12 May 1961, Volume 133, Number 3463 SCIENCE

CURRENT PROBLEMS IN RESEARCH

The Earth's Crust and Upper Mantle

New geophysical methods are beginning to reveal essential details about the earth's outer layers.

Frank Press

An international scientific effort called the "Upper Mantle Project" is being organized by the International Geophysics Committee to stimulate and coordinate studies of the outer layers of the earth. Such a program is particularly appropriate at this time. Some of the major features of the crust and upper mantle have been revealed, and the technology of geophysical exploration has advanced to the point where the essential details can be obtained. Until more is known about the crust and upper mantle, such basic geological problems as the origin of continents and ocean basins, the nature of orogeny and volcanism, and the mechanism of earthquakes will remain unsolved.

The principal tools for exploring the outermost few hundred kilometers are seismological and gravitational ones. Earthquakes and explosions are the sources of seismic energy, and portable as well as permanently installed seismological stations produce the data. Gravity exploration is carried out by means of sensitive gravimeters, which can now operate on sea and land with the precision required for this type of research (about $10^{-6}g$). These methods and the results which have been obtained to date are described here. Some unsolved problems and the methods for attacking them are then discussed.

Seismological Methods and Results

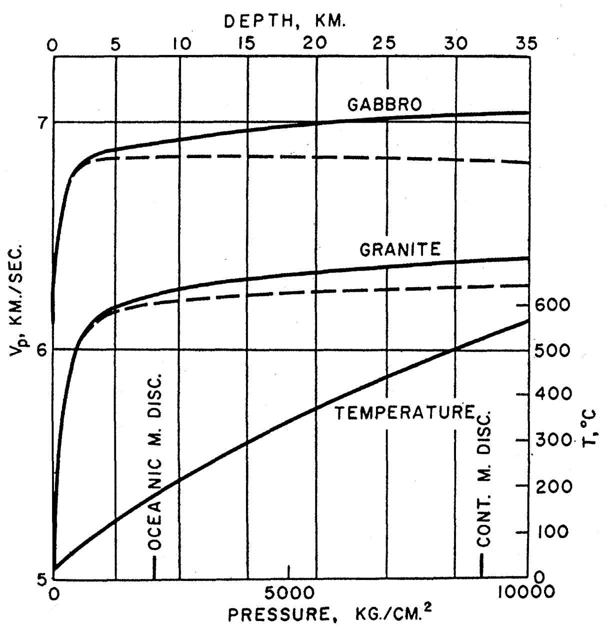
At best, seismological methods reveal the mechanism at the earthquake focus and the variation of elastic velocities with depth. It is only by comparison with laboratory measurements, on rocks, of elastic parameters as functions of temperature and pressure that we can infer which rocks occur at depth. Laboratory results for compressional velocity in granite and gabbro under conditions of pressure and temperature likely to occur in the crust (*1*) are shown in Fig. 1.

Almost 50 years have passed since Mohorovičić first demonstrated that the structure of the earth's outer layers can be deduced from the travel times of refracted seismic waves. His method may be understood from Fig. 2. A source of elastic waves (earthquakes in the early work, explosions in modern methods) radiates energy in all directions. Among the many possible paths available to the waves are three shown in the schematic model of the crust. Waves traversing these paths are characterized by straight-line travel-time curves whose slopes are the reciprocal of the velocities α_i in the corresponding layers and whose intercepts T_i provide the additional information needed to compute layer thicknesses. Appropriate formulas and methods for dealing with complications such as dipping interfaces and velocity inversions may be found in standard textbooks of geophysics.

Field methods consist in recording and timing seismic waves along lines extending from the source, with a sufficient number of observations to determine the travel-time curves. The technique for oceanic investigations differs in that ships are used to detonate explosions and record the sound waves. Either radar or the travel time of sound in water is used to determine the separation between the detonating and the receiving ship, the instant of explosion being transmitted by radio. It is common for both ships to "shoot" and "record," so that the reversed profiles necessary to determine the dip of interfaces can be efficiently obtained. Whereas several hundred to several thousand pounds of explosives are needed for land observations, oceanic observations can be achieved with charges of 300 pounds or less. The "breakthrough" which made oceanic studies feasible was the demonstration by M. Ewing and his co-workers that near-surface explosions and pressure-sensitive detectors are adequate for recording subcrustal refractions.

As a result of the efforts of many investigators in different countries it is possible to make some statements about crustal structure under the continental shields and the ocean basins (Fig. 3). The Mohorovičić discontinuity, which separates the light, silicic crust from the heavier, ultramafic mantle, occurs at a depth of about 35 kilometers under the continental shield areas and about 10 kilometers below sea level in the ocean basins. This fact, together with the known increase in Bouguer gravity from continent to ocean, conclusively demonstrates isostasy as a basic tectonic mechanism on the continental scale (*2*). The continental crustal layer below surficial sediments and uncom-

The author is professor of geophysics and director of the Seismological Laboratory at California Institute of Technology, Pasadena.



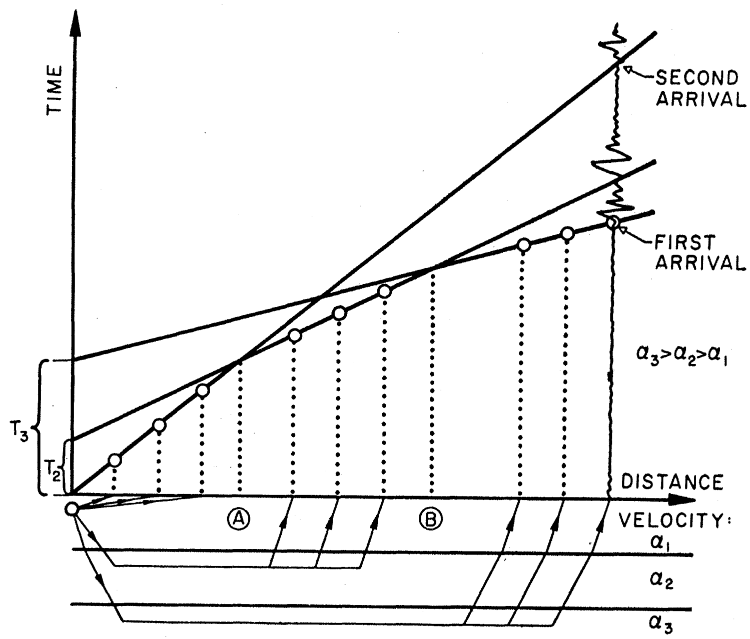


Fig. 1 (left). Changes, with increases in pressure and temperature, in compressional wave velocity in rocks in the continental crust. Solid line, effect of pressure only; dashed line, effects of pressure and temperature combined. Depth and pressure at Mohorovičić discontinuity (M. disc) under oceans and continents is indicated. [After Birch (*1*)] Fig. 2 (right). Schematic diagram of seismic refraction paths in a layered crust underlain by the mantle.

