planes at each point are shown. Model parameters are in Table 4.5.

Figure 4.18 Output of the unsuccessful model Z23, in which the mantle creep activation energy was reduced to 100 kcal/mole. For explanation of symbols see caption above. This change has the effect of causing some shortening in the far southwest, under the Mesopotamian Trough. Model parameters are in Table 4.5.

Figure 4.19 Output of the unsuccessful model Z21, in which the mantle creep activation energy was reduced to 94 kcal/mole. Complete model parameters are in Table 4.5, and figure symbols are explained in caption 4.17.

Figure 4.20 Output of the unsuccessful model Z26. The faulting indicated in the upper 10 km is in the sedimentary layer and is not an essential part of the model. Excepting this, there is little seismicity predicted in the southwestern half of the Zagros. Model parameters given in Table 4.5 and figure symbols explained in caption 4.17.

Figure 4.21 Velocity vectors and deviatoric stresses from the successful model Z25. Note independent shortening of surface sediments at a greater rate than shortening of the basement in the SW half of the range. Model parameters in Table 4.5 and figure symbols explained in caption 4.17.

Figure 4.22 Predicted seismicity and fault planes from model Z25 (top) compared with actual data (bottom). The faulting in the top 10 km of sediments may or may not occur and is not as essential part of the model. For explanation of the seismicity and fault-plane solutions see caption 4.6.

Figure 4.23 Velocity vectors and deviatoric stresses from the successful model Z29. Model parameters in Table 4.5 and explanation of figure symbols in caption 4.17.

Figure 4.24 Predicted seismicity and fault planes from model Z29 (top) compared with actual data (bottom). For explanation of the sources of data see caption 4.6.

Figure 4.25 Temperature and shear stress in the plate suture dipping NE from the Crush Zone. Limestone in the suture causes lower stresses (Z25). If no limestone is present the crustal granites may support higher stresses (Z29). The upper stress limit is fixed by the requirement that creep shall not occur at less than 40 km depth. The onset of creep rapidly reduces the strength.

Figure 4.26 Schematic illustration of the past history of the roots of the Zagros suggested by successful models. Successive positions of the Moho shown as 1, 2, 3. Crust enters from the right, is horizontally shortened, and thickened. Two times vertical exaggeration.

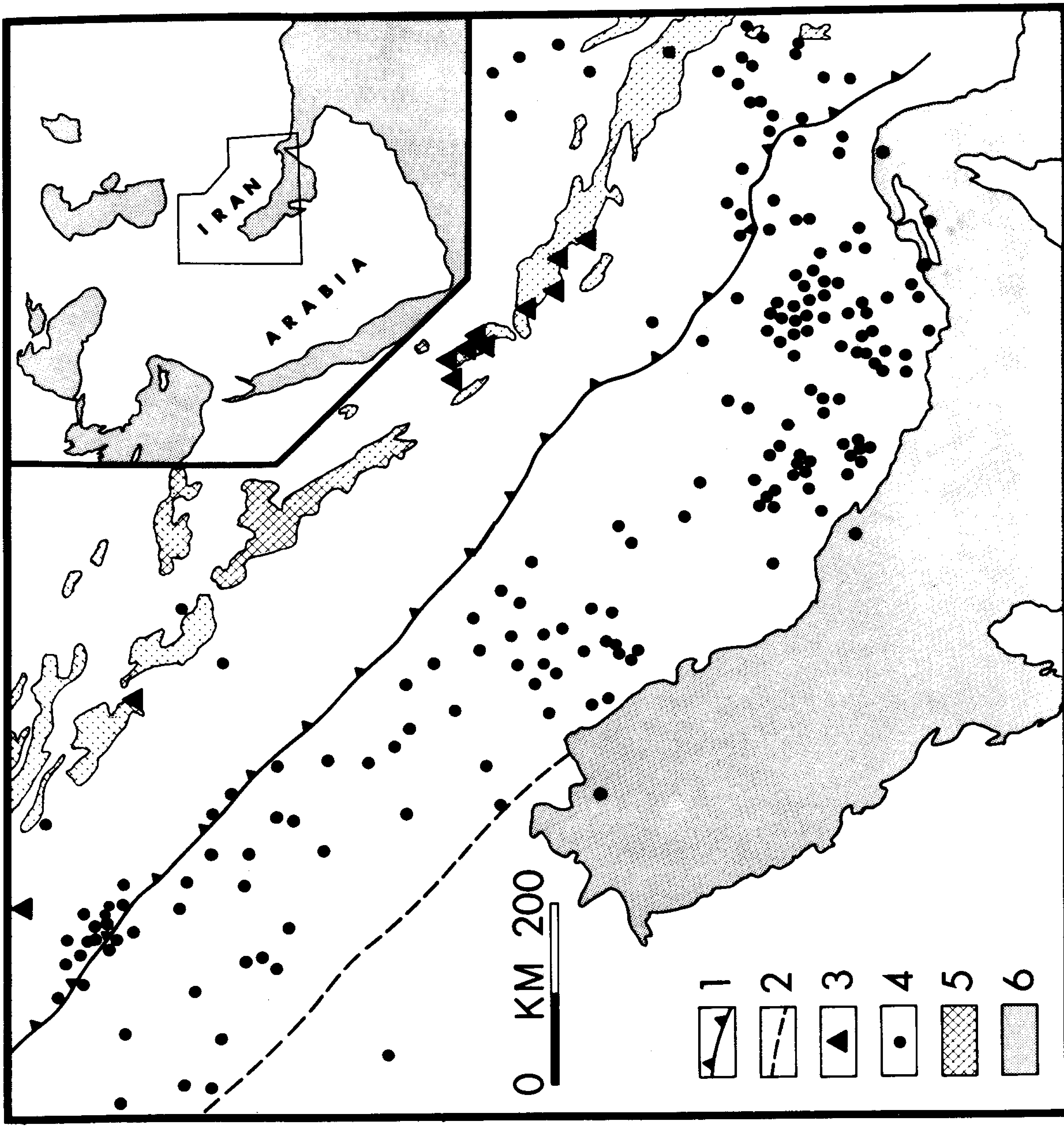


Fig. 4.1

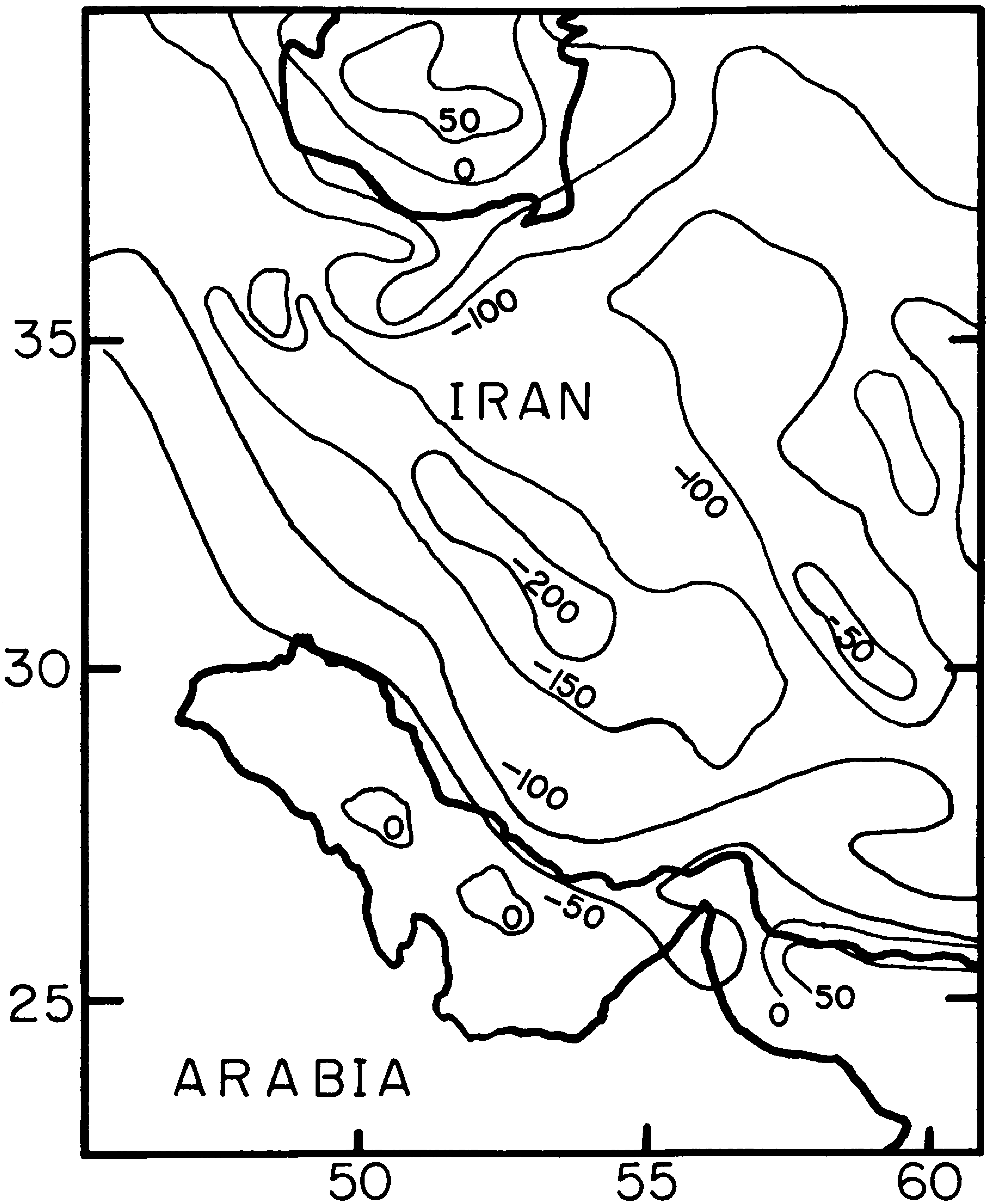


Fig. 4.2


500 KM
400
ZAGROS THRUST
200
100
0
30°
PLATE SUTURE
ASSUMED NORMAL THICKNESS
MOHO
25
50
50 KM
ZAGROS
THRUST
FIRST
DEFORMATION
500 KM
400
200
0
CENTRAL IRANIAN
PLATEAU
CRUSTAL THICKENING
IN ZAGROS
PERSIAN G.
TOTAL BOUGUER ANOMALY
-100
-200
-300
MILLIGALLS

Fig. 4.3

Fig. 4.4

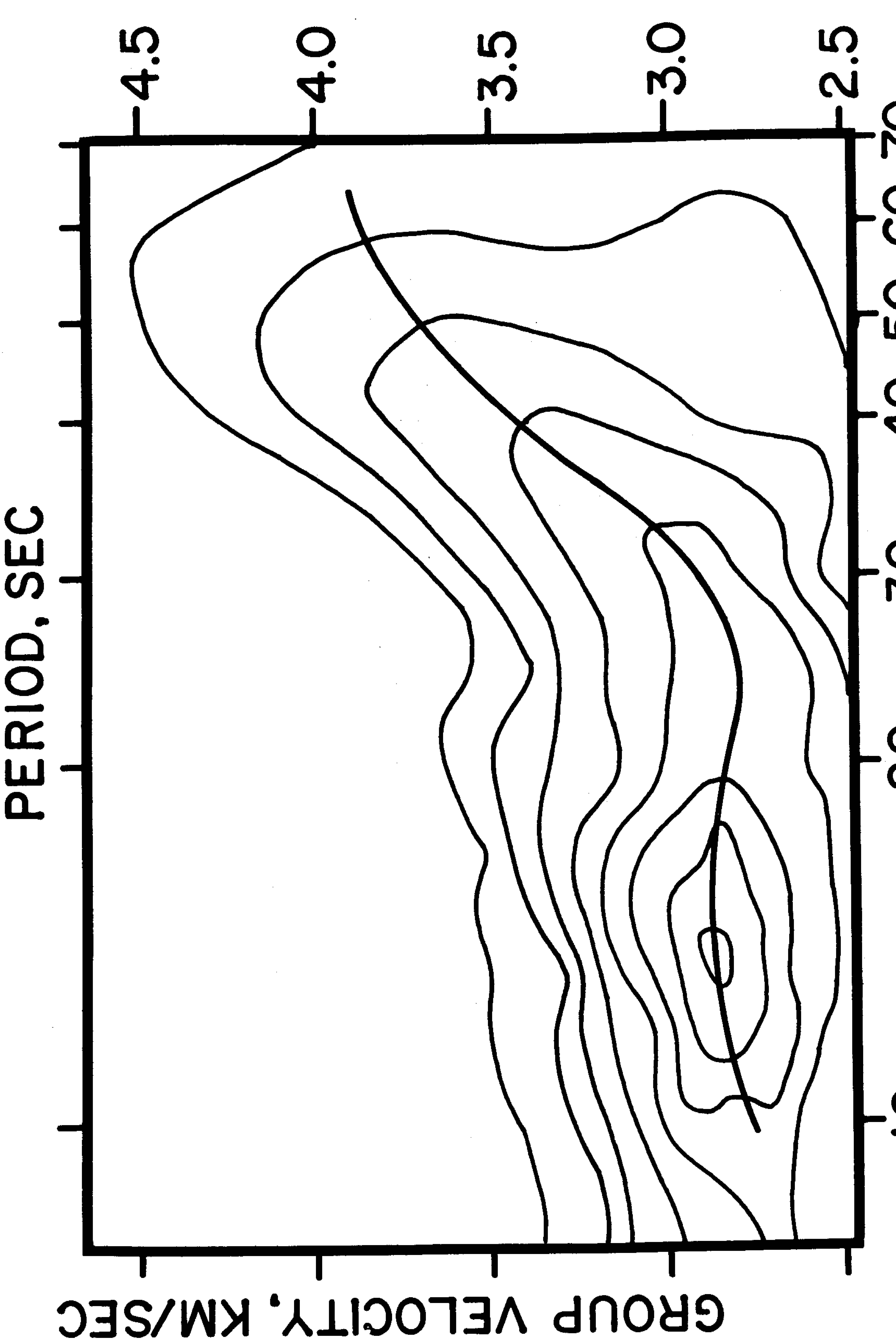

PERIOD, SEC
10
20
30
40
50
60
70
GROUP VELOCITY, KM/SEC
4.5
4.0
3.5
3.0
2.5

Fig. 4.5

0
50
100
KM
Persian
Gulf
SW
A
Mo
Mj
N2
Mf
N22
Mh
Mb
Zagros Thrust
NE

Fig. 4.6


10
20
30
40
50
PERIOD, SEC
0
3
6
9
12
22
32
50
LOG OF SPECTRAL DENSITY
ONE DECADE
JER LPZ
6-23-68
29.8N 51.2E
ZAGROS MODEL

Fig. 4.7


AAE LPZ 6-23-68
ONE DECADE
50
10
20
30
40
50
JER LPZ 6-23-68
ONE DECADE
50
10
20
30
40
50
PERIOD, SEC

Fig. 4.8

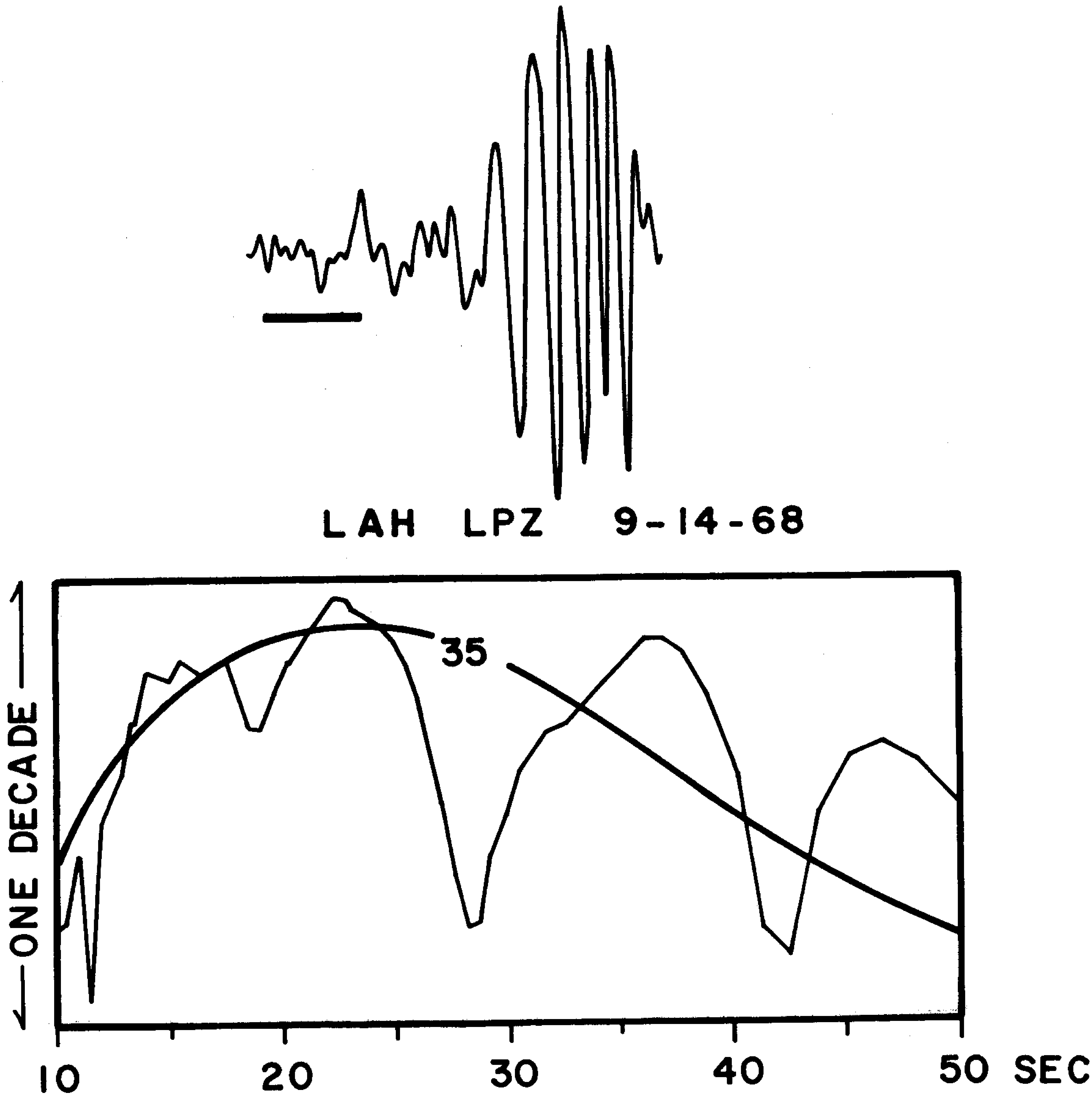


Fig. 4.9

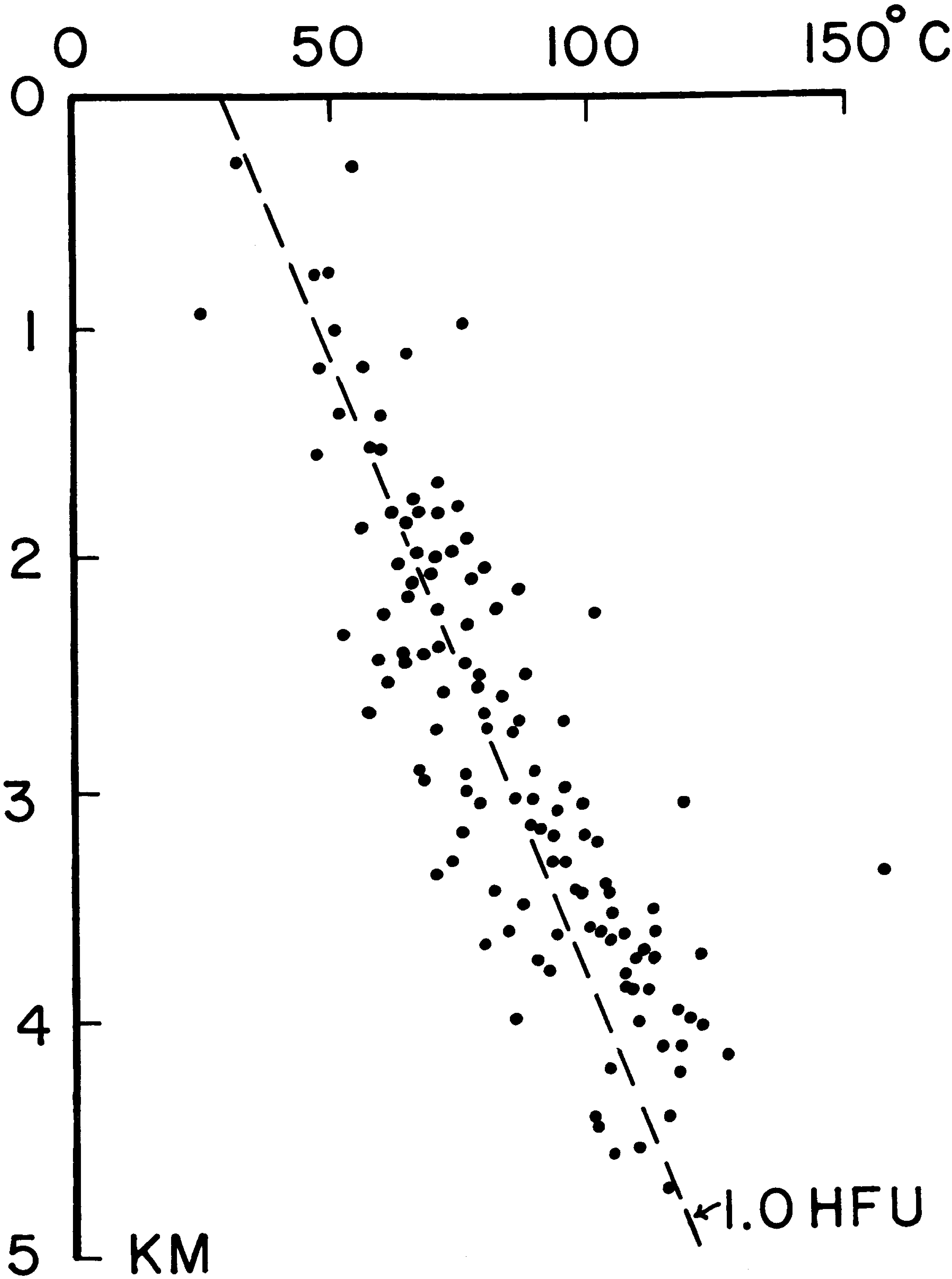


Fig. 4.10

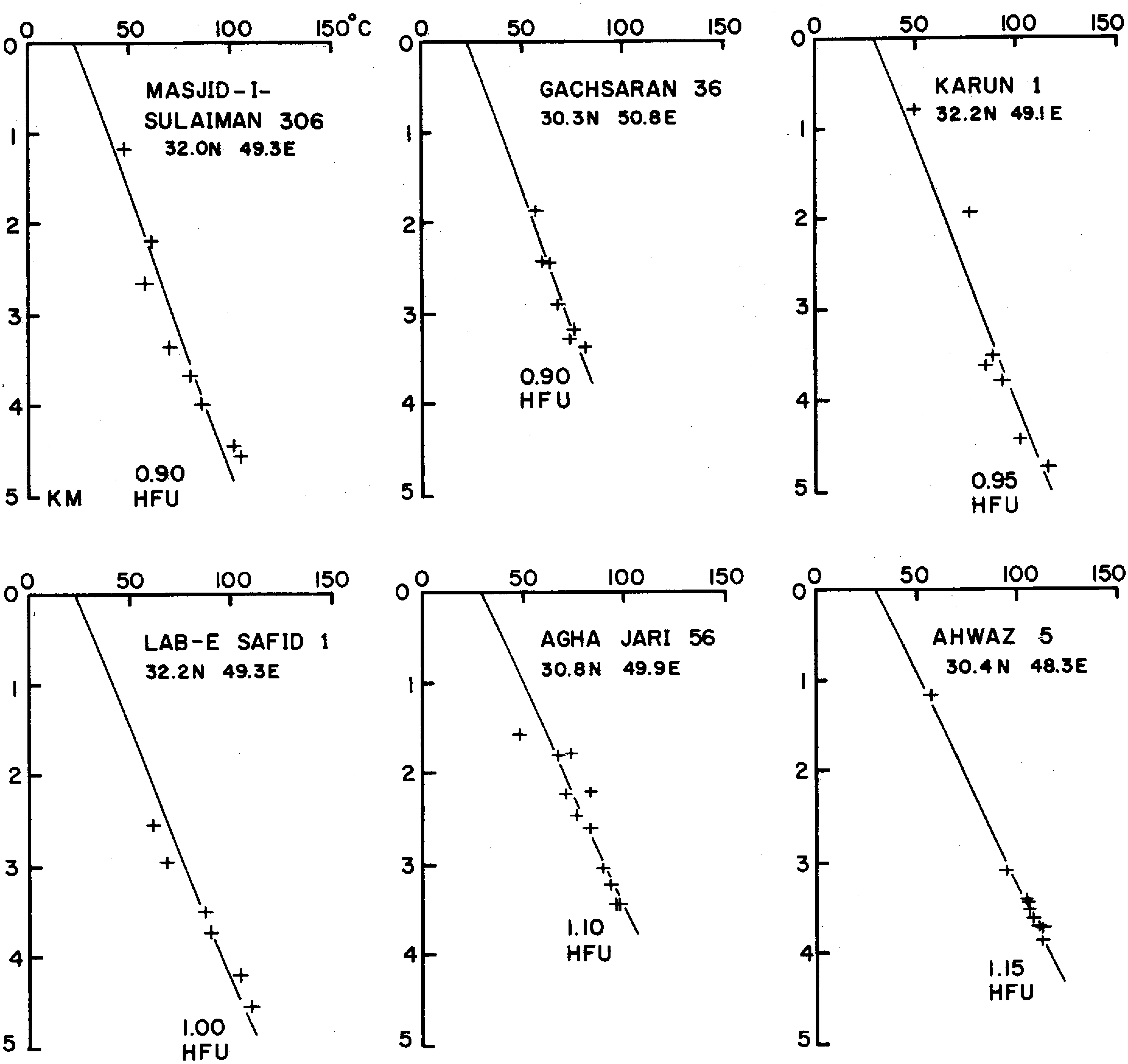
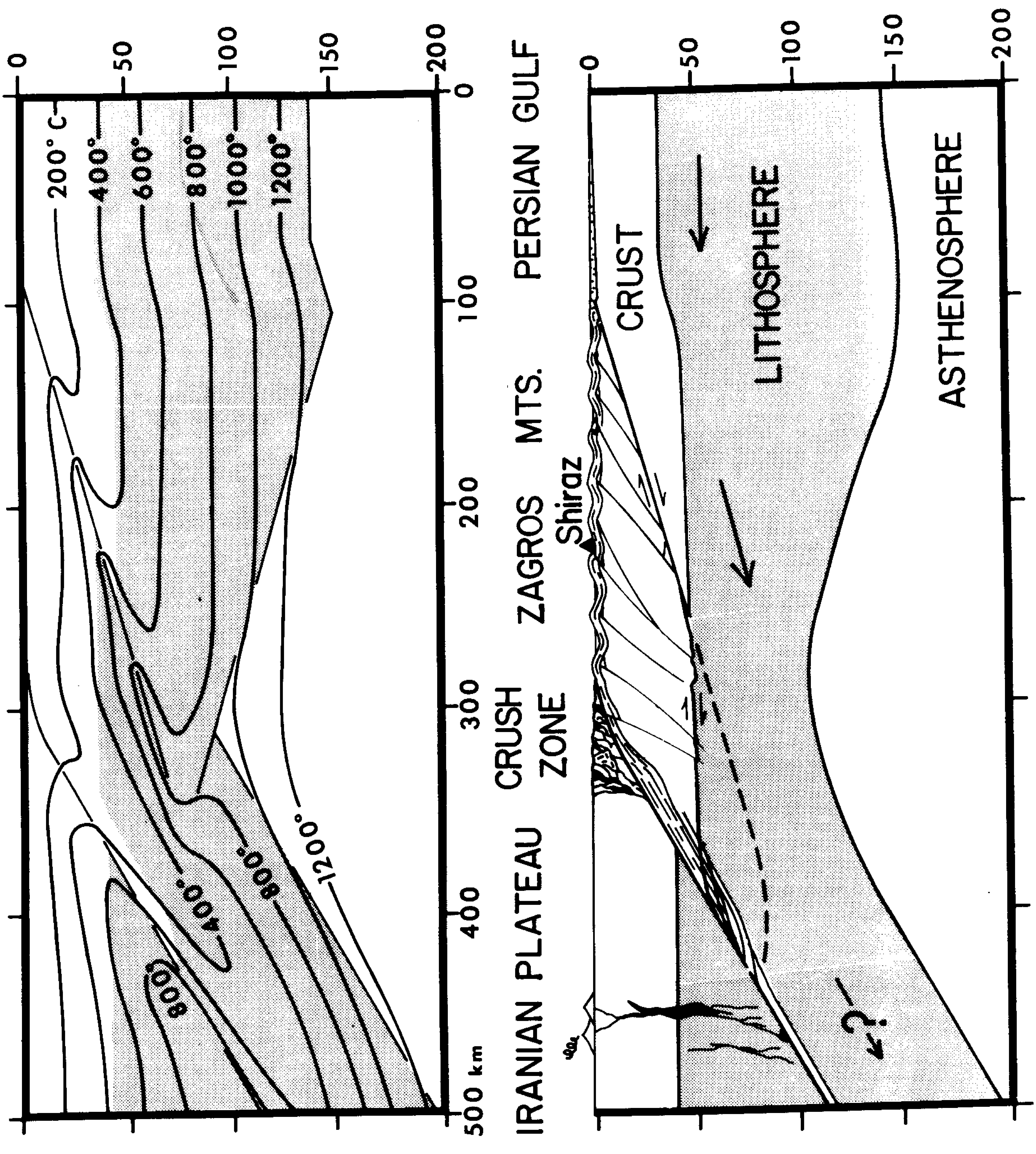

0 50 100 150°C
0 1 2 3 4 5 KM
MASJID-I-
SULAIMAN 306
32.0N 49.3E
0.90
HFU
GACHSARAN 36
30.3N 50.8E
0.90
HFU
KARUN 1
32.2N 49.1E
0.95
HFU
LAB-E SAFID 1
32.2N 49.3E
1.00
HFU
AGHA JARI 56
30.8N 49.9E
1.10
HFU
AHWAZ 5
30.4N 48.3E
1.15
HFU

Fig. 4.11



PERSIAN GULF
ZAGROS MTS.
CRUSH ZONE
IRANIAN PLATEAU
Shiraz
200° C
400°
600°
800°
1000°
1200°
CRUST
LITHOSPHERE
ASTHENOSPHERE
?
0
100
200
300
400
500 km
0
50
100
150
200

Fig. 4.12


H.F.U.
1.5
1.0
0.5
0
AHWAZ
AGHA JARI
GACHSARAN
KARUN
COSTER (1947)
LAB-E SAFID
HEAT FLOW
ERGS / CM2-SEC
60
40
20
0
0
100
200
300
400
500 KM
FIRST DEFORMATION
ZAGROS THRUST

Fig. 4.13


HEAT FLOW
1.5
1.0 HFU
0.5
0
ZAGROS ▼ THRUST
FIRST ▼ DEFORMATION
100°
300°
500°
700°
900°
1100°
1300° C
NE
SW
0
100
200
300
400 KM
0
50
100
150

Fig. 4.14


x
friction
θ
differential
pressure
CRUST
ℓ
creep
stress
LITHOSPHERE
ASTHENOSPHERE
β
friction
z

Fig. 4.15

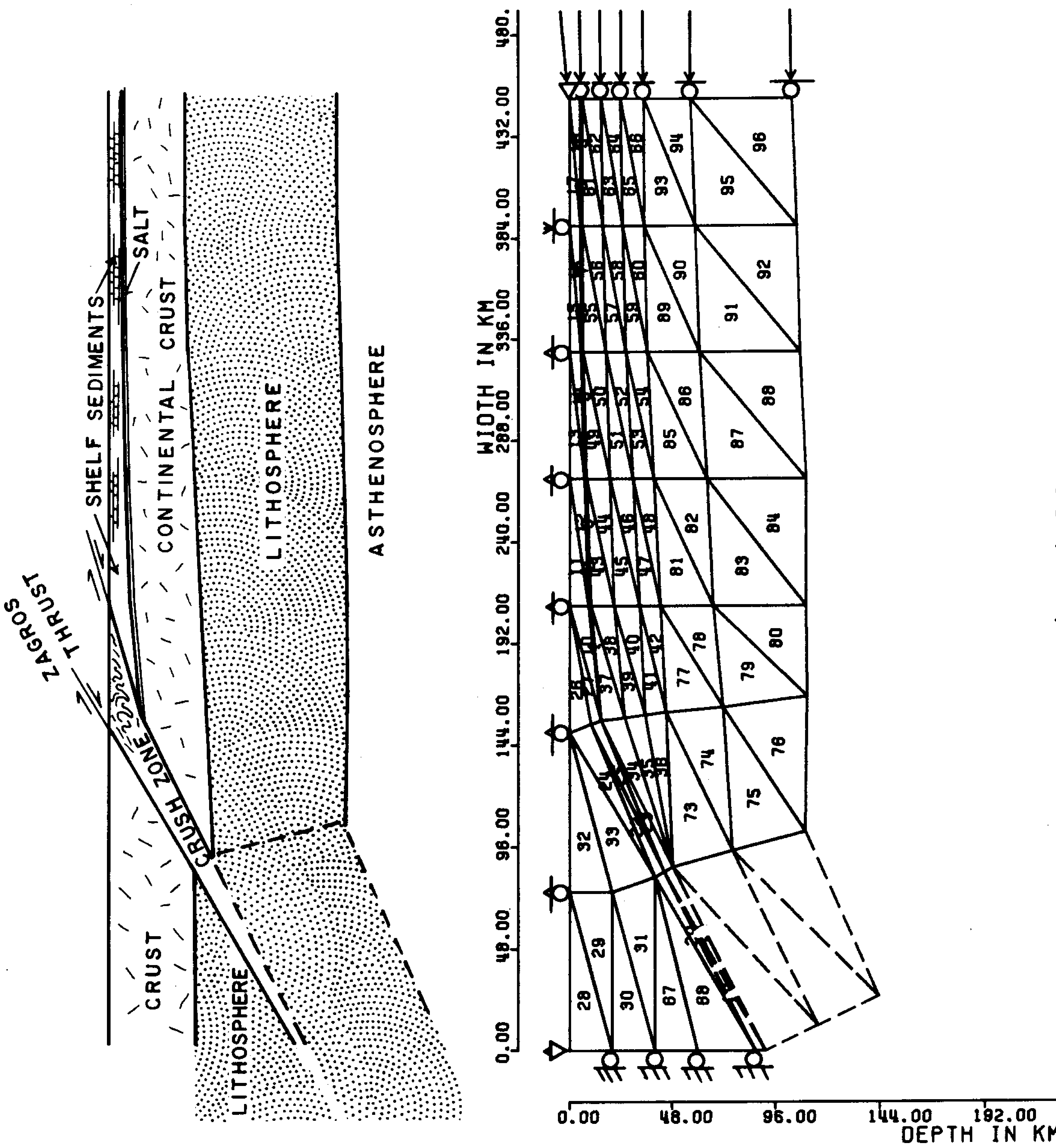

ZAGROS THRUST
SHELF SEDIMENTS
SALT
CONTINENTAL CRUST
CRUSH ZONE
LITHOSPHERE
ASTHENOSPHERE
CRUST
LITHOSPHERE
WIDTH IN KM
DEPTH IN KM

Fig. 4.16

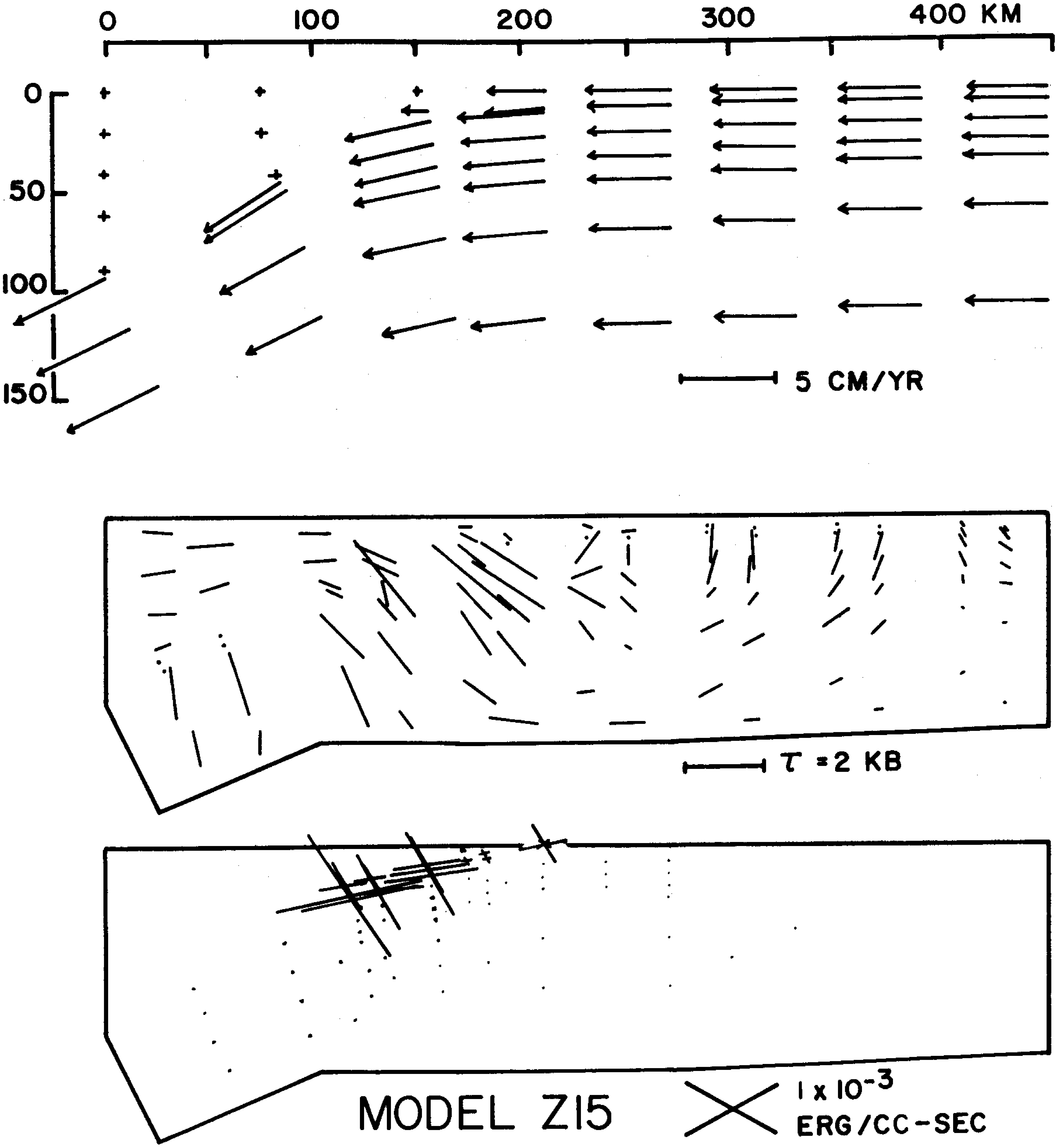

0
100
200
300
400 KM
0
50
100
150
5 CM/YR
τ =2 KB
MODEL Z15
1 x 10^{-3}
ERG/CC-SEC

Fig. 4.17

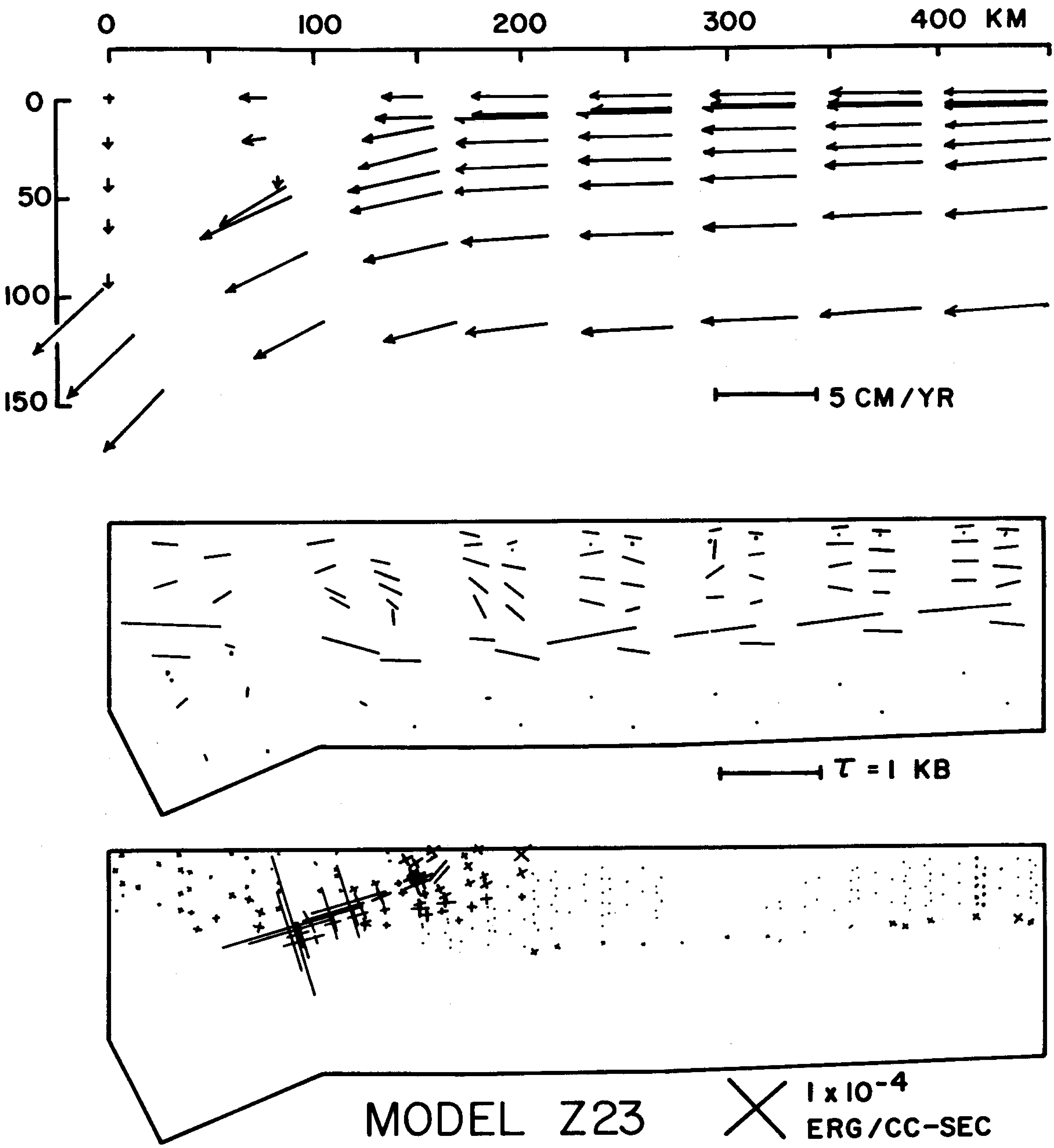

0
100
200
300
400 KM
0
50
100
150
5 CM/YR
τ = 1 KB
MODEL Z23
1 x 10^-4
ERG/CC-SEC

Fig. 4.18

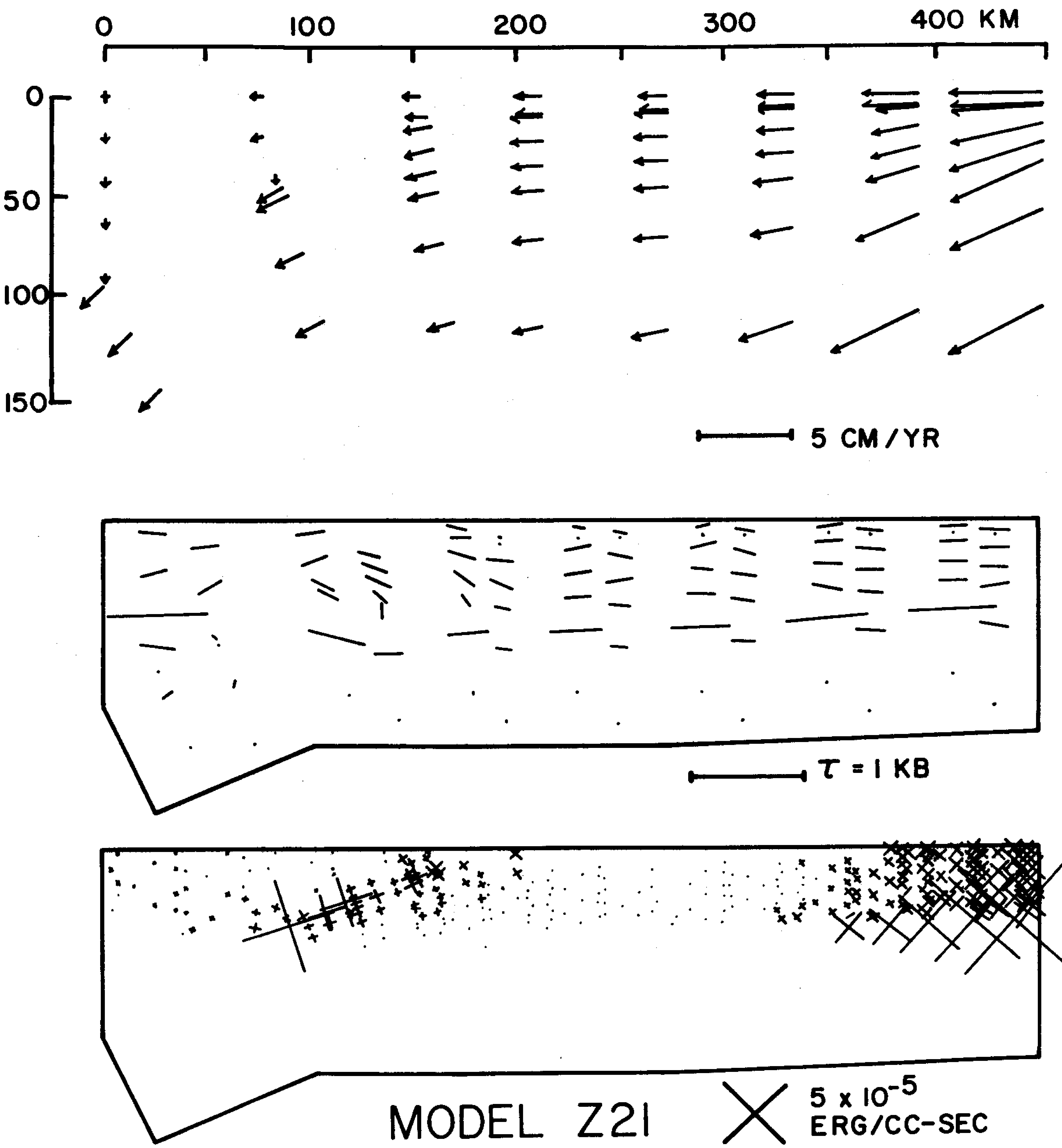
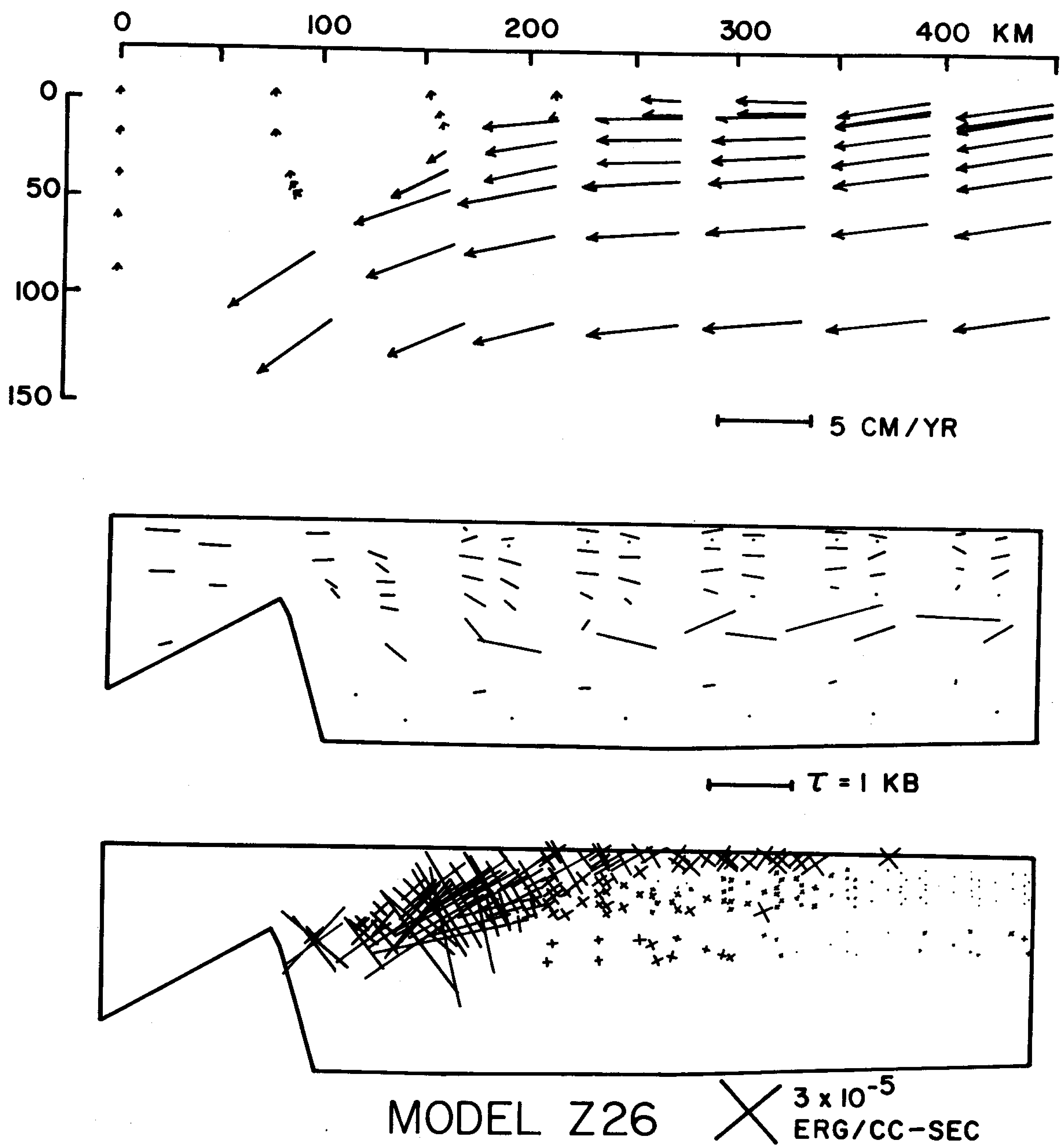
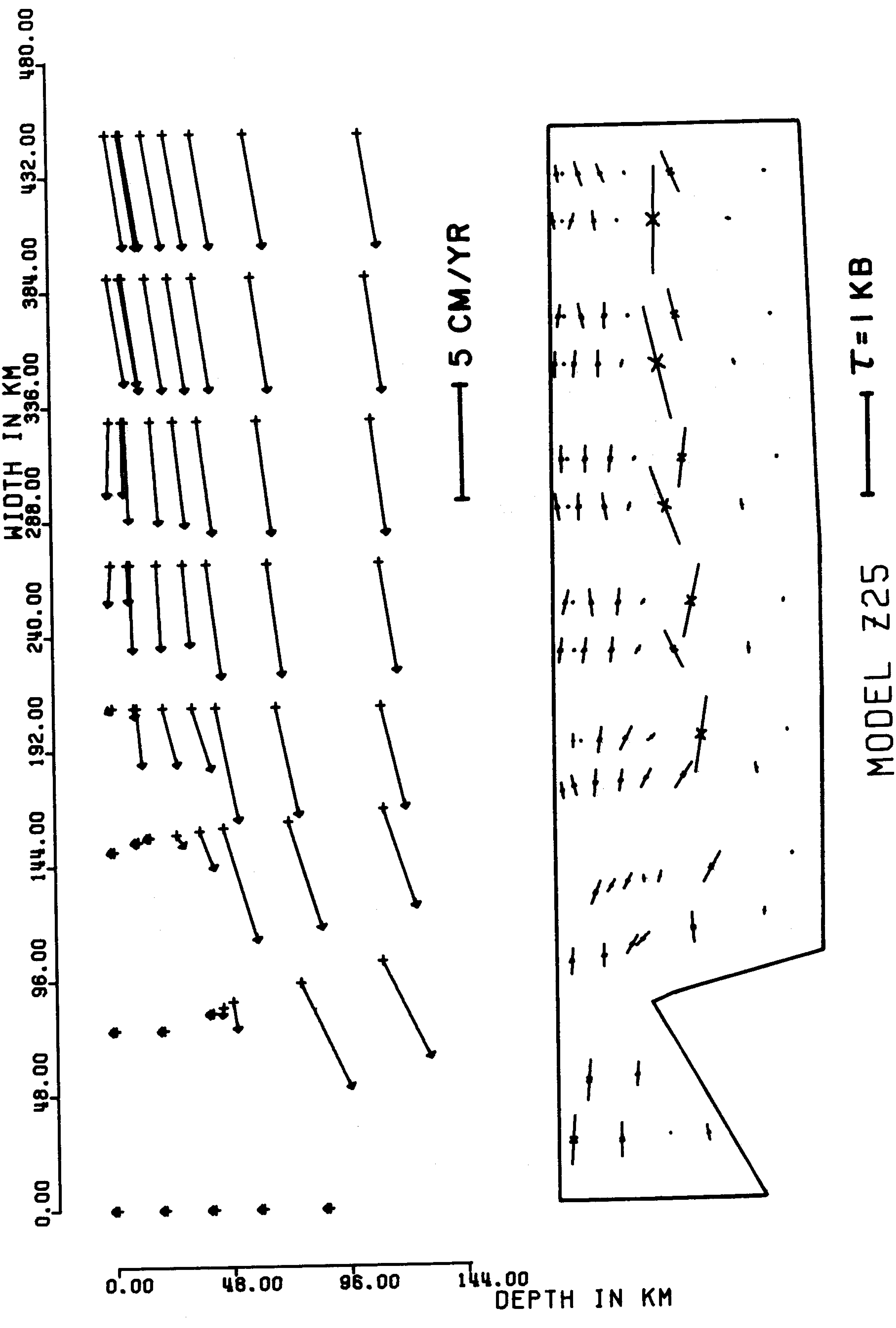

0
100
200
300
400 KM
0
50
100
150
5 CM / YR
τ = 1 KB
MODEL Z21
5 x 10^-5
ERG/CC-SEC

Fig. 4.19



0
100
200
300
400 KM
0
50
100
150
5 CM/YR
τ = 1 KB
MODEL Z26
3 x 10^{-5}
ERG/CC-SEC

Fig. 4.20



WIDTH IN KM
0.00
48.00
96.00
144.00
192.00
240.00
288.00
336.00
384.00
432.00
480.00
DEPTH IN KM
0.00
48.00
96.00
144.00
5 CM/YR
τ=1 KB
MODEL Z25

Fig. 4.21

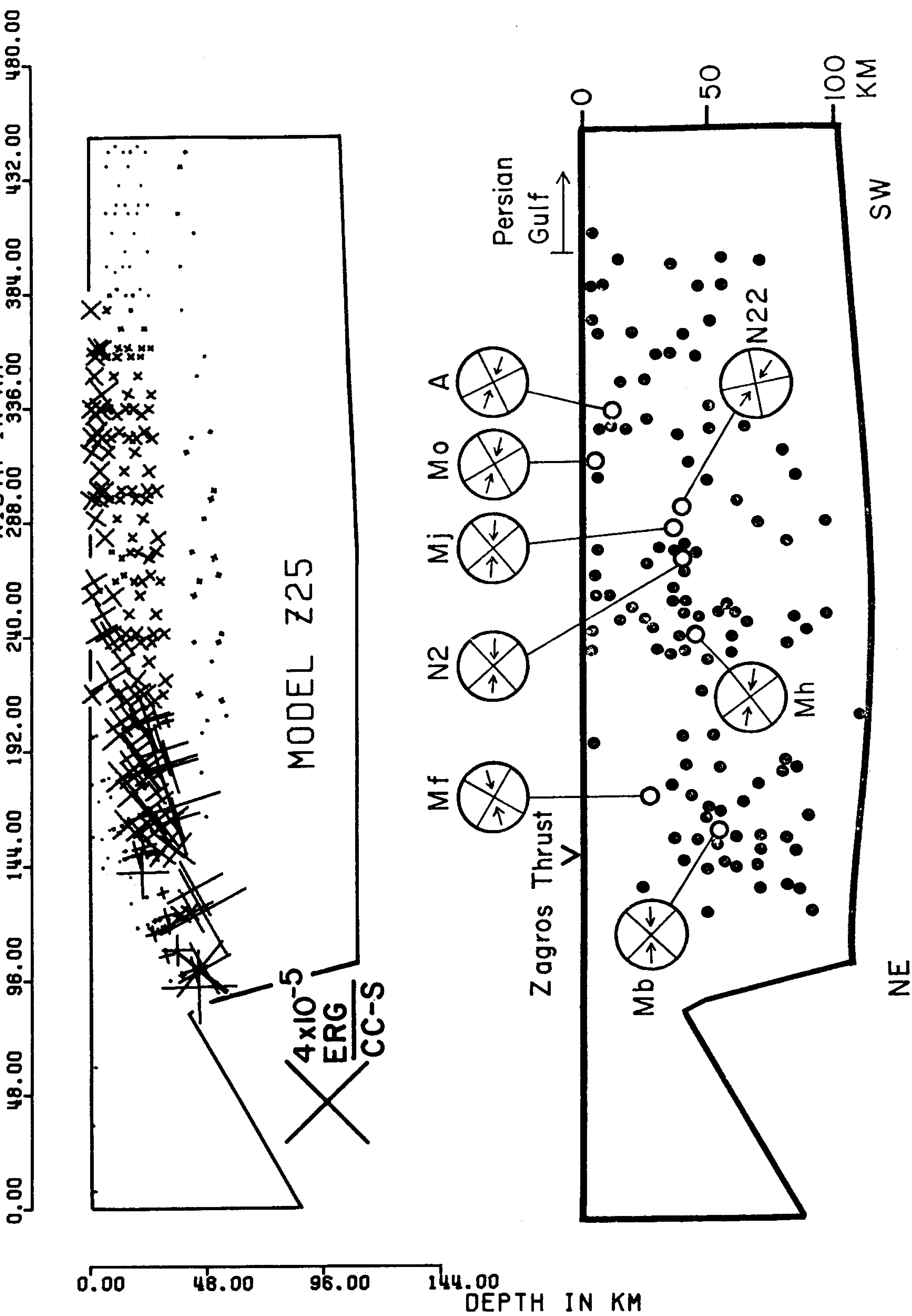

WIDTH IN KM
0.00
48.00
96.00
144.00
192.00
240.00
288.00
336.00
384.00
432.00
480.00
MODEL Z25
4x10-5
ERG
CC-S
0.00
48.00
96.00
144.00
DEPTH IN KM
0
50
100
KM
Persian
Gulf
SW
NE
Zagros Thrust
A
Mo
Mj
N2
Mf
N22
Mh
Mb

Fig. 4.22


WIDTH IN KM
0.00
48.00
96.00
144.00
192.00
240.00
288.00
336.00
384.00
432.00
480.00
DEPTH IN KM
0.00
48.00
96.00
144.00
5 CM/YR
τ = 1 KB
MODEL Z29

Fig. 4.23

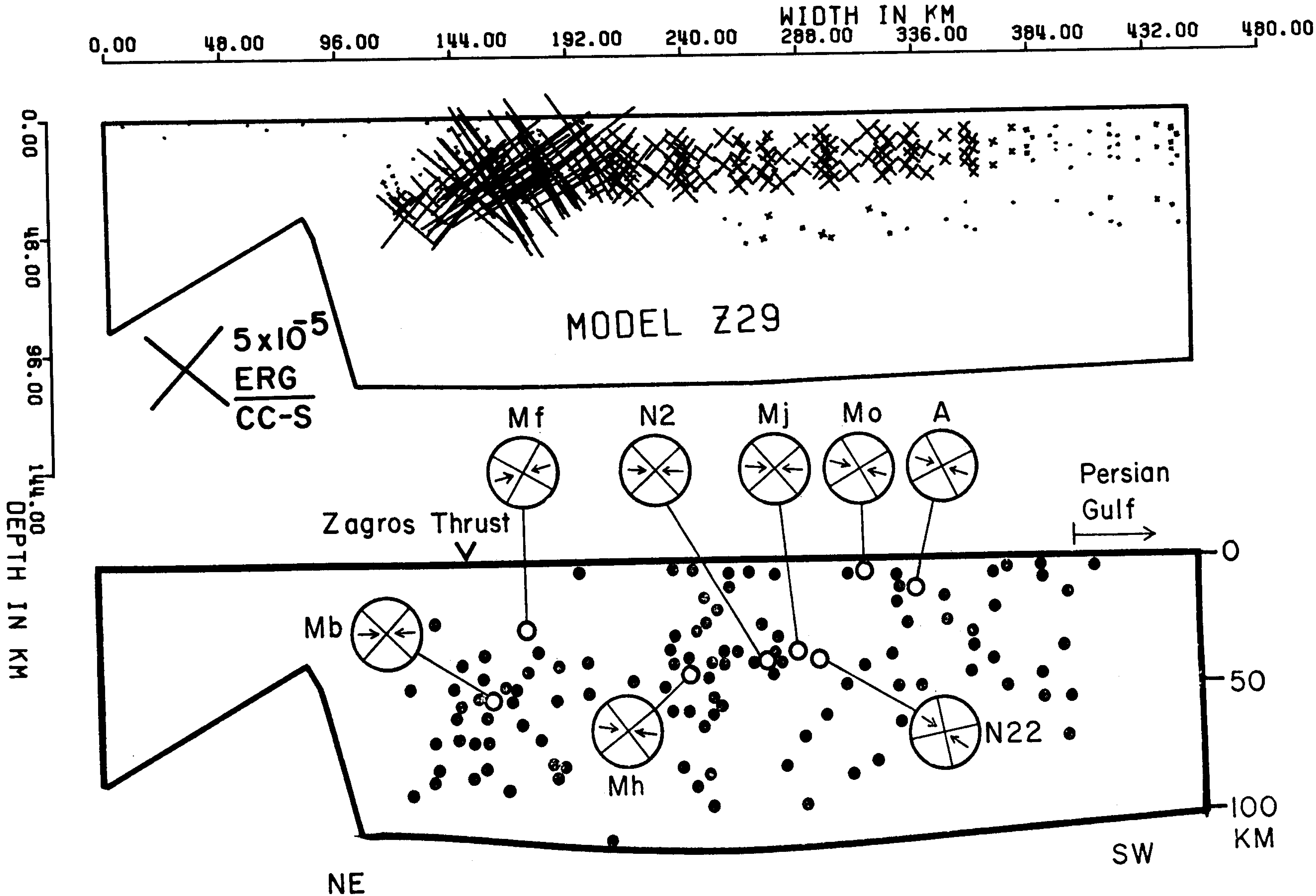

WIDTH IN KM
0.00
48.00
96.00
144.00
192.00
240.00
288.00
336.00
384.00
432.00
480.00
DEPTH IN KM
0.00
48.00
96.00
144.00
MODEL Z29
5x10-5
ERG
CC-S
Mf
N2
Mj
Mo
A
Persian
Gulf
Zagros Thrust
Mb
Mh
N22
0
50
100
KM
SW
NE

Fig. 4.24


800
600
400
200
0
GRANITE
LIMESTONE
SHEAR STRESS IN BARS
0
10
20
30
40
50
500
400
300
200
100
0
GRANITE
LIMESTONE
TEMPERATURE IN °C
DEPTH IN KM

Fig. 4.25

Fig. 4.26

CHAPTER 5: THE HIMALAYAS: ADVANCED CRUSTAL OVERTHRUSTING

The Himalayas today are in the mature stage of continental convergence, with topography, crustal roots, metamorphism and plutonism all well developed. There is general agreement among geologists that the past deformation took the form of detachment of a great sheet of Indian continental crust and its overthrusting to the South over adjacent crust. It is less clear just what is happening at the present; whether the Boundary Fault is merely a branch of the Central Thrust, or represents the start of a new cycle of overthrusting. The rate of convergence in the range is unknown, although this number is important to the understanding of tectonics in China and Siberia. And two major problems of past tectonics remain unresolved: why was the formation of the mountains delayed some 25 m.y. following the continental collision? And what was the heat source that metamorphosed the range and melted the High Himalayan granites?

This chapter and the following one attempt to answer these questions with thermal models of the past and mechanical models of the present. After a brief introduction to the structure of the range, all the available quantitative geologic constraints on its past history are assembled. These are used to develop three different thermal models of the range differing in the mechanical properties and past history of the mid-crustal shear zone. These become inputs into a finite-element treatment of the present deformation which yields the present stress and shear-strain heating in that zone. Concluding that the Himalayan granites were not melted by "uniformitarian" processes (those operating today), the last section examines two "catastrophic" mechanisms: discontinuous plate velocities and disruption of the sub-crustal lithosphere.

5.1 Tectonic Outline of the Range

In broad terms the Himalayas are a diffuse convergent plate boundary between the rigid Indian plate and the deformed Eurasian plate. The range curves in a gentle arc 2500 km long from the Pamir to the Indo-Burman syntaxis (Figure 5.1). At each end the plate boundary turns south and takes on the strike-slip motion of a continental transform fault (Molnar and Tapponnier, 1975). Throughout this length the width of the mountain belt is a uniform 200-250 km, and the continuity of the major tectonic units is remarkable.

One exceptional area is the northwest Himalayas from 73-80° East. Here the Central Thrust and the Boundary Fault become very close and almost merge. Across the plate suture, there is the Ladakh Granite mass which may be surrounded by ultramafic ophiolites (Gansser, 1964) and could have been a separate micro-continent trapped in the collision. Behind this there is the Karakorum fault, the only major strike-slip fault (Molnar and Tapponier, 1975) recognized in the Himalayan region. And finally, there is the possible lithospheric slab dipping steeply to the NW beneath the Hindu Kush (Kisslinger _et al_., 1974; Khalturin _et al_., 1975). The tectonics of this region are very complex, especially as seismic data (Armbruster _et al_., 1975) from the Tarbella array indicate structural trends with no surface expression.

In this thesis, only the simpler central and eastern Himalayas are analyzed, so that the simpler techniques of two-dimensional modelling can be used.

From 80-95° East the Himalayas can be divided into five structural belts which have both geologic and topographic boundaries. Starting on the north, the first is the Trans-Himalayan range of southern Tibet. This terrain consists largely of extrusive dacites, andesites, and basalts mapped on the ground (Hennig, 1915). Large rounded batholithic features recognized on satellite photographs may contain their intrusive equivalents (Kidd, 1975). Both have intruded a sedimentary section of Cretaceous marine sediments consisting of Jura-Neocomian sandstones and phyllites and Gault-Cenomanian limestones. Tectonically, the area was probably a quiet continental margin facing the Tethys on the southwest until the upper Cretaceous, at which time that ocean may have begun to subduct beneath it. Since the collision, Eocene Kailas conglomerates have been deposited in southern Tibet from northern sources. The region averages 5 km in elevation, and is terminated on the south by steeply dipping thrusts.

Next, to the south of Tibet, is the narrow Indus Suture zone, characterized by steeply dipping thrusts, chaotic melanges, and ultramafics. The formation is collectively called the Indus Flysch, and consists of shales, radiolarites, and exotic blocks of limestone with no apparent source

(Gansser, 1964). Although the basic rocks have some intrusive contacts, this does not necessarily imply that they were intruded in their present position. Dewey and Bird (1970) consider them to be ophiolites, scraped from the oceanic lithosphere in the trench of a subduction zone and mixed with a melange of turbidites. The complete mismatch of geology on the two borders of this zone clearly marks the plate suture where the initial collision of continents occurred. The suture zone averages 4.5 km in elevation, but is subjected to heavy local erosion by the Indus and Brahmaputra rivers which follow the fault lines of weakness.

The High Himalayas next to the South include a thick conformable marine sequence of Cambrian through Eocene shallow-water deposits. During deposition this sequence was affected only by mild epeirogenic movements corresponding in time to the Hercynian orogeny (LeFort, 1975). Since the collision they have been mildly deformed into open Jura- or Zagros-type folds with minor thrusting (Figure 5.2). Apparent gravity slides of Indus flysch have been locally emplaced over the sequence (Gansser, 1964). The average regional dip is gently to the North, so that in the southern half of the High Himalayas the pre-Cambrian metamorphic basement is exposed and forms the highest peaks. In many places this basement is intruded by syntectonic tourmaline granites (Figure 5.2) which are associated with a normal metamorphism that dies out in intensity in the lower sedimentary units. The High Himalayas are the former continental shelf of India, and not a part of Eurasia. The evidence for this is summarized by Saxena (1971), and includes the occurrence of Gondwana sediments with the Gangamopteris flora in Kashmir and the Blaini boulder beds (possible tillites from the Permian glaciation of Gondwanaland). The boundary of this zone is the shallow-north-dipping Main Central Thrust, which can be traced continuously from Kashmir to Sikkim.

Below the Central Thrust are the Lesser Himalayas, the foothill belt. Sediments here are almost completely non-fossiliferous and cannot be correlated with those of the High Himalayas. Gansser (1964) interpreted this gap as evidence for subduction of several hundred kilometers of transitional formations into the Central Thrust. According to LeFort (1975) the rocks of the Lesser Himalayas were originally deposited in several shallow tidal, lagoonal, or continental basins up to at least 9 km thickness. The tectonic deformation of this belt is complex and poorly understood, but it clearly involves both severe shortening and gravity sliding. Basement nappes found in this belt are not rooted but appear to be outliers of or gravity slides from the High Himalayan overthrust (Figure 5.2). Metamorphism is restricted to the northern part of the Lesser Himalayas which were once covered by the Central Thrust, and here it is inverted. The southern boundary of the zone is the north-dipping Main Boundary (thrust) Fault.

Beyond this lies the broad, subsiding Ganga Basin, filled with continental alluvium from northern sources. The aeromagnetic depth-to-basement estimates in this belt increase from 1.5 km on the southern border to 9 km under the Main Boundary Fault (Sen Gupta, 1964). Except on the northern border, these mollasse deposits have little structure, and the whole region is covered by meandering rivers which maintain a flat plain within a few hundred feet of sea level. Within 50 km of the Boundary Fault, the sediments are folded and upthrust by the relative motion of the basement shield against the Lesser Himalayas (Fig. 5.2). It is unclear whether the greater portion of sediments are subducted beneath the Boundary Fault or eroded off the axes of the rising folds.

This history of tectonic events in these five regions follows the same order, progressing from North to South.

5.2 Tectonic History

The convergence of the Indian shield with the Eurasian landmass began at least as long ago as the Permian (about 200 m.y.). At that time the Indian shield separated from Gondwanaland and began to move North relative to the magnetic field at an average velocity of roughly 5 cm/year (Dietz and Holden, 1970). There is no evidence to show the position of the trench which accomodated the convergence at this time, and it may have been far off the southern coast of Tibet. Even if there was a volcanic arc in Tibet it would probably have become inactive toward the close of the Jurassic (135 m.y.) when the Northward velocities of India and Asia became approximately equal (Dietz and Holden, 1970).

The thick carbonates deposited in Tibet in the lower Cretaceous indicate a stable continental margin unaffected by tectonics. This sedimentation came to an end approximately 100 m.y. ago (Hennig, 1915), and this event may be related to the formation of a subduction zone beneath the southern Tibet. Such a zone is shown in Figure 5.3, dipping beneath Tibet and consuming the Tethys oceanic lithosphere. The beginning of igneous intrusion into the volcanic arc. can be no later than 79 m.y., the first K/Ar date reported by Chang and Zeng (1973). Beginning at 71 m.y. (and possibly before) this subduction zone had a very high convergence velocity of 14.5 cm/year (Molnar and Tapponnier, 1975) as determined from world-wide paleomagnetic reconstructions. Convergence continued through the lower Eocene (54-49 m.y.) at a slightly reduced velocity (10.6 cm/year), accompanied by voluminous extrusions of andesites in southern Tibet (Hennig, 1915).

The lower Eocene components of the Indus flysch contain brackish and fresh-water components (De Terra, 1935). This might be taken to indicate the approach of the two continents, with the remaining Tethys Ocean being restricted from mixing with surrounding seas. However, the collision had not yet happened, because middle-Eocene pelagic sediments are found in association with ophiolites in the Axial Belt of West Pakistan (Geologic Map of West Pakistan, 1958). The enigmatic Ladahk Granite formed at 45 m.y. just north of the plate suture (Desio and Zanettin, 1970), and its non-intrusive contact with the Indus Flysch implies that this was before the collision. A possible precursor to the event was the transgression of shallow seas depositing Subathu limestones over northern India (Wadia, 1966). Powell and Conaghan (1973) have interpreted this event as post-collisional, but the non-detrital nature of the sediments indicates that an oceanic sediment trap still separated the continents. The downwarping might also have resulted because the thicker continental lithosphere around the margin of India was more resistant to bending and depressed the central part of the shield (Figure 5.3).

A number of lines of evidence combine to place the collision at about 40 m.y. in the upper Eocene. First, plate convergence slowed from 11 to 5 cm/year, accompanied by a dramatic reorganization of plate motions worldwide (Molnar and Tapponnier, 1975). This is also the youngest fossil age in the Indus flysch (Gansser, 1964), which was presumably compressed and elevated above sea level. Basins and grabens in East China began to subside (Chang Ta, 1959), reflecting the new pattern of stress (Tapponnier and Molnar, 1976a).

This date is also consistent with the geography of Pakistan. It seems highly unlikely that this part of the Asian coast was pre-indented in the shape of the Indian shield, especially as volcanic arcs commonly have a gently convex shoreline. More likely, the East-West portion of the continental margin of Pakistan and Afghanistan at 57°-66°E once had a smooth continuation to the East. Making this assumption, the amount of indentation of Asia by India is given by the distance between the present continental shelf break off Karachi and the end of the Indus Suture near Nanga Parbat. This is some 1800 km, measured parallel to the present trend of the strike-slip Quetta-Chaman fault (Figure 5.1). Using Molnar and Tapponnier's reconstructed relative positions of India and Asia, this gives the same collision date of 40 m.y.

Following the collision the shallow seas receded from India in the lower Oligocene (Wadia, 1966) and the southern Tibetan volcanic arc became inactive, giving a final date of 30 m.y. (Chang and Zang, 1973). However, during the entire Oligocene and lower Miocene no mountains were formed on the Indian plate. No Oligocene deposition occurred in any part of the Himalayas except Assam (Powell and Conaghan, 1973). When Murree deposition began in the lower Miocene, the redbeds were derived from weathered southern terrains of low relief (Gansser, 1964; Wadia, 1966). This inactive period is one of the central problems of Himalayan geology.

Powell and Conaghan attempt to relate the pause to slow spreading in the Indian Ocean. But Molnar and Tapponnier show that net convergence still continued at 5-6 cm/year. Thus the shortening must have been north of the Indus suture. Compression of Tibet is one possibility shown in Figure 5.3 and discussed in the next chapter. Reactivation of the Tien Shan and Pamirs by thrust faulting also began in the Oligocene (Molnar and Tapponnier, 1975). It was probably during this time that the Indus, Brahmaputra, and Sutlej rivers established their courses from the elevated Trans-Himalayas south across the future site of the Himalayas to the Indian plains (Holmes, 1965).

Given the surface inactivity of northern India it is hard to understand the subterranean melting of the Himalayan leucogranites during this period. One of the largest of these (Manaslu granite) has been dated by Hamet and Allegre (1976), using accurate whole-rock RB-SR methods, to be 28$\pm$1 m.y. old. The initial 87Sr/86Sr ratio of 0.740 is very high, indicating that the granite originated through crustal melting. Although this is the only reliable date, similar two-mica tourmaline granites are widely dispersed throughout both Higher and Lesser Himalayas on both sides of the Central Thrust (Gansser, 1964; LeFort, 1975) and a general thermal event seems to be implied.

In the chapter on the Zagros some young granites exposed in the Crush Zone were discussed. Because of the shape of the granite solidus (Tuttle and Bowen, 1958) granite magmas do not usually penetrate to the surface, but spread in large concordant sheets at several kilometers depth (LeFort, 1975). Therefore, the apparently greater volume of Himalayan granites may only be the result of deeper erosion. It can be argued that the sequence of events in the Himalayas was identical to recent events in the Zagros, and that both involved disruption of the sub-crustal lithosphere.

The reconstructed plate positions imply that the Indus Suture moved 600-700 km north relative to northern Asia during the Oligocene. Such a motion would tend to bend any downgoing slab backward, downward (Figure 5.3), and away from the buoyant crust. Detachment might easily occur along the lower crustal zone of weakness discussed in section 3.1. In this case the asthenosphere which replaced the detached lithosphere would be much hotter than the crustal solidus. As the asthenosphere transferred heat to the crust it would immediately acquire a chilled margin, free of partial melts and not capable of contaminating the granite or lowering its initial ratio. The alternative to this speculative hypothesis will be presented in section 5.7.

Stratigraphic evidence suggests that Himalayan mountain-building began less than 16 m.y. ago in the middle Miocene. At this time deposition of the coarse clastic Siwalik formations began along the whole length of the range (Gansser, 1964). This terrestrial mollase was not very different from the present alluvial deposits of the Ganga and its tributaries. A gradual increase in the height of the mountains is marked by an up-section increase in grain size (Powell and Conaghan, 1973) and in maximum grade of detrital metamorphic minerals (LeFort, 1975). Rapid sedimentation in the undersea sediment cone of the Indus was already underway 11 m.y. ago, at the greatest age reached by JOIDES hole #218 (Scientific Staff, 1972).

The mountain building of the upper Miocene to present was the result of the formation of the Main Central Thrust. This connection, proposed by Powell and Conaghan, is supported by long-standing evidence for tectonic juxta-position of the Lesser and High Himalayas (Gansser, 1964). It is also consistent with the gravity results of the next section. Mattauer (1976) shows from the ubiquitous shallow dipping cleavage and stretching lineation perpendicular to the range that a crustal-scale overthrust affecting at least 15 km thickness has moved over 100 km horizontally south relative to the Indian shield. Unfortunately, neither the velocity nor the throw of this overthrust can be determined from marine paleomagnetic evidence because of the probability of competitive shortening events in China.

LeFort (1975) has classified the geologic evidence of Tertiary Himalayan deformation into four events, D_1-D_4. Of these, only D_1 and D_2 are major, and D_1 is not observed in the Lesser or the Sub-Himalayas. These tectonic events were respectively slightly before or contemporaneous with the two major metamorphic events, M_1 and M_2. The famous inverted metamorphism around the Central Thrust belongs to M_2, and may possibly be dated by 9-12 m.y. ages (Gansser, 1964) obtained by K-Ar methods on pegmatites above the thrust. Because of the ease of resetting these ages (Hamet and Allegre, 1976) this date gives only the last metamorphic event and perhaps not the age of the pegmatites.

It is suggested that D_1 represents the initial collision of continents in the upper Eocene, with some Zagros-type folding of the future High Himalayan continental margin only. M_1 would be the later upper Oligocene thermal event that melted the leuco-granites and established the normal metamorphism. D_2 and M_2 came together in the middle Miocene, when the overthrusting of a massive crustal slab began the formation of the high mountains. The equivalents of D_2/M_2 may be in the process of formation at depth today, but naturally the exposed rocks which we examine record a frozen or slightly retrograde (M_3) version of the early history of the overthrusting. Considering the delicate balance between crustal shortening in the Himalayas and in China, past instabilities in the orogeny accounting for D_3 and D_4 are not hard to imagine.

5.3 Gravity Data and Crustal Roots

There is some controversy concerning the number of Himalayan subduction zones. While all authors who accept continental drift differentiate the Main Central overthrust from the Indus Suture, only Powell and Conaghan (1973) believe that this completely describes the structure. The French authors LeFort (1975), Mattauer (1975) and Hamet and Allegre (1976) all invoke a third major crustal underthrust dipping from the Main Boundary Fault. Since rooted pre-Cambrian units are not found above the Boundary Fault, this cannot be proven from geologic evidence. However, the two scenarios predict different profiles of the Moho, either with or without a deep bulge under the high peaks region. The question is amenable to resolution with available gravity data. Seismic data published to date are unsatisfactory, since no stations exist in the mountains themselves.

All the published gravity data for the Ganga Plain - Himalayan - Tibetan region through 1972 have been collected by the Gravity Services Division of the Defense Mapping Agency (Wilcox et al., 1972). The data courteously provided by the Agency included 1271 measurements, for which the primary source was the compilation by Gulatee (1956). However, most of this data extends only one-third of the way into the Himalayas. The database was supplemented by two detailed surveys reaching approximately half-way to the suture:

Qureshy et al. (1974) and Kono (1974). The former includes 113 stations between 78-80°E, and the latter contributes 145 points along the approach to the Everest base camp. Kono's data includes carefully-computed terrain corrections of up to 60 mGal, and are estimated to be accurate within 5-15 mGal. The only data from the northern half of the range is that given by Marussi (1964) (only six stations), and the only points on the Tibetan Plateau proper are two pendulum stations occupied by Ambolt (1948).

All of these data were projected onto a single cross-section (Figure 5.4) to permit selection of the average values. Since the range is curved, a novel projection was employed to preserve the relationship of data to tectonic features. For each point, a ray was constructed perpendicular to the range by passing a line through the station and the "pole" of the range (44°N, 89°E in Mercator geometry). This ray intersected both the Indus Suture and the Main Boundary Fault as digitized from the 1:1,848,000 map of Gansser (1964). Using these points of intersection, the fractional distance of the station into the range was calculated and used as the ordinate in the composite plot. This method is not sensitive to the "pole" position chosen.

A problem arises because the data for the back half of the range are all from a region of anomalously low elevations (Marussi, 1964). They do not connect with the projected trend of the rapidly increasing anomaly as determined by

Qureshy et al. (1974) and Kono (1974). In order to estimate a Bouguer anomaly for the rest of the range, I made the assumption that the mechanical processes (and hence the isostatic anomaly) would be more constant from place to place than the crustal thickness (and Bouger anomaly). Thus the isostatic anomaly values (0 to +80 mGal) from the NW region were combined with average elevations (ℓ) from the whole range (see section 7.1) to calculate synthetic Bouguer values according to the formula

$$\Delta g_B \approx -2\pi\,(2.67\ \mathrm{g/cc})\left[\frac{6.67\times 10^{-8}}{\mathrm{cm^3/g\text{-}sec^2}}\right]\ell + \Delta g_i \qquad (5.1)$$

where Δg_B and Δg_i are the Bouguer and isostatic anomalies. These synthetic values are differentiated by circles in Figure 5.4.

As shown, a smooth curve was drawn through the scattered data to allow for downward continuation. A good average representation can be obtained with a smooth curve everywhere except around the Boundary Fault, where anomaly values on the south side are clearly greater by 20% than those on the North side. The short wavelength of this kink (20 km) precludes a deep source, and the sense disagrees with the observed surface thrusting. Therefore, this mismatch was assumed to be caused by low-density sediments of the Ganga Plain.

At the boundary fault these recent sediments with densities as low as 2.3 g/cc (Warsi, 1976) are overthrust by PreCambrian to Triassic sediments with a typical crustal density close to 2.7 g/cc. Thus the average crustal density south of the fault may be about 0.10 g/cc less than to the north. In calculating the depth to Moho, a crust-to-mantle density contrast of -0.53 g/cc was assumed in the plains, as opposed to -0.43 g/cc in the mountains. Then no artificial step in the Moho is produced. No attempt was made to differentiate between dense Paleozoic sediments and basement in this interpretation.

Assuming that the entire Bouguer anomaly results from crustal thickening, the amount of this thickening was obtained by one dimensional downward continuation to the depth of "normal" Moho, or 33 km. The equation relating the Fourier transforms is

$$\mathcal{F}\left(\Delta g_B(x,z)\right) = \exp\left(\frac{2\pi z}{\lambda}\right)\mathcal{F}\left(\Delta g_B(x,0)\right) \qquad (5.2)$$

(Grant and West, 1965), where λ is the wavelength of each component of the Fourier transform. The smoothed curve was digitized at 24 km increments in x, and extended for several hundred kilometers on each side of the Himalayas to avoid end effects. The extension to the South was based on additional data (not shown), and the extrapolation into Tibet linearly approached the -565 mGal value of Ambolt (1948).

It should be emphasized that this inversion method, while "impartial", is also non-unique. The anomaly has a long wavelength and could also be caused by deeper sources. I have assumed that the Himalayas are underlain by "normal" mantle because it transmits S_n waves (Molnar and Oliver, 1969) and has a normal P_n velocity of 8.07 km/sec (Tandon and Dube, 1973). I have therefore ascribed all of the anomaly to deflections of the Moho. Different Moho configurations are possible if the crust contains large density anomalies as well.

The result of this downward continuation, translated into crustal thickness with the assumed density contrasts, is shown in Figure 5.4. At most, the crustal thickness (below sea level) is 60 km under the northern Himalayas. It is slightly less (55 km) under the high peaks region in the center. This is distinctly different from the sections drawn by LeFort (1975) and Mattauer (1975) which show the Moho as deep as 80 km under the high peaks. Their sections were based on the assumption that the Boundary Fault and Central Thrust remain separate at depth with a full crustal layer between them, for a total triple thickness of crust. Their theory of a major new fault connected to the Boundary Fault can only be reconciled with gravity data if the amount of slip is below 20 km, and if this tectonic reorganization has taken place all along the range within the last million years.

The interpretation presented in the figure is based on the idea that the crustal layer which is still attached to the lithosphere will be relatively undeformed. The upper contact of this layer is drawn in 33 km above the determined Moho. The remaining thickness gives the amount of sediments, deformed sedimentary rocks, and the High Himalayan thrust sheet. Even if the Central Thrust dips at a low angle (20°) to the north and flattens with depth, it must intersect the Main Boundary Fault approximately under the high peaks. A similar conclusion was reached by Warsi (1976), who fitted various crustal-underthrusting models to similar gravity data.

This possibility was originally suggested by Powell and Conaghan (1973), as a part of their argument for a single underthrust extending under all of Tibet. The gravity data gives no indication of its leading edge, but only shows that the Moho and presumably the underthrust layer flatten out to the North. The problem of its full extent is discussed in Chapter 6.

An important implication of these results is that the formation of the Boundary Fault was not a discontinuous break from the tectonics that created the Main Central overthrust. Rather, the material between the faults is probably a wedge of highly deformed pre-Tertiary sedimentary rock scraped off the subducting crust. It may even have accumulated gradually, with the current boundary fault moving away from the Central Thrust as more rocks were pasted onto the front of the wedge. The existence of the wedge makes the bend in the subducting lithosphere more gradual than it would otherwise be, and fills a gap that would otherwise be a 12 km-deep trench.

If this is correct, the rocks in the inter-fault wedge are trapped between a fixed northern boundary and a moving lower boundary. If stress on the Boundary Fault is sufficiently low, the wedge can slide over the advancing shield as a solid body. But the mechanical equilibrium may be delicate. If stress on the Boundary Fault becomes as high as the internal strength of the wedge, it might "convect" in an overturning, rolling motion that would be clockwise in a section like Figure 5.4. The complex geology of much of the Lesser Himalayas may reflect past instabilities of this type.

5.4 Thermal Models

At this point the approximate timing and rate of subduction are known from geology, and its approximate geometry is known from gravity anomalies. These would be sufficient to calculate a thermal model by the finite-difference method described in section 2.2, except that the state of stress in the subduction zone and the mechanism of formation of the Himalayan granites are unknown. Accordingly, three different models are calculated to illustrate the range of possibilities. Model H-A assumes simple subduction with maximum shear stress in the fault, according to the faulting and creep laws of "dry" crustal granite. It does not explain the Himalayan metamorphism. In model H-B, a major thermal event is included prior to subduction by removing a portion of the sub-crustal lithosphere, but the subduction process and shear zone flow law are the same. Model H-C also includes the early thermal pulse, but differs from H-B in assuming weaker material and less shear heating during subduction.

All models were calculated on a 41 by 41 point grid with a 5 km vertical increment. In each case the finite difference program was used in conjunction with the shear zone program of Appendix B to obtain mechanically self-consistent models. For each model, subduction without frictional heating was calculated first. These temperatures and an assumed flow law were input to the shear zone program, which is easily modified to treat faults of changing dip. The calculated stresses then gave the correct frictional heating in the final run of the finite-difference program.

Because there is no heat flow data in or even near the Himalayas, the choice of an initial geotherm presents a problem. It is necessary to assume that the northern margin of India prior to the collision was similar to the rest of the continent today. Excluding the anomalous Deccan Trap region in Maharashtra, there are eight heat flow determinations that were rated "good" by Simmons and Horai (1968). From the south (Kolar mine) to the north (New Delhi) they are respectively 0.66, 0.64, 1.20, 2.21, 1.06, 1.45, 1.36, and 1.76 HFU. All are obtained in the stable pre-Cambrian shield, although the last two sites are on basement "islands" surrounded by alluvium, which may tend to focus the heat flow. The average of the eight values is 1.29 HFU (54 mW/m^2). Since this is the same as the heat flow of the model developed in Chapter 2, the same parameters were employed. Below the lithosphere, where temperatures exceed 1350°C, the conductivity is raised by a factor of ten to roughly simulate small-scale convection in the asthenosphere. The transition is spread over three grid points so that equation (2.9) is satisfied.

In each case the underthrusting of crust is made to extend north to the Indus Suture, because (limited) gravity data show a 65 km crust in that region. Unlike Powell and

Conaghan (1973), I do not extend the underthrusting beneath Tibet as well. Reasons for this will be given in the next chapter; briefly, it would be inconsistent with the observed volcanism (Kidd, 1975). The thrusting is distributed in time from the middle Miocene to present, implying an average velocity (1.6 cm/yr) which is only part of the total convergence of India and Asia. Since neither of the numerical techniques is sophisticated enough to model the processes in the Main Central Thrust-Boundary Fault wedge, a single fault was used in the calculations. Probably, as rock was eroded off the top of the overthrusting slab, sedimentary rocks became attached to its base and kept the fault geometry fairly constant; lithosphere rigidity would tend to require this. According to this hypothesis, the original Central Thrust had the geometry of the present Boundary Fault, and the model is fairly accurate.

The sequence of events is shown schematically in Figure 5.5. Models H-B and H-C include a stage from lower Oligocene through lower Miocene in which the crust of the future High Himalayas is exposed to the asthenosphere. This is represented numerically by replacing the sub-crustal lithosphere temperatures in that region by asthenosphere temperatures translated upward 85 km. In the following 24 m.y. the heat equation is solved without further motion. As shown by the inset, lower crustal temperatures exceed the granite solidus (Brown and Fyfe, 1970) below 29 km depth by a small amount at the end of this period. The temperature of 500°C (alumino-silicate triple point) is exceeded at 18 km depth; this agrees nicely with LeFort (1975), who states that silliminite in the normally-metamorphosed High Himalayas is last observed at reconstructed depths of 15-20 km. The maximum granite production predicted by this model is earlier than the maximum metamorphism, however. Melting should be most extensive at the time of initial disruption of the lithosphere, when the Moho temperature temporarily exceeds 900°C. Therefore the granite age of 28 m.y. mentioned above could be interpreted as showing that the thermal event was not immediately after the collision. Or the difference in age might merely result from slow upwelling and cooling of the magma.

Any successful model of the High Himalayas must reproduce the same melting, age, and normal metamorphism to silliminite grade. Whatever its source, the heat must come from below in approximately the same quantity as modelled here. The *ad hoc* part of this model is the assumption of a particular heat source, but since the source is "swept away" to the North by the subsequent underthrusting, the prediction of present temperatures is hardly affected by the details of the assumption. The alternative heat source, shear strain heating, is discussed in the final section of this chapter.

Present-day temperatures predicted by the three models are shown as Figures 5.6, 5.7, and 5.8. In each case the temperatures to the left (north) of the Indus Suture are not calculation results, because the modelling of a 200 m.y. history including induced convection and igneous intrusion is beyond the scope of this thesis. What is drawn is a smooth transition into the estimated thermal structure of Tibet, as determined from seismic studies of the next chapter. While the immediate southern edge of Tibet may have different properties from the average, it is desirable in the mechanical models to include some representation of the weakest part of Tibet, so as to observe the predicted interactions of the two regions.

The differences between models H-A and H-B are very small. A heat flow difference of 17 mW/m^2 (0.4 HFU) in the northern Himalayas is the only thermal remnant of the assumed Oligocene heating event. However, since there are no abnormal temperatures or melting as a result of underthrusting in model H-A it cannot explain the high-grade metamorphism and granites in the High Himalayas. Only a model with some catastrophic event in the Oligocene will meet these tests. Since the predicted differences in heat flow are probably less than the effect of erosion (not included in the models), such petrologic tests are the only ones likely to be available in the near future.

Looking at the details of the shear zone (Figure 5.9), we see that model H-B has only slightly less resistance to underthrusting than H-A. For both the integral of shear stress along the zone is about 1.4×10^{16} dyne/cm. Finite element models will show that this value is much more than the overthrusting slab can support. Since the upper limit is closer to 3×10^{15} dyne/cm, model H-C was developed to incorporate a lower stress.

Of the three free parameters in the flow law (A, B, n), only n has a value supported by theory. Of the remaining two, I chose to reduce the activation energy (B) because this leads to a less temperature-sensitive rheology which can be more accurately modelled. Keeping n=3 and $A=3.49 \times 10^8$ dynes, cm, the activation energy in model H-C becomes 26 kcal/mole (B=4380°K). No physical significance attaches to this value because of the arbitrary pre-exponential constant. However, Figure 5.10 compares the flow law obtained with other rocks studied in the laboratory, and it is most similar to 'wet' artifical quartz. Balderman (1974) determined Q_a=31.6 kal/ mole for his specimens, and if the same value is assumed the properties of the Himalayan shear zone could also be described by $A=4.8 \times 10^7$ dyne/cm, B=5308°K, and n=3.0. This would be at least thirty times weaker than the lowest creep strength required by Zagros models. Accordingly, the alternative possibility of low-stress faulting in the shear zone will be tested with finite elements.

5.5 Estimates of Uplift and Subsidence Rates

It is necessary to study the rates of Himalayan uplift for two reasons. First, the finite element representations of the tectonics will be more accurate with proper boundary conditions. Second, they give the most useful information available at present for determining the rate of convergence within the range. Seismic moment determinations, like those of Chen and Molnar (1976), give only rough lower limits, and direct measurements by satellite- or lasar-ranging have not been attempted.

The available data is scanty, and it will only be possible to assemble a composite profile for the entire range, rather than a two-dimensional map. The assumption that such an approach is valid is based on the continuity of geomorphic features along the length of the range. The known facts are the following:

1. The Ganga Basin as a whole is subsiding. This is shown by the uniform cover of Recent alluvial deposits and lack of older Siwalik outcrops except adjacent to the Boundary Fault (Gansser, 1964). The major rivers have braided channels indicating net deposition. Since the center of the trough is at least 100 m. above sea level, these effects cannot be ascribed to Pleistocene eustatic changes.

2. Relative uplift of the north side of the Boundary Fault averaged 0.44 mm/month over a six-month period in 1969 in the Nahan region. (A tiltmeter was installed by Agrawal and Gaur (1971) in a trench cut across Nahan Thrust at 30°31'N, 77°50'E, where the apparent dip of the fault is 30°.) This is equivalent to a relative uplift rate of 0.53 cm/year.

3. The relative uplift across the Boundary Fault was also 0.5 cm/year over a twelve-year period at Dalhousie. Chugh (1974) describes an 85 km 1st-order levelling line from Pathankot to Dalhousie established in 1960-61 and cutting the Fault at right angles. When the survey was repeated in 1972-73 an uplift averaging 60 mm was found on the northern side. The uplift increased in magnitude in a band 30 km wide behind the fault, and appeared to level off or decrease slightly around Dalhousie.

4. The central parts of the range are also rising. Uplifted and tilted river terrace gravels (Karewas) of Pleistocene age are found, especially in Kashmir (Gansser, 1964).

5. Uplift rates have increased dramatically in the recent past. Many major Himalayan rivers are now found to occupy slit-like vertical-slided gorges cut into the base of V-shaped valleys (Holmes, 1964). This indicated that now only stream-bed erosion is able to keep pace with uplift, whereas the uplift was once slow enough to

equilibrate with down-slope processes as well.

6. The annual suspended load of the Brahmaputra River is 8.46×10^{14} gm (Holeman, 1968), equivalent to an average of .048 cm/year erosion in its source area of 6.66×10^{15} cm^2. This is a useful number because the Brahmaputra closely follows the strike of the range, draining its back half and the southernmost 120 kilometers of Tibet.

7. The combined suspended load of the Indus, Brahmaputra, and Ganga rivers is 2.625×10^{15} g/yr (Holeman, 1968). Making an estimated correction of 6% for suspended load and 11% for bed load (Gregory and Walling, 1973), this figure may be increased to 3.04×10^{15} g/year, or 1.14×10^{15} cc/year of rock eroded. This comes from a drainage area of 2.50×10^{16} cm^2, encompassing the Himalayas, southern Tibet, and the Ganga Plain in a band 800 km wide by 3200 km long. The average erosion rate for the entire region is 0.44 mm/yr. This average includes the deposition in the Ganga Plain, because the sediments left there never arrive at the hydrologic stations at the river mouths.

In addition to these data I propose three assumptions:

1. The Himalayan region as a whole was in isostatic balance before the Pleistocene ice age. Thus the averaged uplift in the whole drainage basin is not less than the averaged erosion rate of 0.44 mm/yr. It may be greater if transient rebound from the removal of Pleistocene ice is

still going on. However, the transient rebound should be much less than in Fennoscandia (1 cm/yr) because only mountain valley glaciers and no continental icecaps were involved.

2. The relative uplift measured on the Boundary Fault are representative average values, not caused by unusual earthquakes. During the operation of the tiltmeter, a portable seismometer recorded no local events greater than magnitude 1 (Agrawal and Gauer, 1971). One earthquake did occur in the region of the repeated levelling (9-20-67) but it was a small one, detected by only twelve stations and not assigned a magnitude by the International Seismological Center. This assumption explains the coincidence of results obtained over different periods of time. It implies that most of the motion on the Boundary Fault takes place by stable sliding rather than stick-slip. This idea is supported by the results of a strainmeter experiment conducted by Sinvhal et al. (1973) across the Nahan section of the Boundary Fault. There, the daily increments of strain were very uniform during the period studied (one month). A similar result for the interplate region as a whole was obtained by Chen and Molnar (1976), who found that only one-fifth to one-third of the paleomagnetic convergence velocity can be detected in earthquake moments.

3. The downwarp under the Ganga Basin is caused by bending of an elastic lithospheric plate in response to edge forces from the Himalayas. This concept is supported by negative isostatic anomalies under the basin and positive values in the range (Qureshy, et al., 1975) as well as the existence of an "outer rise" in the Indian shield (Warsi, 1976). In that case the depth of the alluvial fill on a North-South profile should follow a parabolic law. (The more complex solution of Smith (1974) for a plate on a Winkler foundation is not applicable because the weight of alluvium nearly cancels the restoring force). If this bending is a steady-state effect which is maintained as the plates converge, then the local subsidence rate should be proportional to the spatial derivative of the parabola. That is, the rate of subsidence should increase linearly from zero across the basin in the direction of the Himalayas.