pacted crystalline rocks is characterized by compressional and shear velocities of 6.0 to 6.2 kilometers per second, and of 3.5 to 3.6 kilometers per second in the upper part. These velocities are consistent with findings for granite under pressures equivalent to those found at depths of several kilometers (*1*).

The crustal layer beneath the ocean basins consists of unconsolidated deep-sea sediments a few hundred meters thick, underlain by a layer of mafic rock, with compressional velocity of 6.5 to 7 kilometers per second. Some investigators report a thin layer of compacted sediments or volcanics between the unconsolidated sediments and the mafic rocks. Detailed information about layering in the sediments must await measurement with detectors placed on the sea floor.

The mantle beneath continents and ocean basins is characterized uniformly by compressional velocities of 8 to 8.2 kilometers per second. Shear velocity under continents is 4.6 to 4.8 kilometers per second. No information on shear velocity is available for the suboceanic mantle. These velocities are consistent with a mantle composed of rocks similar to dunite (*1*).

The older notion that elastic velocity increases with depth below the Mohorovičić discontinuity has been abandoned. Gutenberg's observation (*3*) of a shadow zone for earthquake-generated elastic waves led him to postulate the existence of a "low"-velocity layer. Amplitudes of seismic waves from earthquakes show much variability, and the more precise amplitude data obtainable from nuclear explosions (*4*) were required to substantiate the existence of the Gutenberg shadow zone and low velocity region. Figure 4, modified from a figure in Romney's paper (*4*), shows the amplitude variation of compressional waves from the underground nuclear explosions Blanca and Logan. The rapid decrease in amplitude with distance to 1500 kilometers indicates that seismic energy enters this zone by diffraction and not by following geometric paths—a consequence of the decrease in velocity below the Mohorovičić discontinuity. As shown below, surface-wave data offer independent verification and enable us to specify some details of the low-velocity zone. It is of interest that the shadow zone poses a major difficulty in the detection of small nuclear explosions.

Now that these basic features of the crust and upper mantle are understood, geophysicists are turning their attention to the finer details, which are also the crucial details of crustal structure. It is important to know about variations of elastic parameters with depth in the crust. Does the Conrad discontinuity, which separates the crust into two layers, exist? And does it occur uniformly under the continents? Isostasy on a continental scale has been demonstrated. What is the mechanism of compensation of regional topographic highs? How does the continental crust merge with the oceanic crust? Is the contact between crust and mantle an abrupt discontinuity, or is there a zone of gra-

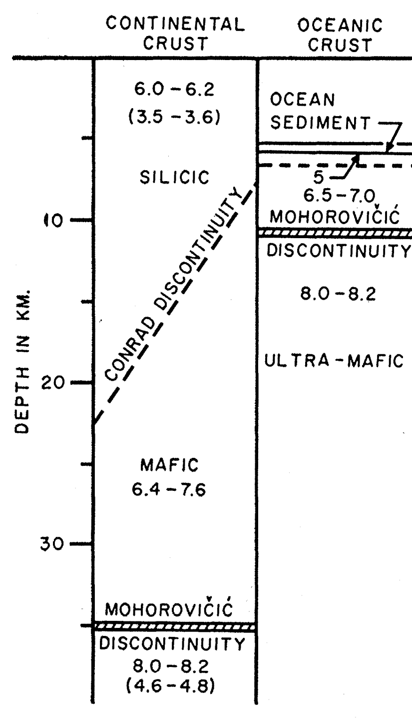


Fig. 3. Seismic indications of crustal layering in continental shields and ocean basins. Compressional and shear velocities (in kilometers per second) are shown. Shear velocities are in parentheses.

dation? Is the mantle a more dense phase of crustal rock, or does the observed difference represent a change in composition? What is the significance of the low-velocity zone in the upper mantle? Does it occur uniformly under continents and oceans? These are some of the basic questions currently faced by geophysicists—questions which can be answered only through refinement of old methods and use of new methods. A review of some of the newer techniques and the results which have been achieved follows.

Correlation Refraction Shooting

The intermediate layer depicted in Fig. 2 is inherently difficult to observe. Detectors located in the limited range between positions *A* and *B* record waves refracted from this layer as "first arrivals." A slight increase in T_2 with respect to T_3 is sufficient to "mask" this layer so that waves refracted from it occur only as "second and third arrivals." It is difficult to identify refracted waves occurring as later arrivals because of the reverberation initiated by the first arrivals. Fortunately, there is often sufficient coherence and character in later-arriving refracted waves that these waves can be identified by correlation across an array of detectors extending along a line extending 1 or 2 kilometers (Fig. 5). This technique of correlation, familiar to petroleum geophysicists, has been successfully applied in crustal investigations by Gamburtsev and his colleagues (*5*), and is called the DSS method. It is to the credit of these workers that they have developed a seismic detection system of high sensitivity and of frequency response such that relatively small explosions are sufficient to complete a profile. Veytsman, Kosminskaya, and Riznichenko (*5*) report that both the Conrad and the Mohorovičić discontinuity have been mapped by this method in regions of complex structure. They report, for example, that the crustal thickness increases to as much as 53 kilometers under the northern Tien Shan Mountains. Compensation for high topography occurs by a mechanism which involves increased depths of the Mohorovičić discontinuity, the Conrad discontinuity behaving in a complicated fashion, sometimes discordant with the former. Typically, however, the intermediate (basaltic) layer is thicker under uplifted areas.