Using all these constraints, the average uplift profile shown in Figure 5.11 can be determined. The northern section (x=0 to 240 km) is fixed at the Brahmaputra erosion rate. The southern end (x=800) is fixed at no vertical velocity. A velocity contrast of 5 mm/year is spread over 30 km (x=355 to 385) around the Boundary Fault. Finally, the absolute value of the whole section is adjusted to give the proper integrated uplift just matching erosion. This is the uplift model U1.

Also shown in Figure 5.11 is a second model (U2) which gives an integrated uplift with a 45% transient glacial-rebound component. Some amount of this rebound is needed to explain the river gorges. Also, this model concentrates the strain associated with different tilting rates into the weak Indus suture, rather than into an arbitrary point in the overthrusting slab (as U1 does). This model was used to initiate the finite element modelling described in the next section, and ultimately the successful models showed it to be mechanically consistent.

5.6 Finite Element Models of Continuing Deformation

To understand the past tectonics of the Himalayas it is necessary first to understand the present. In this section finite element modelling is used to determine the mechanical properties of the shear zone, which properties can be used to greatly restrict the range of past thermal histories. Models are presented in which the major shear zone is the source of Himalayan earthquakes, and others in which it is an aseismic creeping zone. An attempt is made to find a rheology for the whole crust such that an anomalous shear zone is not required. And finally, the present convergence velocity in the range is found, as well as its average profile of uplift and subsidence.

The three major tests which a successful model must pass involve seismicity, fault plane solutions, and differential uplift. Seismicity from the central and eastern Himalayas (between 75-90°E) has been assembled onto a single cross-section with the same projection technique as was used for gravity data in section 5.3. That is, events are plotted according to their fractional distance between the Main Boundary Fault and Indus Suture. This plot, shown in Figure 5.12, includes all those located from 1961-1973 by the Coast and Geodetic Survey. They are heavily concentrated in the southern half of the range, from 20 to 120 km north of the Boundary Fault. Strangely, neither the Indus Suture nor the shallow portions of the Boundary Fault are particularly active. No large events ($M_b > 5.0$) have been located deeper than 70 km. Predicted seismicity from the finite element models is required to match the horizontal seismicity distribution but not the depths, which appear to be randomly distributed around the initial value (33 km) used in the location program of the CGS.

Of the many published fault-plane solutions in the Himalayas, eleven were selected from the work of Fitch (1970); Molnar *et al*. (1973), Rastogi (1974), Tandon and Srivastava (1975), and Banghar (1976). The two criteria for selection were that the dip of the planes be well constained, and that the nodal line (B-axis) be within 30° of the horizontal line parallel to the range in that vicinity. This tended to eliminate solutions from the eastern Himalayas where the direction of relative motion is oblique to the mountain front. The assembled and rotated fault planes are shown in Figure 5.13. No model is successful in matching all of them, especially the two events in the North with near-vertical fault planes and the normal fault under the Ganga Basin. Although the latter is most likely caused by plate bending (Isacks *et al*., 1968; Molnar *et al*., 1973), the grid is apparently not fine enough to reproduce this outer-fiber tensional stress. Another event which could not be matched is the normal earthquake in the northern Himalayas studied by Banghar (1976). Since the strike of the planes is poorly constrained and fails to agree with the aftershocks (which align North-South) it may be that this event actually had a North-South B axis, and indicated out-of-plane stress differences not accommodated by the model. The better models are primarily successful in matching the seven thrusting events in the southern half of the range.

Finally, the models are required to predict a differential uplift velocity of 0.50 cm/year across the Main Boundary Fault, as discussed in the last section. Although the actual fault is very narrow the model has elements of finite size, so a velocity differential of this magnitude spread over one element is considered acceptable.

Figure 5.14 shows the grid of finite elements used in all models, and the geologic cross-section it represents. The average profile of the Moho is taken from the gravity analysis of section 5.3. The shear zone between the subducting and overthrusting crustal layers is represented by a single row of thin elements. This row initially dips 20° North, and then flattens out and remains a constant distance from the Moho. The Indus Suture is not explicitly outlined with elements because there is no evidence that it is still active. All together, the grid has 96 elements, 62 nodes, and 185 degrees of freedom.

The fact that the grid does not extend above sea level requires some explanation. This is primarily a matter of convenience and consistency with the thermal models of section 5.4, which have no provision for topography. Also, it allows the vertical velocity of the surface to be specified without restricting the horizontal component; if the boundary were not straight and flat this would be impossible. It is also not a bad representation of the actual mechanical situation, because most of the range is so deeply dissected by rivers and glaciers that it cannot transmit horizontal stresses for any distance. The primary importance of the topography is that its weight acts to oppose uplift and shortening, and this effect is included in the models with explicit or implicit boundary forces.

No detailed account of the rock parameters assumed is necessary because initial values are the same as in Chapter 4. Upper and lower crust and mantle have the properties listed in Table 4.4. The only new material is the alluvium of the Ganga Plain, which is assigned crustal creep parameters, a coefficient of friction of 0.50, and a density of 2.30 g/cc (section 5.3). Variation of rock properties will be discussed and are summarized in Table 5.1. The mantle is assumed to be dry-olivine peridotite in all models, and a correction for the effect of pressure on creep is made as in Chapter 4. All the models presented were calculated using the accurate variable-viscosity finite element method. Earlier models calculated by initial stress techniques were similar and contributed to the understanding of the effects of various parameters, but are not considered accurate enough to be compared with these.

In all models the horizontal velocity component only was set to zero at X=0 (left-hand side) to provide a reference. On the opposite end a constant convergence velocity was imposed, but the vertical component was again left free. Crustal elements extending below the reference depth of 33 km were assigned a negative density anomaly of -0.43 g/cc in equation (3.23) to represent the uplifting force of the roots. Finally, the resistance of the Indian shield to downwarping is represented by an upward shear force of 2.5×10^{15} dynes/cm applied to the mantle elements of the southern boundary. This amount of force is calculated from the net isostatic anomaly of the region modelled, and is only about one-third the size of the horizontal forces in the successful solutions.

The vertical velocity boundary condition is handled in a novel way. It is desirable to suppress unphysical long-wavelength vertical velocities, both to preserve numerical accuracy and to expedite interpretation. However, directly imposing velocities on the surface can lead to artificial stresses because the near-surface rocks are cold and strong. The problem is solved by beginning with the uplift model U2 imposed on the top. When the first moderately-successful model is obtained (H28) with this method, the stream function values at the base of the lithosphere are found. The basal values are used to stabilize all subsequent models so that the upper boundary may be left free. (Only vertical forces representing the averaged weight of the topography (see section 7.1) are applied to the top). While the conditions on the base may not be perfectly consistent with each model, their effect on stresses and seismicity is small because the lower lithosphere is so weak.

This model H28 which gives the stream function values along the base is shown in Figure 5.15. It was obtained using thermal model H-C, which assumes a weak, creeping material similar to wet quartz in the shear zone. The uplift test is not applicable because of the boundary condition. The seismicity is generally in the right place, although it should extend north to X=250 km. Predicted fault planes are wrong because all dip to the south, and half of the predicted events would be shallow normal ones, in contrast with observations. The best thing about this model is that the velocity field is the same as that assumed in calculating the temperatures; so it is self-consistent if not quite correct.

Next the three thermal models were tested in order with the upper surface free. Model H31 was based on thermal model H-A, and it failed. Seismicity was strongest in Tibet and the Ganga Plain rather than the Himalayas (Figure 5.16). Predicted fault planes had an intermediate (30°) dip which matched none of the solutions. And no relative uplift at all was predicted across the Main Boundary Fault. Thermal model H-B incorporated in finite element model H30 was unsuccessful in very similar ways. Seismicity was concentrated around the Indus Suture and Ganga Plain instead of the Lesser Himalayas (Figure 5.17), thrust planes monotonously dipped 30° and no relative uplift was predicted. Essentially, both models failed because the rocks of the shear zone, with "normal" crustal creep parameters, were too strong. They put more stress on the upper crust than it could support and it failed, either in Tibet or around the suture.

Better results were obtained with thermal model H-C because it incorporates a weak shear zone. Model H29 was essentially a rerun of H28 but with a free upper surface. As shown in Figure 5.18, an uplift of 0.44 cm/yr across the Boundary Fault is predicted. It is encouraging that the general profile of uplift is very similar to the U2 profile used in initial modelling. Fault plane solutions are considerably improved when the boundary is released, covering a range from 5° to 45° North dip. The main flaw of this model is that seismicity dies out too far south, instead of extending to the center of the range.

In order to obtain the proper range of seismicity (X=250-380 km), I have "designed" a material for the shear zone by relaxing the Navier faulting criterion. The material is assumed to have a plastic limit lower than the normal frictional faulting stress and independent of pressure. This plastic limit is taken as one variable and the pre-exponential term A in the creep law

$$\left[\frac{1}{4}(\sigma_{xx}-\sigma_{zz})^2+\tau_{xz}^2\right]^{1/2} = A\exp\left(\frac{B}{T}\right)\left[(\dot{e}_{xx}-\dot{e}_{zz})^2+\dot{e}_{xz}^2\right]^{1/2n} \tag{5.3}$$

as the other. The B value of 7445°K taken from Tullis (1971) is retained. The two constraints are that plastic behavior shall extend 135 km into the shear zone, and that the integrated shear stress be no more than $3.0\text{x}10^{15}$ dyne/cm, as in model H-C.

A solution is obtained using the method of Appendix B, and is termed thermal model H-D. As shown in Figure 5.19, the plastic limit is 200 bars shear stress and creep begins at 380°C. The constant A is found to be $6.5\text{x}10^7$ dynes/cm^2, or 20% of the value for dry-quartz granites. This is the same as the creep strength independently determined for the lower crust of the Zagros in Chapter 4. These shear zone parameters and temperatures formed the input to finite element model H32, shown in Figure 5.20.

This model is very successful in producing a sharp relative uplift (0.38 cm/yr) across the Boundary Fault. It also repreduces the region of maximum seismicity very well, as it was designed to do. However, the match of the fault planes is not as good as in model H29. Instead of a mixture of steep and shallow-dipping planes, only shallow planes parallel to the shear zone are predicted. Therefore only three of the seven crustal thrusting events are matched. Some infrequent mantle thrusting events are predicted, and could possibly correspond to the June 30, 1969 event which was nominally located at 64 km depth (Tandon and Srivastava, 1975).

Having determined that the shear zone was necessarily weaker than the faulting law (Appendix C) predicts, I sought next to determine how much contrast between the shear zone and surrounding crust was required. Rough calculations showed that 300 bars was the maximum permissible average stress in the shear zone. To test this, the plastic limit was raised to 350 bars in model H33. The model failed, as .45 cm/yr of the convergence was taken up in Tibet and .14 cm/yr in the Ganga Plain. Seismicity in these regions was unacceptably high, and the relative uplift at the Boundary Fault disappeared (Table 5.1).

Since the shear zone cannot be strengthened, it remains to be seen if the rest of the crust can be weakened. In model H34 the constant A=6.5E7 was applied to the entire crust as well as the shear zone, without changing the faulting laws. This model failed even more dramatically, because the Indian shield to the south became too weak to sustain the force of convergence. All of the shortening and seismicity fell in the Ganga Plain (Table 5.1), which is clearly unacceptable. This poses a problem in generalizing the results of these models, as different parameters are required in the two ranges.

In a similar way, models H35 and H36 tested the results of lowering the plastic limit in the crust to 1000 and 500 bars respectively without changing the creep strength. The first gave acceptable results very similar to model H29. The principal difference was that the seismicity level in the Ganga Plain increased while the amount of Boundary Fault uplift fell to 0.19 cm/yr. The second clearly failed, as the relative uplift fell to zero and the Ganga Plain seismicity approached Himalayan levels. Thus the shear strength of cold crustal rocks in the Himalayas is apparently not less than 1000 bars. Once again, it appears to be higher than the strengths in the Zagros.

In a final attempt to produce steeply dipping thrust planes, model H37 extended the width of the shear zone to include all the deformed sedimentary rocks between the Boundary Fault and the Central Thrust (Figure 5.4). One additional element in the column x=300-360 km and two elements in the column x=360-420 km were assigned shear zone parameters of lowered creep and faulting strength. The result was that, while some small steep thrust events were predicted, the shortening and seismicity fell too far North, almost at the Central Thrust. Relative uplift on the Boundary Fault fell to 0.11 cm/year; to scale this upward to 0.50 cm/year would require a convergence velocity greater than the total relative motion of India and Eurasia (Molnar and Tapponnier, 1975). The possibility of intermediate parameters in this wedge of sediments exists but was not pursued, in order to keep ad hoc rock parameters to a minimum in the models.

It appears that it is possible to match the seismicity well and the fault planes indifferently (H32), or vice versa (H29), but not both. Both models involve an anomalously weak shear zone in a normal crust, but this zone primarily faults in the former model and creeps in the latter. Either model can be scaled in velocity to produce the proper Boundary Fault uplift, and this was the last step performed. Model H38 is the same as H32 except that the convergence velocity is scaled upward to 2.1 cm/year. Also, a rigid-body translation and rotation rate has been added, which does not affect the stresses. Detailed results are shown as Figure 5.21 and 5.22. Once again, seismicity is correct but only some of the fault planes are matched. The relative uplift is slightly too big (0.57 cm/year) but within the scatter of the data. The best estimate of convergence velocity in this group of models is 2.0±0.2 cm/year. The error range reflects only effects of solution nonlinearity and nonconvergence, not actual errors in the data.

The last model, H39, is the same as H29 (a creeping fault zone), but scaled up to a convergence velocity of 1.8 cm/year. Stresses and seismicity are hardly affected (Figures 5.23 and 5.24), and the proper uplift is obtained. Note that this model implies a large amount of convergence in the sedimentary wedge between the Boundary Fault and Central Thrust, enough to eliminate it in 5-10 m.y. This makes it less appealing than the previous model in which the wedge is a steady-state feature. However, the complex deformation in this zone produces thrust faulting at varying dips (Figure 5.24) and therefore matches more of the fault-plane solutions.

The uplift and subsidence profiles of these two final models are shown in Figure 5.25. Although the baseline velocity and tilting rates are arbitrary and unknown, the shapes are in fairly good agreement with the starting model U2. This suggests that the Himalayas as a whole are tilting toward the North in response to underthrusting and erosion at a rate of about 1° per million years. The fact that models can be scaled to give the proper uplift results from the nonlinear rheology; in warm regions the creep stress varies only as the cube root of velocity, and in plastic areas it is invariant. This means that if new data appear on the relative uplift around the Boundary Fault, the rate of convergence can be redetermined without any further computation. It will merely be four times the rate of relative uplift.

The convergence velocity of 1.8-2.2 cm/year determined here is less than half of the total convergence rate of India and Eurasia. This means that large amounts of strain are presently taking place in China, as suggested by Molnar and Tapponnier (1975). This velocity can also be extrapolated backwards. If the total underthrusting is approximately equal to the width of the range (250 km) then we estimate that the mountains began to form 11-14 m.y. ago in the middle Miocene. This is in good agreement with the stratigraphic date based on the appearance of the Siwalik mollase.

On the other hand, this rate is greater than the rate obtained for the entire convergence zone from tensor summation of seismic moments. Chen and Molnar (1976) determined an average seismic slip of 1.0 cm/year during 1911-1967, and a slightly higher 1.7 cm/year during the active period 1862-1910. This is as it should be, because their figures do not include the contribution of aseismic slip to the total convergence. On the Main Boundary Fault this stable sliding appears to dominate (see section 5.5), just as it apparently does in the Zagros.

5.7 The Frictional Theory of Himalayan Magma Genesis

Finite element models of the last section show that the shear zone beginning at the Main Boundary Fault and dipping through the center of the Himalayan crust is very weak. It cannot at present be producing more than 19 mW/m^2 (0.45 HFU) of shear-strain heating (equation 2.9) on the average, and this is much too little to cause any melting. Therefore, some radically different mechanism acted on the crust in the past. A simple but unconventional mechanism was presented, based on the early detachment of the Tethyan slab. In this section the principal alternate theory, shear-strain heating during overthrusting, is examined. This possibility was suggested both by the author (Bird and Toksöz, 1974) and by LeFort (1975).

It has been shown in thermal models H-A and H-B that the frictional heating would be insufficient at the present rate of convergence even if the shear zone were not weak. However, it has been suggested (Claude Allegre, personal communication, 1976) that plate motions may be discontinuous, with periods of inactivity separating brief periods of very fast convergence. This possibility cannot be tested from paleomagnetic evidence, because the amount of plate convergence between clearly identifiable anomalies (ca. 1000 km) is much more than the amount of underthrusting proposed (80-240 km). Thus it will not be possible to prove or disprove the hypothesis at this time, only to determine the minimum convergence velocity that it would require.

In order to find this minimum, a number of assumptions have been made that tend to favor the hypothesis, as follows:

1. All of the granites formed in the thermal event have been exposed by erosion and mapped. In other words, the amount of melting which the mechanism must explain is no more than the volume of visible granites. An estimate of the exposed area of granites in the well mapped Nahan-Darranga Camp section of the Himalayas was made from Gansser's (1964) map. In this section, extending 1400 km along strike, there are nine major plutons with a total surface area of 8,100 square kilometers.

2. The granites are only 1.5 km thick on the average. If the bodies were isotropically random and eroded down to their centers, the thickness could be estimated as the square root of the surface area of each pluton. For these bodies, such estimates would give 10-45 km thicknesses. However, we assume here that gravity, stratification, and the shape of the solidus act together to spread the plutons into thin sheets. The assumed thickness (M. Mattauer, personal communication, 1976) implies a total volume of 12,100 km^3.

3. The High Himalayan metamorphism affects only a small volume around the shear zone or the rising plutons. It is assumed that the amount of heat required to metamorphose the basement is secondary to the amount required to melt the

plutons.

4. The plutons originate from a large-percentage partial melting of crust. A figure of 30% partial melting is assumed in the source area to minimize the volume of rock that must be heated.

5. The amount of latent heat consumed by dehydration metamorphic reactions in the source area is negligible.

6. The material in the shear zone was once stronger than it is now. I assume the maximum possible crustal strength based on frictional faulting with a friction coefficient of 0.6, and creep only in a 30% fraction of "dry" quartz. This is the initial strength model developed in section 3.2 on the basis of proven deformation mechanisms only.

7. Water-saturation in the source region is assumed, with water pressure close to total pressure. These are the conditions giving the minimum solidus temperature of 643°C (Tuttle and Bowen, 1958). Whether this assumption is or is not inconsistent with the one above depends on the rate of diffusion of water into the quartz crystal lattice, which is unknown.

8. The overthrusting slab and Tibet are both assumed to have been stronger than they are now. Finite element models show that the integral of shear stress along the shear zone does not exceed 3.0×10^{15} dyne/cm. However, this weakness is partially due to the residual heat from the thermal

event under consideration. Therefore, the maximum integral of shear stress, equal to the amount of horizontal force that normal, cool crust can support without faulting, is raised to 6.3×10^{15} dyne/cm. This value is based on frictional behavior without creep down to 20 km depth (as in Figure 3.1) in the overthrusting slab.

9. Deformational energy radiated away as seismic waves is negligable.

10. The present convergence velocity of 2 cm/year represents a reactivation of the range after long quiescence. This assumption necessarily follows from the assumption of large slips in the Oligiocene thermal event and the constraint that total slip is not much more than the width of the range (see Chapter 6).

Using these assumptions, the approximate velocity required for melting can be estimated by a simple physical argument. For each unit of length along the strike of the range, the total work performed is

$$q = \int (\tau \ell)\, ds = \ell \int \tau\, ds \tag{5.4}$$

where ℓ is the slip distance. Since the integral is assumed to be less than 6.3×10^{15} dyne/cm and the slip less than 240×10^{5} cm, the total work is no more than 1.5×10^{23} ergs/cm. If we equate this work with the heating of a layer of thickness W by 400°C,

$$q = (2.4 \times 10^7 \text{cm}) W \rho C_p (400°\text{c}) \tag{5.5}$$

and take C_p=2.2x10^7 erg/cc-°C, then W is 7 kilometers. To find the elapsed time during subduction this is equated to twice the conduction scale distance S (equation 2.7) implying t=275,000 years if κ =.0148 cm^2/sec. This requires a velocity on the order of 90 cm/year. At lower rates the heat would be more broadly dissipated and the highest temperature would be less.

To improve on this rough estimate, the shear zone has been modelled in two dimensions. As in Chapter 2, the temperatures without friction are found with the finite difference program and the total solution with the shear-zone program. The finite difference grid interval was reduced to 2 km for accuracy. The initial geotherm was that of Table 2.1 and, as in previous thermal models, the shear zone was assumed to dip at 20° down to 27 km depth, and then become horizontal.

Results for a number of different plate velocities are shown in Figure 5.26. No melting at all occurs at velocities of 20 cm/year and below. At 40 cm/yr there is some, but it happens too far back into the range. Granites are now seen some 60 km from the Boundary Fault in the Badrinath region. The distance of this point from the original surface exposure of the Central Thrust is unknown, but probably no more than 60 km. Thus only the model with a velocity of 80 cm/year is successful. Interestingly, the integrated shear stress in this model is 5.9×10^{15} dyne/cm, very close to the assumed limit. This means that errors in the creep law assumed have no effect on the conclusions; the frictional heating could not be greater.

These models are geologically naive because they assume all of the possible subduction in one pulse and do not take account of the latent heat of fusion. In a final pair of models I investigated the effects of an earlier orogenic phase by modelling 80 km of subduction (phase D_1) at a rate of 100 cm/year with the maximum frictional heating. An intervening period of 2.5 m.y. without motion was modelled by finite difference, and early in this period the thin heated zone had already dissipated. Then the second phase (D_2) with an assumed 80 km of additional subduction was modelled as before. This time a latent heat of fusion of 27 cal/gm appropriate to 30% partial melting of granite was included in the shear zone program (Claude Allegre, personal communication, 1976). Since this term never amounted to more than 4% of the heat production, no description of the computational technique is given.

Results of the more complex model are very similar. As seen in Figure 5.27, a velocity of 60 cm/year is marginally acceptable, while 40 cm/year is still too low. At the lower velocity not enough melt is produced and it first appears too far back into the range (100 km). At 60 cm/year the amount of melt is twice the minimum required and the first occurrence is 60 km into the range. Since the integrated shear stress in this model is slightly above the limit (6.89×10^{15} dyne/cm), there does not seem to be any way to find a successful model with lower velocities.

Since the rapid-spreading event required by this theory would be geologically brief (130,000 years), it would be unlikely to span more than one reversal of the geomagnetic field. Some evidence independent of paleomagnetism will be required to evaluate the hypothesis. As discussed before, sedimentation in the Ganga Plain tends to refute the hypothesis by showing that molasse deposition followed granite formation by some 12-18 m.y. This is not proof, however, because earlier molasse deposits formed close to the Himalayas might have been subsequently subducted.

The resolution of the debate will most likely come from the study of the Main Central Thrust, in places where up to 15 km of overburden have been removed by erosion (Gansser, 1964). LeFort studied such exposures in Central Nepal and made a possibly critical observation. According to him (LeFort, 1975; Fig. 10) the inverted metamorphic gradient found below the Main Central Thrust <u>continues</u> <u>above</u> the thrust. The maximum metamorphic grade, including incipient melting, is attained 2-3 km into the upper plate, after which the grade again declines because of the increasing proximity of the paleo-surface. This means that heat flowed from the hot upper plate into the underthrusting lower plate, rather than in both directions away from the shear zone. It implies that the metamorphism of the lower plate is due to residual heat of an earlier thermal event predating the thrust. According to Valdiya (1973) the inverted metamorphism of the upper plate is a simple consequence of retrogression:

> The Vaikrita group (Central Crystalline zone) of the great Himalaya of Kumaon is divisible into five lithological formations. The lowermost Munsiari fm. comprising 1,000 to 1,500 m thick assemblage of gneisses and schists, and including grandiorites, bears eloquent testimony of pervasive and extreme cataclasis and attendant retrograde metamorphic transformations which increase in intensity downwards toward the plane of movement--The Main Central Thrust that defines the lower limit of the Vaikrita group. This explains the reverse order of the grade of metamorphism seen in the lower part of the "Central Crystalline zone". It is this mylonitized and degraded lower part, surmounted with less affected or unaffected crystallines of the Munsiari fm. that has been thrust far to the south and now occurs as nappes and their detached klippen over the complexly folded sedimentaries of the Lesser Himalaya.

The nature of the earlier thermal event remains unclear. The effects of slab detachment may again be put forward, but no proof can be given. Or, to preserve the frictional-melting hypothesis, a third deformational phase can be invoked to explain the geologic exposures. At this point the number of weak assumptions supporting the frictional theory reaches twelve. Is it not simpler to conclude that the Main Central Thrust was the result, rather than the cause, of anomalously high temperatures?

5.8 Conclusions

The Himalayas formed subsequent to a continental collision between India and Eurasia about 40 m.y. ago. Crustal thickening was largely the result of crustal subduction beneath another layer of Indian crust at the line of the Main Central Thrust. This subduction did not begin immediately after the collision, but rather in the middle Miocene.

Gravity data does not support the idea of a second major crustal underthrust at the Boundary Fault. If such a new thrust exists its slip must be only about 20 km, which implies a very recent (1 m.y.) formation. Instead, the shape of the Moho suggests that this fault joins the older Central Thrust at depth. The material between the faults is probably a wedge of sedimentary rock scraped off the subducting crust.

Thermal models of this subduction do not predict any melting at a velocity of 1.6 cm/yr regardless of the nature of material in the shear zone. Melting and metamorphism might have been produced by direct contact between crust and asthenosphere prior to the crustal overthrusting event.

Present uplifting of the Himalayas concentrates in the South, where it reaches a maximum of 5 mm/year. Uplift rates exceed erosion rates for the region as a whole, indicating some component of post-glacial rebound.

Finite element models show that the shear zone (Boundary

Fault and Central Thrust) is anomalously weak. The integrated shear stress on this fault from one side of the Himalayas to the other is less than 3.0×10^{15} dyne/cm. It may be weak either because of creep under a constitutive law close to that of artificial quartz with high water content. Or it may deform by faulting at shear stresses of 200 bars, which is far less than laboratory experiments would predict.

The surrounding crust does not participate in this anomalous weakening. Its creep behavior is that of natural "dry" quartz, and it supports shear stresses of at least 1 kilobar without faulting.

The extent of the underthrust crust cannot be determined from seismicity patterns. In either shear zone model, creep begins under the central Himalayas and suppresses seismicity.

The present rate of convergence in the Himalayas is 2.0 ± 0.2 cm/year. This is less than half the total plate velocity, indicating extensive deformation to the north in China. If new evidence shows the Boundary Fault uplift to be different from 5 mm/year, a new convergence rate may be scaled directly from the uplift rate.

The frictional-melting hypothesis for Himalayan granite formation requires plate velocities of over 60 cm/year to be maintained for slips of 80 km or more. This model does not explain the distribution of metamorphism around exposed portions of the Central Thrust. There was a major thermal event leading to melting prior to the formation of the Central Thrust, and the weakening of the crust by heat may have assisted the thrusting. Detachment of the oceanic slab and part of the continental sub-crustal lithosphere is a possible but not unique explanation for this thermal event.

Table 5.1

Inputs and Results of Finite Element Models of the Himalayas

Model	Figure	Thermal Model	Shear zone Parameters*	Crustal Parameters*	Convergence velocity cm/year	Maximum seismicity range, km	Number of fault-planes matched	Main Boundary Fault uplift cm/year	RMS unconverged stress, bars
H28	5.15	H-C	A=3.5 E8 B=4380°K τ_p =.6ρgz	A=3-4 E8 B=7445°K τ_p =.6ρgz	1.6	350-410 Boundary Fault	1/11	0.50 imposed	26
H29	5.18	H-C	A=3.5 E8 B=4380°K τ_p =.6ρgz	A=3-4 E8 B=7445°K τ_p =.6ρgz	1.6	300-385 Lesser Himalayas	6/11	0.44	18
H30	5.17	H-B	A=3.5 E8 B=7445°K τ_p =.6ρgz	A=3-4 E8 B=7445°K τ_p =.6ρgz	1.6	80-180 Indus Suture	0/11	-.05	14
H31	5.16	H-A	A=3.5 E8 B=7445°K τ_p =.6ρgz	A=3-4 E8 B=7445°K τ_p =.6ρgz	1.6	0-120 Tibet	1/11	-.06	20
H32	5.20	H-D	A=6.5 E7 B=7445°K τ_p =200b	A=3-4 E8 B=7445°K τ_p =.6ρgz	1.6	250-390 Lesser Himalayas	4/11	0.38	66
H33	-	H-D	A=6.5 E7 B=7445°K τ_p =350b	A=3-4 E8 B=7445°K τ_p =.6ρgz	1.6	250-480 Himalayas and Ganga Plain	3/11	0.03	14

Table 5.1 (continued)

Model	Figure	Thermal Model	Shear zone Parameters*	Crustal Parameters*	Convergence velocity cm/year	Maximum Seismicity range, km	Number of fault planes matched	Main Boundary Fault uplift cm/year	RMS unconverged stress, bars
H34	-	H-D	A=6.5 E7 B=7445°K τ_P =150bar	A=6.5 E7 B=7445°K τ_P =.6ρgz	1.6	370-480 Ganga Plain	1/11	-0.27 (wrong way)	19
H35	-	H-D	A=6.5 E7 B=7445°K τ_P =200bar	A=3-4 E8 B=7445°K τ_P=1000bar	1.6	250-390 Lesser Himalayas	4/11	0.19	35
H36	-	H-D	A=6.5 E7 B=7445°K τ_P =200bar	A=3-4 E8 B=7445°K τ_P =500bar	1.6	260-480 Himalayas & Ganga Plain	4/11	-0.04 (wrong way)	36
H37	-	H-D	Extended N to Main Central Thrust	A=3-4 E8 B=7445°K τ_P =.6ρgz	1.6	250-340 Lesser Himalayas	4/11	0.11	35
H38	5.21 5.22	H-D	A=6.5 E7 B=7445°K τ_P =200bar	A=3-4 E8 B=7445°K τ_P =.6ρgz	2.1	250-390 Lesser Himalayas	4/11	0.57	56
H39	5.23 5.24	H-C	A=3.489E8 B=4380°K τ_P =.6ρgz	A=3-4 E8 B=7445°K τ_P =.6ρgz	1.8	300-390 Lesser Himalayas	6/11	0.52	16

* Parameters of creep law

$$\left[\frac{(\sigma_{xx}-\sigma_{zz})^2}{4}+\tau_{xz}^2\right]^{1/2}=A\,e^{\left(\frac{B}{T}\right)}\left[(\dot{e}_{xx}-\dot{e}_{zz})^2+\dot{e}_{xz}^2\right]^{1/2n}$$

Chapter 5 - Figure Captions

Figure 5.1 Tectonic setting of the Himalayas. Figure taken from Molnar and Tapponnier (1975). Fault traces in China are based principally on satellite photo interpretation, and sense of motion on fault-plane solutions.

Figure 5.2 Three cross-sections of the Himalayan range. Figure taken from Gansser (1964). No vertical exaggeration. Crimson bodies (formation 3) are synorogenic tourmaline granites. Orange-tan formation (#4,5) is pre-Cambrian basement.

Figure 5.3 Schematic cross-sections of Himalayan geologic history. No vertical exaggeration. Rocks at the left (NW) end are held fixed. Continental crust is stippled, igneous rocks are black. Arrows show roughly the amount of motion in a period of one million years.

Figure 5.4 All Bouguer gravity anomaly values in the Himalayan region projected onto a single cross-section relative to Main Boundary Fault and Indus Suture. Circles represent synthetic values as explained in the text. At bottom, anomalous crustal thickness determined by one-dimensional downward continuation to 33 km, assuming 0.43 g/cc density contrast in mountains and 0.53 g/cc in the plains. Dotted line shows one possible geologic explanation for this crustal thickening. Vertical exaggeration by 2.4X in the lower plot.

Figure 5.5 Schematic diagram of events leading to the thermal models H-A, H-B, and H-C. Lithosphere is shaded and crust marked by dashes. In models H-B and H-C the crust is exposed to the asthenosphere prior to underthrusting. This results in the temperature increase shown by the inset. Dashed line is the assumed initial geotherm.

Figure 5.6 Present-day Himalayan temperatures predicted by model H-A. Two hundred fifty kilometers of subduction has occurred at 1.6 cm/year on a fault dipping 21° North and becoming horizontal at 27 km depth. Temperatures to left of Indus Suture (IS) are from results of Chapter 6, not the calculation. Temperatures to right of Main Boundary Fault (MBF) according to initial model of Table 2.1. In this model there is maximum (dry-quartz-creep) frictional heating but no previous thermal pulse.

Figure 5.7 Present-day Himalayan temperatures predicted by model H-B. This model differs from H-A by an early thermal pulse (Figure 5.5). See additional remarks above.

Figure 5.8 Present-day Himalayan temperatures predicted by model H-C. This model differs from H-B in lower frictional heating. See additional remarks in caption 5.6.

Figure 5.9 Details of shear zone stress (top) and temperature (bottom) in the three thermal models H-A, H-B, and H-C. Each was obtained by method of Appendix B with shear zone thickness limited to 10 km. Detailed parameters in the text.

Figure 5.10 Comparison of creep law assumed in model H-C (arrow) with other published results at a common temperature.

Figure 5.11 Two model profiles of average uplift and subsidence for the regions drained by the Indus, Ganga, and Brahmaputra rivers. Model U1 is matched to the measured erosion rate. Model U2 assumes that uplift exceeds erosion because of post-Pleistocene rebound.

Figure 5.12 Composite plot of earthquake epicenters in the Himalayas 1961-1973, between longitudes 75-90°E. Concentration of events at 33 km depth is a result of insufficient information to determine depth. Events plotted according to position within the range; 0.0 is at the Main Boundary Fault and 1.00 at the Indus Suture. Size of symbol is proportional to the body wave magnitude minus 3.0.

Figure 5.13 Composite plot of earthquake fault-planes in the Himalayas. Scale is the same as in Figure 5.12. B-axis has been rotated perpendicular to figure by less than 30°. Sources include Fitch (1970) [F], Molnar et al. (1973) [M], Rastogi (1974) [R], Tandon and Srivastava (1975) [T], and Banghar (1976) [B]. The normal fault solution at 0.76 may not be representative.

Figure 5.14 Geologic cross-section of the central Himalayas and the grid of finite elements used to represent it. No vertical exaggeration. Dark shading indicates mantle elements; light dashes mark shear zone elements; and light shading indicates alluvium Other elements have crustal properties. Horizontal velocity boundary conditions are shown with engineering symbols. Vertical velocity boundary conditions explained in text.

Figure 5.15 Output of finite element model H28 in which vertical velocity was imposed at the upper surface. At top, instantaneous velocity vectors of nodes plotted with the tail on the nodal position. At center, deviatoric shear stress is proportional to length of line in each element; orientation of line gives direction of maximum compression. At bottom, predicted seismicity is shown as magnitude and orientation of plastic dissipation. A cross symbol is plotted at each integration point in a plastic state. The size of the symbol is proportional to shear stress and strain rate. The two arms of the symbol show the two possible fault planes.

Figure 5.16 Output of the unsuccessful model H31, based on thermal model H-A. Model parameters in Table 5.1, explanation of symbols above.

Figure 5.17 Output of unsuccessful model H30, based on thermal model H-B. Model parameters in Table 5.1, symbols explained in caption 5.15.

Figure 5.18 Output of model H29, based on thermal model H-C. Similar to figure 5.15 except that top was left free (topographic loads only).

Figure 5.19 Shear stress (top) and temperature (bottom) in the Himalayan shear zone according to thermal model H-D. Material creep strength is 19% that of dry crustal granite, but an upper limit of 200 bars is placed on the shear stress as a plasticity condition. Creep begins when temperature exceeds 380°C.

Figure 5.20 Output of model H32, based on thermal model H-D shown in Figure 5.19. Model parameters are in Table 5.1, and explanation of symbols is in caption 5.15. This model assumes a low faulting stress of 200 bars in the shear zone.

Figure 5.21 Velocity vectors (top) and deviatoric stresses (bottom) from one of the final models, H38. This model has a weak, faulting shear zone. Symbols explained in caption 5.15.

Figure 5.22 Predicted seismicity from model H38 (center) compared with actual seismicity (top) and fault-plane solutions (bottom). See captions 5.12, 5.13, and 5.15 for symbols.

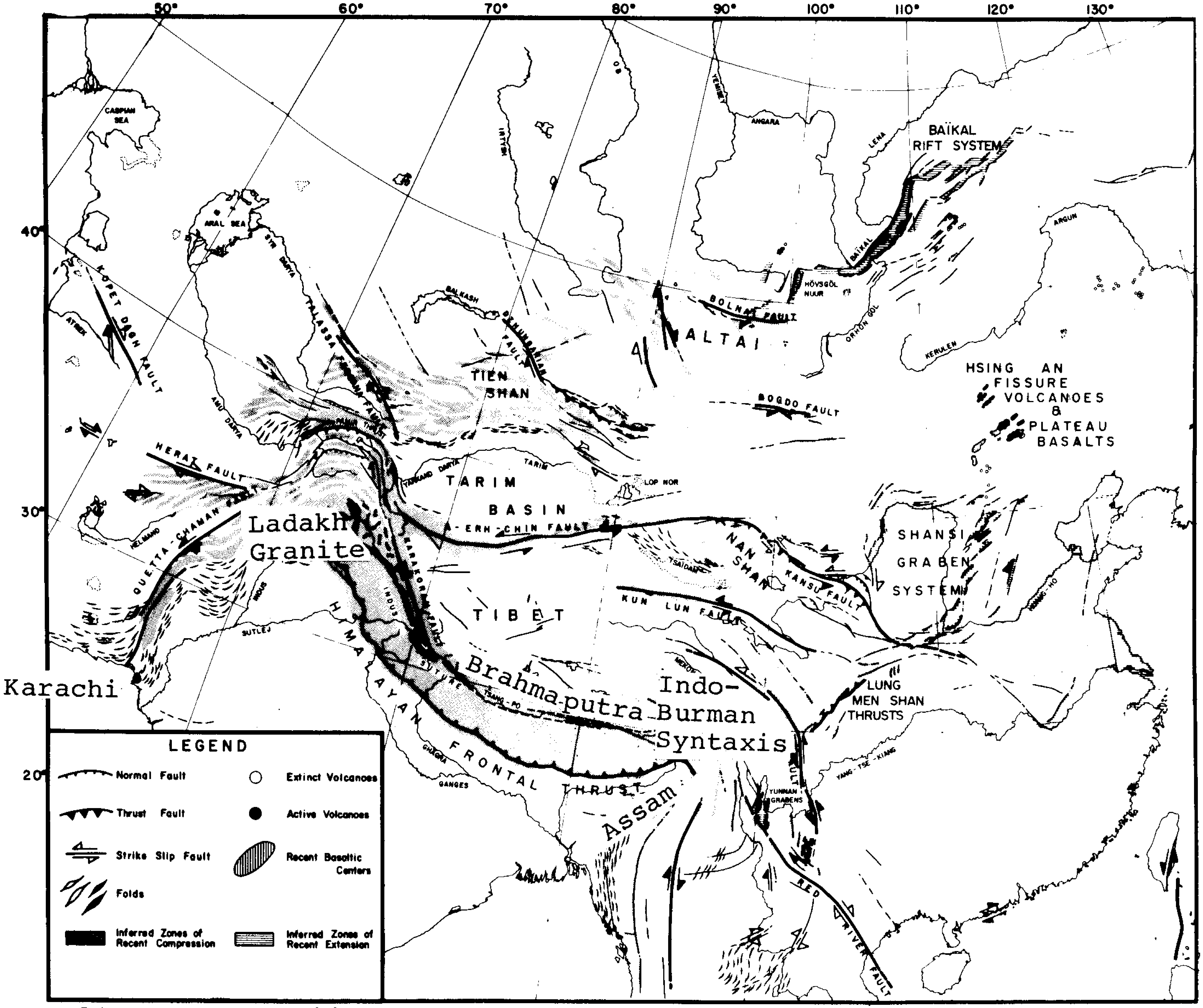
Figure 5.23 Velocity vectors (top) and deviatoric stresses (bottom) from the alternate final model H39. This model has a weak creeping shear zone. Symbols explained in caption 5.15.

Figure 5.24 Predicted seismicity from model H39 (center) compared with actual seismicity (top) and fault-plane solutions (bottom) Inset shows North-dipping fault planes without auxiliary planes to reduce clutter; both shallow and steeply dipping thrusts are predicted in a small area. See captions 5.12, 5.13, and 5.15 for symbols.

Figure 5.25 Predicted uplift rates of the two final models compared with starting model U2 developed in section 5.5 (Figure 5.11). Velocity varies parabolically between nodes (dots). Note that any amount of regional uplift or tilting could be added to the model results without affecting stresses and seismicity.

Figure 5.26 Stress and temperature in a shear zone which dips at 20° down to 27 km depth and then flattens out, after 240 km of slip. Five curves represent different subduction velocities and are labelled in cm/year. These calculations include no heat of fusion term, and temperature is limited to the solidus.

Figure 5.27 Shear zone conditions in the Himalayas during stage D_2 according to the model of discontinuous high velocities. Amount of slip is 80 km in a zone which dips 20° down to 27 km depth and then flattens out. Curves are shown for velocities of 40 and 60 cm/year. At 40 cm/year an insufficient amount of melt would be formed, too far back into the range.



BAÏKAL RIFT SYSTEM
HÖVSGÖL NUUR
BOLNAI FAULT
ALTAI
BOGDO FAULT
HSING AN FISSURE VOLCANOES & PLATEAU BASALTS
KERULEN
ARGUN
ORHON GOL
LENA
BAIKAL
ANGARA
YENISEY
OB
IRTYSH
BALKASH
DZHUNGARIAN FAULT
TIEN SHAN
TALASSA FAULT
CASPIAN SEA
ARAL SEA
SYR DARYA
AMU DARYA
KOPET DAGH FAULT
ATREK
HERAT FAULT
HELMAND
QUETTA-CHAMAN FAULT
Ladakh Granite
Karachi
TARIM
TARIM BASIN
LOP NOR
ERH-CHIN FAULT
NAN SHAN
TSAIDAM
KANSU FAULT
SHANSI GRABEN SYSTEM
KUN LUN FAULT
TIBET
KARAKORAM FAULT
INDUS
SUTLEJ
SUTURE
TSANG-PO
Brahma-putra
Indo-Burman Syntaxis
MEKONG
LUNG MEN SHAN THRUSTS
YUNNAN GRABENS
YANG-TSE-KIANG
HWANG-HO
RED RIVER FAULT
HIMALAYAN FRONTAL THRUST
GHAGRA
GANGES
Assam
LEGEND
Normal Fault
Thrust Fault
Strike Slip Fault
Folds
Inferred Zones of Recent Compression
Extinct Volcanoes
Active Volcanoes
Recent Basaltic Centers
Inferred Zones of Recent Extension
50°
60°
70°
80°
90°
100°
110°
120°
130°
40°
30°
20°

Fig. 5.1


LESSER HIMALAYA
HIGH HIMALAYA
TRANS-HIMALAYAS
BOUNDARY FAULT
MAIN CENTRAL THRUST
INDUS SUTURE
N
GENERALIZED SECTIONS THROUGH THE HIMALAYAS

Fig. 5.2

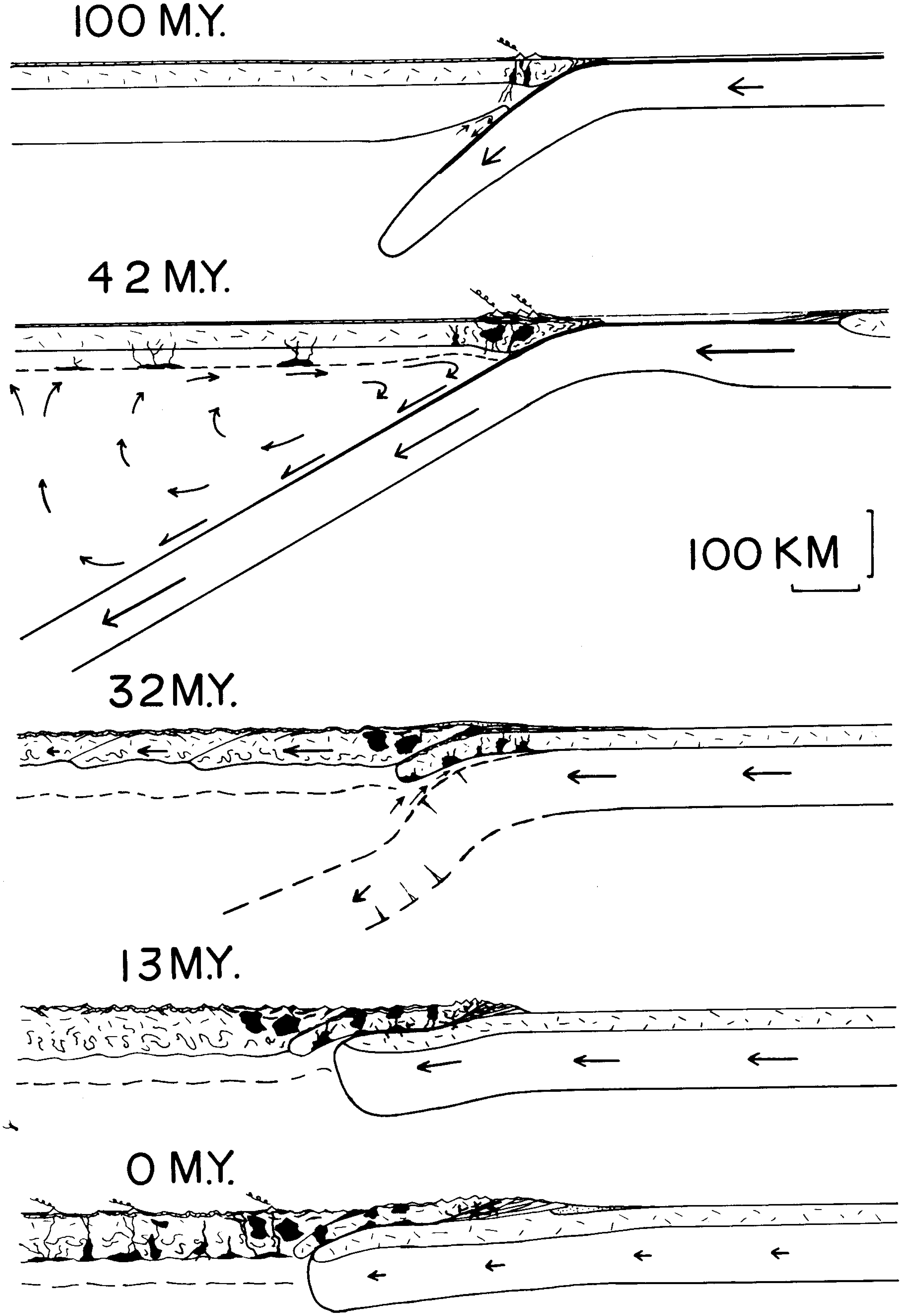


Fig. 5.3

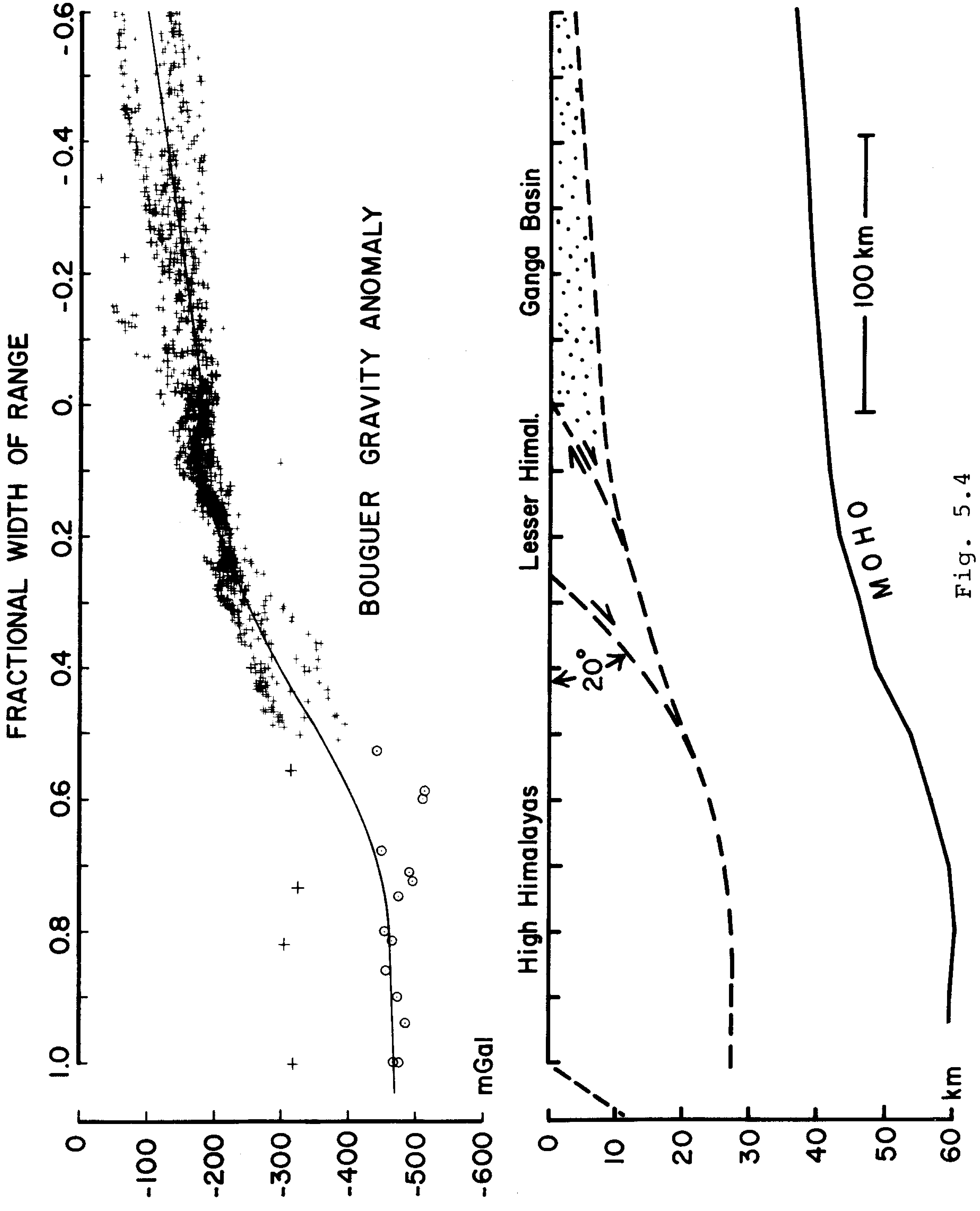

FRACTIONAL WIDTH OF RANGE
1.0
0.8
0.6
0.4
0.2
0.
-0.2
-0.4
-0.6
BOUGUER GRAVITY ANOMALY
mGal
0
-100
-200
-300
-400
-500
-600
High Himalayas
Lesser Himal.
Ganga Basin
20°
MOHO
100 km
km
0
10
20
30
40
50
60