In the United States, correlation methods have been applied by Woollard, Meyer, and Steinhart (*6*), and by Aldrich and Tuve. In the midcontinent region these investigators found normal crustal thickness (35 to 40 kilometers), with evidence which favored the presence of the Conrad discontinuity at a depth of about 15 kilometers. On the eastern flank of the Rocky Mountains a crustal thickness of 50 to 55 kilometers was found, with evidence of three crustal layers with velocities (in kilometers per second) and thicknesses (in kilometers) as follows: (i) velocity, 6.08; thickness, 15 to 20; (ii) velocity, 6.97; thickness, 17; and (iii) velocity, 7.58; thickness, 10 to 20.

It is still too early to generalize from these newer results. Some common features are apparent, but differences occur which may be unreal and related to differences in method of interpretation. The presence of the intermediate (basaltic) layer is widely accepted, although the velocity in this layer shows great and possibly real variability (from 6.4 to 7.6 km/sec). Crustal thickening under major uplifts seems to occur generally. Exceptions which have been reported may be due to masking of the intermediate layer or to mistaking the Conrad for the Mohorovičić discontinuity. When the intermediate layer occurs with velocities of 7.6 kilometers per second the question arises as to which interface separates the crust from the mantle. When more correlated refraction experiments have been made and when common interpretive procedures have been established under the aegis of the Upper Mantle Project, then more meaningful evaluation and generalization will be possible.

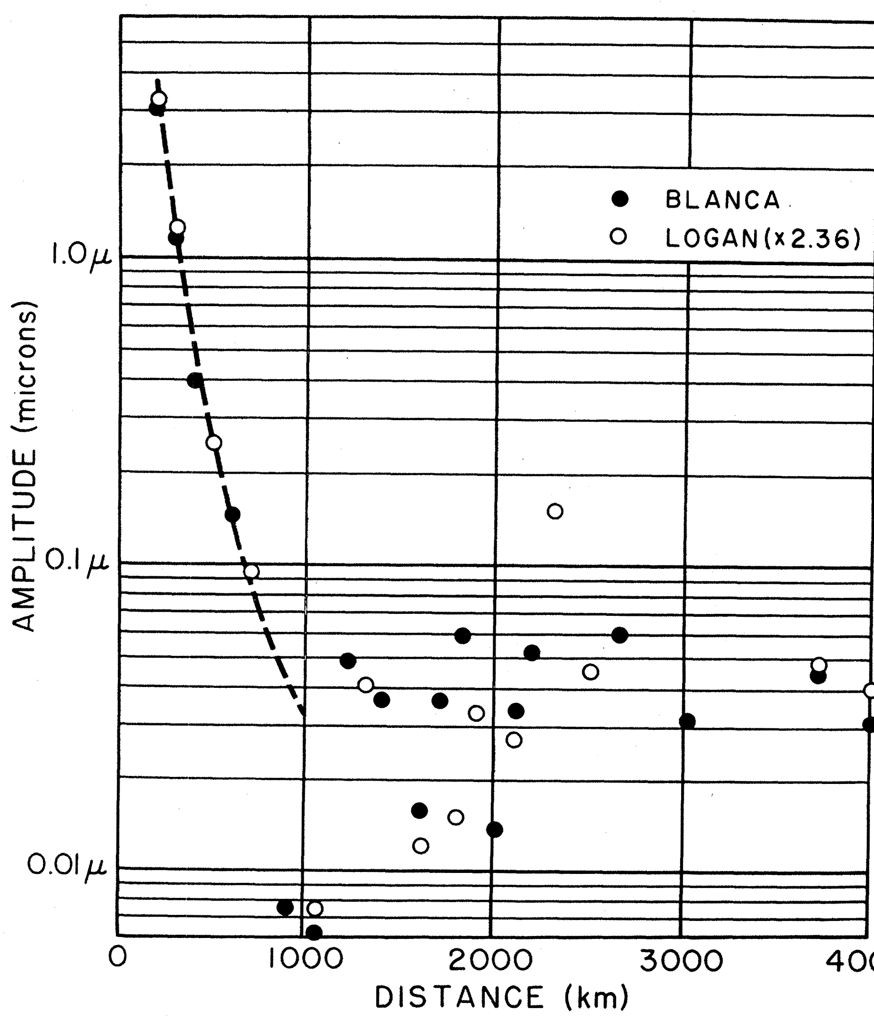


Fig. 4. Amplitude of compressional waves from underground nuclear explosions in Nevada. The rapid decrease in amplitude between 300 and 1500 kilometers defines the shadow zone. [After Romney (*4*)]

Surface Wave Methods

Propagation of seismic surface waves over the earth occurs in the wave guide formed by the free surface, the ocean (for oceanic paths), the crustal layers, and the velocity gradients in the mantle. The wave guide is dispersive because the longer the wave, the greater the depth of penetration. Thus, the initial seismic transient undergoes phase and amplitude distortion during propagation to form a long train of waves. With the advent of high-speed computers and sensitive, long-period seismographs, much progress has been made in deducing the properties of the wave guide (that is, the structure of crust and mantle) from an analysis of the phase distortion.

Two methods of analysis are available. The variation of group velocity with frequency can be obtained from seismograms, if the time of origin and the location of the seismic source are known. Group-velocity dispersion is controlled by the structure of the crust and mantle between epicenter and seismograph. Phase velocity is obtained from an array of at least three seismographs and is sensitive to structure beneath the array. Only relative times at each station are required to derive phase velocity. Figure 6 shows three seismograms (aligned with respect to time) from a triangular array of California stations approximately 100 kilometers apart. These are records of Rayleigh waves from a distant Pacific Ocean earthquake. Note the increase in frequency of the waves with increasing time—an increase which is controlled by the group velocity dispersion. The time required for a given crest or trough to cross the array depends on the phase velocity.