Fig. 5.4

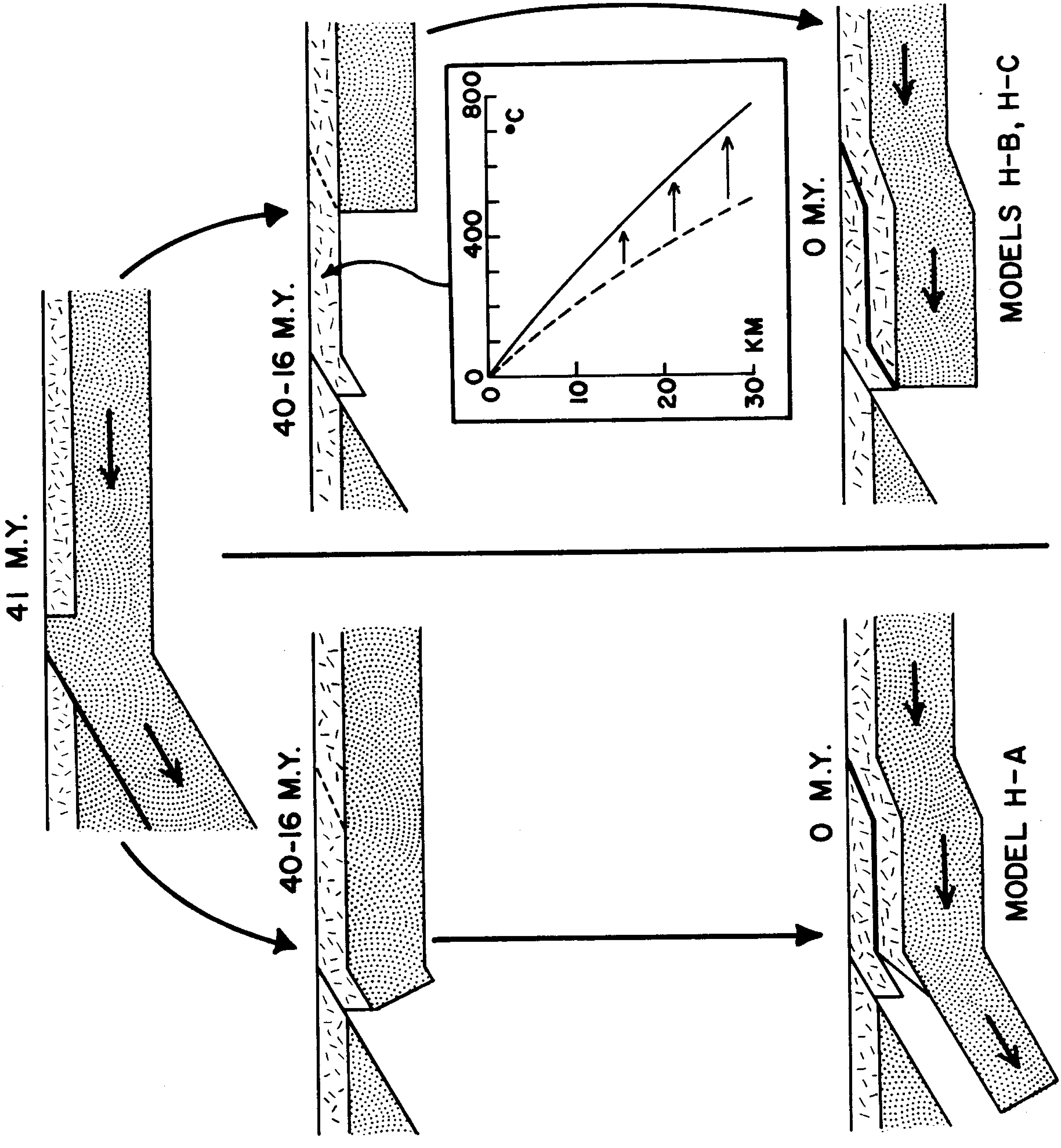

41 M.Y.
40-16 M.Y.
40-16 M.Y.
0 M.Y.
0 M.Y.
MODELS H-B, H-C
MODEL H-A
0
400
800
°C
0
10
20
30
KM

Fig. 5.5


100 mW/m²
50
0
HEAT FLOW
M.B.F.
I.S.
0
50
100
150
200
200
400
600
800
1000
1200
1400°C
400 KM
300
200
100
0
MODEL H-A

Fig. 5.6

100 mW/m2
50
0
HEAT FLOW
M.B.F.
I.S.
0
50
100
150
200
200
400
600
800
1000
1200
1400°C
400 KM
300
200
100
0
MODEL H-B

Fig. 5.7


100 mW/m2
50
0
HEAT FLOW
M.B.F.
I.S.
0
50
100
150
200
200
400
600
800
1000
1200
1400°C
400 KM
300
200
100
0
MODEL H-C

Fig. 5.8

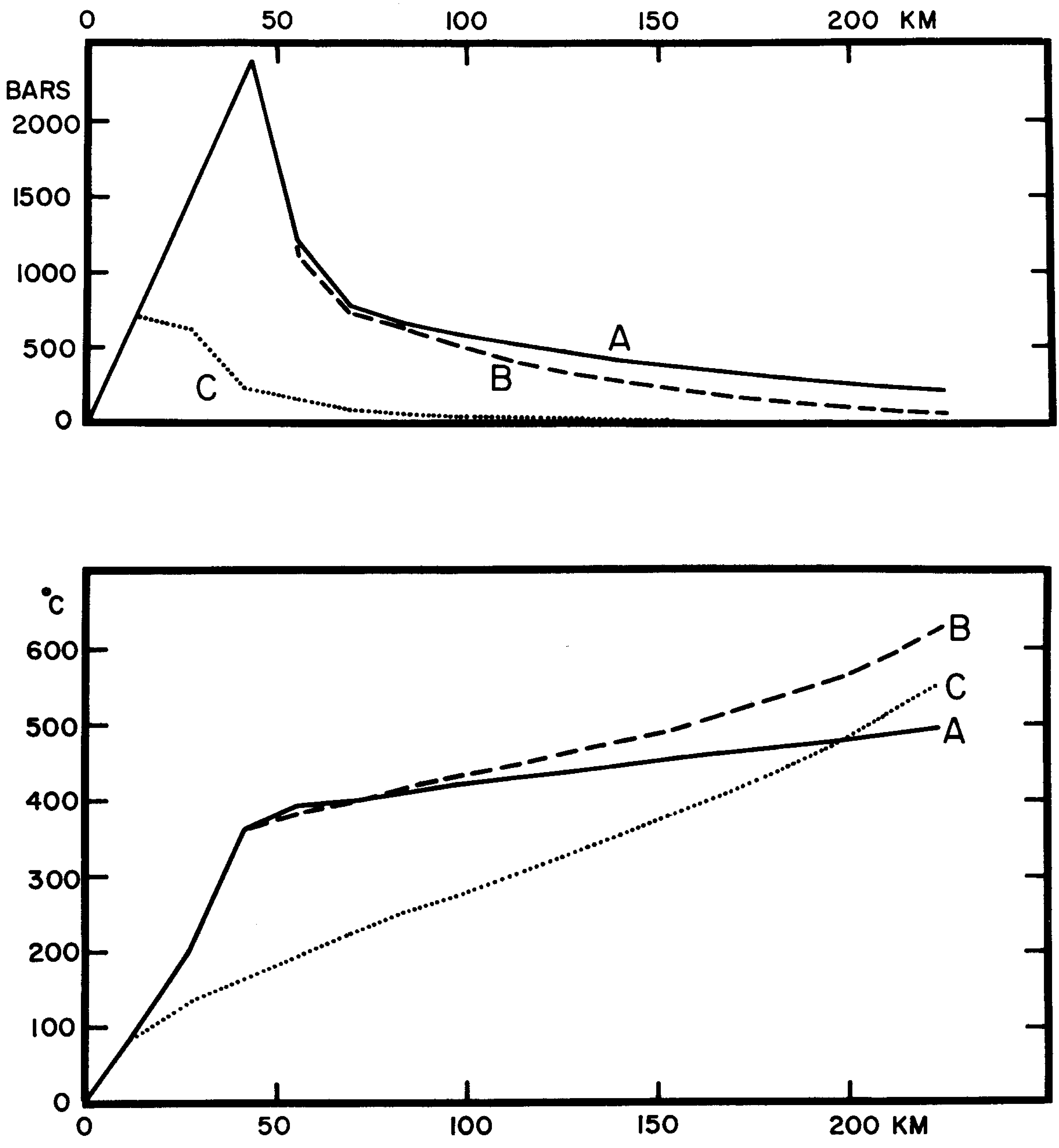
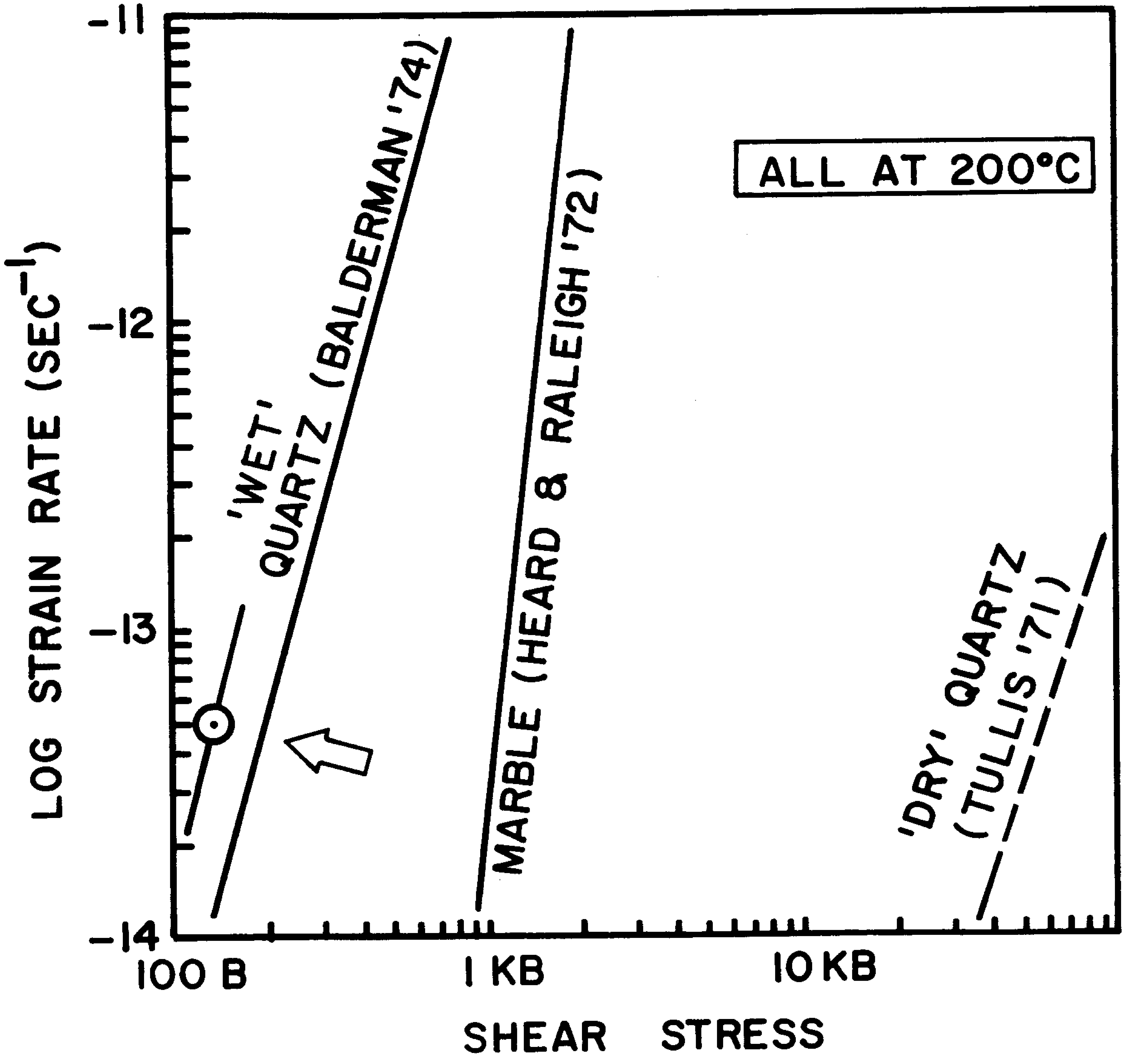

0
50
100
150
200 KM
BARS
2000
1500
1000
500
0
A
B
C
°C
600
500
400
300
200
100
0
B
C
A
0
50
100
150
200 KM

Fig. 5.9



ALL AT 200°C
'WET' QUARTZ (BALDERMAN '74)
MARBLE (HEARD & RALEIGH '72)
'DRY' QUARTZ (TULLIS '71)
LOG STRAIN RATE (SEC⁻¹)
-11
-12
-13
-14
100 B
1 KB
10 KB
SHEAR STRESS

Fig. 5.10


MM/YEAR
5
4
3
2
1
0
-1
U2
U1
INDUS
SUTURE
MAIN
BOUNDARY
FAULT
100
200
300
800
KM

Fig. 5.11


FRACTION OF WIDTH OF RANGE
-0.39
-0.19
0.00
0.20
0.40
0.59
0.79
0.98
1.18
1.38
1.57
0.00
48.00
96.00
144.00
192.00
DEPTH OF EVENT IN KM
MAGNITUDE
8
7
6
5
4

Fig. 5.12


NORTH
FRACTION OF WIDTH OF RANGE
SOUTH
1.57
1.38
1.18
0.98
0.79
0.59
0.40
0.20
0.
−0.20
−0.40
0
48
96
144
KM
T
2-11-66
M
3-14-67
B
1-19-75
M
9-26-64
M
10-21-64
M,R
1-12-65
R
6-01-65
F,M,R
6-27-66
T
6-30-69
M
2-19-70
M
8-15-68

Fig. 5.13


TIBETAN PLATEAU
HIMALAYAS
GANGA BASIN
ALLUVIUM
UPPER CRUST
LOWER CRUST
LITHOSPHERE
ASTHENOSPHERE
0
100
200
300
400
KM

Fig. 5.14

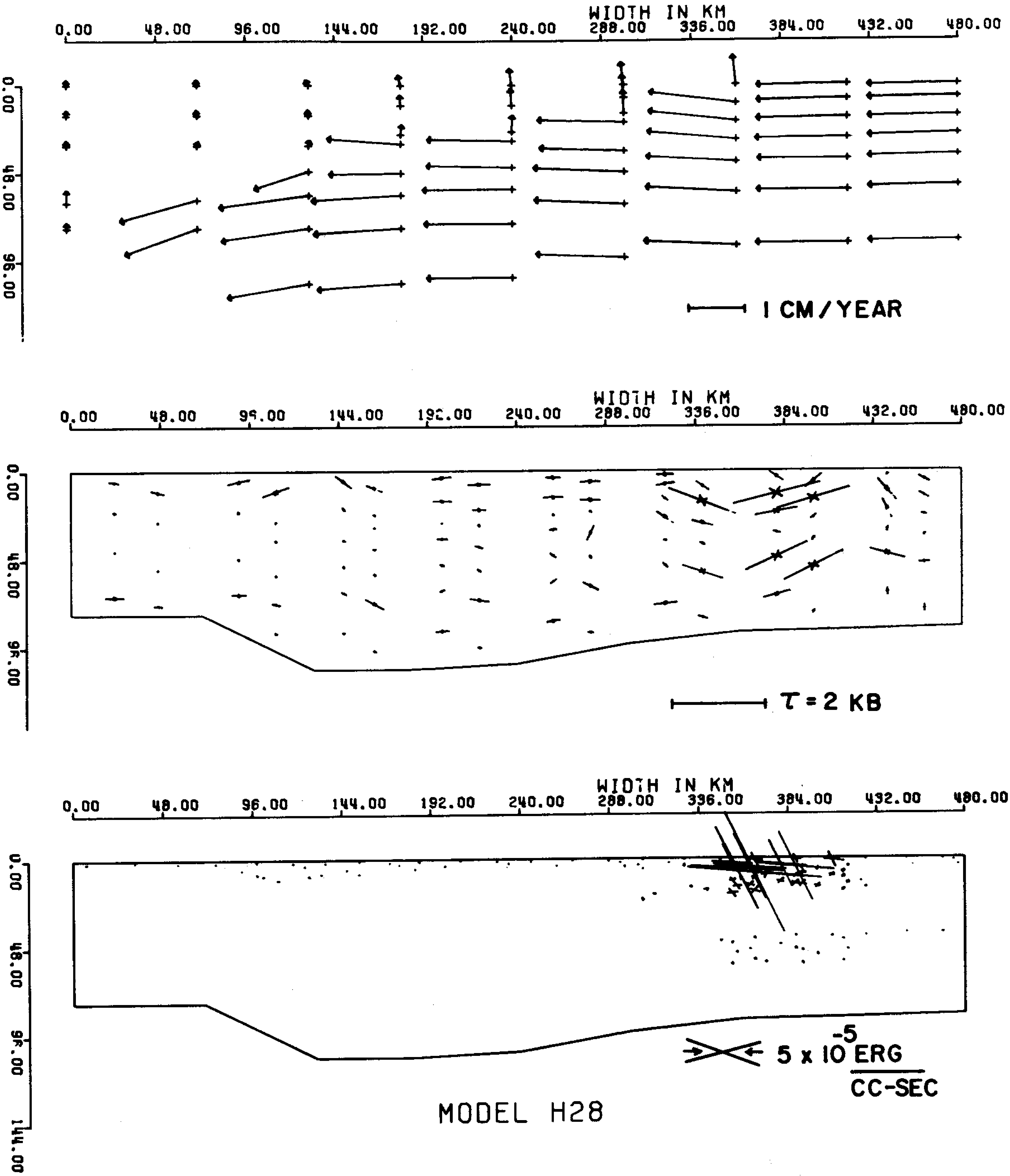
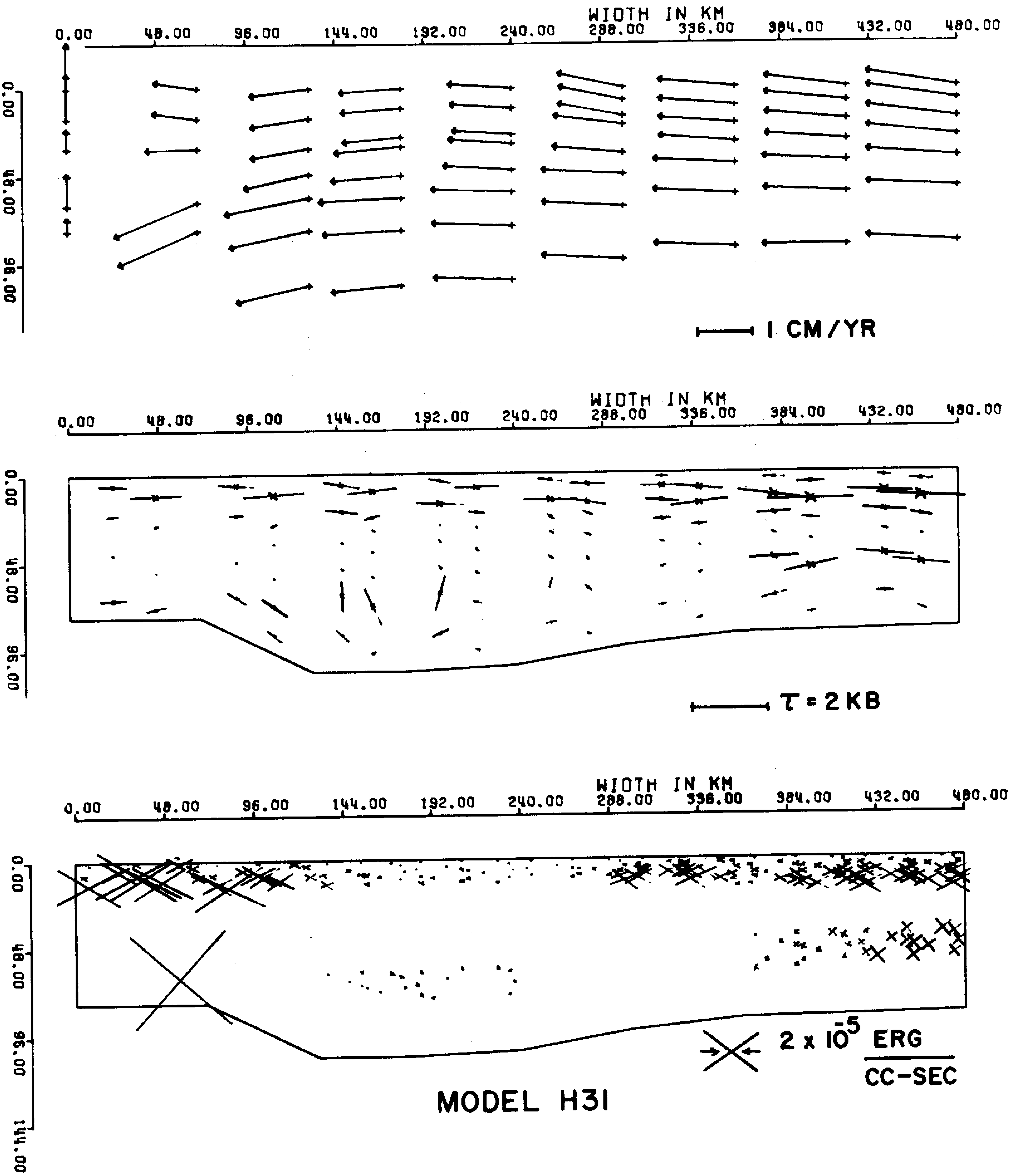
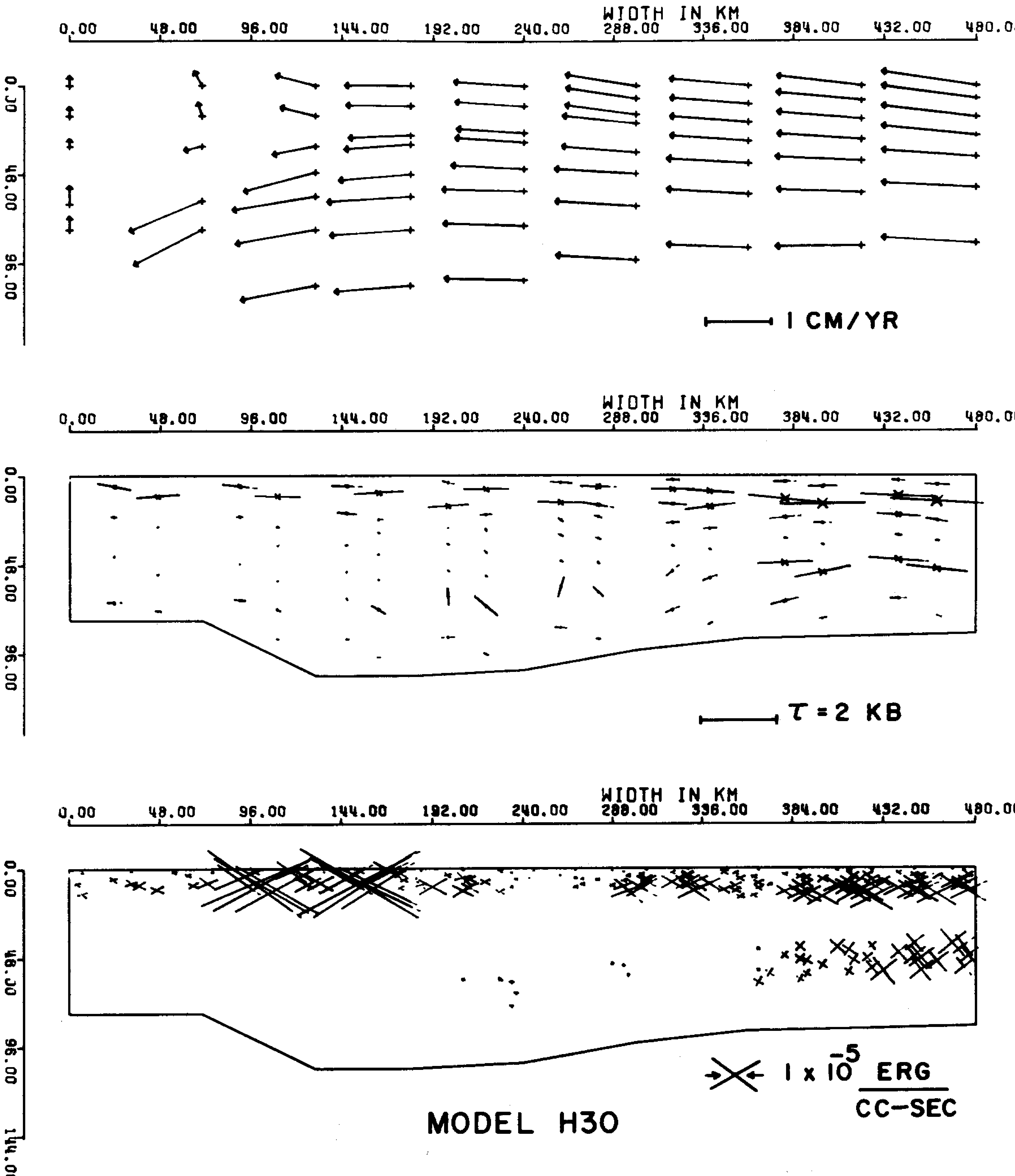

WIDTH IN KM
0.00
48.00
96.00
144.00
192.00
240.00
288.00
336.00
384.00
432.00
480.00
1 CM/YEAR
τ = 2 KB
5 x 10 -5 ERG
CC-SEC
MODEL H28

Fig. 5.15



WIDTH IN KM
0.00
48.00
96.00
144.00
192.00
240.00
288.00
336.00
384.00
432.00
480.00
0.00
48.00
96.00
1 CM/YR
WIDTH IN KM
0.00
48.00
96.00
144.00
192.00
240.00
288.00
336.00
384.00
432.00
480.00
0.00
48.00
96.00
τ = 2 KB
WIDTH IN KM
0.00
48.00
96.00
144.00
192.00
240.00
288.00
336.00
384.00
432.00
480.00
0.00
48.00
96.00
144.00
2 x 10⁻⁵ ERG / CC-SEC
MODEL H31

Fig. 5.16



WIDTH IN KM
0.00
48.00
96.00
144.00
192.00
240.00
288.00
336.00
384.00
432.00
480.00
0.00
48.00
96.00
1 CM/YR
WIDTH IN KM
0.00
48.00
96.00
144.00
192.00
240.00
288.00
336.00
384.00
432.00
480.00
0.00
48.00
96.00
τ = 2 KB
WIDTH IN KM
0.00
48.00
96.00
144.00
192.00
240.00
288.00
336.00
384.00
432.00
480.00
0.00
48.00
96.00
144.00
1 x 10⁻⁵ ERG / CC-SEC
MODEL H30

Fig. 5.17

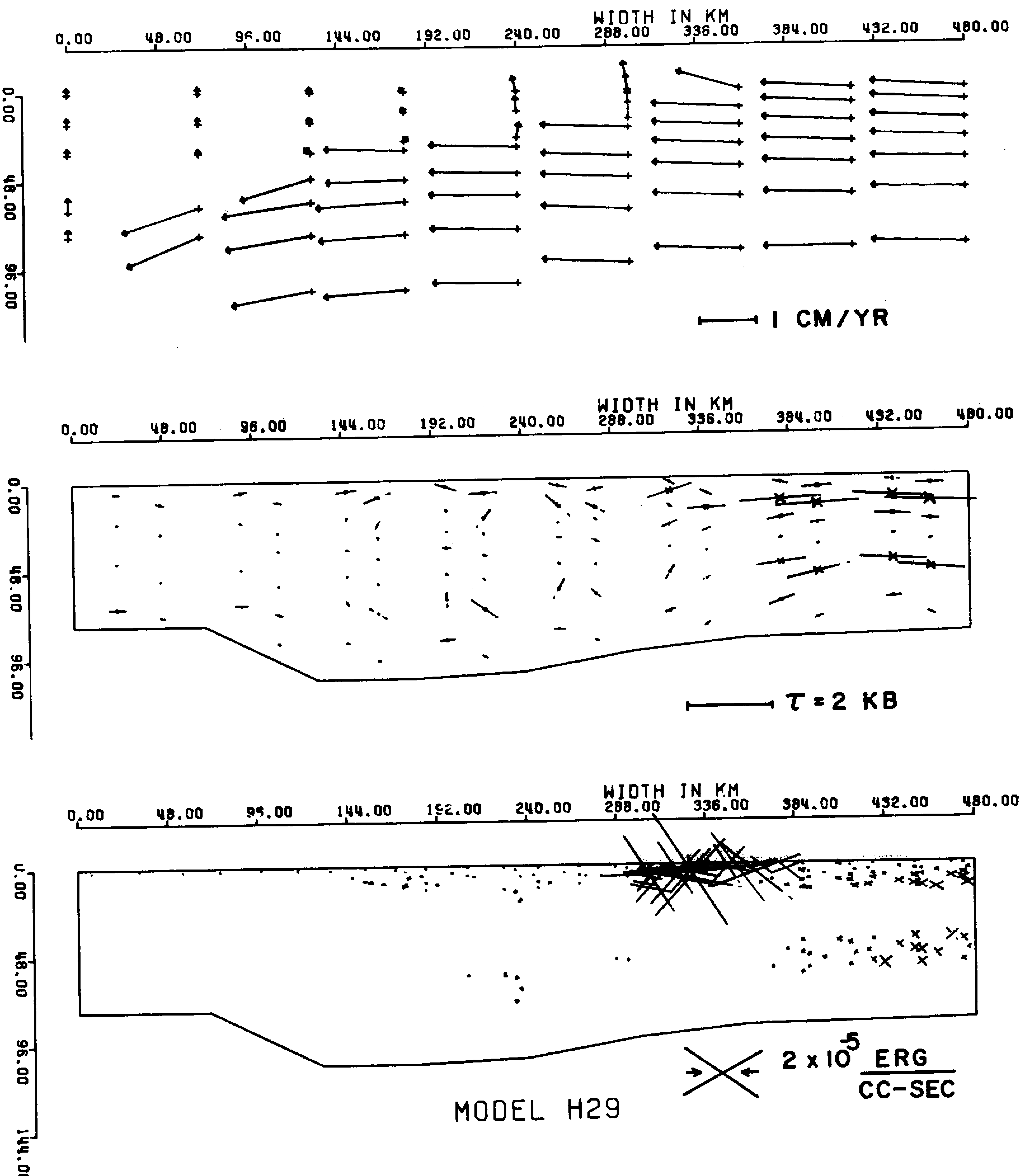

WIDTH IN KM
0.00
48.00
96.00
144.00
192.00
240.00
288.00
336.00
384.00
432.00
480.00
1 CM/YR
WIDTH IN KM
τ = 2 KB
WIDTH IN KM
2×10^{-5} ERG / CC-SEC
MODEL H29
144.00

Fig. 5.18

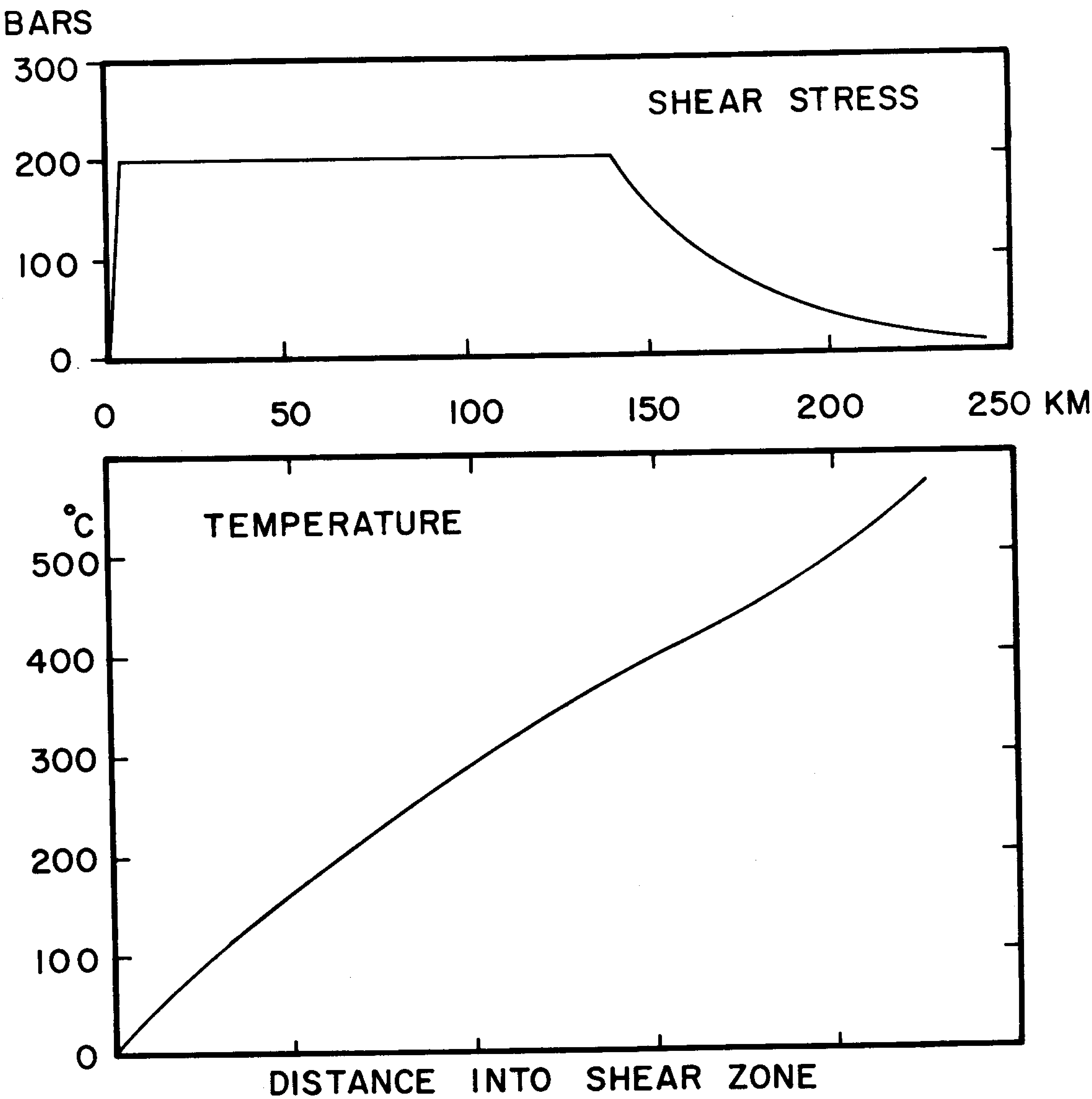

BARS
300
200
100
0
SHEAR STRESS
0
50
100
150
200
250 KM
°C
TEMPERATURE
500
400
300
200
100
0
DISTANCE INTO SHEAR ZONE

Fig. 5.19

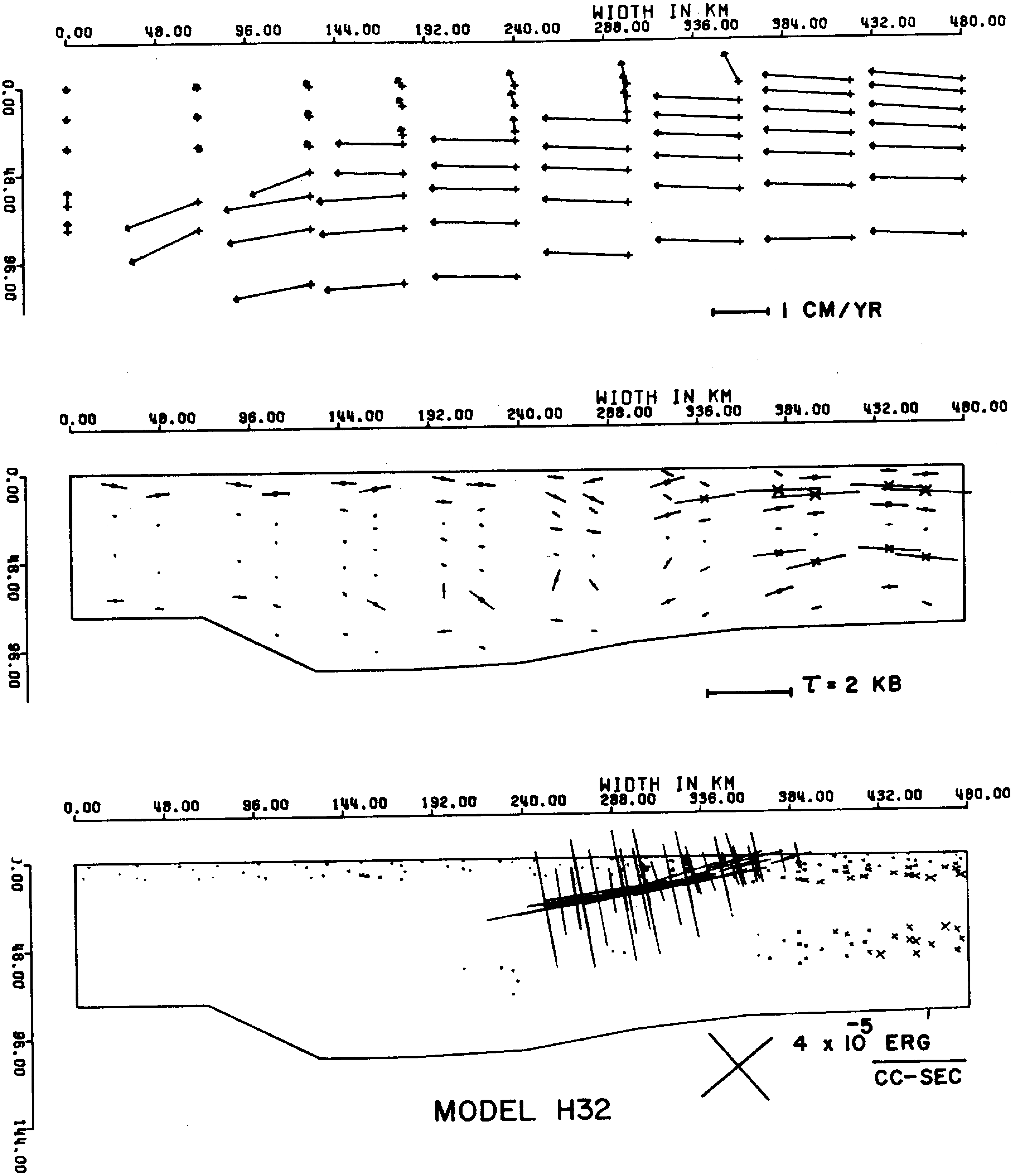
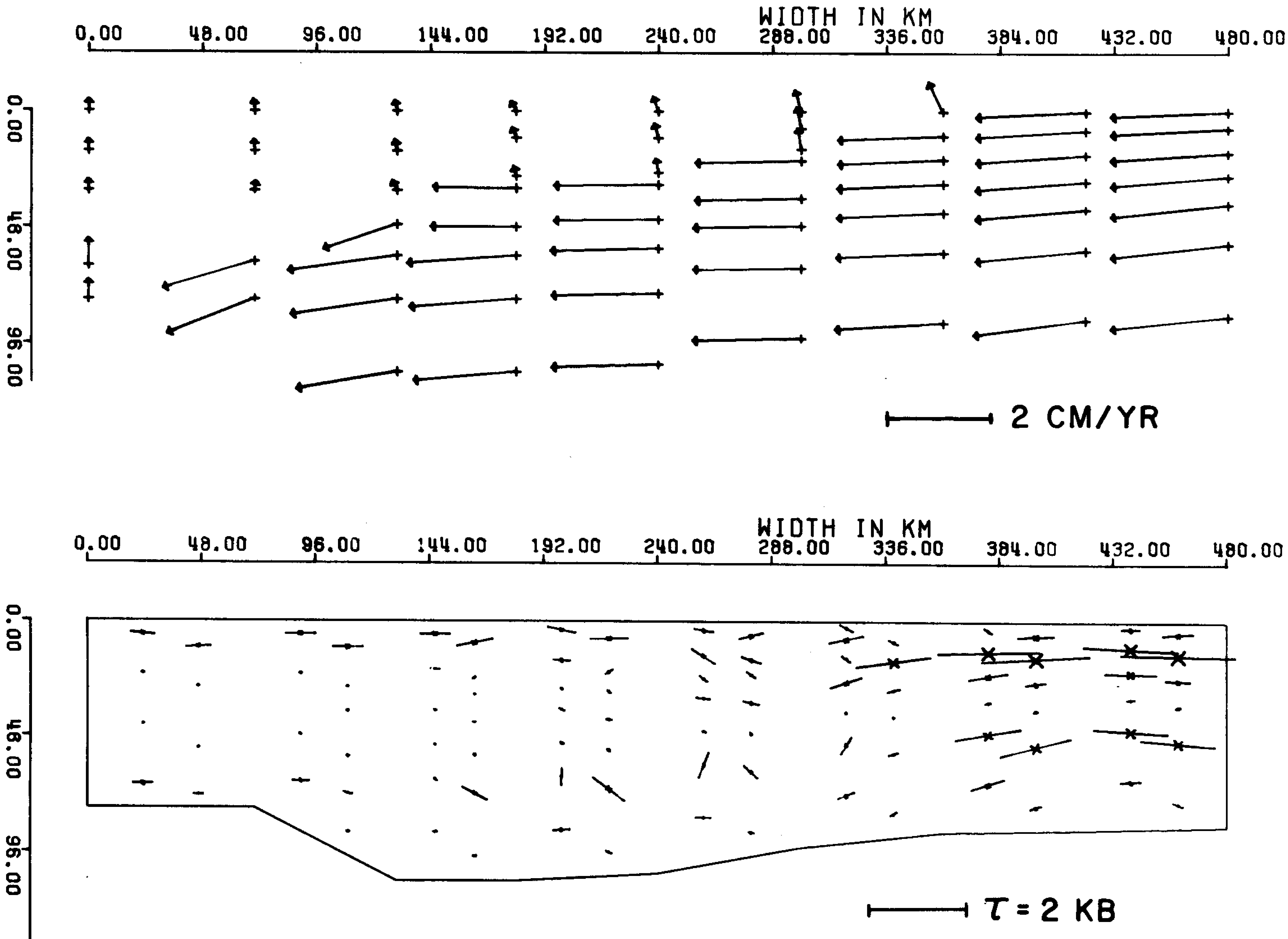
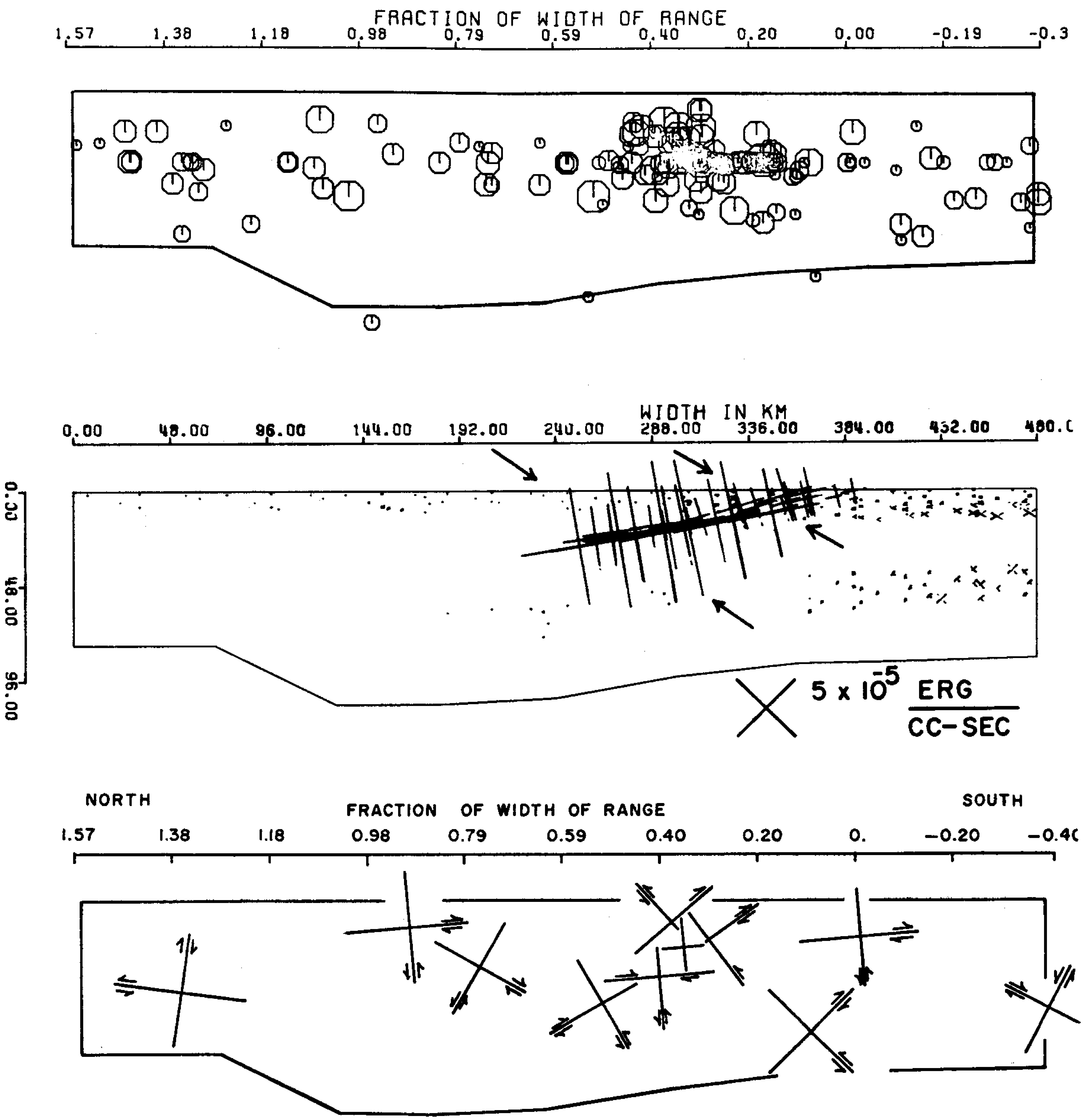

WIDTH IN KM
0.00
48.00
96.00
144.00
192.00
240.00
288.00
336.00
384.00
432.00
480.00
1 CM/YR
τ = 2 KB
4 x 10^{-5} ERG/CC-SEC
MODEL H32