Figure 7 shows the observed curve for group-velocity dispersion of Rayleigh waves for the period range 10 to 400 seconds (*7*). Between 10 and 100 seconds, two curves are observed, one for the oceanic crust, the other for the continental crust. Beyond 100 seconds the two curves merge into a single curve, indicating that for these more deeply penetrating, longer waves the upper mantle controls the dispersion. Observed curves for dispersion are interpreted by computing theoretical curves for various models and selecting

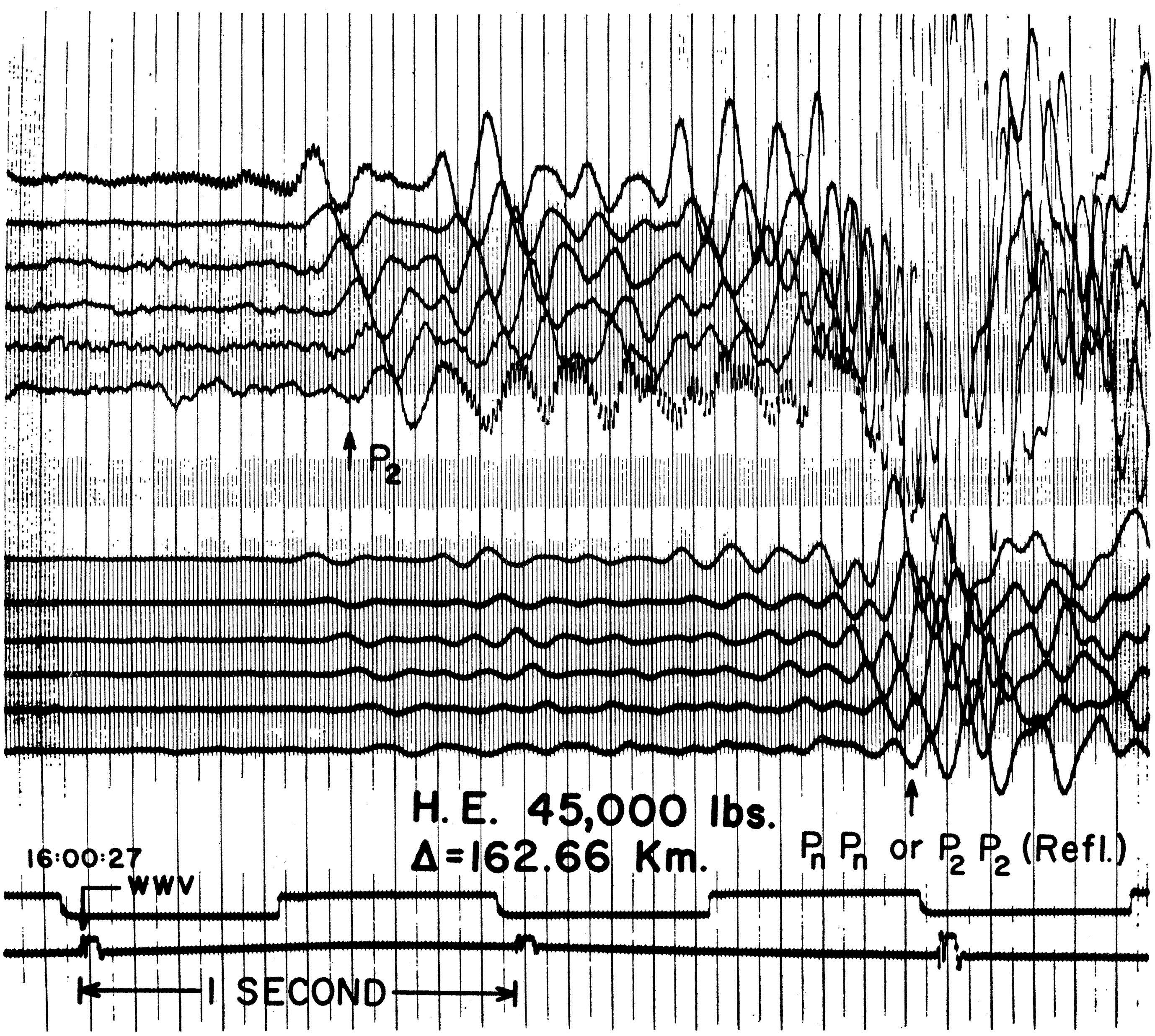


Fig. 5. Arrivals of refracted waves at a distance of 162 kilometers from the explosion. The lower six traces are a version of the upper six made with techniques of lower sensitivity. The detectors were separated by 1000 feet.