Fig. 5.20



MODEL H38

Fig. 5.21



MODEL H38

Fig. 5.22

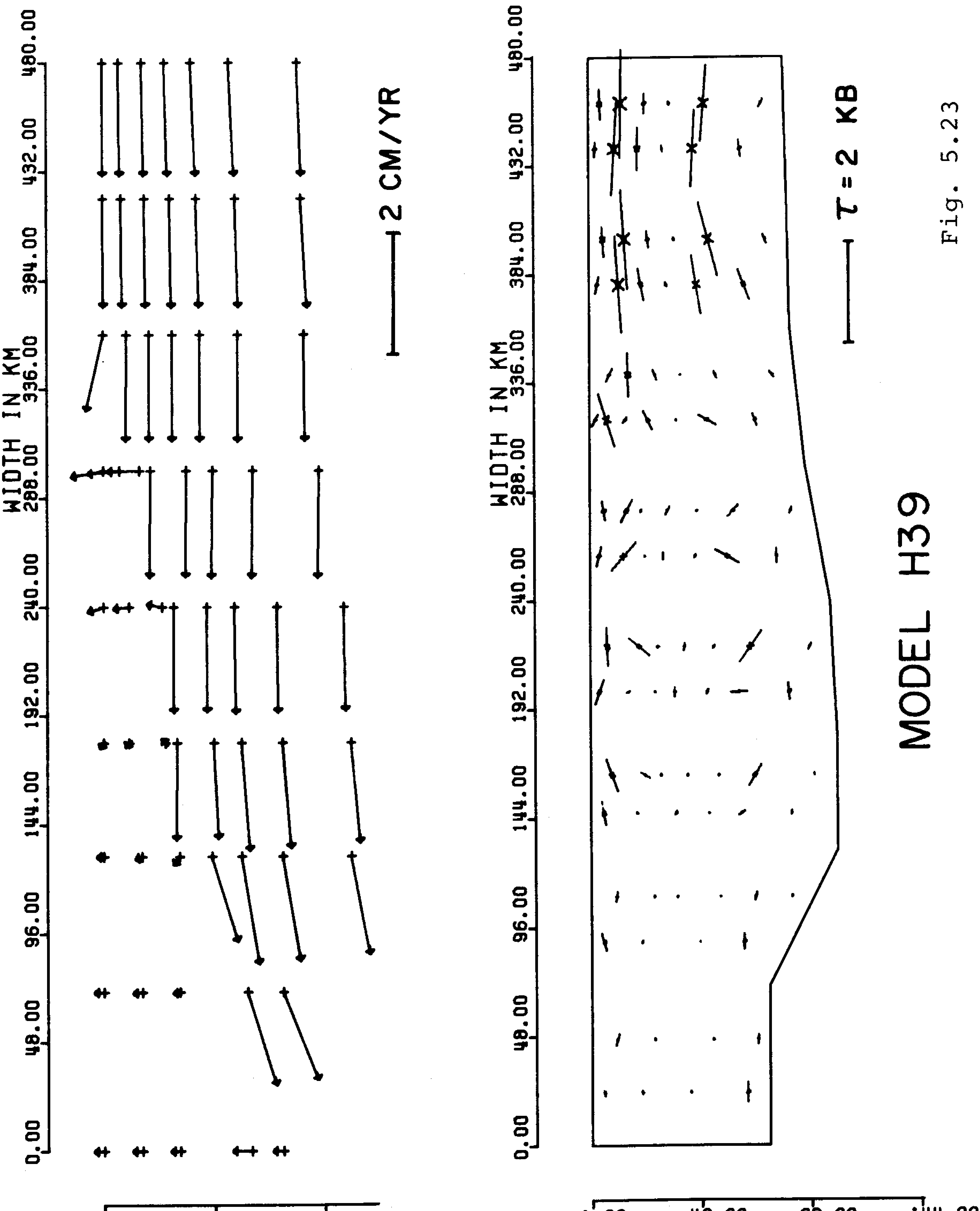

WIDTH IN KM
0.00
48.00
96.00
144.00
192.00
240.00
288.00
336.00
384.00
432.00
480.00
2 CM/YR
τ = 2 KB
MODEL H39

Fig. 5.23

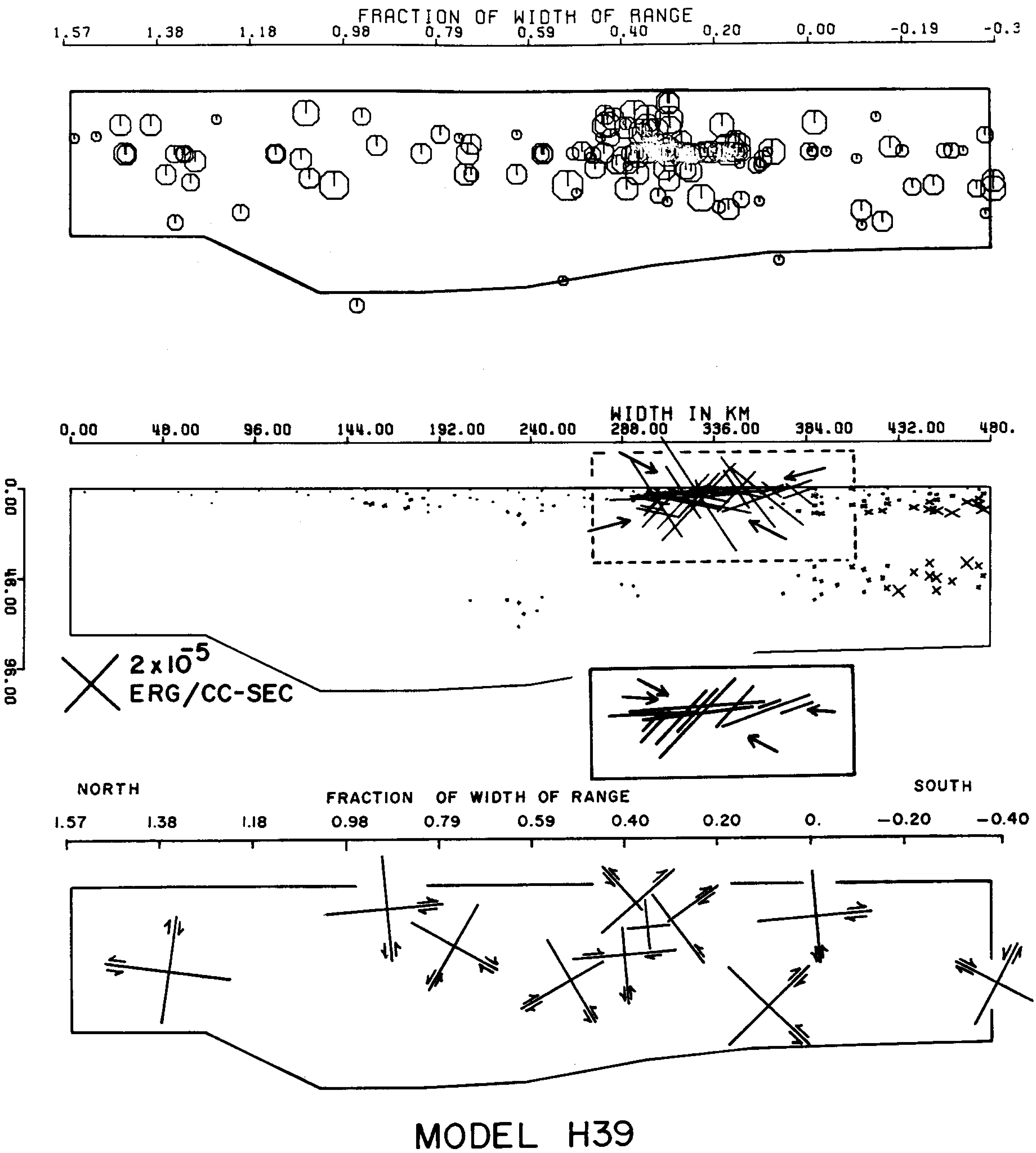
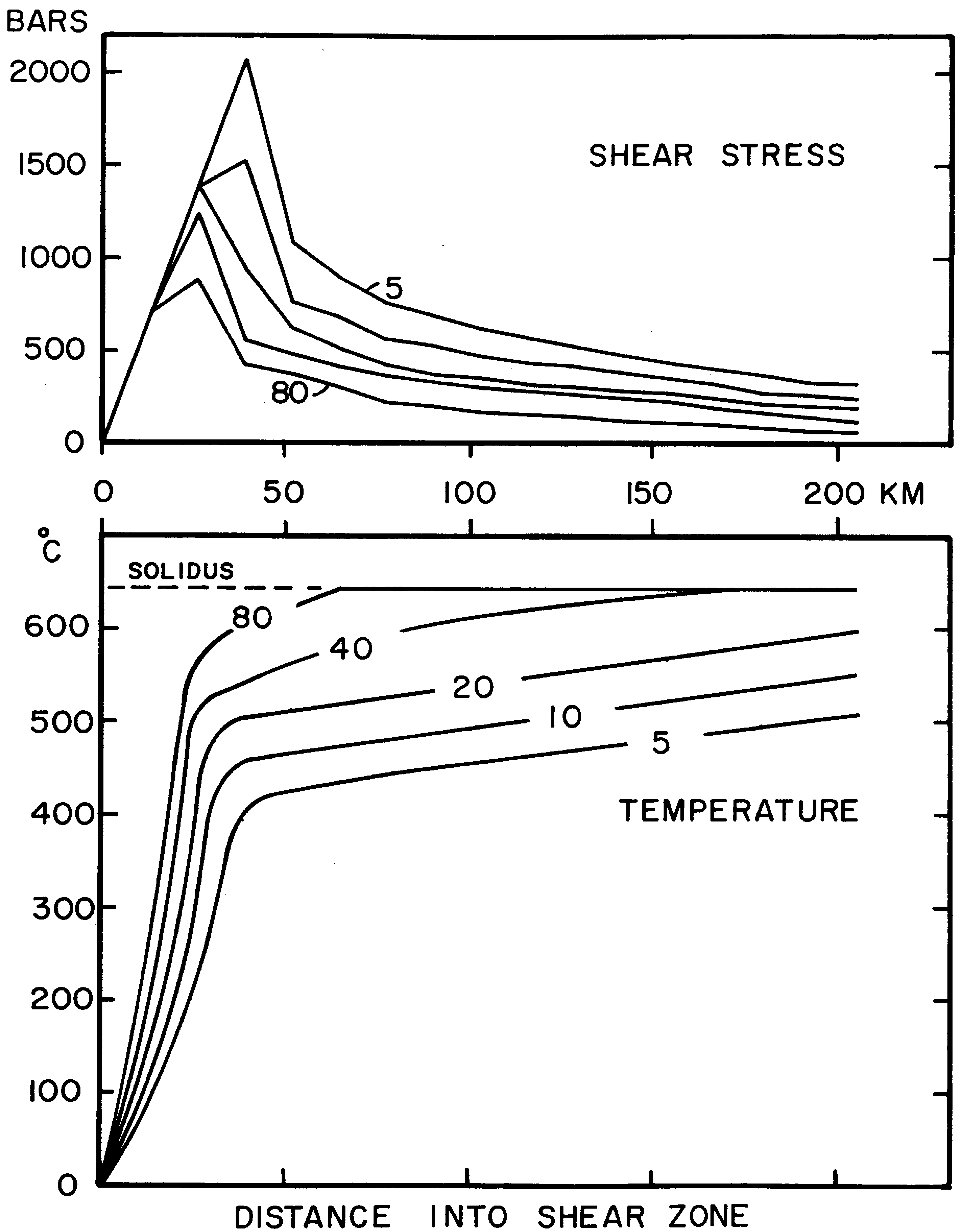

FRACTION OF WIDTH OF RANGE
1.57
1.38
1.18
0.98
0.79
0.59
0.40
0.20
0.00
-0.19
-0.3
WIDTH IN KM
0.00
48.00
96.00
144.00
192.00
240.00
288.00
336.00
384.00
432.00
480.
0.00
48.00
96.00
2×10^{-5}
ERG/CC-SEC
NORTH
SOUTH
FRACTION OF WIDTH OF RANGE
1.57
1.38
1.18
0.98
0.79
0.59
0.40
0.20
0.
-0.20
-0.40
MODEL H39

Fig. 5.24


MM/YEAR
H38
H39
U2
5
4
3
2
1
0
-1
-2
0
100
200
300
400
500 KM
INDUS SUTURE
BOUNDARY FAULT

Fig. 5.25



BARS
2000
1500
1000
500
0
SHEAR STRESS
5
80
0
50
100
150
200 KM
°C
SOLIDUS
80
40
20
10
5
600
500
400
300
200
100
0
TEMPERATURE
DISTANCE INTO SHEAR ZONE

Fig. 5.26

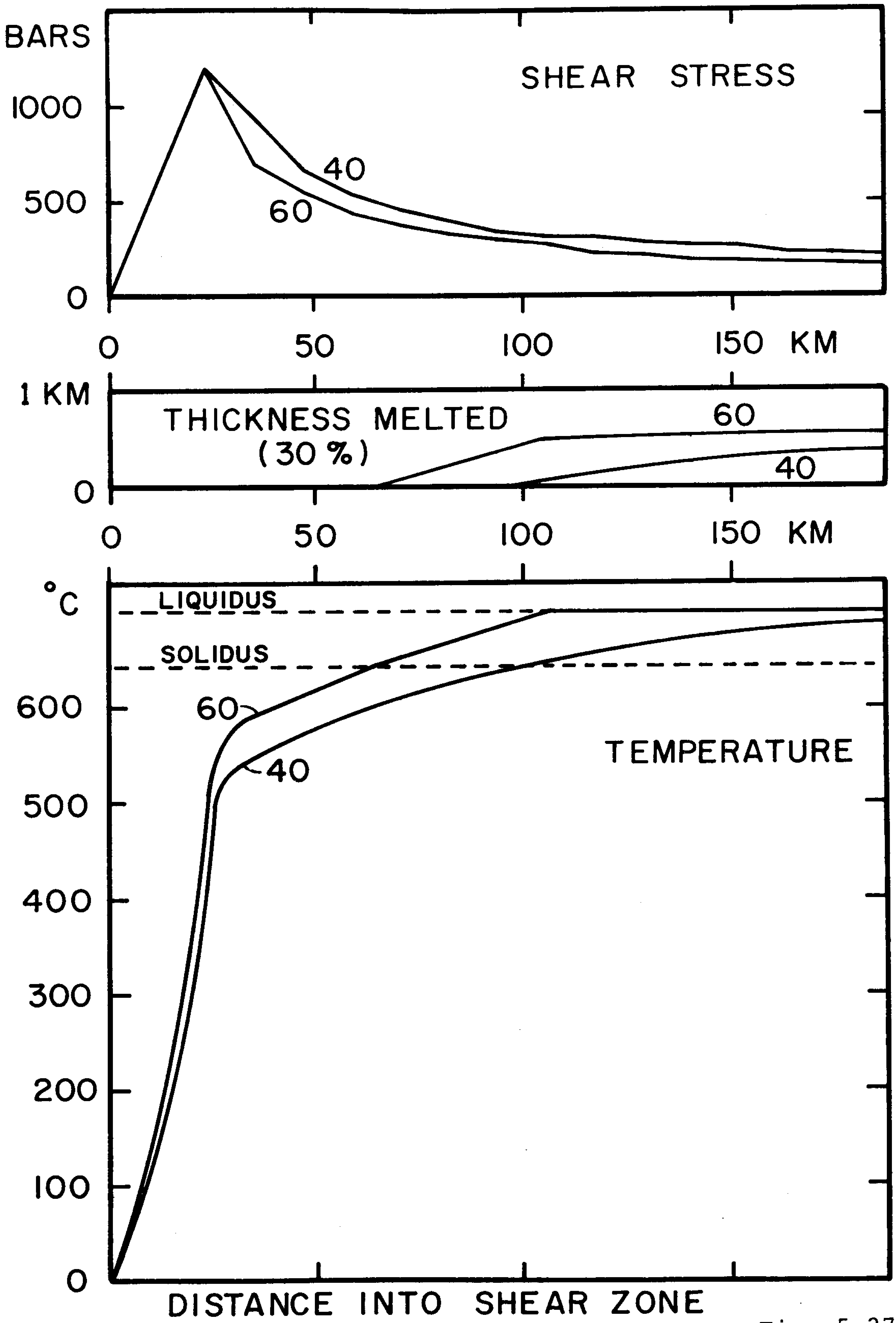

BARS
1000
500
0
SHEAR STRESS
40
60
0
50
100
150 KM
1 KM
THICKNESS MELTED
(30 %)
60
40
0
0
50
100
150 KM
°C
LIQUIDUS
SOLIDUS
600
60
40
TEMPERATURE
500
400
300
200
100
0
DISTANCE INTO SHEAR ZONE

Fig. 5.27

CHAPTER 6: THE TIBETAN PLATEAU: TECTONICS BEHIND A PLATE SUTURE

Ever since the famous lecture of Argand (1924), geologists have regarded the tremendous high plateau of Tibet as a product of the Himalayan orogeny. This much can be inferred from its position alone, lying immediately north of the Indian-Eurasian plate suture and between the meridians of the strike-slip faults that bound the Indian shield on both sides. But many basic questions remain unresolved: How thick is the crust? What supports the high elevations? What source of heat is producing the volcanism across its surface? And what is the temporal and mechanical relation of the Tibetan orogeny to the Himalayan?

Of the five hypotheses discussed by Powell and Conaghan (1975) for the genesis of Tibet, only three are consistent with our present geologic knowledge. The first is that Tibet was elevated by a single great underthrust of the Indian continental crust under the entire plateau, creating a double thickness (Argand, 1924; Carey, 1955; Holmes, 1965; Powell and Conaghan, 1973; Powell and Conaghan, 1975). The second is that the crust was thickened by lesser underthrusts as a series of three microcontinents successively collided with Asia prior to the Himalayan orogeny (Chang and Zeng, 1973). The last is that Tibet was deformed by a combination of mantle heating during the subduction of the Tethys and horizontal compression following the Himalayan collision (Dewey and Burke, 1973; Burke et al. 1974; Bird and Toksöz, 1975; Toksöz and Bird 1976).

The purpose of this chapter is to critically review the existing geologic and geophysical knowledge of Tibet in terms of these three hypotheses. New data on Rayleigh wave dispersion and attenuation will be used to show that the crust is twice as thick as normal and highly attenuating near the Moho. Volcanism in Tibet, examined in the light of preceeding thermal models, will be shown to require anomalous temperatures prior to the continental collision. The hypothesized source of these temperatures, induced convection above downgoing slabs, may be a universal feature of subduction zones.

6.1 Geomorphology and Geology of Tibet

The Tibetan plateau is a lens-shaped region lying between the Indus-Brahmaputra ophiolite belt and the Paleozoic Kun Lun range (Figure 6.1). On the west it pinches out where the two ranges converge; on the east its boundary is poorly defined because of heavy erosion destroying its plateau character. The total area is 700,000 km^2 in which the average elevation is 5.1 km.

Although Tibet contains many mountain chains of Himalayan proportions and rugged topography, the elevation of the valley floors is surprisingly uniform. According to the Operational Navigational Charts G-7, G-8, H-9, and H-10 published by the Defense Mapping Agency, the highest non-glacial lake or salt playa is at 6.0 km elevation and the lowest at 4.3 km. The base level is generally about 500 m higher in northern Tibet (the Chang Tang "platform") than in the south (Nyenchhen Thanghla "basin"). This minor contrast is somewhat obscured by apparently random swells and dips of long wavelength. There is no indication in the topography of a northward dip such as might be expected above the leading edge of an underthrust sheet of continental lithosphere. If that theory is correct, the shape of the leading edge of India must correspond closely to the shape of the old Kun Lun chain, either by coincidence or because of deformation of the Kun Lun.

Except on its southern and eastern margins, Tibet has no exterior drainage. Instead, its streams flow into some 1250 lakes and playas from which the scant rainfall evaporates again. Thus the average area integrated into a single drainage system is only 600 km^2, and the RMS radius of the system is only 14 km. This youthful geomorphology is partially a product of the Himalayan rain-shadow, but still argues for a very recent uplift of the plateau. It was found that there is no significant change in the extent of drainage integration (as measured by density of lakes and playas) from South to North. Barring meteorological complications this might be taken to mean that the uplift was everywhere of equal age. In contrast, the multiple-collision model would require the northern portions of the uplift to be formed in the Paleozoic, and the single-underthrust theory would predict earlier uplift in the south.

The geology is dominated by extensive Cretaceous marine carbonates, mapped as far north as 35°N, 84°E by Hennig (1915). As discussed in Chapter 5, these carbonates place an upper limit on the age of the uplift. This limit is far too recent to be accommodated by the multiple-collision theory of Chang and Zeng. The limestones show that the events which Chang and Zeng have dated (by K/Ar methods) were followed by periods of stability, and are not related to the present uplift. Furthermore, there is no geologic evidence for additional plate sutures between the Himalayas and the Kun Lun (Maurice Terman, personal communication, 1976), and thus no independent confirmation of multiple collisions.

In the few published surface photographs of Tibet (Anonymous, 1974), these sedimentary rocks are everywhere seen to be tilted, folded, and strongly deformed. The lack of visible fault scarps along valley margins suggests that their flat floors are depositional rather than graben structures. Compressional tectonics appear to dominate. Unfortunately, these pictures represent only the populous southeastern portion of the Nyenchhen Thanghla, and may not be representative.

In LANDSAT images compiled by Tapponnier and Molnar (1976b) there are no mesas and canyons evident. These would be expected in a dry elevated region (like the Colorado Plateau) if the strata were still horizontal. Instead there is a pervasive "tectonic fabric" generally parallel to the bounding Himalayas and Kun Lun, which converges toward two vergations at the eastern and western ends of the plateau. A few grabens with a NNE trend are superimposed. No major throughgoing faults are recognized. Although there is no proof that deformation and uplift were simultaneous, Tibet has clearly been subjected to fairly homogeneous North-South compression in Tertiary times.

Overlying these folded rocks in the eastern half of Tibet (80°-93°E) are two major bands of Neogene-Quaternary volcanics, mapped from LANDSAT photographs by Kidd (1975). While the southern band adjacent to the Indus-Brahmaputra suture probably began to form prior to the collision, reports of boiling springs suggest continuing intrusion. The northern band is believed to be very young and active on the basis of well-preserved cones and flows, and a possible eruptive cloud in one image. Samples of these volcanics were collected by Hedin's expedition, and are principally andesites, dacites, and rhyolites with very little basalt (Dewey and Burke, 1973). On the basis of chemical and isotopic analysis (Kevin Burke, personal communication, 1975) these volcanics are believed to originate from partial melting of the lower continental crust of Tibet, where the composition would be predominantly dioritic. Unfortunately, no dates have been obtained.

These young volcanics present a major problem for the theory of crustal thickening by underthrusting. As was shown in model III of Chapter 2, an underthrust of this type will not produce any melting of crust, even after radioactive heat has accumulated for 40 m.y. (the maximum possible age according to this theory). The temperature should be even lower than in that model, because the extreme geometry of the thrust proposed by Powell and Conaghan (1973, 1975) implies little or no frictional heating along the fault. If the crust was thickened in this way following the Himalayan collision, it should be cool, thick and underlain by normal lithosphere.

6.2 Geophysical Evidence

The only gravity information in Tibet results from two pendulum stations (#463, #467) and 33.5°N by 80.2° and 84.8°E occupied by Ambolt (1948). The Bouguer anomalies at these sites are some of the largest ever recorded (-566 and -565 mGal) and indicate a mass deficit of 1,350 kg/cm^2 below sea level. However, free air anomalies were very small (0 and +18 mGal) and show very close isostatic compensation of the plateau. It is not surprising that Tibet is compensated, since otherwise a shear stress of 3 kb down to 100 km depth would be required all around its border to support it.

Assuming that the compensation is general, it is still unclear how much results from thick crust and how much from low-density mantle. A minimum crustal thickness can be calculated if it is assumed that the mantle under Tibet is no lighter than the mantle under mid-ocean ridges. There, according to plate tectonic theory, the hot and mobile asthenosphere extends almost to the surface. Since ridges are isostatically compensated within 20 mGal (Lambeck, 1972) and rise about 2700 m above old sea floor (Sclater et al., 1971) the mantle mass deficit may be as much as 600 kg/cm^2. Subtracting this figure from Tibet's, we find that the Moho must still be depressed some 17 km (assuming a 0.43 g/cc contrast across it). Considering that 5 km of the Tibetan crust is above sea level, its total thickness is not likely to be less than 55km. If it is this thin, then the implied high temperatures in the upper mantle should be detectable by geophysical means. Otherwise the crustal thickness might be as much as 70 km, if the mantle makes no contribution.

One indication of high temperatures at depth is the lack of magnetic anomalies. According to Marussi (1964), both the vertical and horizontal components of the anomalous magnetic field fall to zero on the northern, western, and southern boundaries of the plateau. The simplest explanation is that the Curie point isotherm is upwarped and the magnetized layer is unusually thin. For magnetite the Curie point is about 580°C, which would normally be attained just below the Moho (Chapter 2). If this temperature is actually reached at shallower depths in Tibet where the Moho is unusually deep, then the crustal underthrusting model is greatly weakened.

Another indication is poor transmission of seismic waves in the crust and upper mantle. Using the records from the station at Lhasa, Detrick (1976) determined an average P_n velocity of 7.9 km/sec for Tibetan paths in the distance range of 2-10°. Similar low P_n velocities in the western U.S. are linked to high heat flow and attenuation (Herrin, 1970). According to the underthrusting model, P_n velocities should fall in the normal range 8.1-8.3 km/sec. Furthermore, Molnar and Oliver (1969) found that the shear wave phase S_n

does not propagate across Tibet, which indicates a generally low Q_b and probably high temperatures. The possibility that high temperatures extend into the crust is suggested by the discovery that the guided crustal wave L_g also will not cross Tibet (Ruzaikin et al., 1976). However, the interpretation of this observation is controversial, because changes in crustal thickness may also scatter the energy at the boundaries of Tibet.

Present seismicity in Tibet does not support any of the theories on its mode of uplift. It is moderately intense and well distributed, without any particular lineations such as the multiple-suture model would suggest. Fault-plane solutions with very shallow or horizontal planes have been published by Rastogi (1974) and by Tandon and Srivastava (1975), but in each case the sense of motion disagrees with the underthrusting model (lower block moves south). The horizontal-compression model is only slightly more successful; only two solutions in Tibet have North-South compressional axes (Tapponnier and Molnar, 1976b). Apparently more representative are normal-faulting events with East-West extension; four of these were identified by Tapponnier and Molnar, 1976b). In these events the North-South direction is the intermediate or average stress direction. Perhaps the simplest interpretation of the wide variety of fault-plane solutions is that the crustal thickening mechanism is no longer active. The recent earthquakes may represent only local readjustment in response to viscous relaxations, erosional stresses, and distant earthquakes on adjacent plate boundaries.

6.3 Rayleigh Wave Group Velocities

Because Tibet appears to be very uniform in topography, geology, and geophysics, it can reasonably be represented as a horizontally-layered structure. Surface waves are ideal for studying such structures, and Rayleigh waves are preferable to Love waves in this case because they cannot be contaminated even if lateral refraction causes waves to arrive at an unexpected angle of incidence. Unfortunately, since no long-period seismometers are operated in western China, only group velocities of the surface waves can be determined.

Both the underthrusting and the horizontal-compression model have been simplified to crude horizontally layered velocity-density models in Figure 6.2. Here the underthrusting model is expressed in two variations, one in which the weaker portion below 60 km (section 3.2) is "scraped off" during the underthrusting. Also shown is a non-isostatic model in which normal continental shield is uplifted by deep forces. Because Rayleigh wave group velocities are very sensitive to the position of the Moho(s) they can easily be used to distinguish between models.

Previous workers have determined that group velocities are low in Tibet (Santo, 1965; Gupta and Narain, 1967; Tung and Teng, 1974). However, resolution of the Tibetan velocities has been poor because long paths were used. Santo assumed the shape of the dispersion curve and attempted to find its magnitude by regionalization. Gupta and Narain made an implicit regionalization by assuming that all of the travel-time "anomaly" relative to normal shield between the Arctic and India was caused in the Himalayas and Tibet. Tung and Teng used long paths divided into Tibetan and East China regions and other paths not crossing East China. However, two of their "Tibetan" paths lie mostly in the Himalayas and another crosses the Takla Makan desert, Assam, and Burma.

To avoid the uncertainties introduced by regionalization, this study employs a large number (20) of short paths. All were chosen to include at least 50% Tibetan Plateau, as shown in Figure 6.3. The parameters of the earthquakes are given in Table 6.1. An attempt to correct for path segments outside of Tibet will be described below.

In each case the long-period vertical component record of a WWSSN station was digitised at an interval of approximately 1.4 sec. Missing points were detected and interpolated, the average and linear trend of the date were removed, and a cosine bell taper was applied at both ends. Then seismograms were subjected to the multiple-filter analysis of Dziewonski et al. (1969). This technique produces a plot of spectral amplitude (kinetic plus potential energy of the wave in a spectral window) versus time and period. The maximum amplitude at each period is used to define the group arrival time and hence the group velocity. The frequency window employed was one of constant relative bandwidth, with a Gaussian amplitude response of

$$L(\omega) = \exp\left(-\lambda\left(\frac{\omega-\omega_0}{\omega_0}\right)^2\right) \tag{6.1}$$

and the variable parameter λ was set to 50.3.

Figure 6.4 shows all the clear spectral amplitude maxima from the twenty seismograms. Results are similar to those of Tung and Teng (1974), but somewhat slower because of the larger precentage of Tibet in the paths. Also these records show a velocity minimum at 30-40 second periods which was not found by Tung and Teng. At 40 seconds the waves travel at only 2.6-3.1 km/sec, or at least 25% slower than in normal shield. As the model curves show, these velocities cannot be matched to any structure with high-velocity mantle material in the upper 60 km. Thus the two underthrusting models of Figure 6.2 are ruled out. The possibility of an extreme model in which all the lithosphere is removed during the underthrusting (as suggested by Powell and Conaghan, 1973) is investigated below.

As low as they are, these group velocities may still be artificially high because of the inclusion of non-Tibetan path segments. An attempt was made to correct for these, by measuring the path fractions in the Indian shield and in non-Tibetan parts of China and subtracting off the travel times determined from standard curves. The group velocity standard in China was the dispersion curve on Tung and Teng (1974); for India I assembled results of Tandon and Chaudhury (1964), Chaudhury (1965) and Chaudbury (1966) into a single curve. No attempt was made to distinguish the Himalayas from Tibet, because their thick crust produces a similar dispersion. However, the correction was a failure because it did not decrease the scatter of the data. Instead, the standard deviation relative to the best-fitting cubic polynomial was increased from 0.14 km/sec to 0.21. Similar scattering occurred when each correction was applied separately. Either the standard curves used were very inaccurate or the assumption of straight great-circle paths is faulty. The latter seems more likely, although a long-period Indian array would be needed to resolve the question. In any case, it is the uncorrected group velocities which are used here in the interpretation, with the understanding that they give only minimum values for the Tibetan velocity anomalies and crustal thickness.

In order to investigate the trade-off between crustal thickness and mantle velocity a set of 35 model dispersion curves was calculated in which only these two parameters were varied. The simple starting model is given in Table 6.2; it has lower crustal velocities appropriate for diorite

(Press, 1966), and a 5 km sedimentary layer on top with velocities chosen to match the short-period dispersion observed by Chen and Molnar (1975). When crustal thickness was changed the change was divided evenly between layers 2-7. When lowered shear velocities were tested, the anomaly was applied to the lower crustal and upper mantle layers 5-8. P-wave velocities in these layers were also reduced by a smaller amount, employing the assumption that the bulk modulus remained constant. From each model group velocities were predicted using the flat-Earth program of Harkrider based on the method of Haskell (1951). These results were tested against the average group velocity curve of the twenty seismograms, which has been determined by stacking their spectral amplitudes and finding the maximum power versus time.

The mismatch between models and observations is expressed as an RMS group velocity error and contoured in Figure 6.5. An error of less than 0.10 km/sec is considered acceptable, although models in the center of the shaded band matched within 0.03 km/sec. Essentially, any crustal thickness allowed by isostasy (55-70 km) is also consistent with the observed group velocities, if the smaller thicknesses are accompanied by low-velocity mantle and the larger ones are not. Of the models matching group velocity observations, the one with a 55 km crust has a P_n velocity of 7.9 km/sec as observed, a low velocity consistent with volcanism, low magnetic field, and attenuation of S_n. It also gives the most conservative estimate of the amount of tectonic deformation that went into thickening the crust. But larger crustal thicknesses are also possible if the crust and mantle velocity anomalies are uncoupled.

Because of economic considerations these preliminary models were rather coarse, with few layers and not extending very deep. The successful models were subdivided and extended using upper mantle velocities of Toksöz et al. (1967). During the recomputation it was found that upper mantle velocities were slightly too high and they were adjusted down to a shear velocity of 4.26 km/sec. Complete parameters of these final models are given in Table 6.3. The fit of the models to the group velocities is very good (Figure 6.6). They also match excellently some phase velocities determined by Patton (1975) for paths between the Hindu Kush and Shillong or Changmai (Table 6.3).

Also shown in the figure (6.6) is a similar model with the same mantle structure but with a 66-km crust created by superimposing two normal layers of the Gutenberg continental Earth model. The fit of this model is distinctly worse even with the anomalous mantle. Considering the topographic, thermal, and seismic data already discussed, it appears that the hypothesis of large scale crustal doubling by underthrusting in Tibet can be eliminated.

6.4 Rayleigh Wave Attenuation

The observation that Rayleigh wave attenuation is strong in Tibet grew out of the difficulty encountered in trying to collect good long-period records for group velocity analysis. Many events produced good Rayleigh waves at western stations but poor ones in India. For example, Figure 6.7 contains the Rayleigh waves from the Chinese subaerial nuclear test at Lop Nor on Oct. 14, 1970. The record at Kabul of waves passing through the Takla Makan, Karakorum, and Hindu Kush is of classic form and clearly contains energy with periods of 12 to 60 seconds. Yet on paths to New Delhi or Shillong, which include the Tibetan Plateau, almost all of the energy above 20 seconds is absorbed.

To investigate this effect further, a total of seven events were found for which a good Rayleigh wave record was available from at least two stations. In each case one station (KBL or HKC) has a path which does not cross Tibet and which serves as a reference (Figure 6.8). Parameters of all the events are listed in Table 6.4. All together there are 16 paths crossing Tibet.

As before, each seismogram (LPZ component) was digitized at 1.4 sec intervals in a group velocity window of 4.0-2.5 km/sec. This excludes the S waves and most of the undesirable laterally-refracted Rayleigh waves. In most cases higher modes were not noticed; when they were, they were of shorter period (ca. 5 sec.) than the range of interest, and the digitizing interval was reduced to prevent aliasing. Each digitized string was checked for missing points, detrended, tapered at the ends, and Fourier analysed for its amplitude spectra $A(\omega)$. Because multipathing and digitizing errors can both produce artifical lows in the spectrum, attenuation analysis was not based on direct amplitude spectra, but on the smoothed spectra resulting from convolution with a normalized Gaussian smoothing operator L such that:

$$\bar{A}(\omega) = L(A(\omega)) = \frac{\int_0^{\omega_N} A(\xi) \exp\left(-50\left(\frac{\xi-\omega}{\omega}\right)^2\right) d\xi}{\int_0^{\omega_N} \exp\left(-50\left(\frac{\xi-\omega}{\omega}\right)^2\right) d\xi} \quad (6.2)$$

Differences between the values of $\bar{A}(\omega)$ at different stations may have many causes, including instrument response, Earth structure at the station, distance from the source, source radiation pattern, finite size and directional rupture of the source, refraction and reflection along the path, and attenuation. In the attempt to isolate the effects of attenuation, it is necessary to assume that some of these are unimportant and find ways to cancel out others. Instrument response is no problem in this case because all records are from standard WWSSN instruments, and only a correction for magnification is needed. Earth structures under most of the stations are likely to be similar, as NIL, LAH, NDI, SHL, and HKC are all located on stable continental crust of 40 km thickness (Tandon and Chaudhury, 1964; Tung and Teng, 1974). Kabul is suspected to have an abnormal crustal thickness on the basis of its elevation (1920 m), and therefore I have principally selected events to the east of Tibet and used HKC as a reference station (Figure 6.8). Source finiteness is apt to contribute distortions of less than 1 db in amplitude at periods over 20 sec since all the events used were of magnitude 6.0 or less (Tsai and Aki, 1970).

In the case of the nuclear test, which theoretically has a uniform radiation pattern, a simple distance correction proportional to $(\sin\Delta)^{1/2}$ is sufficient to reveal the combined effects of attenuation and lateral inhomogeneity. These effects are shown in Figure 6.9, where the difference between the power spectra of the Lop Nor test records are plotted against period. An almost identical attenuation/scattering spectrum is recorded on the two paths across Tibet. The difference between these two spectra and the one at Kabul is not likely to be due to lateral refraction or defocusing of the waves. Tibet's phase velocities are low just as its group velocities are low (Table 6.3), and since it is shaped like a great converging lens, it should tend to focus energy on the Indian stations. If anything, this will cause us to underestimate the attenuation.

All the other events are natural earthquakes with complex radiation patterns. We cannot apply the usual technique of observing the same "ray" on opposite sides of the attenuating region, because no stations are available north of Tibet. Instead, it is necessary either to predict the radiation pattern theoretically or to normalize the observed spectra at some reference frequency. The first course is difficult, because only two events have known mechanisms (with only one plane determined), their depths are unknown, and lateral refraction makes the appropriate take-off angle of the ray uncertain. Instead, I chose to normalize the observed spectra to the reference spectrum at a period of 20 seconds. This period is chosen because it is the average spectral peak on the seven reference seismograms and also a peak on many of the Tibetan seismograms. In making this correction we implicitly assume no attenuation of 20 second waves, which is probably not correct. Any attenuation which does occur will cause the Tibetan spectrum to be shifted upward too much, and the calculated attenuations will be less than the actual values. Thus only a lower limit to the attenuation spectrum is obtained. On the other hand, this method also gives a first-order correction for the effects of lateral refraction, boundary reflections, and Earth structure under the station. If attenuation occurs on the reference path as well, then Tibetan attenuation will again be underestimated. But this is probably a negligible effect, since Rayleigh wave Q at 20 seconds period is rarely below 200 (Tsai and Aki, 1969), whereas the values under investigation are about 40.

With these assumptions, Rayleigh wave attenuation ($1/Q_R$) is defined by

$$\frac{1}{Q} \equiv \frac{-1}{2\pi}\frac{\Delta E}{E} \tag{6.3}$$

where ΔE is the energy dissipated in one cycle and E is the total energy. Its minimum value can be obtained as a function of frequency:

$$\min\left(\frac{1}{Q_R}\right) = \frac{PG}{\pi\Delta}\ln\left(\frac{\bar{A}(\omega)\cdot\bar{A}_{RF}(.05)}{\bar{A}_{RF}(\omega)\cdot\bar{A}(.05)}\right) \tag{6.4}$$

where P is period and G is group velocity. For the Lop Nor event no normalization is necessary and we use

$$\min\left(\frac{1}{Q_R}\right) = \frac{PG}{\pi\Delta}\ln\left[\frac{\bar{A}(\omega)}{\bar{A}_{RF}(\omega)}\sqrt{\frac{\sin\Delta}{\sin\Delta_{RF}}}\right] \tag{6.5}$$

This still gives only a minimum value because $A_{RF}(\omega)$ may be affected by attenuation as well.

The spectral differences normalized at 20 seconds are shown in Figure 6.10. Despite considerable variations, it is clear that some effect is strongly reducing the amplitude of 40 second waves on all but two of the sixteen paths. The two paths which do not show differential attenuation are not spatially separated from the others (Figure 6.8) and so it seems most likely that the 20-second reference component of these records has probably been reduced by multipath interference.

The major remaining problem is that attenuation within Tibet is inseparable from loss of the waveguide effect and reflection at the boundaries. The difference between the fundamental mode waveform of one successful Tibetan group velocity model and that of a "normal" Gutenberg continental Earth model is shown in Figure 6.11. There is little difference at 20 seconds. At 40 seconds the Tibetan structure is actually a better waveguide, retaining more of the energy in the crust. However, because of the different rigidities, there is a bad mismatch of either displacement or stress amplitudes where the edge of Tibet abuts against more normal crust. This may cause reflections of some of the long-period energy and produce an appearance of attenuation. On the other hand, Tibet is bounded on both sides by mountain ranges with dipping Mohos, which act to spread the transition over one or more wavelengths. In a similar Love-wave problem of a dipping interface, Boore (1970) found that after two-way transmission through such a boundary the amplitude of surface waves was not measurably reduced. That is, amplitude changed on entering the region of thicker crust because of differences in the eigenfunctions. But very little reflection occurred and the amplitude changes were therefore reversible.

Using equations (6.3) and (6.4) the observed spectra have been converted to attenuation in Figure 6.12. Although standard deviations are as large as the average values, attenuation seems to actually peak at about 50 seconds. In the following interpretation I assume that Q is independent of period, and that this peak is a result of the depth of the attenuating layer. Making the assumption that purely compressional strains are fully recoverable (Anderson and Archambeau, 1964), the attenuation contribution from any layered structure can be calculated as:

$$1/Q_R = \int_0^{\infty} \left[\left(\frac{\beta}{c} \frac{\partial c}{\partial \beta} \right) \frac{1}{Q_\beta} + \left(\frac{\alpha}{c} \frac{\partial c}{\partial \alpha} \right) \frac{1}{Q_\alpha} \right] dz$$

$$= \sum_{i=1}^{N} \left[\left(\frac{\beta}{c} \frac{\partial c}{\partial \beta} \right)_i + \frac{4}{9} \left(\frac{\alpha}{c} \frac{\partial c}{\partial \alpha} \right)_i \right] l_i \qquad (6.5)$$

(Lee and Solomon, 1976), where

α = P-wave velocity

β = S-wave velocity

C = Rayleigh wave phase velocity

and the "attenuation thickness" ℓ is defined by

$$\ell \equiv \frac{\Delta Z}{Q_\beta} \tag{6.6}$$

the layer thickness times its shear wave attenuation. The "attenuation kernel", or contribution to $1/Q_R$ of a unit attenuation thickness at various depths, has been calculated from partial derivatives of the Tibetan model of Table 6.3 and plotted in Figure 6.13. As this clearly shows, a peak at 40-50 seconds can only be produced by a layer centered at 70 km depth. The model curve in Figure 6.12 is for such a layer with an ℓ of 2 km. That is, either a layer of 1 km with a Q of 0.5, or a 20 km layer with a Q of 10., etc. There is also a second upturn in attenuation at 70 seconds which suggests the presence of a deeper, asthenospheric attenuating layer. This has been very tentatively matched to a layer at 250 km with ℓ=14 km, but no great confidence is placed in these values because of the low long-period power in the seismograms.

Such extremely strong attenuation is almost certainly due to a large percentage of partial melting near 70 km depth. There are good reasons for believing that this melting is in the lower crust rather than in the upper mantle. First, diorite melts at 830-1000°C at these pressures (Brown and Fyfe, 1970) and such melting would initiate crustal convection that would buffer the Moho temperature and prevent melting in the uppermost mantle. Second, the observed volcanics indicate a crustal source by their high-silica, low-magnesium composition.

Therefore it appears that the Tibetan crust is close to 70 km thick, with partial melting extensive near the Moho. It is not possible to rule out attenuation (or melting) at shallower depths because of the way in which the surface wave spectra are normalized at 20 seconds period. The low magnetic anomaly field suggests that high temperatures are anomalously close to the surface, and it may be that the entire lower half of the crust is in a semi-molten state.

6.5 Tectonic Shortening of Tibet

Having found that the horizontal-compression hypothesis for crustal shortening is one of the three most compatible with geologic and geophysical evidence, it remains to find the cause of shortening and relate it to the Himalayan orogeny. Assuming a present thickness range of 55-70 km, Tibet must have been shortened to only 45-60% of its original width. This shortening might have consumed as much as 1,000 km of the convergence between India and Asia. Using the convergence velocities determined by Molnar and Tapponnier (1975), this could represent the total convergence over an 18 m.y. period. This would help greatly in explaining the time gap between 40 m.y. when the continents collided and 16 m.y. when the Himalayas began to form. During that time, the Indian plate pushed the plate suture northward at its own velocity, and suffered no internal deformation.

Yet even this hypothesis does not explain the present melting of Tibetan crust. Simple uniform strain in a uniform continental plate will depress the isotherms and thicken the plate. Warming by conduction through the base of a plate thickened to 240 km would take about 450 m.y., which is obviously not available in this case. Radioactive heating in the thickened crust will make some contribution (Dewey and Burke, 1973), but because of the lowered mantle geotherm this would proceed even more slowly than in models II and III of Chapter 2. It appears inescapable that the Moho region in Tibet was already anomalously hot before the collision took place.

This conclusion makes the mechanism more plausible, because it explains the concentration of deformation in the Eurasian plate (Dewey and Burke, 1973; Molnar and Tapponnier, 1975; Tapponnier and Molnar, 1976a,b). With the upper mantle at 800°C or more, the long-term strength of olivine is a few hundred bars or less (section 3.2). Resistance to shortening in Tibet would have initially come mainly from a shallow brittle layer at the top of the crust, probably about 13 km thick, above the depth where quartz creep would significantly relieve stress. The shortening of such a layer requires only 2.65×10^{15} dynes/cm even if the anomalous weakness discussed in previous chapters does not apply. And this is less than the driving force presently acting in the Zagros Mountains at an early stage of convergence.

Developing the Zagros analogy further, anomalous high temperatures in the Iranian plateau northeast of the plate suture would also be expected. In fact, there is seismic evidence for this in the form of attenuation. Molnar and Oliver (1969) found that S_n waves are transmitted poorly across the Iranian Plateau to Shiraz. And Nezihi Canitez has found seismograms (Figure 6.14) from a nuclear test in Novaya Zemlya that show strong attenuation of Rayleigh waves with periods over 20 seconds. This indicates that the attenuation is probably distributed across the plateau, and not only in the volcanic arc behind the suture.

Toksöz and Bird (1977) have speculated that slow heating of the lithosphere above oceanic subduction zones may be a general process, which is responsible for marginal basins in oceanic lithosphere and plateaus (like the Altiplano) in continental lithosphere. The proposed mechanism is one in which induced convection above a downgoing slab concentrates the upwelling of hot asthenospheric material into a band parallel to the subduction zone (Figure 6.15). As the slab descends, it drags asthenospheric material down with it. If the material is unable to pass through the olivine-spinel phase change (Shubert and Turcotte, 1971) it will be deflected upwards in a broad return flow. This would bring hotter mantle material from depth in contact with the base of the lithosphere and warm it. Also, the flow would tend to shear off the weaker parts of the lithosphere, accelerating the thermal erosion. Eventually the crust will melt, and it might also be thickened by intrusions if the mantle melt fraction becomes concentrated into intrusions. The induced flow begins immediately with the beginning of subduction, but thermal and mechanical erosion of the lithosphere is slow. Marginal basins are not formed in continental plates because the lithosphere is thicker (Toksöz et al., 1967) and takes longer to disrupt.

One line of evidence supporting this theory is the observation that intensity of deformation behind the suture zone varies as the amount of ocean subducted prior to collision. In the Alps, the known positions of Africa and Europe (Dewey et al., 1973) restrict the possible oceanic subduction to less than 500 km. This amount was not sufficient to produce even a chain of volcanism in northern Italy, although such a chain would be expected above a mature Andean-type subduction zone. In the Zagros convergence zone some 3000 km of ocean have been destroyed, and here an andesitic volcanic chain has been active behind the suture since the Eocene. The interior of the Iranian plateau has been heated and slightly uplifted, but not weakened to the extent that it cannot resist the forces of the Zagros orogeny. Finally, in the Himalayan convergence zone as much as 8000 km may have been subducted since the breakup of Pangaea (Dietz and Holden, 1970). As a result there was a fully developed volcanic arc in the Nyenchhen Tanglha, melting of the lower crust in Tibet, and destruction of the strength of the lithosphere, so that it collapsed and shortened drastically in the subsequent collision. As the finite element models of Chapter 5 showed, the lack of present-day convergence in Tibet does not imply an ability to support large deviatoric stresses. Its present resistance is almost entirely due to the opposing pressures of its topography and deep crustal roots.

Table 6.1

Events used in study of Rayleigh wave group dispersion in Tibet.