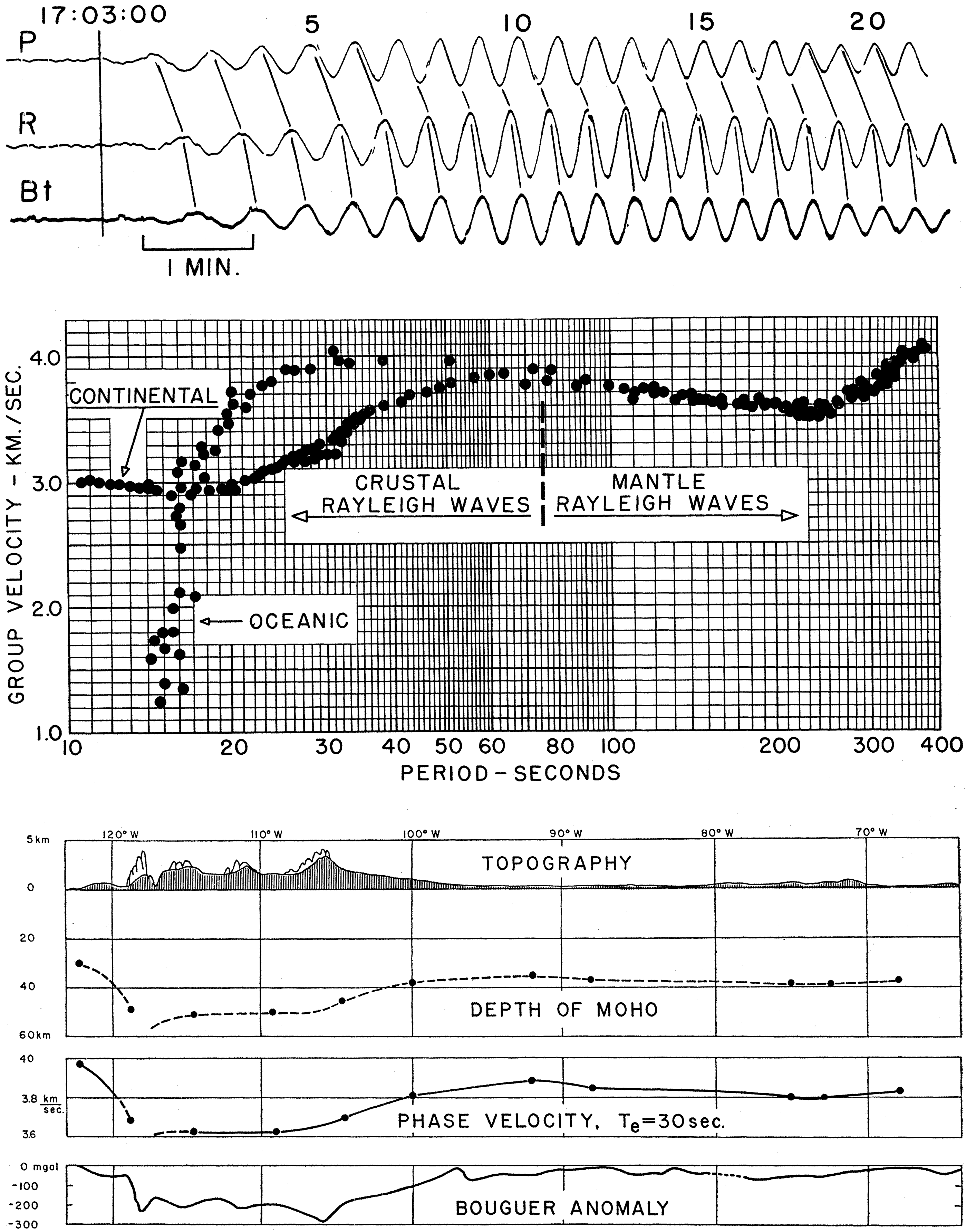


Fig. 6 (top). Rayleigh waves from an earthquake in the Solomon Islands, recorded by a triangular array of seismographs at Pasadena, Riverside, and Barrett, California, respectively. The seismograms are aligned with respect to time, and corresponding crests are indicated. Fig. 7 (middle). Observed dispersion of Rayleigh waves in the period range 10 to 400 seconds. [After Ewing and Press (*7*)] Fig. 8 (bottom). Profile across the United States, showing the relation of Rayleigh-wave phase velocity to topography and Bouquer anomaly. Depths to the Mohorovičić discontinuity are inferred from the phase velocity. [After Ewing and Press (*9*)]

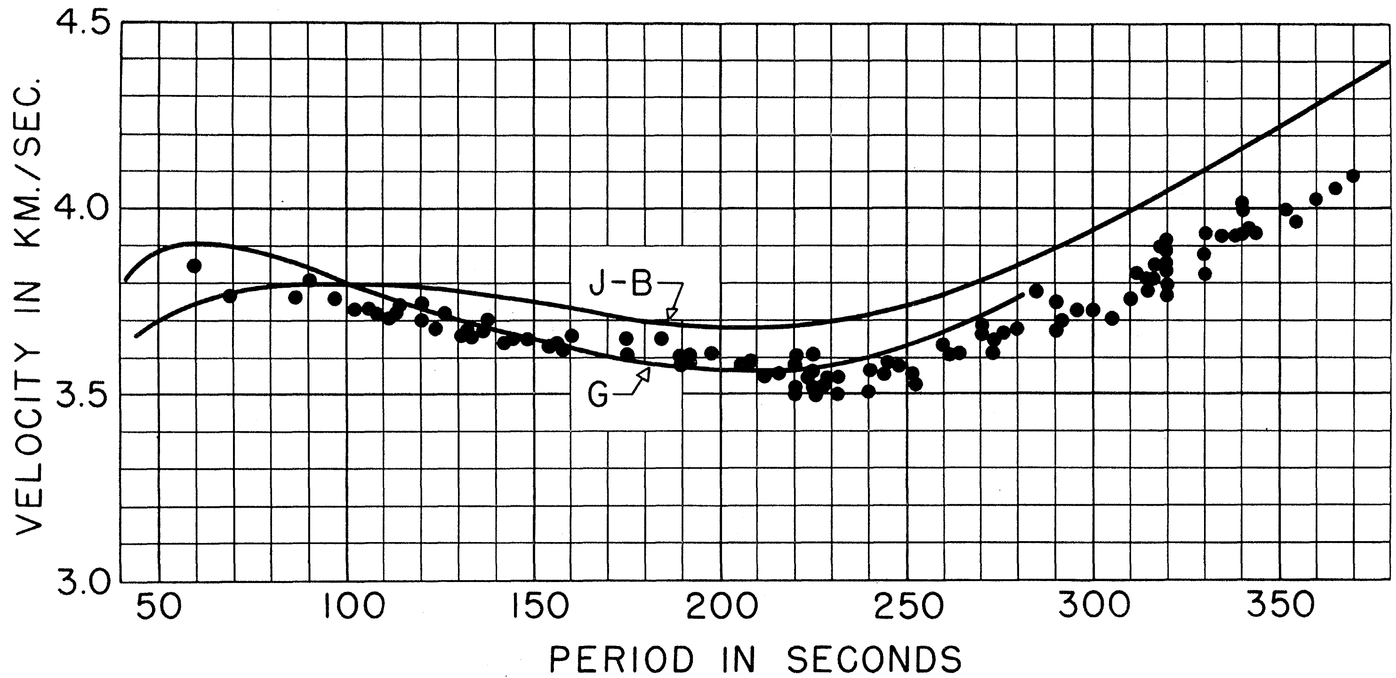


Fig. 9. Theoretical dispersion curves, showing that the curve based on the Gutenberg model of the mantle (which includes the low-velocity zone) fits the experimental data better than that based on the Jeffreys-Bullen model. The Jeffreys-Bullen model does not contain velocity inversion. [After Dorman *et al.* (*13*)]

that model which best fits the data and which is consistent with auxiliary information such as is available from refraction shooting.

A number of studies by these methods have recently been reported (*8*). In general, there is agreement that surface-wave data require an increase in velocity with depth in the continental crust which is consistent with the presence of the Conrad discontinuity. The continental crust thickens under topographically high regions and thins at continental margins in a manner consistent with regional isostatic compensation. Figure 8 shows a transcontinental profile with a remarkable correlation between phase velocity and gravity anomaly (*9, 10*). Surface-wave results are in better agreement with gravity anomalies in revealing isostatic compensation than are the data from explosion seismology. This may follow from the greater influence of local heterogeneities on the waves of short wavelength generated by explosions and from the regional "averaging" effect of the surface waves of longer wavelength. The masking of refracted waves from intermediate layers is probably also a factor in this discrepancy.