Date	Time	Latt.	Long.	(M_b) Magnitude	Station	Dist. km
Mar. 16, 1964	01:05:18	36.9	95.5	5.9	SHL	1306
Jun. 17, 1965	20:14:50	32.0	87.7	5.1	NDI	1071
					LAH	1264
Feb. 5, 1966	15:12:33	26.2	103.2	5.6	LAH	2864
Sept.28, 1966	14:00:21	27.5	100.1	5.7	LAH	2525
Oct. 14, 1966	01:04:44	36.4	87.4	5.2	SHL	1278
					NDI	1282
					LAH	1319
Mar. 14, 1967	06:58:03	28.5	94.3	5.7	LAH	1952
May 27, 1967	19:05:48	36.1	77.8	5.4	SHL	1785
Aug. 30, 1967	04:22:05	31.6	100.3	6.1	LAH	2450
Feb. 24, 1970	02:07:37	30.6	103.0	5.8	KBL	3200
Jul. 31, 1970	13:10:47	28.6	103.6	5.3	KBL	3318
Aug. 16, 1971	18:53:59	28.9	103.7	5.3	KBL	3320
Aug. 16, 1971	22:37:32	28.8	103.6	5.3	KBL	3322
Feb. 7, 1973	16:06:25	31.5	100.3	5.9	NDI	2237
					NIL	2538
Mar. 23, 1973	19:14:53	53.1	31.9	5.3	NDI	2218
					NIL	2508
					KBL	2890

Table 6.2

Velocity Standard for Reconnaisance

Group Velocity Dispersion Models of Tibet

(Crustal thickness = 65 km; Shear velocity anomaly = 0%)

Layer	Thickness, km	V_p, km/sec	V_s, km/sec	ρ, g/cc
1	5.	4.50	2.60	2.50
2	10.	6.09	3.46	2.66
3	10.	6.07	3.44	2.65
4	10.	6.08	3.44	2.65
5	10.	6.26	3.72	2.82
6	10.	6.30	3.70	2.83
7	10.	6.34	3.67	2.84
8	60.	8.10	4.67	3.20
9	80.	8.10	4.55	3.25
10	∞	9.18	5.30	3.80

TABLE 6.3A: Tibetan Rayleigh Wave Group Velocity Model (55 KM Crust)

Layer	Depth, km	V_p km/sec	Vs km/sec	ρ, g/cc	Period sec	Group velocity km/sec	Phase velocity km/sec
1	0 - 5	4.50	2.60	2.50			
2	5 - 13.33	6.09	3.46	2.66	10	2.804	2.986
3	13.33-21.66	6.07	3.44	2.65	15	2.883	3.057
4	21.66-30.0	6.08	3.44	2.65	20	2.860	3.123
5	30. - 38.33	6.19	3.534	2.82	25	2.813	3.203
6	38.33-46.66	6.23	3.515	2.83	30	2.787	3.298
7	46.66-55.	6.269	3.487	2.84	35	2.811	3.401
8	55. - 65.	7.94	4.261	3.125	40	2.881	3.499
9	65. - 75.	7.94	4.261	3.125	45	2.978	3.587
10	75. - 85.	7.94	4.261	3.125	50	3.084	3.662
11	85. - 95	7.94	4.261	3.125	55	3.187	3.724
12	95. -105	7.94	4.261	3.125	60	3.282	3.776
13	105. -115	7.94	4.261	3.125	65	3.366	3.820
14	115 -140	8.10	4.55	3.25	70	3.438	3.856
15	140 -180	8.10	4.55	3.25	75	3.498	3.887
16	180 -220	8.10	4.55	3.25	80	3.547	3.914
17	220 -280	8.10	4.55	3.25			
18	280 -340	8.10	4.55	3.25			
19	340 -410	8.227	4.75	3.60			
20	410 - ∞	9.18	5.30	3.80			

TABLE 6.3B: Tibetan Rayleigh Wave Group Velocity Model (70 KM Crust)

Layer	Depth, km	V_p km/sec	V_s km/sec	ρ, g/cc	Period sec	Group velocity km/sec	Phase velocity km/sec
1	0. - 5.	4.50	2.60	2.50	10	2.809	2.987
2	5. - 15.8	6.09	3.46	2.66	15	2.873	3.059
3	15.8- 26.7	6.07	3.44	2.65	20	2.845	3.131
4	26.7- 37.5	6.08	3.44	2.65	25	2.828	3.215
5	37.5- 48.3	6.728	3.841	2.82	30	2.833	3.305
6	48.3- 59.2	6.771	3.820	2.83	35	2.861	3.396
7	59.2- 70.0	6.814	3.790	2.84	40	2.908	3.485
8	70. - 75.	7.94	4.261	3.125	45	2.975	3.568
9	75. - 80.	7.94	4.261	3.125	50	3.056	3.643
10	80. - 85.	7.94	4.261	3.125	55	3.146	3.708
11	85. - 95.	7.94	4.261	3.125	60	3.237	3.763
12	95. -105.	7.94	4.261	3.125	65	3.323	3.810
13	105. -115.	7.94	4.261	3.125	70	3.401	3.850
14	115. -140.	8.10	4.55	3.25	75	3.468	3.883
15	140. -180.	8.10	4.55	3.25	80	3.523	3.912
16	180. -220.	8.10	4.55	3.25			
17	220. -280.	8.10	4.55	3.25			
18	280 -340	8.10	4.55	3.25			
19	340 -410	8.227	4.75	3.60			
20	410 - ∞	9.18	5.30	3.80			

Table 6.4

Events and paths used for attenuation study of Tibet

Date	Time	Latt.	Long.	Station	Distance km
Mar. 16, 1964	01:05:18	36.9	95.5	HKC	2419
				SHL	1306
				NDI	1933
				LAH	2030
Feb. 5, 1966	15:12:33	26.2	103.2	HKC	1192
				LAH	2864
Sept.28, 1966	14:00:21	27.5	100.1	HKC	1533
				LAH	2525
Oct. 14, 1970	07:29:59	40.9	89.4	KBL	1919
				NDI	1752
				SHL	1720
Mar. 24, 1971	13:54:18	35.5	98.2	KBL	2648
				SHL	1259
Aug. 16, 1971	18:53:55	28.9	103.7	HKC	1277
	AND	AND	AND	KBL	3319
	22:37:34	28.8	103.6	NIL	2929
				NDI	2574
				SHL	1223

CHAPTER 6 Figure Captions

Figure 6.1 Three dimensional diagram of Earth's topography in the Zagros-Himalayan region with the oceans drained away. View is at a low angle toward the northwest, and vertical exaggeration is 200x. Tic marks on axes represent 1000 km. Plot created by the Versatec Corporation.

Figure 6.2 Four possible tectonic models to explain the high elevation of Tibet. Each model is reduced to a horizontally-layered velocity and density model for preliminary group velocity analysis. Figures in each layer are P-velocity and S-velocity in km/sec and density and g/cc. Lithosphere is shaded and crust is stippled. Vertical dimensions drawn to scale but not horizontal.

Figure 6.3 Paths used for measurement of Rayleigh wave group velocities in the Tibetan Plateau (shaded). All paths are at least 50% Tibetan. Events are marked by circles and stations by triangles.

Figure 6.4 Model group velocity dispersion curves for the models of Figure 6.2 compared with data. Tung and Teng (1974) data are averages over six long paths. Black dots are maxima of individual spectral amplitude plots obtained from multiple-filtering of seismograms whose paths are shown in Figure 6.3.

Figure 6.5 Trade-off diagram for crustal thickness versus lower-crust and upper-mantle velocity anomaly. Each dot represents a calculated model. Contours show the RMS group velocity difference between model and observations at 10-second intervals of period between 10 and 80 seconds. Acceptable models fall in the shaded region.

Figure 6.6 Composite stacked plot of Rayleigh wave spectral amplitudes versus group velocity (time) and period. Contours are at intervals of 5 decibels down from the maximum at lower left. Shaded region is within 2 db of the maximum at each period and provides a measure of resolution. Solid model curves are for 55 km and 70 km crust models with low shear velocities in the mantle. Dashed curve is for a model with two normal crustal layers on top of one another and the same lithosphere beneath.

Figure 6.7 Attenuation of Rayleigh waves from the Lop Nor nuclear test in crossing the Tibetan plateau (shaded). All records are long-period vertical component of WWNSS stations corrected to equal magnification. Note greater attenuation of longer (30-80 sec) periods.

Figure 6.8 Diagram of paths employed for Rayleigh wave attenuation measurements (dashed and dotted) and reference paths to KBL or HKC used to estimate the source spectrum (solid lines). The two paths which failed to show attenuation are not separated spatially from the fourteen which did.

Figure 6.9 Difference between the Rayleigh power spectra of the Lop Nor event at New Delhi and Shillong and the spectrum of the same event at Kabul. Difference is corrected for radial spreading and instrument magnification. To obtain amplitude spectral differences, divide the scale values by two.

Figure 6.10 Power spectral differences relative to the reference station for Rayleigh waves used in the attenuation study. Difference is set to zero at 20 second periods to correct for distance, source radiation function, magnification, and to help correct for lateral refractions, reflections, and differences of earth structure on the different paths.

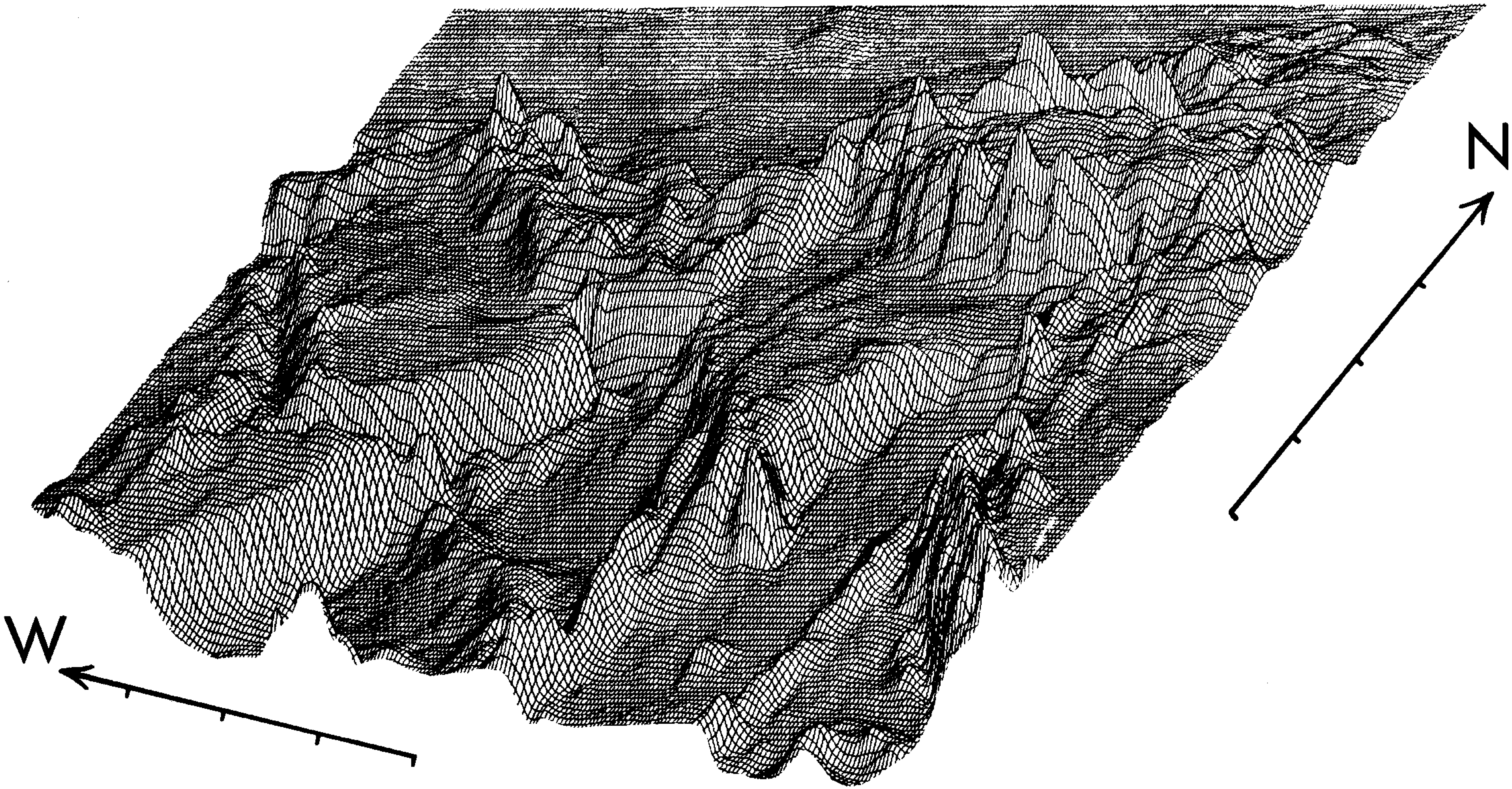
Figure 6.11 Rayleigh wave fundamental mode eigenfunctions for two periods in the 55 km crust Tibetan model discussed in text (solid lines). Also shown is the waveform in a Gutenberg continental earth model (dashed lines). Note that depth scale is expanded two times in the lower half. Eigen-functions are normalized to equal vertical displacement W_o at the surface.

Figure 6.12 Attenuation data for the Lop Nor event (top) and for all data combined (bottom) compared to model curves. Two angular lines at top correspond to two curves of Figure 6.9. At bottom the average (dot) and standard deviation (bar) of attenuation are plotted at several periods. Attenuation artificially goes to zero at 20 sec period because of the normalization used. Dashed line shows effect of a layer centered at 70 km depth with an attenuation thickness of 2 km. Dotted line shows additional effect of an asthenospheric attenuating layer at 250 km with an attenuation thickness of 14 km.

Figure 6.13 Rayleigh wave attenuation kernel of a Tibetan earth model. Vertical axis corresponds to $1/Q_R$, and two horizontal axes to period of the Rayleigh wave and depth of the attenuating layer.

Figure 6.14 Rayleigh wave (LPZ) seismograms from three WWNSS stations of a nuclear test in Novaya Zemlya. Equal magnification and time scale (bar) for each record. Long period Rayleigh waves are strongly attenuated in passing through the Iranian Plateau adjacent to the Zagros suture zone. Seismograms were collected and interpreted by Canitez.

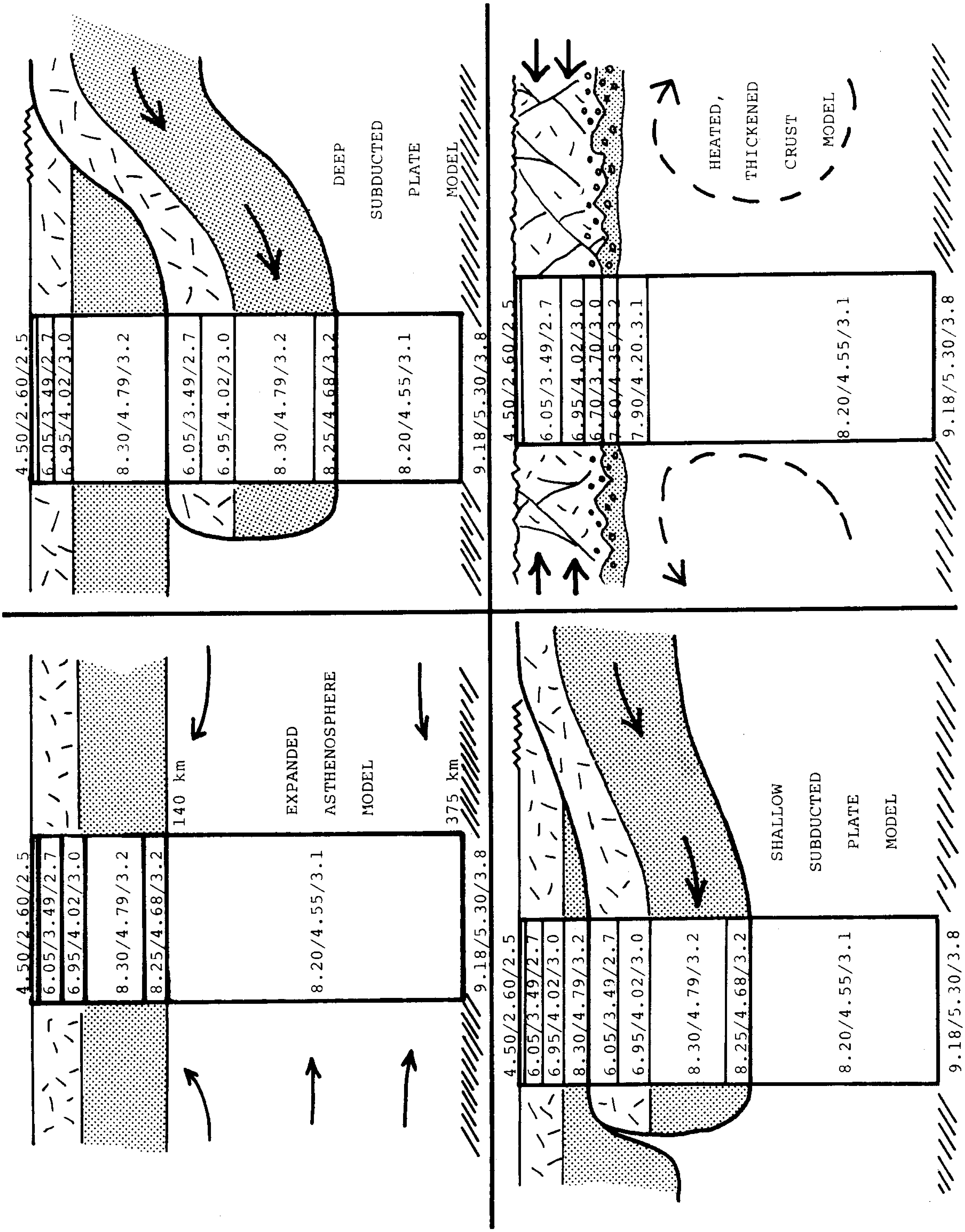
Figure 6.15 Schematic diagram of convection beneath a continental plate caused by a downgoing oceanic slab. The asthenospheric flow (arrows) is stopped at a phase change (dashes) and returns upward. The lithosphere is eroded, the crust (stippled) may be melted and thickened by mantle intrusions. Tibet may have passed through this phase before the Himalayan collision, 40 m.y. ago.



TOPOGRAPHY OF THE ZAGROS —
HIMALAYAN REGION
at 200 X v.e.

Created by VERSATEC

Fig. 6.1



DEEP SUBDUCTED PLATE MODEL
4.50/2.60/2.5
6.05/3.49/2.7
6.95/4.02/3.0
8.30/4.79/3.2
6.05/3.49/2.7
6.95/4.02/3.0
8.30/4.79/3.2
8.25/4.68/3.2
8.20/4.55/3.1
9.18/5.30/3.8
EXPANDED ASTHENOSPHERE MODEL
140 km
375 km
4.50/2.60/2.5
6.05/3.49/2.7
6.95/4.02/3.0
8.30/4.79/3.2
8.25/4.68/3.2
8.20/4.55/3.1
9.18/5.30/3.8
HEATED, THICKENED CRUST MODEL
4.50/2.60/2.5
6.05/3.49/2.7
6.95/4.02/3.0
6.70/3.70/3.0
7.60/4.35/3.2
7.90/4.20.3.1
8.20/4.55/3.1
9.18/5.30/3.8
SHALLOW SUBDUCTED PLATE MODEL
4.50/2.60/2.5
6.05/3.49/2.7
6.95/4.02/3.0
8.30/4.79/3.2
6.05/3.49/2.7
6.95/4.02/3.0
8.30/4.79/3.2
8.25/4.68/3.2
8.20/4.55/3.1
9.18/5.30/3.8

Fig. 6.2

Fig. 6.3

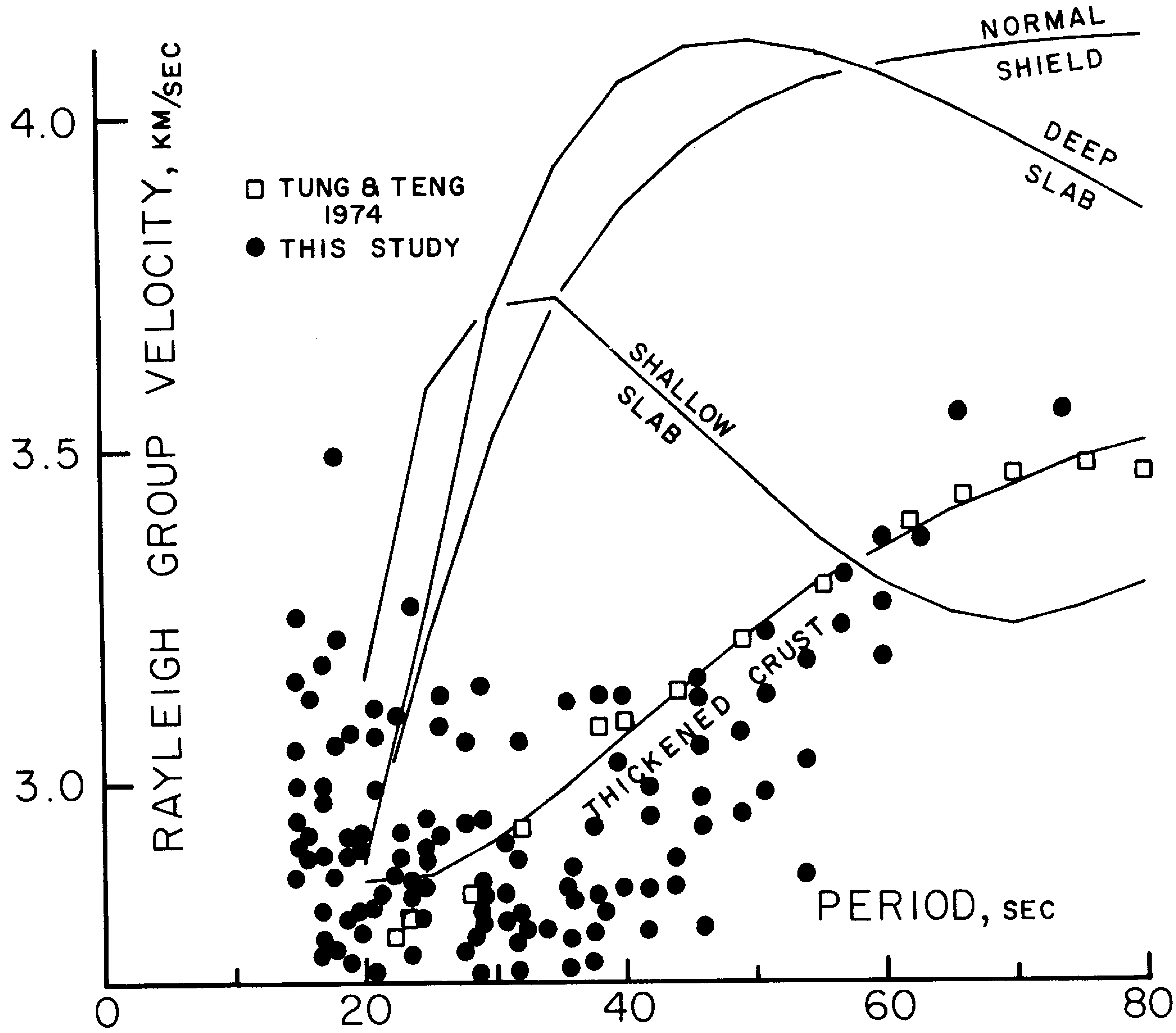

NORMAL
SHIELD
DEEP
SLAB
SHALLOW
SLAB
THICKENED CRUST
□ TUNG & TENG
1974
● THIS STUDY
RAYLEIGH GROUP VELOCITY, KM/SEC
PERIOD, SEC
4.0
3.5
3.0
0
20
40
60
80

Fig. 6.4

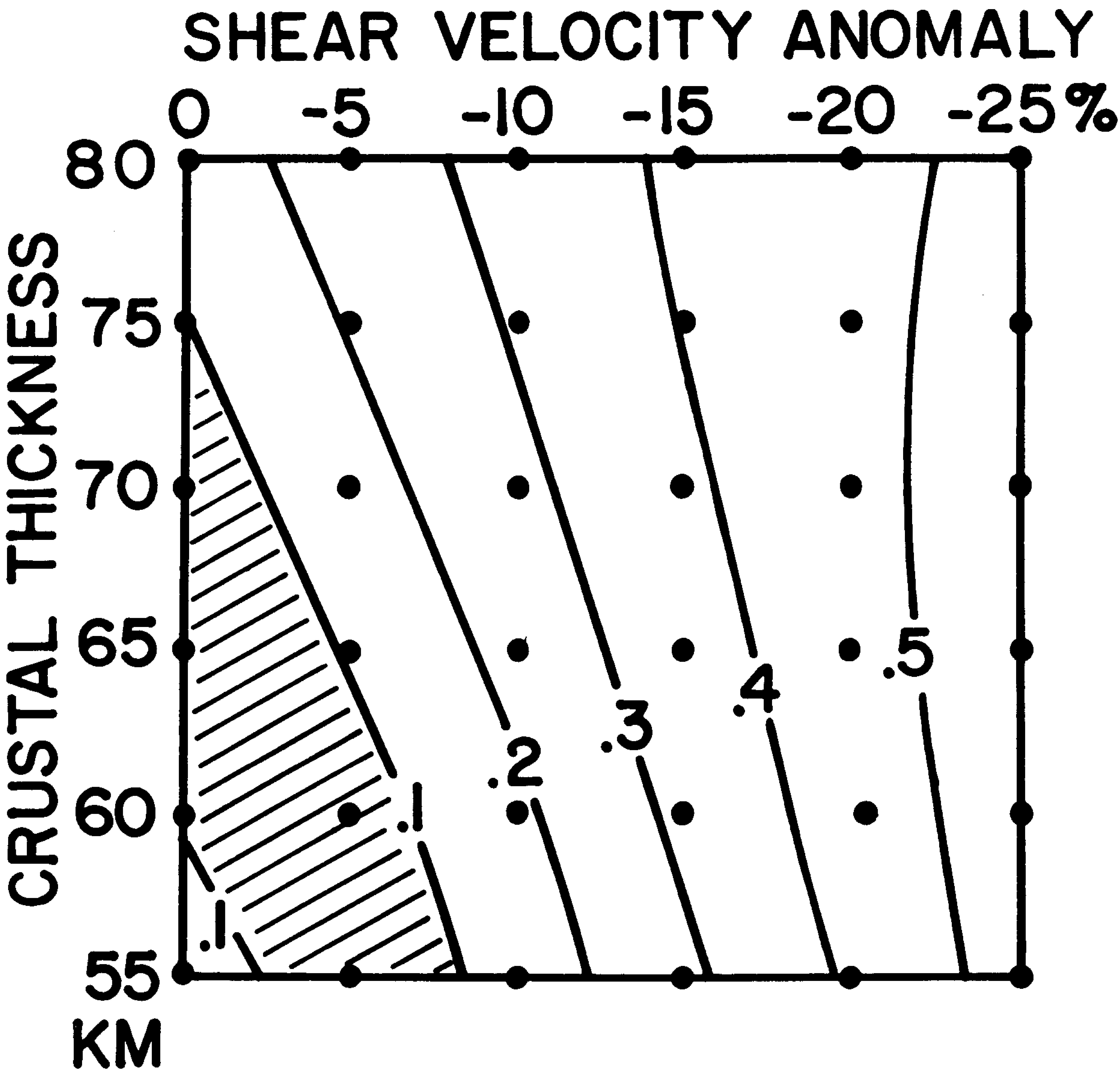

SHEAR VELOCITY ANOMALY
0
-5
-10
-15
-20
-25%
80
75
70
65
60
55
KM
CRUSTAL THICKNESS
.1
.1
.2
.3
.4
.5

Fig. 6.5

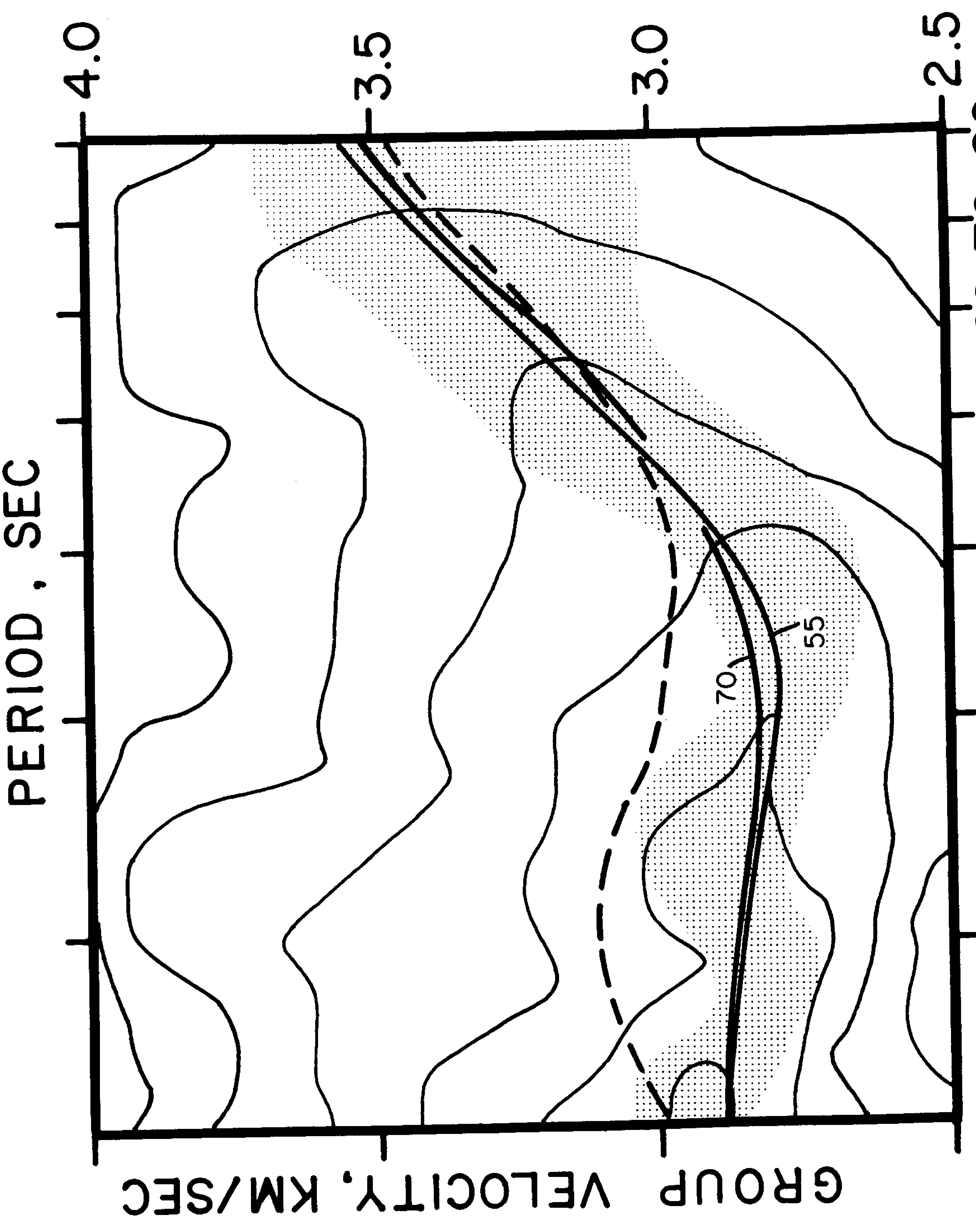

PERIOD , SEC
GROUP VELOCITY, KM/SEC
4.0
3.5
3.0
2.5
20
30
40
50
60
70
80
70
55

Fig. 6.6

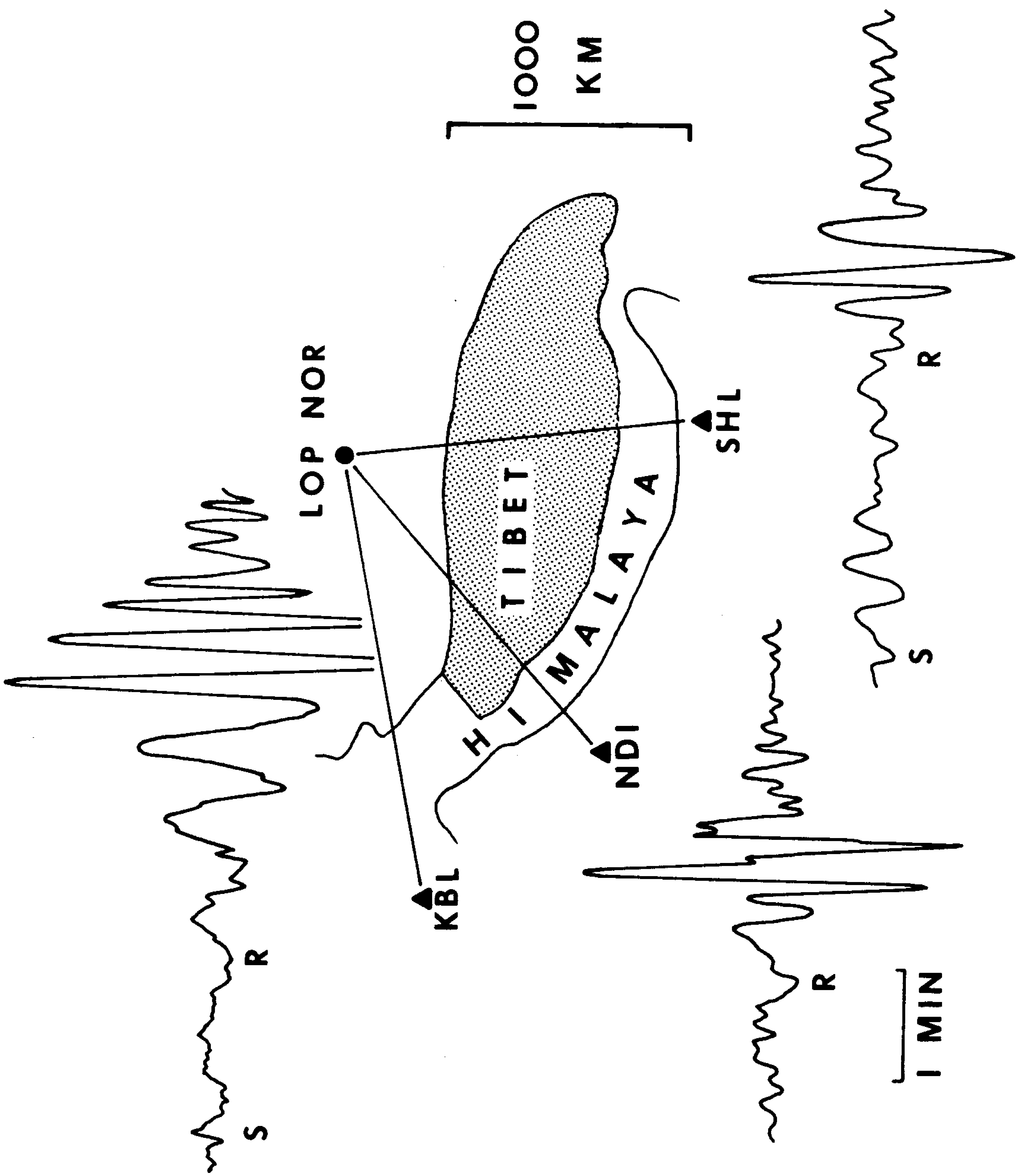


Fig. 6.7

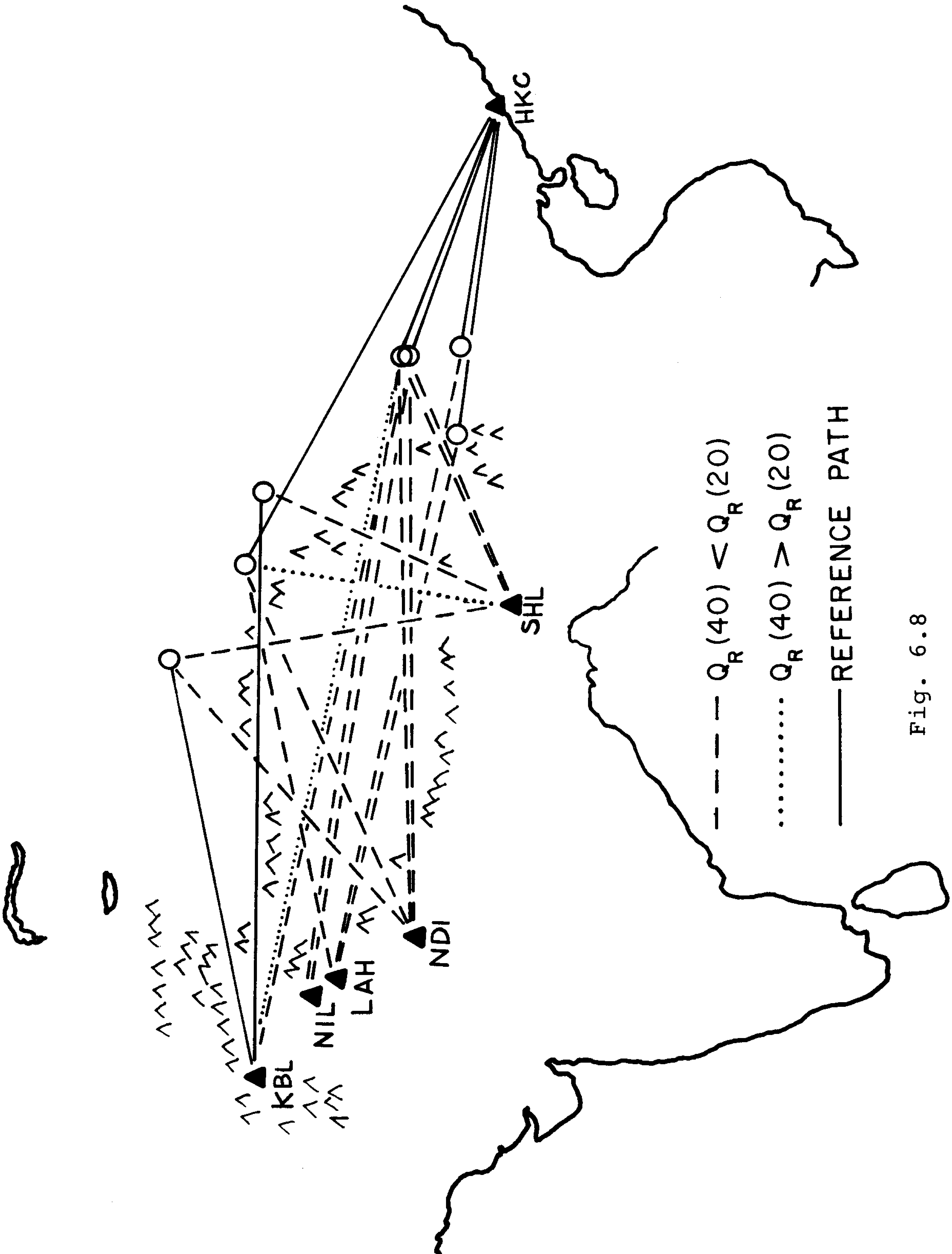


Fig. 6.8


10
DB
0
-10
OCT. 14, 1970 (LOP NOR)
Reference : KBL
PERIOD
20
40
60
80 SEC
NDI
SHL

Fig. 6.9

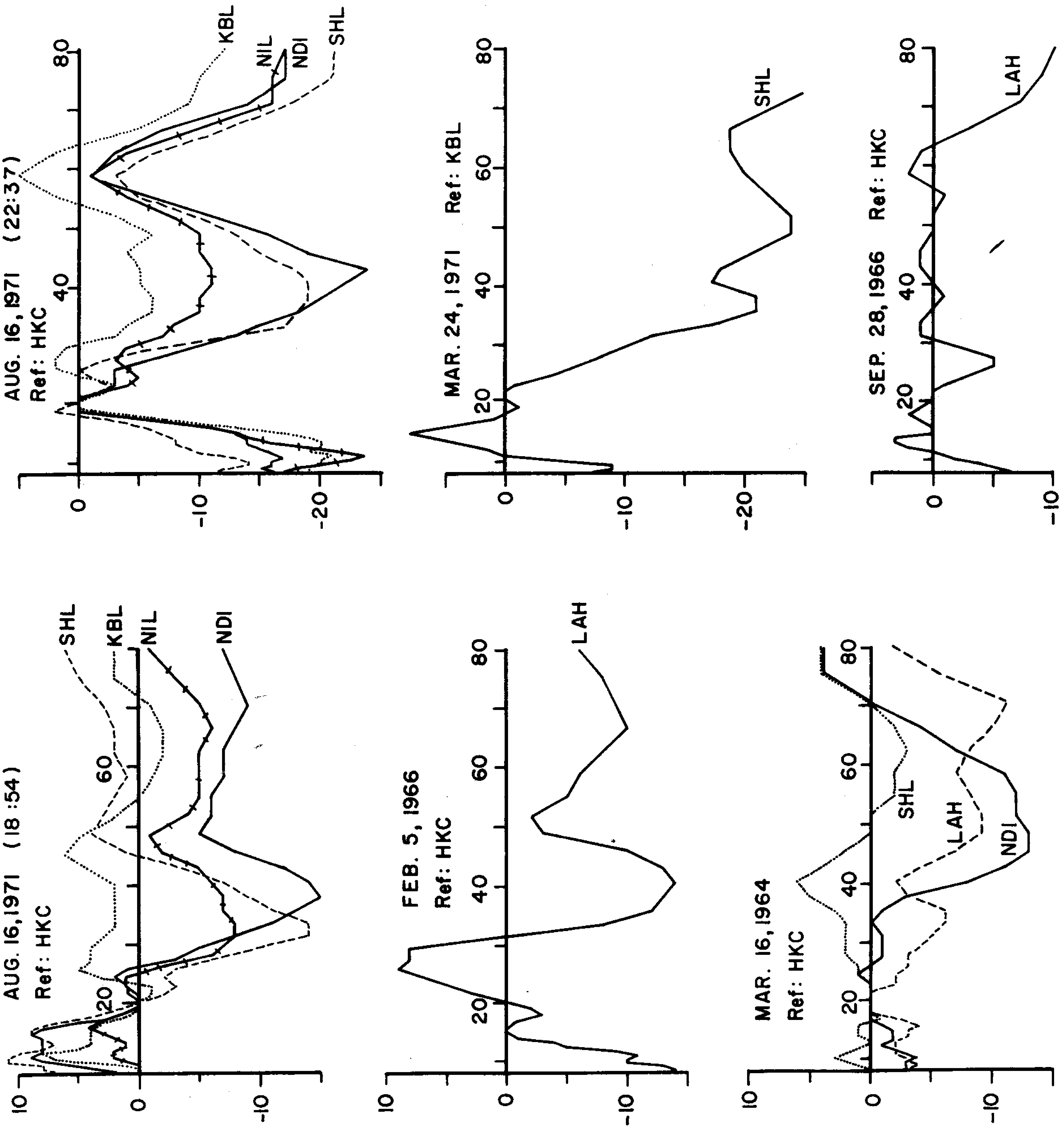

AUG. 16, 1971 (18:54)
Ref: HKC
SHL
KBL
NIL
NDI
FEB. 5, 1966
Ref: HKC
LAH
MAR. 16, 1964
Ref: HKC
SHL
LAH
NDI
AUG. 16, 1971 (22:37)
Ref: HKC
KBL
NIL
NDI
SHL
MAR. 24, 1971 Ref: KBL
SHL
SEP. 28, 1966 Ref: HKC
LAH

Fig. 6.10


T = 20 sec
0
.25
.50
.75
1.
W/Wo
DEPTH
0
.25
.50
.75
1.
-U/Wo
BARS
0
1
2
3
σzz/cm
BARS
0
1
2
3
τxz/cm
— 0
— 20
— 40
— 60
— 80
— 100 KM
T = 40 sec
0
.25
.50
.75
1.
W/Wo
DEPTH
0
.25
.50
.75
1.
- U/Wo
BARS
0
1
2
3
σzz/cm
BARS
0
1
2
3
4
τxz/cm
— 0
— 40
— 80
— 120
— 160
— 200

Fig. 6.11

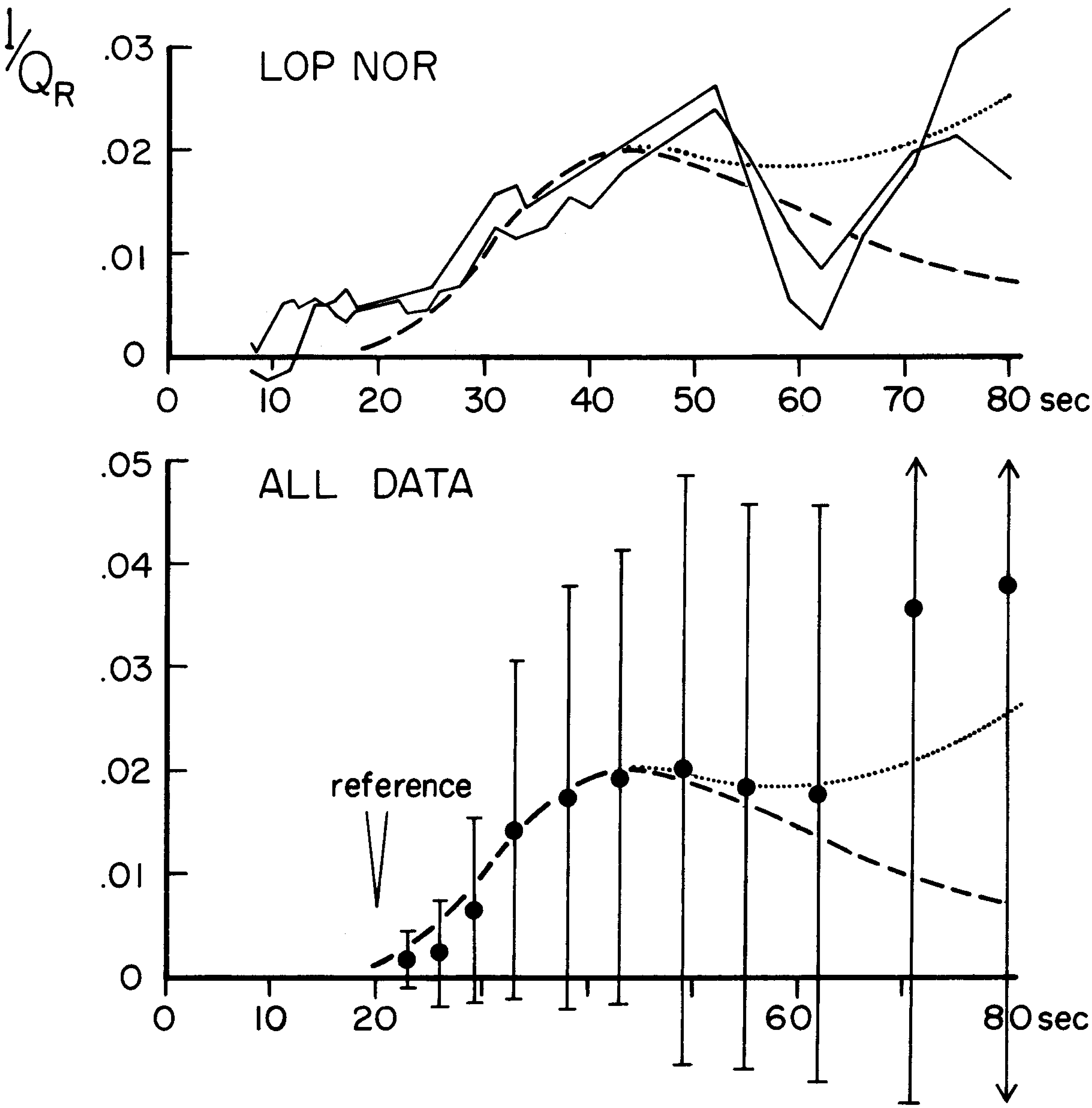
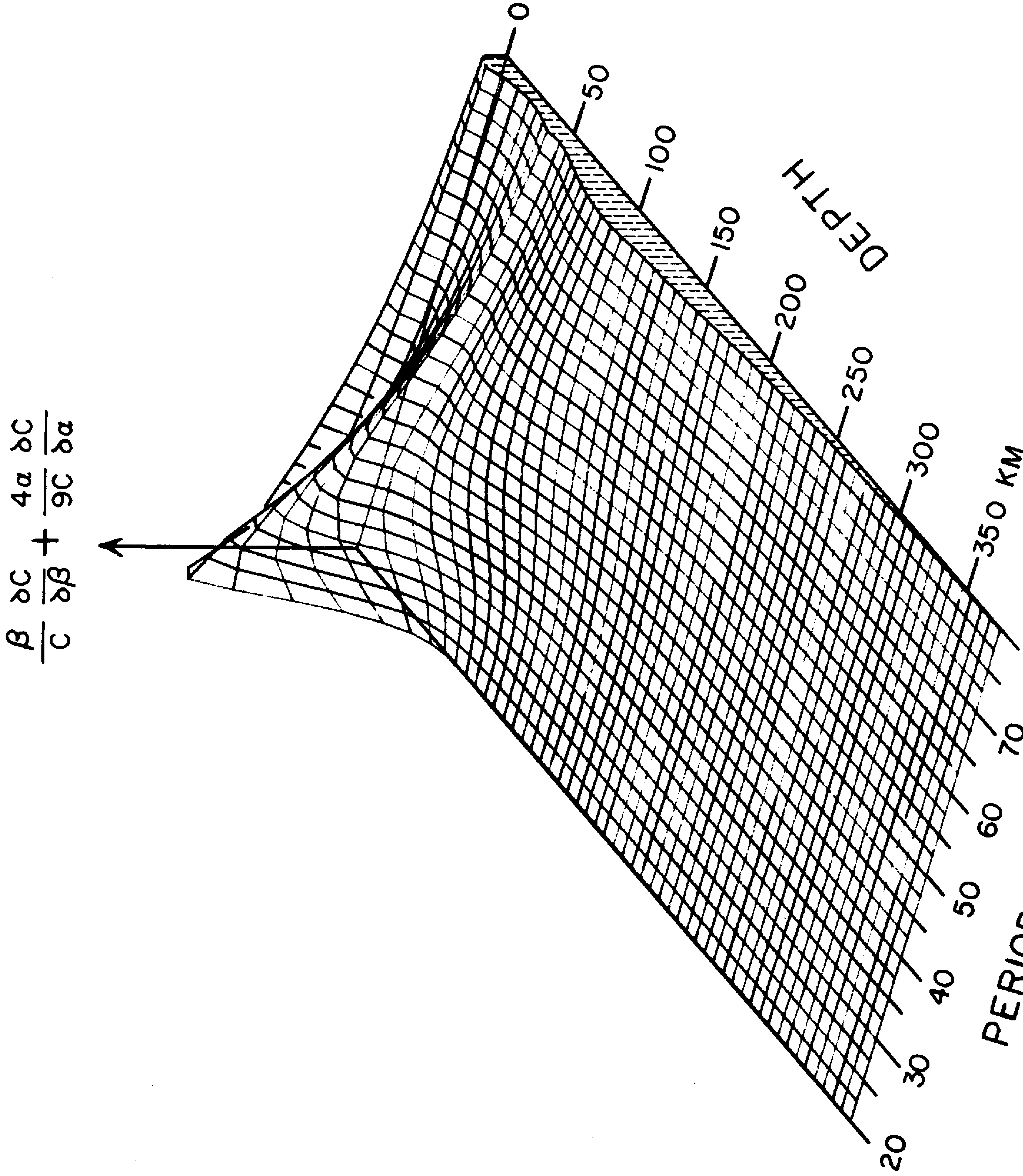

1/Q_R
.03
.02
.01
0
LOP NOR
0
10
20
30
40
50
60
70
80 sec
.05
.04
.03
.02
.01
0
ALL DATA
reference
0
10
20
60
80 sec

Fig. 6.12



$\frac{\beta}{C}\frac{\partial C}{\partial \beta}+\frac{4\alpha}{9C}\frac{\partial C}{\partial \alpha}$
DEPTH
0
50
100
150
200
250
300
350 KM
PERIOD
20
30
40
50
60
70
80 SEC

Fig. 6.13

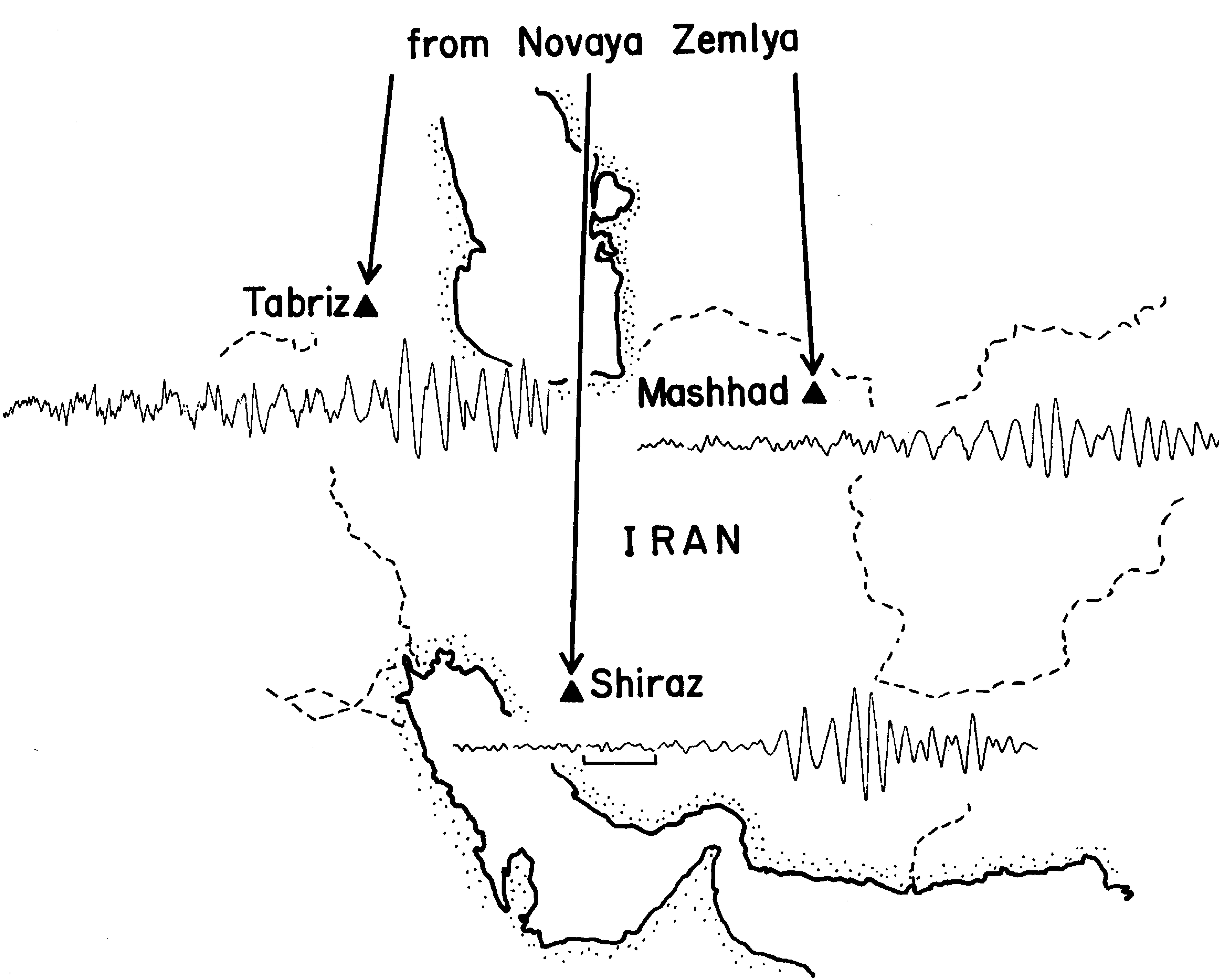


Fig. 6.14

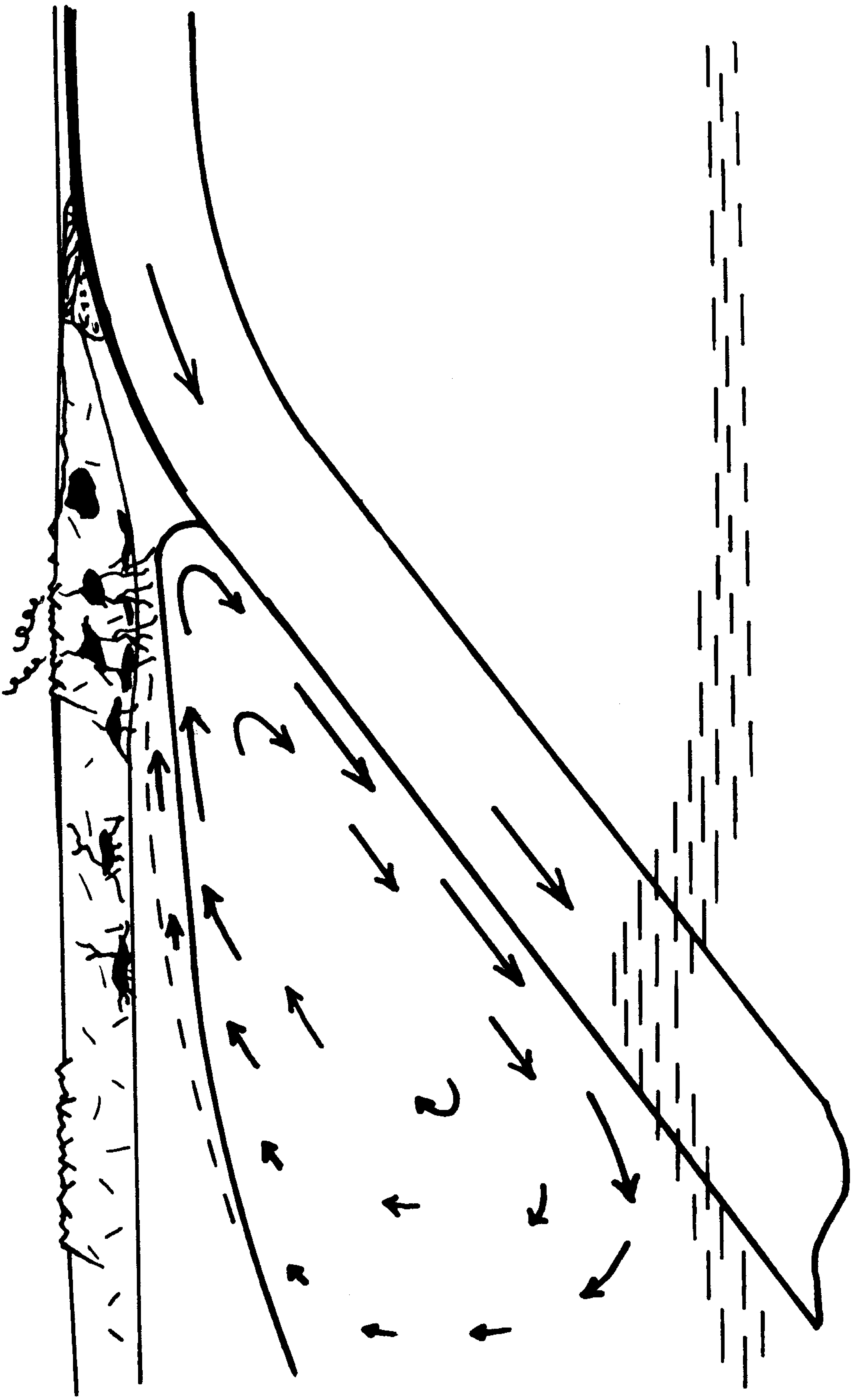


Fig. 6.15

CHAPTER 7: SPECULATIONS ABOUT MOUNTAINS, DRIVING FORCES, AND EARTHQUAKES

The study of tectonophysics can be divided into three parts. The first is the study of the state of stress, particularly deviatoric stresses and shear stresses, in the Earth. The second is the identification of total force systems, in which the loops of stress are closed and the total entropy changes are positive. The third is the elucidation of the mechanisms by which these stresses result in strain, heating, and seismicity. Corresponding to this subdivision are the three segments of this short chapter, which attempts to identify features common to the Zagros, Himalayas, and Tibet. This first discusses the importance of topographic forces as a buffer and indicator of tectonism. Next the forces presently acting in the two ranges are related to the known driving forces of plate tectonics. Finally, a hypothesis concerning the origin and propagation of large earthquakes is put forward to explain the surprising weakness of parts of the continental crust.

7.1 The Significance of Mountain Topography

Since Pierre Bouguer discovered the crustal roots of the Andes from pendulum deflections in 1735, the concept of isostatic equilibrium between topographic mass and buoyancy of roots has been gradually accepted. The minor departures from compensation are themselves tectonically interesting but will not be discussed here. Instead, I wish to point out two different forms of topographic equilibration that are important in active continental convergence zones: the existence of topographic slopes above thrust faults, and the stress regulating effect of hot mobile interiors.

The topographic data to which I will refer is presented in Figure 7.1. A topographic survey of the Himalayas and southern Tibet was made, using the Operational Navigational Charts published by the Defense Mapping Agency. In the central Himalayan region (26-32°N by 77-95°E), the topography was sampled on a regular rectangular grid at an interval of 1/3° in latitude and longitude. This produced 550 points with estimated errors ranging from 100 feet (30 m) in the south to 1300 feet (400 m) in Tibet. These points were assembled onto a common cross-section by the technique outlined in section 5.3, in which points are plotted according to their fractional position between the two limits of the mountain range. Also shown is the curve of averaged topography within 10% intervals of the width of the range.

The topography of the southern half of the range is dominated by a fairly uniform gradient up from sea level to 5 km, with an average slope of 1.7°. If we temporarily ignore the special character and temperature of shear zone rocks and treat the medium as a rigid-plastic continuum undergoing deformation, then the theory of Chapple (1975) and Elliott (1976) is applicable: Assume that everywhere in the medium

$$\tau^* = \sqrt{\tfrac{1}{4}(\sigma_{xx} - \sigma_{zz})^2 + \tau_{xz}^2} \qquad (7.1)$$

is constant. If we treat the mass above sea level as contributing negligible stiffness, then the two components of shear stress will be independent of x:

$$(\sigma_{xx} - \sigma_{zz}) = m(z)\ ; \quad \tau_{xz} = r(z) \qquad (7.2a,b)$$

This implies

$$\frac{\partial \sigma_{xx}}{\partial x} = \frac{\partial \sigma_{zz}}{\partial x} = -g\rho\,\frac{\partial h(x)}{\partial x} \qquad (7.3)$$

where h is the average topographic height. Using the stress equilibrium equation this yields the shear stress on a horizontal surface at depth:

$$\tau_{xz} = \int_0^z - \frac{\partial \sigma_{xx}}{\partial x} dz = g\rho z \frac{\partial h}{\partial x} \tag{7.4}$$

According to section 5.3, the depth of the active fault in this region varies from zero to 25 km. This simple analysis based on topography along predicts shear stresses from zero to 200 bars; previously, finite element models showed that 200 bars was likely to be the maximum average stress in the entire fault zone.

Earlier, in Chapter 4, it was demonstrated that a balance exists in the Zagros between the shear stress deforming the base of the crust and the dip of the Moho (or, equivalently, of the surface). In the Zagros the shear stress is less (100 bars) on the surface of sliding, and the topographic gradient is also lower (0.5°). The difference in the proportionality factor is due to the greater depth of the sliding surface in the Zagros.

It appears that the principle developed by Chapple and Elliott to describe the tectonics of sedimentary rocks is equally applicable to crustal overthrusting on a grander scale. If the topographic gradient becomes too low, then overthrusting stops and the upper plate is deformed and thickened, against whatever barrier is supporting it, until the gradient is increased. Should the slope become too high, the rate of overthrusting will exceed the rate of convergence as the upper plate slumps downward and outward over the subducting plate, reducing the gradient.

If this hypothesis is correct, then topographic gradients should be accompanied by shear stress at depth and vice versa. The correlation is strongly supported by the distribution of seismicity in the two ranges. It is quite uniform in the Zagros, where the topographic slope is constant. In the Himalayas it is confined to the southern half, between the foothills and the high peaks.

This principle suggests an answer to the question which has previously been deferred; that of how the diffuse crustal shortening such as we see in the Zagros developed into the clearcut crustal doubling and slab-type overthrusting of the Himalayas. This could be explained if there had been an uplift of the margin of the Indian plate, creating a steep topographic gradient in the region where the Main Central Thrust was yet to form. Since it has been concluded that the main phase of metamorphism and plutonism preceded the formation of the thrust, these two events may well be linked. If the sub-crustal lithosphere under the margin of India were removed by any mechanism, the subsequent influx of lighter asthenosphere to replace it could cause uplift and metamorphism simultaneously. It is even possible that molten crustal material of Tibet flowed across the suture as Tibet was shortened and its Moho was depressed. This hypothesis might be tested by a comparison of the trace element chemistry of Tibetan and Himalayan plutons.

The concept of Tibet as a hot, mobile region and the concept of Tibet as the foundation supporting the over-thrusting slab in the Himalayas are not at all contradictory. Once again, topograhic forces balance the system. Even if the crust of Tibet were entirely liquid and confined between the two "dams" of the Himalayas and Kun Lun, the regional tectonics would be no different. The 5 km elevation of Tibet creates an excess vertical stress of 1300 bars at sea level. In the absence of strength, this pressure is hydrostatically distributed in all directions, including horizontally North and South. This component provides most of the resistance to the Himalayan convergence in the successful models of Chapter 5.

The proof of this principle is also in Figure 7.1. In this diagram, the plate suture would be hard to detect if not labelled. There is a slight dip there caused by the rapid erosion of the entrenched Indus and Brahmaputra rivers following the fault zone. But basically the average Tibetan elevation of 5 km extends straight across into the Himalayas as far south as the central peaks. Like Tibet, the northern Himalayas need not be strong because they have reached an elevation where their topographic mass and crustal roots alone can transmit the driving forces as lithostatic pressure.

This concept has two important consequences. First, it explains the timing of the formation of the Himalayas. At the time of the collision, Tibet was weak, and absorbed all of the convergence velocity. This caused it to shorten, the crust was thickened, and isostasy caused the elevation to rise. As it rose, the hyrostatic forces opposing further convergence rose more than linearly with time, until about 16 m.y. ago they became high enough to fracture the margin of the Indian shield and push up the Himalayas.

Second, Tibet today can be used as a "pressure gauge" for the entire Asian continental orogeny. The confusion of fault-plane solutions in Tibet is fortunate, because it shows that no consistent stress is transmitted across the plateau in the form of stress differences. The lithostatic term must be dominant. This tectonic overpressure of 1300 bars must continue from sea level down to the beginnings of the crustal roots, at about 33 km. If we assume that it then decays linearly to zero at the minimum compensation depth of 70 km, we obtain the integral of anomalous horizontal stress with depth as no less than 6.95×10^{15} dynes/cm. This is the value of the "driving force" of the Himalayan orogeny, equal to a kilobar of horizontal compression acting over a thickness of 70 km. It acts radially outward in all directions from Tibet, reactivating the old mountain ranges of Asia and driving the great strike-slip faults identified by Molnar and Tapponnier (1975). If the elevation of Tibet in earlier times could be determined (perhaps from the study of fossil pollens in lake sediments), it would be valuable for the reconstruction of Asian tectonics.

The mathematics behind the relationship of topography and "driving force" is given more systematically in the next section.

7.2 The Driving Forces in Continental Convergence

The idea that horizontal density contrasts cause deviatoric stresses even in the isostatic case was perhaps first applied to geology by Frank (1972) in his discussion of the horizontal compression originating in mid-ocean ridges. A similar treatment was given by Artyushkov (1973), although a different geologic structure was assumed. In this section the work of these authors is extended, with simplifying assumptions, to a definition of driving force for two-dimensional problems and a means for calculating it. The forces required to drive the Zagros and Himalayas are related to known sources of force.

Three assumptions are required:

1. Isostasy is general. At a constant compensation depth (Z_c) the integral of weight $\int_0^Z g\rho(s)ds$ is everywhere equal. Lambeck (1972) has shown that any different assumption would imply stresses of tens of kilobars in the lithosphere. Because there is evidence for deep-seated compensation under mid-ocean ridges, the depth of compensation is assumed to be at the base of the thickest lithosphere. Toksöz *et al*. (1967) place this at approximately 120 km.

2. Deviatoric and shear stresses vanish at the compensation depth. (This assumption implies the one above). This is equivalent to assuming that the plates are free to move

without drag. If the viscosity of the asthenosphere is about 1 x 10^{20} poise as determined by Smith (1974) and Cathles (1975), then a plate in motion at 5 cm/year with respect to the deep mantle would experience a drag of 1-2 bars. Since the asthenosphere is likely to be in a state of convection, forces of this magnitude may be applied in many directions on different parts of the plate. In order to formulate a first-order theory, I will consider their net effect negligible.

3. The distribution of mass is nearly two-dimensional with a vertical plane of symmetry. Since the region under consideration is a plate, we can not require that it be infinitely long, or flat. But a small plate like the Arabian, bounded by two transform faults, fits this criterion as long as the shear stress on the plane of the transforms is not large. Stress on transforms is an open question at present, but Lachenbruch and Thompson (1972) have argued that the stresses must be small to allow perpendicular arrangements of ridges and transform faults.

It is now possible to define driving force (Ω) as a scalar function of position, as an integral of stress and a double integral of density:

$$\Omega(x) \equiv \int_{h(x)}^{z_c} \left[\sigma_{xx}(x,z) - \int_{h(x)}^{z} g(s)\rho(s)\,ds \right] dz \tag{7.5}$$

where x is the horizontal dimension in the mirror plane, h(x) is the elevation of the land or sea surface, and $\rho(x,z)$ is the density distribution. The physical interpretati of Ω is the vertical integral of stress differences supported by the strength of the lithosphere. Over a broad area, the average vertical stress

$$\langle \sigma_{zz}(x,z) \rangle = \int_{h(x)}^{z} g(s)\rho(s)\,ds \tag{7.6}$$

so the definition above becomes

$$\Omega(x) = \int_{h(x)}^{z_c} \left[\sigma_{xx}(x,z) - \langle \sigma_{zz}(x,z) \rangle \right] dz \tag{7.7}$$

This also implies that in a region where there is no long-term strength, there can be no stress differences and hence no "driving force".

According to plate-tectonic theory, mid-ocean spreading ridges are places where the weak, partially-molten asthenosphere rises almost to the surface. In such places the rising magma is probably under lithostatic stress because it is too hot and fluid to sustain long-term stress differences. If a ridge exists in the domain, at x = 0, then it follows that $\Omega(0) = 0$. According to the present definition, "driving force" does not come "from" ridges or anywhere else but is an intrinsic property of a point. Since the stress σ_{xx} is generally unknown, some other means is needed to compute Ω.

Because $\Omega(0) = 0$, it follows that

$$\Omega(x) = \Omega(x) - \Omega(0) =$$

$$\int_{h(x)}^{z_c} \left[\sigma_{xx}(x,z) - \sigma_{xx}(0,z)\right] dz + \int_{h(x)}^{z_c} \int^{z} g(s)\left[\rho(x,s) - \rho(0,s)\right] ds\, dz \qquad (7.8)$$

Now, because of the two dimensional assumption, the stress component, τ_{xy} is everywhere zero. Because of assumption (2) the component τ_{xz} is zero at the base of the domain $z = z_c$. It follows that in the absence of horizontal body forces that the integrated horizontal stresses must balance to prevent a net acceleration of the domain:

$$\int_{h(0)}^{z_c} \sigma_{xx}(0,z)\, dz = \int_{h(x)}^{z_c} \sigma_{xx}(x,z)\, dz \qquad (7.9)$$

This causes the first term of (7.8) to drop out. Freeing the coordinate system, and calling the density profile under a ridge (or any other hot, strength-free zone) $\rho_R(z)$, we obtain the theorem:

$$\Omega(x) = \int_{h(x)}^{z_c} \int^{z} g(s)\left[\rho(x,s) - \rho_R(s)\right] ds\, dz \qquad (7.10)$$

The assumption of isostasy has not been employed in deriving this, but is is a necessary constraint on estimates of $\rho(x,y)$, to avoid a conflict with assumption (2).

In the case of the Zagros Mountains, finite element models Z25 and Z29 showed that the value of Ω in the Arabian shield is 2.8 - 5.5 x 10^{15} dyne/cm, by means of equation (7.7). If the value is also calculated relative to zero in the spreading Red Sea, by means of (7.10), it will be possible to determine whether the force of the spreading ridge is sufficient to create the mountains.

The calculation of Ω for the Arabian shield is illustrated in Figure 7.2, in a type of diagram introduced by Frank (1972). The depth of the Red Sea is taken as 1.5 km, an average value for the spreading central region, and the density of seawater as 1.03 g/cc. The average density of the magma below is obtained by assuming that it is the same material as rises under other mid-ocean ridges, that these other ridges are isostatically compensated, that the difference in density between lithosphere and asthenosphere is entirely due to thermal expansion, and that the density of cold lithosphere is 3.3 g/cc. All of the assumptions are used to derive and evaluate the equation

$$\rho_{MAGMA} = \frac{\rho_{H_2O}(d_B - d_R) + \frac{1}{2}\rho_{LITH}(Z_C - d_B)}{\frac{Z_C}{2} - d_R + \frac{d_B}{2}} \qquad (7.11)$$

where d_B is the depth of old ocean basins and d_R is the depth of the spreading ridge. (A value of 5.7 km for basins is taken from the survey of Sclater et al. (1971)). Then, working upward into the Arabian shield from the pressure at the compensation depth, the density is assumed to increase linearly to 3.30 g/cc at the Moho due to the geothermal gradient. The loop is closed at sea level if the average density of the 33 km crust is 2.80 g/cc, a reasonable figure.

The driving force exerted on the Arabian shield is equal to the area between the curves in Figure 7.2. It is 2.96×10^{15} dyne/cm if the values above are correct. This figure matches very well the driving force of 2.8×10^{15} dyne/cm required by model Z26. Thus I conclude that no downgoing slabs, asthenospheric tractions, or shear stresses on transforms are required to explain the formation of the Zagros. This coincidence of values tends to confirm the model in which the Crush Zone contains weak limestones and the crust is deformed and seismic at a low regional shear stress of 200 bars. A possible mechanism for such earthquakes is given in the last section.

The situation in the Indian plate is much more complex. The plate has no symmetry, and mid-plate earthquake mechanisms solutions by Fitch et al. (1973) show that shear stresses are dominant in some regions. Under these conditions Ω cannot be defined. However, it is easily shown that no unknown driving forces are required to explain the convergence.

Comparing a mid-ocean ridge (2.6 km depth) to an old ocean basin (5.7 km depth), an additional 690 bars of topographic load is present in the former. Assuming that the isostatic compensation of this difference is distributed evenly over the upper 80 km of the mantle, the Ω value of the old ocean basin can be estimated as 2.8 x 10^{15} dyne/cm. This is equivalent to a horizontal deviatoric compression of 560 bars distributed over 50 km. It is also, coincidentally, almost the same as the value for the continental shields, so there will be little change in the amount of driving force across the coast of India.

The value for Tibet is higher; about 7 x 10^{15} dyne/cm as estimated in the last section. This mismatch merely shows that the assumption of plane stress is not valid, and that lines of horizontal deviatoric compression converge on the Himalayas from a wide region of the Indian plate. The same result was obtained by Fitch et al. (1973) on the basis of mid-plate fault-plane solutions. Pressure axes for Indian earthquakes were horizontally North-South, indicating that driving stress is transmitted from the South. Also, left-lateral strike-slip events were observed on the Ninety East Ridge, showing that the eastern part of the plate is exerting northward forces on the western part. The authors speculated that some of the stress results from the pull of the downgoing Indonesian slabs.

The plateau of Tibet is about 2000 km wide from west to east, so a total deviatoric force of about 1.4×10^{24} dynes is required to continue the Himalayan convergence. However, the Indian Plate has also some 12,000 km of spreading ridges and 6,000 km of downgoing slabs. The potential driving force of the ridges alone is 3.4×10^{24} dynes. And Richardson and Solomon (1976) estimate that the driving force exerted on oceanic lithosphere by downgoing slabs is 70-150% of that supplied by ridges. Clearly, the Himalayan convergence event is only opposing about one-third of the known driving forces acting on the Indian plate.

7.3 Speculations of the Mechanisms of Earthquakes

In the chapter on the Zagros it was determined that rocks of the continental crust fail and produce earthquakes at shear stresses of 100-800 bars, and that the highest value that can be reconciled with known driving forces is 200 bars. The major thrust faults of the Himalayas are seismic at shear stresses of less than 350 bars. Both of these values are an order of magnitude below the failure stresses predicted by laboratory studies on deformation of similar rocks at crustal pressures (Stesky et al., 1974). A similar problem in a strike-slip regime is the lack of a heat flow anomaly above the San Andreas fault in California (Brune et al., 1969). This places an upper limit of about 250 bars on the shear stress there, which cannot be reconciled with the laboratory behavior of local rocks (Stesky and Brace, 1973).

Presumably the difference in scale between the laboratory and the Earth is responsible. One earthquake parameter which is virtually impossible to recreate in the laboratory is the amount of slip, which is proportional to the amount of frictional work converted into heat on the fault surface. Assuming the same slip velocity, the work rises linearly with the duration of slip, but the distance through which heat is conducted rises only as the square root. This means that the temperature rise during natural earthquakes may be several orders of magnitude greater than in the lab. McKenzie and Brune (1972) demonstrated that for large earthquakes melting on the plane of faulting is a strong possibility.

In support of their numerical model McKenzie and Brune presented photographs of a shallow-dipping thrust fault in the Nepal Himalayas in which there is a 2 cm layer of fresh, bubbly rock glass. Francis (1972) and Sibson (1975) have described other terrains in which rock glass is found on fault planes. There is little doubt that such melting is a common process, at least in continental collision belts.

If melting occurs, it is possible to envision a process in which the laboratory frictional law and the low tectonic stress can be reconciled. The sequence of events is shown in Figure 7.3.

The curves of shear stress and displacement in the figure are adapted from the calculations of Ida (1972; case 1), who showed that for a dynamic longitudinal shear crack the concentration of stress around the tip may persist after the initiation of slip. Similar results have been obtained numerically by Andrews (1976) for the case of a dynamic plane-strain shear crack. The innovation in this section is conceptual rather than numerical; I would replace the "cohesive force" discussed by the above authors with the frictional faulting stress. The "surface energy" of the dislocation would be interpreted as the energy necessary for melting, rather than the energy expended in breaking interatomic bonds.

At time A, the dislocation is initiated at a distant point, and shear waves are radiated to the point in question. These waves create a stress concentration around the edges of the dislocation loop. Between times A and B the edge approaches and the shear stress rises. At time B the frictional stress is reached and the fault begins to slip as the edge of the dislocation passes through. Regardless of the initial tectonic stress, this slip takes place at the high frictional stress, so temperature on the fault rises rapidly. At time C the temperature reaches the solidus and the fault melts. Stress will drop sharply, seeking the level which maintains a thin film of melt. With the stress relieved, large amounts of slip are possible, and the energy radiated takes the form of seismic waves which in turn apply load to adjacent unfaulted regions. Finally, at time D when the potential static slip has been consumed or even exceeded (Madariaga, 1976) the rate of slip will drop to zero. This removes the heat source, and freezing of the film of melt will be rapid. After this time the fault does not move, and temperature gradually decays back to normal.

For this process to operate it must be initiated and it must obey conservation of energy. The initial source of the earthquake must be a region of stress concentration (perhaps near the end of an old break) where the tectonic stress becomes equal to the frictional stress. Such a region could be very small. As soon as the dislocation propagates out of the highly stressed region, it releases energy as heat and seismic waves which must be less than the total energy stored in the area as elastic strain. This constraint determines how far the event can propagate.

Yet even if this hypothesis is correct, it does not entirely remove the problem posed by the low strength of continental crust. The melting mechanism can only operate in rapidly-propagating dislocations (earthquakes), and gives no assistance in explaining observations of aseismic creep. Such creep may even dominate the deformations, as it apparently does in the shallow parts of the Himalayan Boundary Fault, and perhaps even in the interplate region as a whole (Chen and Molnar, 1976). In this area at least there is still some basic mechanism operating which we have not yet discovered.

Chapter 7 Figure Captions

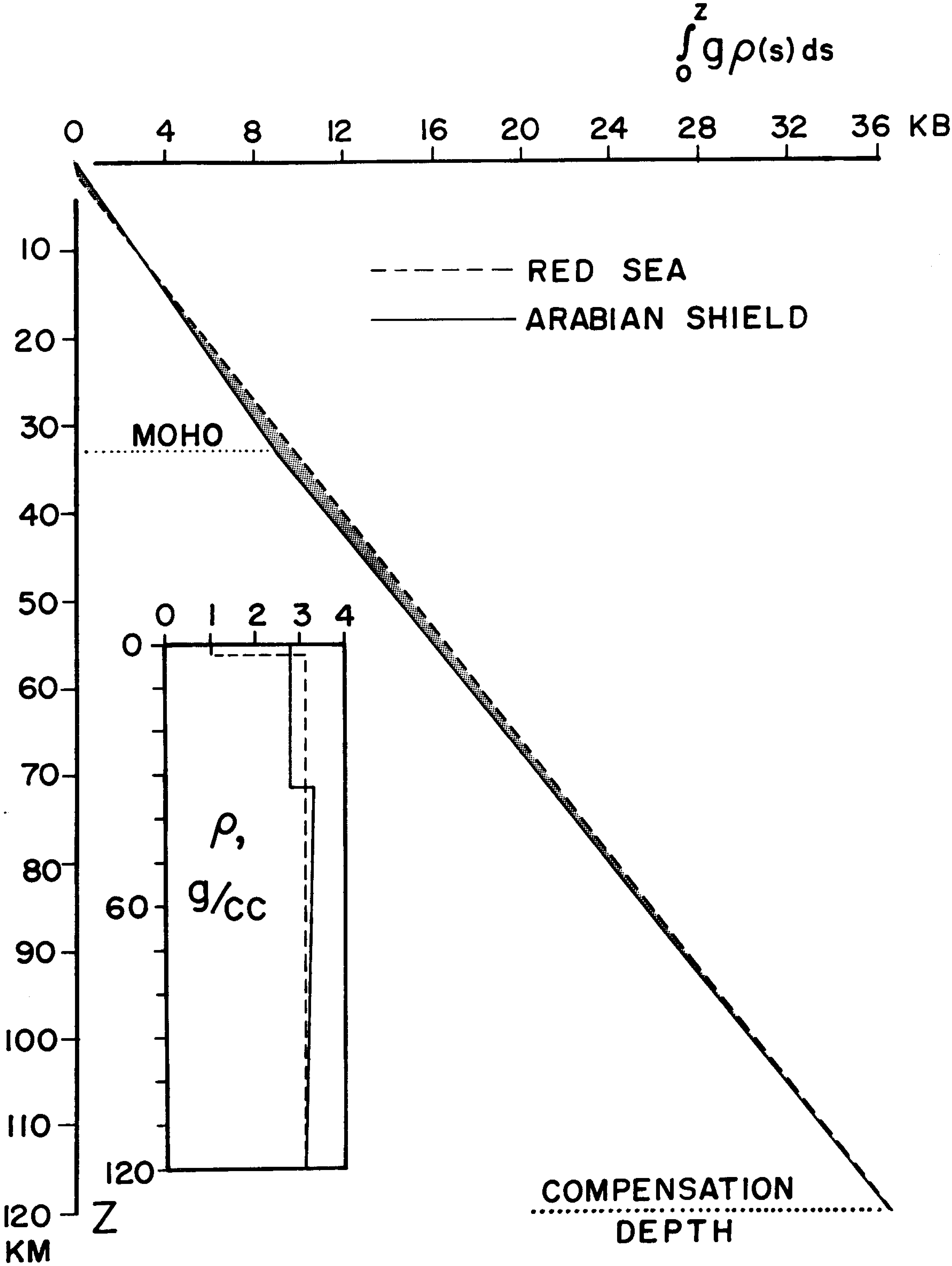
Figure 7.1 Averaged topography of the Himalayas and Trans-Himalayas. Crosses are land elevations sampled on a 1/3° by 1/3° grid in the region 26-32°N by 77-95°E. Points have been plotted according to their relative position between the Boundary Fault and Indus Suture. Curve consists of straight-line segments connecting the average elevations in each 0.1 division of the width of the range.

Figure 7.2 Profile of vertical stress versus depth in two adjacent parts of the earth. Inset shows the assumed profiles of density. Although the integrated mass becomes equal at the compensation depth (isostasy) there is a net difference in the integrals of the two curves (shaded area). Since there are no deviatoric stresses below the Red Sea, this difference gives the integral of deviatoric stress (driving force) in the Arabian shield.

Figure 7.3 Schematic diagram of stress, slip, and temperature at one point on a fault during the propagation of an earthquake dislocation through that point. At time A, elastic waves begin to arrive from the hypocenter. At B, stress reaches frictional limit, slip begins, and temperature rises due to friction. At C, melting occurs, temperature stabilizes, and stress decreases as the slip accelerates. At time D, slip stops, temperature begins to fall, and stress may reverse as stopping phases radiate in from the ends of the fault. Stress and slip plots are after Ida (1972).


SOUTH
NORTH
8 KM
6
4
2
0
Indus Suture
Boundary Fault
100 KM
-0.2
0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
Fractional Width

Fig. 7.1



$\int_0^z g\rho(s)\,ds$
0 4 8 12 16 20 24 28 32 36 KB
RED SEA
ARABIAN SHIELD
MOHO
10 20 30 40 50 60 70 80 90 100 110 120 KM
Z
0 1 2 3 4
0
60
120
ρ, g/cc
COMPENSATION DEPTH

Fig. 7.2

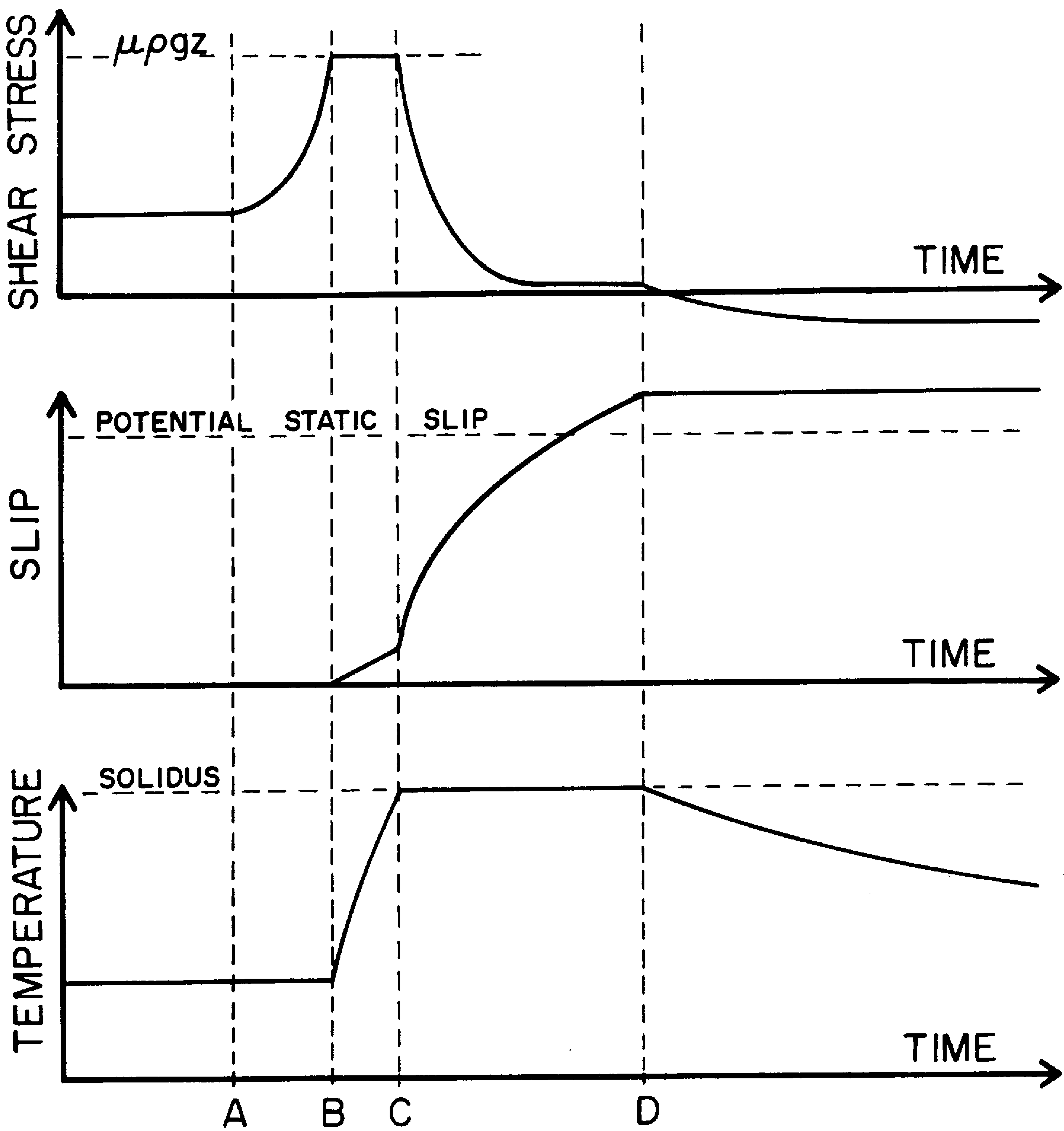

SHEAR STRESS
$\mu\rho g z$
TIME
SLIP
POTENTIAL STATIC SLIP
TIME
TEMPERATURE
SOLIDUS
TIME
A
B
C
D

Fig. 7.3

REFERENCES

Abdalian, Sedrak, (1963), Seismo-tectonique de l'Ivan, Institute of Geophysics of the University of Tehran Publication #16, 105 pp.

Agrawal, P.N. and V.K. Gaur, (1971), Study of Crustal deformation in India, Tectonophysics, 15, 287-296.

Akasheh, Bahram, (1972), Seismizität und Tektonik von Iran, Journal of the Earth and Space Physics, I, #1, 8-23.

Akasheh, Bahram, (1973), Mechanism of the earthquake of April 10, 1972 in Qir (Iran), Zeitschrift für Geophysik, 39, 1055-1061.

Akasheh, Bahram, (1975), Travel time residuals in the Iranian Plateau, Journal of Geophysics, 41, 281-288.

Allègre, C.J., F. Albarède, M. Grunenfelder, and V. Koppel, (1974), 238U/206Pb-235U/207Pb-232Th/208Pb zircon geochronology in alpine and non-alpine environments, Contributions to Mineralogy and Petrology, 43, 163-194.

Ambolt, Nils, (1948), Relative Schwerkrafts-Bestimmungen mit Pendeln in Zentralasien, Reports from the Scientific Expedition to the Northwestern Provinces of China under the Leadership of Dr. Sven Hedin, Vol. II, Stockholm.

Anderson, Don L. and C.B. Archambeau, (1964), The anelasticity of the Earth, Journal of Geophysical Research, 69, 2071-2084.

Andrews, D.J., (1972), Numerical simulation of sea-floor spreading, Journal of Geophysical Research, 77, 6470-6481.

Andrews, D.J., (1976), Rupture velocity of plane-strain shear cracks, submitted to Journal of Geophysical Research.

Anonymous, (1974), Tibet Today, Foreign Languages Press, Peking, 117 pp., illus.

Argand, Emile, (1924), La tectonique de l'Asie, 13th International Geological Congress, Brussels, 1922, Comptes rendu, I, part 5, 171-372.

Armbruster, J.G., L. Seeber, and K.H. Jacob, (1975), Tectonics of the Lesser Himalayas in North Pakistan based on microearthquake observations, EOS Transactions of the American Geophysical Union, 56, #6, 396.

Artyushkov, E.V., (1973), Stresses in the lithosphere caused by crustal thickness inhomogeneities, Journal of Geophysical Research, 78, #32, 7675-7708.

Azizbekov, Sh.A, (1971), Magmatism in the Caucusus, Iran, and Turkey (English translation from the Russian), International Geological Review, 13, #10, 1464-1470.

Balderman, M.A., (1974), The effect of strain-rate and temperature on the yield point of hydrolitically weakened synthetic quartz, Journal of Geophysical Research, 79, 1647-1652.

Banghar, A.R., (1976), Mechanism solution of Kinnaur (Himachal Pradesh, India) earthquake of January 19, 1975, Tectonophysics, 31, T5-T11.

Bazeley, G.P., Y.K. Cheung, B.M. Irons, and O.C. Zienkiewicz, (1965), Triangular elements in bending-conforming and non-conforming solutions, Proceedings of the Conference on Matrix Methods in Structural Mechanics, Air Force Institute of Technology, Wright Patterson AFB, Ohio.

Berggren, William A., (1972), A Cenozoic time-scale; some implications for regional geology and paleobiogeography, Lethaia, 5, #2, 195-215.

Birch, Francis, (1942), Handbook of Physical Constants, Geological Society of America Special Paper, 36.

Bird, Peter and M. Nafi Toksöz , (1974), Thermal Regimes in continental convergence and effects on seismicity, Earthquake Notes, 46, 48.

Bird, Peter and M. Nafi Toksöz (1975), Structure and evolution of the Tibetan Plateau, EOS Transactions of the American Geophysical Union, 56, 397.

Bird, Peter, M.N. Toksöz, and N.H. Sleep (1975), Thermal and mechanical models of continent-continent convergence zones, Journal of Geophysical Research, 80, #32, 4405-4416.

Bird, R. Byron (1960) New variational principle for incompressible non-Newtonian flow, The Physics of Fluids, 3, #4, 539-541.

Blacic, James D., (1975), Plastic-deformation mechanisms in quartz: the effect of water, Tectonophysics, 27, 271-294.

Boore, David M., (1970), Love waves in nonuniform wave guides: finite difference calculations, Journal of Geophysical Research, 75, #8, 1512-1527.

Brown, G.C. and W.S. Fyfe, (1970), The production of granitic melts during ultrametamorphism, Contributions to Mineralogy and Petrology, 28, 310-318.

Brune, J. and J. Dorman, (1963), Seismic waves and earth structures in the Canadian Shield, Bulletin of the Seismological Society of America, 53, 167-210.

Brune, J.N., T. Henyey, and R. Roy, (1969), Heat flow, stress, and rate of slip along the San Andreas Fault in California, Journal of Geophysical Research, 74, 3821-3827.

Burke, K.C.A., Dewey, J.F., and W.S.F. Kidd, (1974), The Tibetan plateau; its significance for tectonics and petrology, Geological Society of American Abstracts with Programs, 6, #7, 1028-1029.

Byerly, J.D., (1966), The frictional characteristics of Westerly Granite, Ph.D. thesis, MIT, Cambridge, Mass.

Canitez, Nezihi, (1969), The focal mechanisms in Iran and their relations to tectonics, Pure and Applied Geophysics, 75, 76-87.

Carey, S.W., (1955), The orocline concept in geotectonics, Proceedings of the Royal Society of Tasmania, 89, 255-288.

Carslaw, H.S. and J.C. Jaeger, (1959), Conduction of Heat in Solids, Clarendon Press, Oxford, 510 pp.

Carter, Neville, (1971), Static deformation of silica and silicates, Journal of Geophysical Research, 76, 5514-5540.

Carter, Neville, (1975), High temperature flow of rocks, Reviews of Geophysics and Space Physics, 13, #3, 344-349, 379-381.

Carter, N.L. and H.G. Ave'Lallement, (1970), High temperature flow of dunite and peridotite, Geological Society of America Bulletin, 81, 2181-2202.

Cathles, L.M., III, (1975), The viscosity of the Earth's mantle, Princeton University Press, Princeton, N.J.

Chang, C.F. and S.L. Zeng, (1973), Tectonic features of the Mount Jolmo Lungma region in southern Tibet, Scientia Sinica, I, 1-12.

Chang, Ta, (1959), The Geology of China (English translation), United States Department of Commerce, 1963,

Chapple, William M., (1975), Mechanics of thin-skinned fold and thrust belts, EOS Transactions of the American Geophysical Union, 56, 457.

Chaudhury, H.M., (1965), Seismic surface wave dispersion and the crust across the Grangetic Basin, Indian Journal of Meteorology and Geophysics, 17, 385-394.

Chaudhury, H.M., (1966), Sedimentary Rayleigh waves across the Grangetic Basin-a preliminary study, Indian Journal of Meteorology and Geophysics, 18, 273-278.

Chen, Wang-Ping and Peter Molnar, (1975), Short-period Rayleigh wave dispersion across the Tibetan plateau, Bulletin of the Seismological Society of America, 65, 1051-1058.

Chen, Wang-Ping and Peter Molnar, (1976), Seismic moments of major earthquakes and the average rate of slip in central Asia, preprint.

Christensen, N.I. and D.M. Fountain, (1975), Constitution of the lower continental crust based on experimental studies of seismic velocities in granulite, Geological Society of America Bulletin, 86, 227-236.

Chugh, R.S., (1974), Study of recent crustal movements in India and future programme, International Symposium on Recent Crustal Movements 1974, Zurich.

Clark, Sydney P., Jr., (1966), Thermal Conductivity, in Handbook of Geophysical Constants, Geological Society of America Memoir, 97, 459-482.

Clark, Sydney P., Jr. and A.E. Ringwood, (1964), Density distribution and constitution of the mantle, Reviews of Geophysics, 2, #1, 35-88.

Coster, H.P., (1947), Terrestrial heat flow in Persia, *Monthly Notices, Royal Astronomical Society, Geophysical Supplement*, *5*, 131-145.

Crawford, A.R., (1972), Iran, continental drift, and plate tectonics, *Proceedings of the 24th International Geological Congress*, section 3, 106-112.

Davis, B.T.C. and J.L. England, (1964), The melting of forsterite up to 50 kilobars, *Journal of Geophysical Research*, *69*, #6, 1113-1116.

Desio, A. and B. Zanettin, (1970), *Geology of the Baltoro Basin*, volume 2, Brill, Leiden, 308 pp.

DeTerra, H., (1935), Geological studies in the northwest Himalaya between the Kashmir and Indus valleys, *Memoir of the Connecticut Academy of Arts and Sciences*, *8*, p.18.

Detrick, Robert, (1976), Pn arrivals, surface wave dispersion, and crustal structure of the Tibetan plateau, unpublished manuscript, Massachusetts Institute of Technology, Cambridge.

Dewey, John F. and John M. Bird, (1970), Mountain belts and the new global tectonics, *Journal of Geophysical Research*, *75*, #10, 2625-2647.

Dewey, John F. and John M. Bird, (1971), Origin and emplacement of the ophiolite suite: Appalachian ophiolites in Newfoundland, *Journal of Geophysical Research*, *76*, 3179-3206.

Dewey, J.F. and K.C.A. Burke, (1973), Tibetan, Variscan, and preCambrian basement reactivation: products of continental collision, *Journal of Geology*, *81*, 683-692.

Dewey, John F., Walter C. Pitman, III, William B.F. Ryan, and Jean Bonnin, (1973), Plate tectonics and the evolution of the Alpine System, *Geological Society of America Bulletin*, *84*, 3137-3180.

Dietz, R.S. and J.C. Holden, (1970), Reconstruction of Pangaea: Breakup and dispersion of continents, Permian to present, *Journal of Geophysical Research*, *75*, 4939-4956.

Diment, W.H., T.C. Urban, and F.A. Revetta, (1970, Some geophysical anomalies in the Eastern United States, in *The Nature of the Solid Earth*, McGraw-Hill, Inc., cop. 1972, 544-572.

Dorman, J., M. Ewing, and J. Oliver, (1960), Study of shear velocity distribution in the upper mantle using Rayleigh waves, Bulletin of the Seismological Society of America, 50, 87-115.

Dziewonski, A., S. Bloch, and M. Landisman, (1969), A technique for the analysis of transient seismic signals, Bulletin of the Seismological Society of America, 59, #1, 427-444.

✓ Elliott, David, (1976), The motion of thrust sheets,
[sic] Journal of Geophysical Research, 81, #5, 949-963.

Emery, K.O., E. Uchupi, J.D. Phillips, C.O. Bowen, E.T. Bruce, and S.T. Knott, (1970), Continental rise off eastern North American, Bulletin of the American Association of Petroleum Geologists, 54, 44-108.

Falcon, Norman Leslie, (1969), Problems of the relationship between surface structure and deep displacements illustrated by the Zagros Range, in Time and Place in Orogeny, Geological society of London, London; pp. 9-22.

Fitch, T.J., (1970), Earthquake mechanism in the Himalayan, Burmese, Andaman regions and continental tectonics in Central Asia, Journal of Geophysical Research, 75, 2699-2708.

Fitch, T.J., M.H. Worthington, and I.B. Everingham, (1973), Mechanisms of Australian earthquakes and contemporary stress in the Indian Ocean plate, Earth and Planetary Science Letters, 18, #2, 345-356.

Forsyth, Donald and Seiya Uyeda, (1975), On the relative importance of the driving forces of plate motion, Geophysical Journal of the Royal astronomical Society, 43, 163-200.

Frank, F.C., (1972), Plate tectonics, the analogy with glacier flow, and isostasy, in Flow and Fracture of Rocks, AGU Geophysical Monograph, 16, 285-292.

Gansser, A., (1964), Geology of the Himalayas, Wiley-Interscience, N.Y., 289 pp.

Goetze, C. and W.F. Brace, (1972), Laboratory observations of high-temperature rheology of rocks, Tectonophysics, 13, 583-600.

Goetze, C. and D.L. Kohlstedt, (1973), Laboratory study of dislocation climb and diffusion in olivine, Journal of Geophysical Research, 78, 5961-5971.

Grant, F.S., and G.F. West, (1965), Interpretation Theory in Applied Geophysics, McGraw-Hill Book Co., New York, 584 pp.

Gregory, K.J. and D.E. Walling, (1973), Drainage Basin Form and Process: A Geomorphological Approach, Wiley, New York, 456 pp.

Grosvenor, M.B., W. Chamberlin, J.M. Darley, A.D. Grazzini, F. Shor, and J.B. Billard, (1966), National Geographic Atlas of the World, National Geographic Society, Washington, 343 pp.

Gulatee, B.L., (1956), Gravity data in India, Technical Paper #10, and Addenda #1,2,3,4,5. Geological Survey of India, Delhi.

Gupta, H.K. and H. Narain, (1967), Crustal structure in the Himlayan and Tibet Plateau region from surface wave dispersion, Bulletin of the seismological Society of America, 57, 235-248.