The combined use of seismic refraction and surface-wave methods, together with gravity methods, offers a powerful tool with greater potential than any single method for arriving at a unique solution in regions of complicated structure (*11*).

Surface-wave results for the upper mantle are currently of great interest. Dispersion of mantle Rayleigh waves with periods from 100 to 400 seconds has been cited to demonstrate that the low-velocity zone of the upper mantle is a universal feature under continents and oceans, the minimum velocity occurring at a depth of about 140 kilometers (*12*). Figure 9, taken from the work of Dorman, Ewing, and Oliver (*12*), shows that the theoretical curve based on Gutenberg's model of the mantle fits the dispersion data better than the curve based on the Jeffreys-Bullen model. The Gutenberg model includes the low-velocity zone in the upper mantle, whereas the Jeffreys-Bullen model shows a continuous increase in velocity with depth. Dorman *et al.* (*13*) made the important discovery that the upper mantle differs under continents and under the Pacific Ocean. They favored the interpretation that the low-velocity zone is shallower under the Pacific Ocean. Aki and Press (*14*) reported that the Pacific mantle differs from the continental mantle in a way that would be consistent with the theory that a low-velocity zone occurs at the same depth under ocean and continent but that the velocity under the ocean is lower. They further presented evidence of differences between the mantle under the Pacific Ocean and under the Atlantic and Indian oceans—differences which could be explained by a reduction in shear velocity at the top of the mantle under the Atlantic and Indian oceans.

Thus, for the first time, data are available which show that the mantle is not a symmetrical shell and that the oceanic and continental structures extend to depths of several hundred kilometers beneath the Mohorovičić dis-

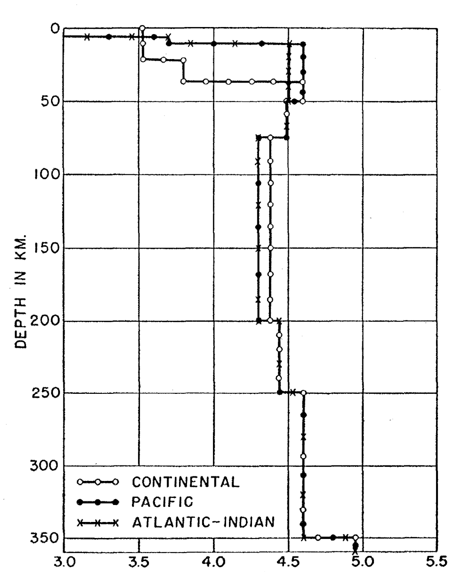


Fig. 10. Velocity-depth functions consistent with dispersion data for Pacific Ocean, Atlantic and Indian oceans, and continental paths. [After Aki and Press (*14*)]

continuity. Velocity-depth structure for the mantle under the Pacific Ocean, the Atlantic and Indian oceans, and the continents that is consistent with dispersion data is shown in Fig. 10. The layered structure is artificial, resulting from coding of the data for computation with a digital computer; the velocity-depth function is probably continuous. These are not unique solutions, but the differences between the curves represent real differences in the mantle of the type required by the data.

These results have important implications with respect to the genesis of continents and to continental drift. A complete physical explanation of the low-velocity zone has yet to be given. It is not unreasonable to expect that the velocity reversal is associated with a "softening" of rock due to the greater influence of temperature than of pressure at these depths. The low velocity could signify that the rocks at these depths are near the melting point. Verhoogen (*15*), among others, has pointed out that at depths between 100 and 200 kilometers the temperature comes close enough to that required for partial melting of basalt from a meteoritic material to make this a reasonable possibility. To speculate even further, the low-velocity zone may coincide, not only with the source of the primary basaltic magma, but also with the level of reduced strength, along which movements due to isostatic adjustment and polar wandering occur.

Gravity Observations

Bouguer gravity anomalies reveal an excess or deficiency of mass at depth, according to whether the sign of the anomaly is positive or negative. Since major density differences occur at the Conrad and Mohorovičić discontinuities, it is reasonable to expect that the anomaly may be explained in terms of variations in the depths of these interfaces, typically of the Mohorovičić. Geodesists have long recognized isostasy as a fundamental mechanism, from a study of gravity anomalies and their topographic association (*2*). Woollard (*16*) has recently reviewed the subject, making use of new seismic data. He finds that the relationship of Bouguer gravity anomaly with elevation (Fig. 11) and with crustal thickness (Fig. 12) is firmly established and reinforces the concept of isostatic compensation on a regional basis. A further conclusion can be reached from the form of the curve in Fig. 12—namely, that as the crust becomes thicker, its mean density increases. This is consistent with seismic observations that increased depth to the Mohorovičić discontinuity is usually associated with a thickening of the more dense, intermediate, basaltic layer rather than a thickening of the less dense, silicic top layer.