Gutenberg, B., (1956), The energy of earthquakes, Quarterly Journal of the Geological Society of London, 112, 1-14.

Hallam, Anthony, (1974), "The Zagros Suture Zone of Iran: A problem in paleotectonic reconstruction," address to the Department of Geological Sciences, Harvard University, Cambridge, Mass., April, 1974.

Hamet, Jean and Claude-Jean Allegre, (1976), RB-SR Systematics in granite from central Nepal (Manaslu). Significance of the Oligiocene age and high $^{87}Sr/^{86}Sr$ ratio in Himalayan orogeny, Geology, 4, 470-472.

Baskell, N.A., (1951), The dispersion of surface waves on multilayered media, Bulletin of the Seismological Society of America, 43, 17-34.

Haynes, S.J. and H. McQuillan, (1974), Evolution of the Zagros suture zone, southern Iran, Geological Society of America Bulletin, 83, 739-744.

Heard, H.C., (1972), Steady-state flow in polycrystalline halite at pressure of 2 kilobars, American Geophysical Union Monograph, 16, 191-209.

Heard, H.C. and N.L. Carter, (1968), Experimentally induced 'natural' intragranular flow in quartz and quartzite, American Journal of Science, 266, 1-42.

Heard, H.C. and C.B. Raleigh, (1972), Steady-state flow in marble at 500°C, Geological Society of America Bulletin, 83, 935-956.

Hennig, Anders, (1915), Zur Petrographie und Geologie von Südwesttibet in Southern Tibet: Discoveries in Former Times compared with my own Researches in 1906-1908, vol. 5, Lithographic Institute of the General Staff of the Swedish Army, Stockholm.

Herrin, Eugene, (1970), A comparative study of upper mantle models: Canadian shield and Basin and Range provinces, in Frank Press, editor, The Nature of the Solid Earth, McGraw-Hill, cop. 1972, pp. 217-231.

Hobbs, B.E., A.C. McLaren, and M.S. Paterson, (1972), Plasticity of single crystals of synthetic quartz, American Geophysical Union Monograph, 16, 29-53.

Holeman, J.N., (1968), The sediment yield of major rivers of the world, Water Resources Research, 4, 737-747.

Holmes, Arthur, (1965), Principles of Physical Geology (revised edition), Ronald Press Company, New York, 1288 pp.

Hubbert, M.K. and W.W. Rubey, (1959), Role of fluid pressure in the mechanics of overthrust faulting, Geological Society of America Bulletin, 70, 115-206.

Ida, Yoshiaki, (1972), Cohesive force across the tip of a longitudinal-shear crack and Griffith's specific surface energy, Journal of Geophysical Research, 77, #20, 3796-3805.

Ingersoll, L.R., O.J. Zobell, and A.C. Ingersoll, (1954), Heat conduction; with engineering, geological, and other applications, University of Wisconsin Press, Madison, 325 pp.

Isacks, Bryan, Jack Oliver, and Lynn R. Sykes, (1968), Seismology and the new global tectonics, Journal of Geophysical Research, 73, #18, 5855-5899.

Jessop, A.M., M.A. Hobart, and J.G. Sclater, (1975), World heat flow date compilation-1975, Geothermal Series, #5, Earth Physics Branch, Department of Energy, Mines, and Resources, Ottawa.

Kay, R., N.J. Hubbard, and P.W. Gast, (1970), Chemical characteristics of oceanic ridge volcanic rocks, Journal of Geophysical Research, 75, 1585-1613.

Khalturin, V.I., T.G. Rautian, and P. Molnar, (1976), The spectral content of Pamir-Hindu Kush earthquakes; evidence for a high Q zone in the mantle, preprint.

Kidd, W.S.F., (1975), Widespread late Neogene and Quaternary calc-alkaline volcanism on the Tibetan Plateau, EOS Transactions of the American Geophysical Union, 56, 453.

King, P.B., (1950), Tectonic framework of southeastern United States, Bulletin of the American Association of Petroleum Geologists, 34, 635-671.

Kisslinger, C., E.R. Engdahl, G. Boucher, and R. Ganse, (1974), Seismic wave propagation and earthquake characteristics in Asia, Semi-Annual Technical Report #2, Cooperative Institute for Research in Environmental Sciences, Boulder, Colorado.

Knopoff, L. and A.A. Fonda, (1975), Upper-mantle structure under the Arabian Peninsula, Tectonophysics, 26, 121-134.

Kohlstedt, D.L. and C. Goetze, (1974), Low-stress high-temperature creep in olivine single crystals, Journal of Geophysical Research, 79, #14, 2045-2051.

Kono, Masaru, (1974), Gravity anomalies in East Nepal and their implications to the crustal structure of the Himalayas, Geophysical Journal of the Royal Astronomical Society, 39, 283-299.

Lachenbruch, A.H. and G.A. Thompson, (1972), Oceanic ridges and transform faults: their intersection angles and resistance to plate motions, Earth and Planetary Science Letters, 15, 116-122.

Lambeck, Kurt, (1972), Gravity anomalies over ocean ridges, Geophysical Journal of the Royal Astronomical Society, 30, 37-54.

Lee, Wook Bae and Sean C. Solomon, (1975), Inversion schemes for surface wave attenuation and Q in the crust and mantle, Geophysical Journal of the Royal Astronomical Society, 43, 47-71.

Lees, George Martin, (1952) Foreland folding, The Quarterly Journal of the Geological Society of London, 108, 1-33.

Lees, G.M. and N. L. Falcon, (1952) The geographical history of the Mesopotamian Plains, Geographical Journal, 118, 24-39.

LeFort, Patrick, (1975) Himalayas: The collided range. Present knowledge of the continental arc, American Journal of Science, 275-A, 1-44.

LePichon, Xavier, (1968) Sea floor spreading and continental drift, Journal of Geophysical Research, 73, #12, 3661-3697.

McClintock, Frank A, (1971) Plasticity aspects of fracture, in Harold Liebowitz, editor, Fracture: An Advanced Treatise, vol. 3, 48-227, Academic Press, New York.

MacDonald, G.J.F., (1959) Calculations on the thermal history of the Earth, Journal of Geophysical Research, 64, 1967-2000.

McKenzie, Dan P., (1972) Active tectonics of the Mediterranean region, Geophysical Journal of the Royal Astronomical Society, 30, 109-185.

McKenzie, Dan, and James N. Brune, (1972) Melting on fault planes during large earthquakes, Geophysical Journal of the Royal Astronomical Society, 29, 65-78.

McKenzie, D.P., (1969) Speculations on the consequences and causes of plate motions, Geophysical Journal of the Royal Astronomical Society, 18, 1-32.

Madariaga, Raul, (1976) Dynamics of an expanding circular fault, Bulletin of the Seismological Society of America, 66, #3, 639-666.

Marcal, P.V. and I.P. King, (1967) Elastic-plastic analysis of two-dimensional stress systems by the finite element method, International Journal of Mechanical Science, 9, #3, 143-156.

Marussi, Antonio, (1964) Geophysics of the Karakorum, Scientific Reports of the Italian Expedition to the Karakorum (K2) and Hindu Kush, part II (Geophysics), Vol. I, E.J. Brill, Leiden.

Mattauer, M., (1975) Sur le mécanisme de formation de la schistosité dans l'Himalaya, Earth and Planetary Science Letters, 28, 144-154.

Minear, J.W. and M.N. Toksöz, (1970a) Thermal regime of a downgoing slab and new global tectonics, Journal of Geophysical Research, 75, 1397-1419.

Minear, J.W. and M.N. Toksöz, (1970b), Thermal regime of a downgoing slab, Tectonophysics, 10, 367-390.

Minear, J.W. and M.N. Toksöz (1971) Reply, Journal of Geophysical Research, 76, 610-612.

Molnar, P. and J. Oliver, (1969) Lateral variations of attenuation in the upper mantle and discontinuities in the lithosphere, Journal of Geophysical Research, 74, 2648-2682.

Molnar, Peter, T.J. Fitch, and F.T. Wu, (1973) Fault plane solutions of shallow earthquakes and contemporary tectonics in Asia, Earth and Planetary Science Letters, 19, 101-112.

Molnar, P., and P. Tapponnier, (1975), Tectonics of Asia: consequences and implications of a continental collision, Science, 189, 419-426.

Naylor, R.S., (1971) Acadian orogeny: an abrupt and brief event, Science, 172, 558-559.

Nowroozi, Ali A., (1971), Seismotectonics of the Persian Plateau, eastern Turkey, Caucasus, and Hindu-Kush regions, Bulletin of the Seismological Society of America, 61, #2, 317-341.

Nowroozi, Ali A., (1972) Focal mechanism of earthquakes in Persia, Turkey, West Pakistan, and Afghanistan and plate tectonics of the Middle East, Bulletin of the Seismological Society of America, 62, #3, 823-850.

Oliver, Jack, and Bryan Isacks, (1967) Deep earthquake zones, anomalous structures in the upper mantle, and the lithosphere, Journal of Geophysical Research, 72, #16, 4259-4275.

Osborn, E., (1959) The complimentariness of orogenic andesite and alpine peridotite, Geochimica et Cosmochimica Acta, 33, 307-324.

Oxburgh, E.R., (1972) Flake tectonics and continental collision, Nature, 239, 202-204.

Pakiser, L.C. and J.S. Steinhardt, (1953) Explosion seismology in the Western Hemisphere, in Research in Geophysics vol. 2, MIT Press, Cambridge, Mass., pp. 123-147.

Parmentier, E.M. and D.L. Turcotte, (1975) Numerical experiments on the structure of plumes, Journal of Geophysical Research, 80, #32, 4417-4424.

Parrish, D.K., A. Kriva, and N.L. Carter, (1976) Finite element folds of similar geometry, Tectonophysics, 32, 183-207.

Patton, Howard J., (1975) Source and path effects for Rayleigh waves from central Asian earthquakes, EOS Transactions of the American Geophysical Union, 56, 397.

Pavoni, N., (1975) Recent crustal movements, in International Geodynamics Project: First Report of Switzerland, Swiss Committee for the International Geodynamic Project, Bern; 62 pages.

Peaceman, D.W. and H.H. Rachford, Jr., (1955) The numerical solution of parabolic and elliptic differential equations J. Soc. Ind. Appl. Math., 3, 28-41.

Phillips, J.D., (1970) Magnetic anomalies in the Red Sea Philosophical Transactions of the Royal Society of London, vol 267, 205-217.

Post, Robert L., Jr., (1973) The flow laws of Mt. Burnett Dunite, Ph.D. thesis, University of California at Los Angeles.

Post, Robert L., Jr. and David T. Griggs, (1973) The Earth's mantle; evidence of non-Newtonian flow, Science, 181, #4106, 1242-1244.

Powell, C. McA. and P.J. Conaghan, (1973) Plate tectonics and the Himalayas, Earth and Planetary Science Letters, 20, 1-12.

Powell, C. McA. and P.J. Conaghan, (1975) Tectonic models of the Tibetan plateau, Geology, 3, 727-731.

Press, Frank, (1966) Seismic velocities, in Sydney P. Clark ed., Handbook of Physical Constants, Geological Society of America Memoir, 97, 195-218.

Press, Frank, (1970) Earth models consistent with geophysical data, Physics of Earth and Planetary Interiors, 3, 3-22.

Qureshy, M.N., S. Venkatachalam, and C. Subrahmanyam, (1974) Vertical tectonics in the Middle Himalayas: an appraisal from recent gravity data, Geological Society of America Bulletin, 85, 921-926.

Rastogi, B.K., (1974) Earthquake mechanisms and plate tectonics in the Himalayan region, Tectonophysics, 21, 47-56.

Richardson, Randall M., and Sean C. Solomon, (1976) Intra-plate stress as an indicator of plate tectonic driving forces, Journal of Geophysical Research, 81, #11, 1847-1856.

Ross, John V. and K.C. Nielsen, (1975) Some observations and experimental data on experimentally deformed orthopyroxene EOS Transactions of the American Geophysical Union, 56, 911.

Roy, R.F., D.D. Blackwell, and F. Birch, (1968) Heat generation of plutonic rocks and continental heat-flow provinces, Earth and Planetary Science Letters, 5, 1-12.

Rutischauser, H., (1966) Algorithms 1 - Linear equation systems with symmetric positive definite band matrices by Cholesky's method - Computing, Archives for Electronic Computing, 1, 77-78.

Rutter, E.H. and S.M. Schmid, (1975) Experimental study of unconfined flow of Solnhofen Limestone at 500° to 600°C, Geological Society of America Bulletin, 86, 145-152.

Ruzaikin, A.I., I.L. Nersesov, V.I. Khalturn, and P. Molnar, (1976), Propagation of Lg and lateral variations in crustal structure in Asia, preprint.

Rybach, L., (1975) Geotraverse Basle-Chiasso in International Geodynamics Project: First Report of Switzerland, Swiss Committee for the International Geodynamics Project, Bern, 62 pages.

Salisbury, M.H. and D.M. Fountain, (1976) Continent/continent obduction zones: geological and geophysical case studies, EOS Transactions of the American Geophysical Union, 57, 335.

Santo, Tetsuo, (1965) Lateral variation of Rayleigh wave dispersion character, Part II: Eurasia, Pure and Applied Geophysics, 62, 67-80.

Sarne, Y., (1974) Material nonlinear time dependent three dimensional finite element analysis of reinforced and prestressed concrete structures, Ph.D. thesis, MIT, Cambridge, Mass.

Saxena, M.N., (1971) The crystalline axis of the Himalaya: the Indian shield and continental drift, Tectonophysics, 12, 443-447.

Schatz, J.F., (1971) Thermal conductivity of earth materials at high temperatures, Ph.D. thesis, MIT, June, 1971.

Schatz, J.F., and G. Simmons, (1972) Thermal conductivity of Earth materials at high temperatures, Journal of Geophysical Research, 77, 6966-6983.

Schmid, S.M., (1976) Rheological evidence for changes in the deformation mechanism of Solnhofen limestone towards low stresses, Tectonophysics, 31, T21-T28.

Schubert, G., and D.L. Turcotte, (1971), Phase changes and mantle convection, Journal of Geophysical Research, 76, 1424-1432.

Scientific Staff, The (1972) Deep sea drilling project, leg 22, Geotimes, 17, #6, p.15.

Sclater, John G., Roger N. Anderson, and M.Lee Bell, (1971) Elevation of ridges and evolution of the central eastern Pacific, Journal of Geophysical Research, 76, #32, 7888-7915.

Sen Gupta, Supriga, (1964) Possible subsurface structures below the Himalayas and the Gangetic Plains, Proceedings of the International Geological Congress, New Delhi, part 11, 334-351.

Sibson, Richard H., (1975) Generation of pseudotachylyte by ancient seismic faulting, Geophysical Journal of the Royal Astronomical Society, 43, 775-794.

Simmons, Gene and Ki-iti Horai, (1968) Heat flow data 2, Journal of Geophysical Research, 73, #20, 6608-6629.

Sinvhal, H., P.N. Agrawal, G.C.P. King, and V.K.Gaur, (1973) Interpretation of measured movement at a Himalayan (Nahan) thrust, Geophysical Journal of the Royal Astronomical Society, 34, 203-210.

Skeels, D.C., (1940) Gravity in sedimentary basins, Transactions of the American Geophysical Union, 21, 187-202.

Skinner, B.J., (1966) Thermal expansion, in Handbook of Physical Constants, Geological Society of America Memoir 97, 75-96.

Sleep, Norman Harvey, (1973) Deep structure and geophysical processes beneath island arcs, Ph.D. thesis, MIT, Cambridge, Mass.

Smith, Albert T., Jr., (1974) Time-dependent strain accumulation and release at island arcs: implications for the 1946 Nankaido earthquake, Ph.D. thesis, MIT, Cambridge, Mass.

Sorokhtin, O.G., (1973) A possible mechanism of formation of regional overthrusts and geosyncline folding, Akademia Nank SSSR, Izvestia: Physics of the Solid Earth, 415-423.

Stesky, R.M., and W.F. Brace, (1973) Estimation of frictional stress on the San Andreas Fault from laboratory measurements, in R.L. Kovach and A. Nur, editors, Proceedings of the Conference on Tectonic Problems of the San Andreas Fault System, Stanford, 1973, 206-214.

Stesky, R.M., W.F. Brace, and P.-Y.F. Robin, (1974) Friction in faulted rock at high temperature and pressure, Tectonophysics, 23, 177-203.

Stöcklin, J., (1968) Structural history and tectonics of Iran: Bulletin of the American Association of Petroleum Geologists, 52, #7, 1229-1258.

Stöcklin, J. and M.H. Nabavi, (1973) Tectonic map of Iran, 1:2,500,000, Geological Survey of Iran, Tehran.

Strang, Gilbert and George J. Fix, (1973) An Analysis of the Finite Element Method, Prentice-Hall, Englewood Cliffs, N.J., 306 pages.

Tandon, A.N. and H.M. Chaudhury, (1964) Thickness of the Earth's crust between Delhi and Shillong from surface wave dispersion, Indian Journal of Meteorology and Geophysics, 15, 467-474.

Tandon, A.N., and R.K. Dube, (1973) A study of crustal structure beneath the Himalayas from body waves, Pure and Applied Geophysics, 111, 2207-2215.

Tandon, A.N., and H.N. Srivastava, (1975) Focal mechanisms of some recent Himalayan earthquakes and regional plate tectonics, Bulletin of Seismological Society of America, 65, #4, 963-969.

Tapponnier, Paul and Peter Molnar, (1976a) Slip-line field theory as a quantitative method for understanding large-scale continental tectonics, preprint.

Tapponnier, Paul and Peter Molnar, (1976b) Active faulting and Cenozoic tectonics in China, preprint.

Tazieff, Haroun, (1972) The Afar Triangle, in Continents Adrift: Readings from Scientific American, W.H. Freeman and Co., San Francisco, 172 pages.

Toksöz, M. Nafi, Michael A. Chinnery, and Don L. Anderson, (1967) Inhomogeneities in the Earth's mantle, Geophysical Journal of the Royal Astronomical Society, 13, 31-59.

Toksöz, M.Nafi, J.W. Minear, and B. Julian (1971) Temperature field and geophysical effects of a downgoing slab, Journal of Geophysical Research, 76, 1113-1138.

Toksöz, M. Nafi, N. H. Sleep, and A. T. Smith, (1973) Evolution of the downgoing lithosphere and the mechanisms of deep-focus earthquakes, Geophysical Journal of the Royal Astronomical Society, 35, 285-310.

Toksöz, M. Nafi, and P. Bird, (1976) Temperatures in continental convergence zones, Tectonophysics, in press.

Toksöz, M. Nafi, and P. Bird, (1977) Formation and evolution of marginal basins and continental plateaus, Proceedings of the M. Ewing Symposium, Lamont-Doherty Geological Observatory, in press.

Tsai, Yi-Ben, and Keiiti Aki, (1969) Simultaneous determination of the seismic moment and attenuation of seismic surface waves, Bulletin of the Seismological Society of America, 59, 275-287.

Tsai, Yi-Ben and Keiiti Aki, (1970), Precise focal depth determination from amplitude spectra of surface waves, Journal of Geophysical Research, 75, #29, 5729-5743.

Tullis, J.A., (1971), Preferred Orientations in Experimentally Deformed Quartzites, Ph.D. thesis, University of California at Los Angeles, 362 pp.

Tullis, Jan, R.A. Yund, and T. Tullis, (1976), Experimental deformation of westerly granite and a polycrystalline albite rock, EOS Transactions of the American Geophysical Union, 57, 322.

Tung, James P. and Ta-liang Teng, (1974), Surface wave studies on crustal and mantle structure of China, University of Southern California Geophysical Laboratory Technique Report Number 74-2, Los Angeles, 36 pp., illus.

Turcotte, D.L. and E.R. Oxburgh, (1969), A fluid theory for the deep structure of dip-slip fault zones, Physics of Earth and Planetary Interiors, I, 381-386.

Tuttle, O.F. and N.L. Bowen, (1958), Origin of granite in the light of experimental studies in the system $NaAlSi_3O_8$-$KAlSi_3O_8$-SiO_2-H_2O, Geological Society of America Memoir 74.

Valdinge, K.S., (1973), Lithological subdivision and tectonics of the "Central Crystalline Zone" of Kumaon Himalaya, in Harsh K. Gupta (ed.), Seminar on Geodynamics of the Himalayan Region, March 5-8, 1973, National Geophysical Research Institute, Hyderabad.

Van Bemmelen, R.W. and H.P. Berlage, (1935), Versuch Einer Mathmatischen Behandlung Geotektonischer Bewegungen Unter Besouderer Berücksichtegung der Undationstheorie, Gerlands Beiträge Zur Geophysik, 43, 19-55.

Veevers, J.J., J.G. Jones, and J.A. Talent, (1971), Indo-Australian stratigraphy and the configuration and dispersal of Gondwanaland, Nature, 299, 383.

Wadia, D.N., (1966), Geology of India, 3rd ed. revised, English Language Book Society, London, 536 pp.

Warsi, Waris E.K., (1976), Plate tectonics and the Himalayan orogeny: a modelling study based on gravity data, M.S. thesis, Massachusetts Institute of Technology, Cambridge, Mass.

Weertman, Johnannes and Julia R. Weertman, (1975), High temperature creep of rock and mantle viscosity, Annual Review of Earth and Planetary Sciences, 3, 293-315.

Wegener, Alfred L., (1915), The origin of Continents and Oceans (translated from the 4th revised German edition by John Biram), New York, Dover Publication, 1966.

Wells, A.J., (1969), The crush zone of the Iranian Zagros mountains, and its implications, Geological Magazine, 106, 385-394.

Wilcox, L.E., W.J. Rothermel, and J.T. Voss, (1972), A geophysical geoid of Eurasia, EOS Transactions American Geophysical Union, 53, 343.

Woollard, George P., (1959), Crustal structure from gravity and seismic measurements, Journal of Geophysical Research, 64, 1521-1544.

Woollard, George P., (1968), The interrelationship of the crust, the upper mantle, and isostatic gravity anomalies in the United States, in The Crust and Upper Mantle of the Pacific Area, A.G.U. Geophysical Monograph, 12, 312-341.

Worzel, J.L. and G.L. Shurbet, (1955), Gravity anomalies at continental margins, Proceedings of the National Academy of Sciences, 41, 458.

Wyss, Max, (1970), Stress estimates for South American shallow and deep earthquakes, Journal of Geophysical Research, 75, 1529-1544.

Zienkiewicz, O.C., (1971), The Finite Element Method in Engineering Science, McGraw-Hill, London, 521 pp.

Appendix A

Table of Symbols Employed

A	stress constant in nonlinear rock flow law, $\frac{\text{dyne(sec)}^{1/n}}{\text{cm}^2}$
$A(\omega)$	amplitude spectrum of seismic wave, cm-sec
A_{el}	area of a finite element, cm^2
a	depth of brittle faulting/nonlinear creep transition in the crust, cm
B	temperature constant in nonlinear rock flow law, °K
b	Burger's dislocation vector, cm
C_p	specific heat at constant pressure, erg/gm-°C
D	constant of empirical experimental flow law, $[(\text{kilobar})^n\ \text{sec}]^{-1}$
$\underset{\sim}{D}_{ij}$	anisotropic plastic viscosity tensor, poises
d	total slip distance in a shear zone, cm
E	constant of plain-strain flow law (2.23), $\left[\frac{\text{dyne}^n}{\text{sec-cm}^{2n}}\right]^{1/2}$
"	ALSO, energy density of seismic waves, erg/cc
e_{ij}	component of strain tensor, dimensionless
$\dot{e}_{ij}$	component of strain rate tensor, sec^{-1}
F	heat flux, $\text{ergs/cm}^2\text{-sec}$
$\vec{F}_i$	applied load vector in finite element method, dynes
f	deviatoric boundary forces, dynes/cm^2

G	elastic shear modulus, dynes/cm^2
"	ALSO, group velocity of Rayleigh waves, km/sec
"	ALSO, the universal gravitational constant, $\frac{\text{dyne-cm}^2}{g^2}$
g	gravitational accelleration due to the Earth, cm/sec^2
H	heat production rate, ergs/cm^3-sec
h	applied boundary forces, dynes/cm^2
K	thermal conductivity, ergs/cm-sec-°C
$\underset{\sim}{K}_{ij}$	stiffness matrix in finite element method, $\frac{\text{dynes-sec}}{\text{cm}}$
k	Boltzmann constant, erg/°K-mole
k_i	constant of thermal history in (B.6), $\frac{°C}{sec^{3/2}}$
L	amplitude response function of filter, dimensionless
ℓ	various length dimensions, cm
M	bandwidth of stiffness matrix $\underset{\sim}{K}_{ij}$, integer
m	function determining stress from strain rate dynes/cm^2
N	number of variables in finite element approximation, integer
n	stress exponent in nonlinear rock flow law, dimensionless

P	pressure, dynes/cm^2
"	ALSO, period of wave, sec
P_{ij}	components of total stress tensor, dynes/cm^2
Q	heat energy density, ergs/cm^3
Q_a	activation energy, ergs/mole
Q_α	quality factor for seismic P-waves, dimensionless
Q_β	quality factor for seismic S-waves, dimensionless
Q_R	quality factor for seismic Rayleigh waves, dimensionless
q	a quantity of energy, ergs
R	radius of Mohr's circle, dynes/cm^2
r	function determining stress from strain rate, dynes/cm^2
S	scale distance for heat conduction, cm
s	distance down-dip in a shear zone, cm
T	temperature, °C or °K
ΔT	temperature change resulting from a specific mechanism °C or °K
T_o	reference temperature, °C or °K
T_m	melting temperature, °C or °K
t	time, sec
Δt	time interval, sec
u	velocity in the x-direction, cm/sec
V	relative velocity of adjacent plates, cm/sec or cm/year

V_a	activation volume, cm^3/mole
V_d	dislocation velocity, cm/sec
W	width of shear zone, cm
w	velocity in the z-direction, cm/sec
x	horizontal dimension, cm
z	depth from Earth surface, cm
α	thermal expansion coefficient, $°C^{-1}$
"	ALSO, P-wave velocity, km/sec
β	miscellaneous proportionality constants
"	ALSO, shear wave velocity, km/sec
γ	miscellaneous proportionality constants
Δ	distance between source and station, degrees
ϵ_i	non-equilibrium stress residual averages, $\frac{dynes}{cm^2}$
η	viscosity, poises
θ	angle of dip below horizontal of planar features, radians
κ	thermal diffusivity, cm^2/sec
λ	constant of Gaussian filter relative bandwidth, dimensionless
μ	coefficient of rock friction, dimensionless
$\vec{\nu}$	outward normal vector of a surface
ρ	rock densities, gm/cm^3
ρ_o	reference density at a particular depth, gm/cm^3

$\Delta\rho$	density anomalies, gm/cm^3
ρ_d	dislocation density, cm^{-2}
σ_{ij}	components of deviatoric stress tensor, $dynes/cm^2$
τ	shearing stress, $dynes/cm^2$
ϕ	stream function, cm^2/sec
Ω	plate driving force, dyne/cm
ω	frequency, Herz

Appendix B

Shear Zone Modelling Technique

In every shear zone the temperature, stress, and width are determined by stress and energy equilibria. These physical principles are the following:

$$\tau V = F_{up} + F_{down} + \frac{1}{2}\rho C_p V \frac{\partial}{\partial s}(W\Delta T) \qquad \text{(B.1)}$$

$$\frac{\partial \tau}{\partial W} = 0 \qquad \text{(B.2)}$$

$$\tau \leq A\left(\frac{V}{W}\right)^{1/n} \exp\left(\frac{B}{T_o + \Delta T}\right) \qquad \text{(B.3)}$$

$$T_o + \Delta T \leq T_m(z) \qquad \text{(B.4)}$$

$$\tau \leq \frac{\rho g z \sin(\arctan(\frac{1}{\mu}))}{\sqrt{1+1/\mu^2} - \cos(2\theta - \arctan(\frac{1}{\mu}))} \qquad \text{(B.5)}$$

where τ = shear stress on fault plane, dynes/cm^2

V = relative plate velocity, cm/sec

F = heat flux due to friction, ergs/cm^2-sec

ρ = density, g/cc

C_p = specific heat, ergs/gm-°C

s = distance from trench down-dip, cm

W = thickness of shear zone, cm

ΔT = temperature increase caused by friction, °C

A = constant for rock creep, (sec)$^{1/n}$-dynes/cm^2

B = constant derived from creep activation energy, °K

T_o = temperature without any frictional heating, °K

T_m = melting temperature, °K

g = gravitational accelleration, cm/sec^2

μ = coefficient of friction, nondimensional

θ = angle of dip of subduction zone, radians

n = empirical stress-activation exponent, dimensionless

z = depth, cm

Equation (B.1) expresses the conservation of energy by equating the work done by friction (eq. 2.9) to the additional heat flux in both directions and the additional heating of the shear zone which friction causes. Equation (B.2) is the variational form of the minimum-dissipation principle, stating that the shearing will adjust itself so as to maintain only the minimum possible stress. Equation (B.3) is an empirical formula (Goetze and Brace, 1972; Weertman and Weertman, 1975) for the nonlinear creep of minerals by intragranular dislocation mechanisms. Equation (B.4) states that temperature is not allowed to exceed the solidus, because melting is assumed to both reduce shear stress and increase heat transport by several orders of magnitude. Lastly, equation (B.5) is the frictional sliding or earthquake law which limits stress according to a simple Navier criterion. The trigonometric terms involving the angle of dip are derived from Mohr's circle in Appendix C.

The problem is posed by holding constant V, ρ , C_p, A, B, n, T_o, T_m, g, μ , and θ . Then s (or equivalently z) is taken as the independent variable, and at any given point we solve for W, ΔT, and τ . While there are five equations to satisfy, only (B.1) and (B.2) are always imposed. The other three are inequalities all tending to put an upper limit on stress, and so the most stringent requirement only is used at any particular depth.

Because terms on the right-hand side of (B.1) require a knowledge of conditions at lesser values of s, it is necessary to solve (B.1) through (B.5) for all depths from the surface down to $z=s \cdot \sin(\theta)$. In practice, the best that can be done is to solve for W, ΔT, and τ at a finite number of equally-spaced points, and assume that interpolation will give reasonable estimates for values of ΔT between solutions. In the shallow part of the subduction zone, equation (B.5) requires that τ increase linearly with s, so a linear variation of τ between solution points is the basic assumption at all depths.

In calculating the flux F_{down} into the slab, it is not necessary to consider any flux that is not due to frictional heating. It is also not necessary to consider the finite thickness of the slab if that thickness is more than several times the scale distance (eq. 2.7). Thus the problem reduces to one of finding the flux into a halfspace of zero initial temperature after a specified thermal history has been imposed on the surface. This thermal history is determined by the values of ΔT at solution points with lesser values of s than the one in question. From the reference point of the slab which is subducting, the assumed linear variation of frictional heating in s suggests an approximately linear variation of F_{down} with time. According to Carslaw and Jaeger (1959), the result of a linearly-varying flux into a half-space is a surface temperature variation proportional to the 3/2 power of time. Thus the consistent representation of the past thermal history of the slab surface is as a summation of such terms:

$$\Delta T\underset{\substack{\uparrow \\ \text{slab} \\ \text{frame}}}{(t)} = \Delta T\underset{\substack{\uparrow \\ \text{fixed} \\ \text{frame}}}{\left(\frac{s}{V}\right)} = \sum_{i=1}^{N} k_i \left[t + (1-i)\frac{\Delta s}{V}\right]^{3/2} \tag{B.6}$$

Here the N values of k_i are set to give the correct value of ΔT at lesser values of s:

$$k_i = \left[\Delta T_i - \sum_{j=1}^{i-1} (i-j+1)^{3/2} k_j\right] \left(\frac{\Delta s}{V}\right)^{-3/2} \tag{B.7}$$

Having expressed the thermal history at the surface as a sum of staggered and weighted time exponentials, we can obtain the surface flux from an analytical space differentiation of the individual temperature solutions (Carslaw and

Jaeger, p.63).

$$(F_{down})_i = K \sum_{j=1}^{i} \frac{\partial}{\partial z}\Bigg|_{z=0} \left\{ k_j \, \Gamma\left(\tfrac{5}{2}\right)\left[\frac{4\Delta s}{V}(1+i-j)\right]^{3/2} i^3\mathrm{erfc}\left[\frac{z/2}{\sqrt{\kappa\frac{\Delta s}{V}(1+i-j)}}\right]\right\}$$

$$= 1.329\, K \frac{\Delta s}{V\sqrt{\kappa}} \sum_{j=1}^{i} (1+i-j)k_j \qquad \text{(B.8)}$$

This treatment assumes that heat conduction into the slab is perpendicular to the fault, as in Figure 2.2. If this is not the case, the formula will be in error locally but will still give the correct total flux for the subduction zone as a whole.

The calculation of F_{up} into the overriding plate is simpler because the flux is constant with time. However, the effect of the surface boundary condition cannot be ignored. If a constant flux F is applied to a plate of thickness z_o and initial temperature zero, the temperature of the surface is given by (Carslaw and Jaeger, p. 113)

$$\Delta T(z_o, t) = \frac{F z_o}{K}\left[1 - \frac{8}{\pi^2}\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)^2} e^{\left[\frac{-(2n+1)^2 \kappa \pi^2 t}{4 z_o^2}\right]}\right]$$

(B.9)

As shown in Figure B.1 this temperature increase toward the asymptotic limit Fz/K almost exactly balances the loss of temperature due to the removal of frictional heating at the beginning of the finite difference calculation. Thus the value of t in equation (B.9) is the length of the transition period between oceanic and continental subduction. For a given depth and ΔT, we can now evaluate

$$(F_{up})_i = \frac{K \Delta T_i}{Z_i} \left[1 - \frac{8}{\pi^2} \sum_{n=0}^{\infty} \frac{(-1)^2}{(2n+1)^2} e^{\left[\frac{-(2n+1)^2 \kappa \pi^2 t}{4 Z_i^2}\right]} \right]^{-1} \cos\theta \tag{B.10}$$

The summation converges after a few terms. The factor of cos (θ) is to correct for the fact that the fault and free surface are not parallel. It has been assumed in using the plate analogy that this heat flows straight up. If it does not, then (as with F_{down}) the local estimate is in error but the global estimate is still good.

The final problem is to evaluate the internal heating term $\rho C_p \frac{\partial}{\partial s}(W \cdot \Delta T)$ from (B.1) when W changes from point to point. For this it is necessary to know the temperature of the material that is about to be warmed to ΔT_i and incorporated into the shear zone. The amount that this material has already been warmed by friction should in principle be calculated from the complete thermal history expressed in (B.6) and (B.7). However, this term is small, and it has been found that it can be evaluated to better than ±5 mW/m^2 accuracy using the approximation

$$\Delta T(s) \approx c\sqrt{s} \tag{B.11}$$

leading to

$$\frac{V}{2}\rho C_p \frac{\partial}{\partial s}(W\Delta T) \approx \rho C_p \left\{ W_{i-1}\left(\Delta T_i - \Delta T_{i-1}\right)\frac{V}{2\Delta S} + \right.$$

$$\left. \frac{3V}{4\Delta S}\Delta T_i \int_{W_{i-1}}^{W_i}\left[1-\left(e^{\frac{-Vr^2}{4\kappa S}} - \frac{r\sqrt{\pi V}}{2\sqrt{\kappa S}}\operatorname{erfc}\left(\frac{r\sqrt{V}}{2\sqrt{\kappa S}}\right)\right)\right]dr \right\}$$

(B.12)

The shear zone is assumed to expand at the expense of the subducting plate, on the assumption that its sedimentary or crustal upper layers will be weaker than the overriding lithosphere. Equation (B.12) evaluates the internal heating term by a backwards difference. While a centered difference would certainly be more accurate, it would require an iteration of the solution for each depth, which has not been attempted.

The values of ΔT and W at a particular depth are found by a two-dimensional search technique. A number of different values of W are investigated, in a geometric progression from the smallest to the largest allowed values. For each value of W, (B.1) is used to evaluate τ for different values of ΔT. Only one value of ΔT (at each width) gives a value of τ which also satisfies (B.3). The level of stress needed to achieve the melting temperature is also found. Finally, all possible widths of creeping fault zones are compared with the possibility of faulting (B.5) or melting (B.4) to find the mechanism which requires the least stress (B.2).

No great accuracy can be claimed for a technique using so many approximations. It is in essence only an optimized trial-and-error approach to the complex thermo-mechanical equilibrium problem in subduction zones. At worst, it still gives a good relative estimate of the effects of creep strength and plate velocity, as shown in Chapter 2.

Figure B.1 Temperature history of an arbitrary point in the overthrusting plate during a continental collision. Subduction of oceanic lithosphere occurs prior to time A. Subduction of the continental rise occurs from A to B, followed by subduction of the continental margin. The decay of former frictional heat is calculated by finite-differences. The effect of frictional heat after time t=A is given by equation (B.9). Final temperature could be either higher or lower than initial depending on parameters.

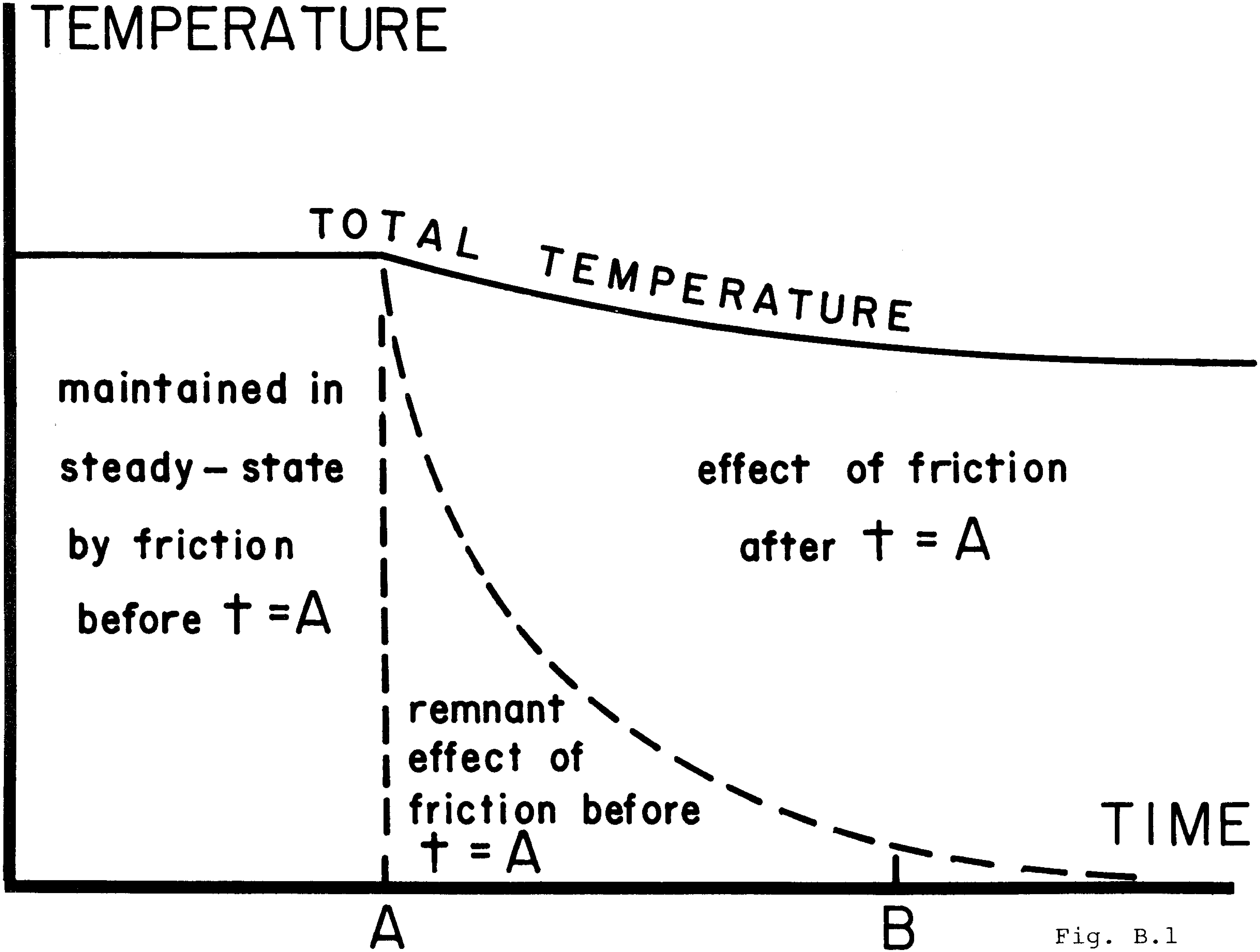

TEMPERATURE
TOTAL TEMPERATURE
maintained in steady-state by friction before t = A
effect of friction after t = A
remnant effect of friction before t = A
TIME
A
B

Fig. B.1

Appendix C

A plasticity criterion for faulting

Formulas in this appendix are relevant to cases of plain strain in a vertical plane, where the intermediate principal stress (σ_2) axis is horizontal. Given the coefficient of friction (μ) and the orientation of stresses and the weight of the overburden, either the difference of principal stresses or the shear stress on the plane of failure are found. The failure criterion assumed is that shear stress on some plane be equal to the coefficient of friction times the effective normal stress on that plane. Effective stresses are those from which the pore pressure has been subtracted. Rather than carrying this correction through the math, I have assumed that pore pressure is hydrostatic as in an open system. In that case the correction can be made by subtracting the fluid density from the rock density and calling the difference "effective density" (ρ).

In a Mohr diagram (Figure C.1) the state of stress at a point is represented by a circle with center at pressure P, equal to the mean normal stress

$$P = (\sigma_1 + \sigma_3)/2 \tag{C.1}$$

The radius (R) of the circle is equal to the maximum shearing stress on any plane.

$$R = (\sigma_1 - \sigma_3)/2 = \sigma_1 - P \tag{C.2}$$

Faulting occurs when the circle just touches the failure criterion at one point; and the radius to the point of tengency has the same direction as the normal to the fault plane.

Consider a case in which only the effective vertical stress is known:

$$\sigma_{zz} = \rho g z \tag{C.3}$$

where ρ is effective density, g is gravitational accelleration, and z is depth from the free surface. Define the angle γ as the angle between the vertical and the most compressive (σ_1) axis. This defines the position of horizontal and vertical directions on the Mohr plot, as in the Figure.

Using this information we can write

$$\rho g z = P + R\cos(2\gamma) \tag{C.4}$$

Also, P can be related to R by the trigonometric relation

$$P = \frac{R}{\sin(\arctan(\mu))} \tag{C.5}$$

Consequently,

$$R = \frac{\rho g z}{\cos(2\gamma) + \dfrac{1}{\sin(\arctan(\mu))}} \tag{C.6}$$

which can be slightly simplified to

$$(\sigma_1 - \sigma_3) = \frac{2\rho g z}{\cos(2\gamma) + \sqrt{1 + 1/\mu^2}} \tag{C.7}$$

This is the relation needed to correct stresses down to the plastic limit in the finite element program of Chapter 3.

A consequence of this equation is that the stress difference must be greater when γ approaches 90°. The function (C.7) is shown in Figure C.2 for different values of μ. This corresponds to the well-known principal in geology that thrust faults require a larger stress difference than normal faults. However, the effect of stress orientation disappears as μ approaches zero.

Another problem which arises is to determine the shear stress on a fault plane of known orientation if the frictional coefficient and σ_{zz} are known but principal stress directions are not. If the dip of the fault plane from the horizontal is θ, then we observe from Figure C.1 that the angle β is :

$$\beta = 2\theta - \arctan\left(\frac{1}{\mu}\right) \tag{C.8}$$

This helps to find that

$$2\gamma = 180^\circ - \beta = 180^\circ - 2\theta + \arctan\left(\frac{1}{\mu}\right)$$

(C.9)

Since a shift by 180° simply changes the sign of the cosine function, the new form of (A.3.7) becomes:

$$(\sigma_1 - \sigma_3) = \frac{2\rho g z}{\sqrt{1 + 1/\mu^2} - \cos\left(2\theta - \arctan\left(\frac{1}{\mu}\right)\right)}$$

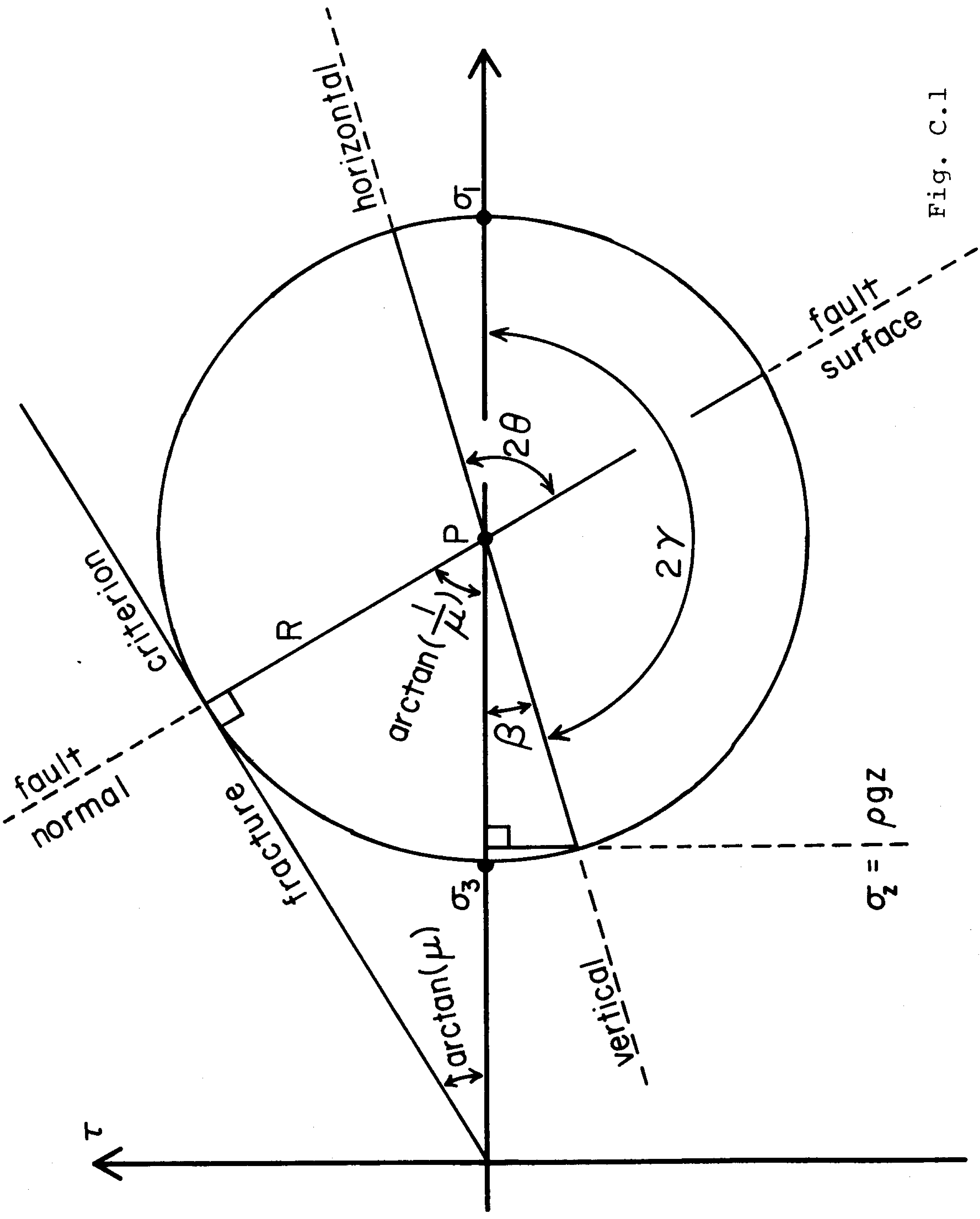
(C.10)

Since R is now known, it is easy to find the τ-coordinate of the point of tengency as

$$\tau = \frac{\rho g z \sin\left(\arctan\left(\frac{1}{\mu}\right)\right)}{\sqrt{1 + 1/\mu^2} - \cos\left(2\theta - \arctan\left(\frac{1}{\mu}\right)\right)}$$

(C.11)

This involves only the known depth, friction, and dip of the fault plane. It is identical to equation (B.5) which is used to evaluate shear-strain heating in a shear zone of known dip.



horizontal
vertical
fault surface
fault normal
fracture criterion
R
P
σ_1
σ_3
2θ
2γ
β
$\arctan(\frac{1}{\mu})$
$\arctan(\mu)$
$\sigma_z = \rho g z$
τ

Fig. C.1

$\left(\dfrac{\sigma_1 - \sigma_3}{2\rho g z}\right)$
1.0
0.8
0.6
0.4
0.2
$\mu = .8$
$\mu = .6$
$\mu = .4$
0
30
60
90
Dip of σ_1 Axis

Fig. C.2

Biographical Note

The author was born in [redacted] Massachusetts on [redacted] 1951. After several years of annual commuting to Houston, his family settled in Concord, Mass., where he attended public school. He studied geology at Harvard College, and met Jean Campbell, whom he married in March, 1972. He graduated in that year and moved to the Earth and Planetary Sciences department at M.I.T. After completion of this work, he will be at the University of California, Los Angeles.

Publications

- Bird, Peter, and J.D. Phillips (1975) Oblique spreading near the Oceanographer Fracture, Journal of Geophysical Research, 80, 4021-4027.
- Bird, Peter, M.N. Toksöz, and N.H. Sleep (1975) Thermal and mechanical models of continent-continent convergence zones, Journal of Geophysical Research, 80, 4405-4416.
- Toksöz, M. Nafi, and P. Bird, (1976) Temperatures in continental convergence zones, Tectonophysics, in press.
- Toksöz, M. N., and Peter Bird (1976) Formation and evolution of marginal basins and continental plateaus, in the Evolution of Island Arcs, Deep Sea Trenches, and Back-Arc Basins, American Geophysical Union, Washington, D.C.