Isostatic compensation is regional

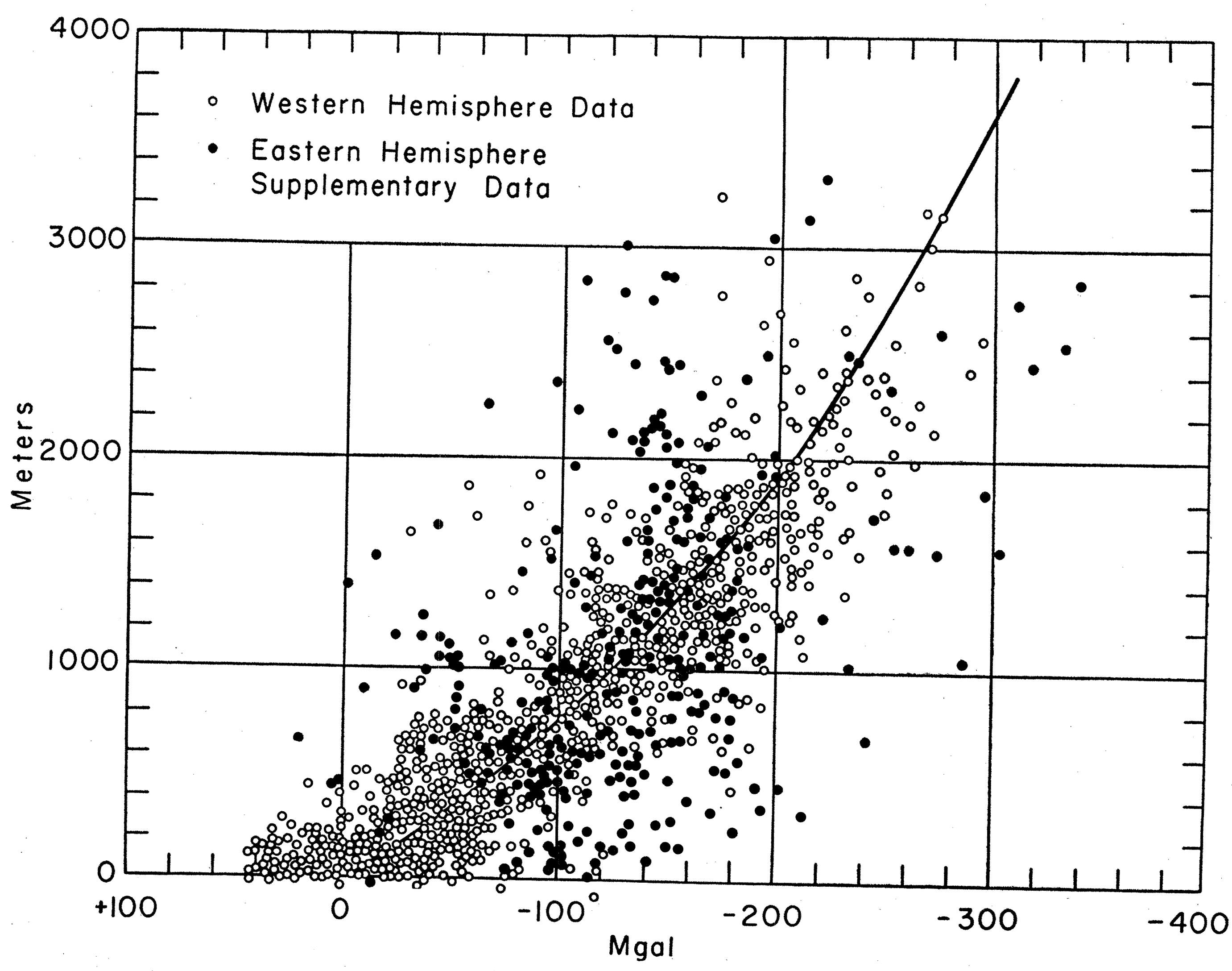


Fig. 11. Relation between the Bouguer gravity anomaly and topographic elevation. [After Woollard (*16*)]

rather than local. This can be demonstrated, for example, by noting that individual mountains in an uplifted region show little correlation between local topography and gravity (the main scarp of the Sierra Nevada Mountains is not evidenced in the gravity profile across this feature, for example).

It is noteworthy that the major release of tectonic energy in the form of earthquakes is associated with deep-sea trench-island arc structures in which the largest isostatic unbalance is known to occur. A recent result of great interest (*17*) demonstrates that a sharp downward flexure of the Mohorovičić discontinuity occurs at the seaward side of the trench. The stresses needed to maintain this texture in a state of isostatic unbalance are probably those also responsible for the seismic activity.

Remarks

A schematic summary of the status of our knowledge of the crust and upper mantle is shown in Fig. 13. The continental crust (*D*) shows a change in elastic parameters with depth in a way that would be consistent with the occurrence of light silicic rock (granite) at the top and more dense mafic rock (gabbro) at the bottom (*1*). In many places the contact is sharp and defines the Conrad discontinuity. The silicic layer shows less variability in velocity (*18*) than the mafic layer. The crust thins at the continental margin (*C*); it thickens under uplifts, mainly because of the increased thickness of the mafic layer (*E*). The oceanic crust (*B*) consists of rocks that correspond in density and elastic velocity with rocks found at the top of the continental mafic layer.

Island arcs and deep-sea trenches are associated with crustal thickening, the downward flexure in the Mohorovičić discontinuity occurring at the outer margin of the trench. The trench is isostatically uncompensated and tectonically very active, as evidenced by the high seismicity.

The velocity inversion in the upper mantle produces a minimum velocity at a depth of about 140 kilometers. The low-velocity zone of the upper mantle differs somewhat under continents and oceans, in a way that would be consistent with proposed higher temperatures (and "softer" rocks) under the oceans than under the continents. There is evidence which suggests that the top of the mantle has lower shear velocity (higher temperature?) under the Atlantic and Indian oceans than it has under the Pacific. The low-velocity zone is probably a result of temperature. Rocks near the melting point can occur in this zone, which may be the source of the primary basaltic magma.

The possibility that the Conrad and Mohorovičić discontinuities represent changes in phase rather than in composition has been much discussed. None of the measurements reported here enable us to take a firm position on this hypothesis. Plausible arguments can be made on both sides of the question; it will probably be resolved only by drilling through the oceanic crust to the top of the mantle. (*19*).

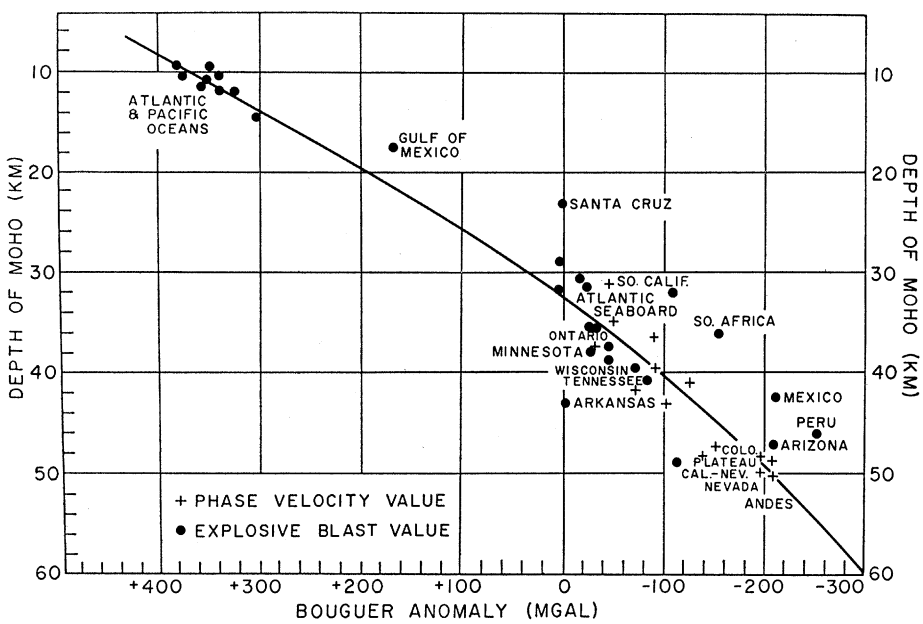


Fig. 12. Relation between the Bouguer gravity anomaly and thickness of the crust. Circles, points plotted from seismic refraction data; crosses, points plotted from data on phase velocity of surface waves. [After Woollard (*16*)]

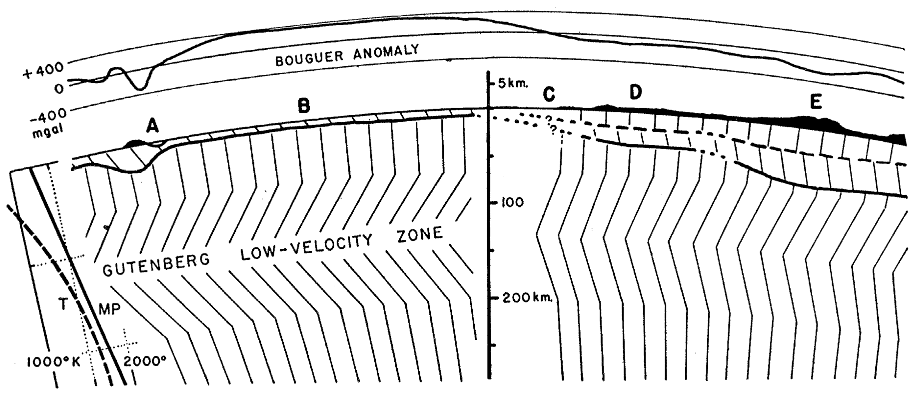


Fig. 13. Schematic diagram of the structure of crust and mantle. Solid curve (at left), melting point; dashed curve (at left), temperature.

References and Notes

1. F. Birch, *Contributions in Geophysics* (Pergamon, New York, 1958), p. 158; *J. Geophys. Research* **65**, 1083 (1960).
2. For an extensive analysis of gravity findings bearing on isostasy, see W. A. Heiskanen and F. A. Vening Meinesz, *The Earth and Its Gravity Field* (McGraw-Hill, New York, 1958).
3. B. Gutenberg, *Bull. Seismol. Soc. Am.* **38**, 121 (1948); *Bull. Geol. Soc. Am.* **65**, 337 (1954); *Science* **131**, 959 (1960).
4. C. Romney, *J. Geophys. Research* **64**, 1489 (1959).
5. For a summary, see Y. V. Riznichenko, *Bull. Acad. Sci. U.S.S.R., Geophys. Ser.* **2**, 1 (1957); P. S. Veytsman, I. P. Kosminskaya, Y. V. Riznichenko, "Report 11th Assembly, International Union of Geodesy and Geophysics" (Academy of Sciences, U.S.S.R., Moscow, 1957), p. 34.
6. For an excellent summary of older work as well as a presentation of new data, see the doctoral thesis of J. S. Steinhart, University of Wisconsin (under the direction of G. P. Woollard) (1960).
7. M. Ewing and F. Press, *Trans. Am. Geophys. Union* **37**, 213 (1956); *Handbuch der Physik* (Springer, Berlin, 1956), vol. 47, pp. 119, 246.
8. A partial list follows: R. L. Kovach, *J. Geophys. Research* **64**, 805 (1959); M. Bath and A. Vogel, *Geofis. pura e appl.* **38**, 10 (1957); F. Press, M. Ewing, J. Oliver, *Bull. Seismol. Soc. Am.* **46**, 97 (1956); J. Oliver and M. Ewing, *ibid.* **48**, 33 (1958); F. Press, *Bull. Geol. Soc. Am.* **67**, 1647 (1956); E. F. Savarensky and D. I. Sikharulidze, *Izvest.*

Akad. Nauk S.S.S.R. Ser. Geofiz. **6**, 620 (1959) (English edition).
9. M. Ewing and F. Press, *Bull. Geol. Soc. Am.* **70**, 229 (1960).
10. For Soviet work with the phase velocity method, see E. F. Savarensky and S. S. Ragemon, *Izvest. Akad. Nauk S.S.S.R. Ser. Geofiz.* **12**, 866 (1958) (English edition).
11. F. Press, *J. Geophys. Research* **65**, 1039 (1960); J. Oliver, R. Kovach, J. Dorman, *ibid.* **66**, 215 (1961).
12. This was shown in an approximate calculation by H. Takeuchi, F. Press, and N. Kobayashi [*Bull. Seismol. Soc. Am.* **49**, 355 (1959)]. It was elegantly and conclusively demonstrated by J. Dorman, M. Ewing, and J. Oliver (*13*).
13. J. Dorman, M. Ewing, J. Oliver, *Bull. Seismol. Soc. Am.* **50**, 87 (1960).
14. K. Aki and F. Press, *Geophysical J.*, in press.
15. J. Verhoogen, *Am. Scientist* **48**, 134 (1960).
16. G. P. Woollard, *J. Geophys. Research* **64**, 1521 (1959).
17. M. Talwani, G. Sutton, J. Worzel, *ibid.* **64**, 1545 (1959).
18. This conclusion is based not only on the uniformity of compressional velocities for the upper crust but also on the remarkable constancy in all continents of the velocity of short-period surface shear waves (L_g). See, for example, F. Press, *Trans. Am. Geophys. Union* **37**, 615 (1956).
19. This article is contribution No. 1024 of the Division of Geological Sciences, California Institute of Technology, Pasadena.

Patents and Inventive Effort

The evidence is insufficient to prove or disprove the claim that patent protection promotes inventive effort.

Fritz Machlup

A session of the December 1960 meeting of the AAAS was devoted to "The patent system and the advancement of knowledge." This issue, discussed for centuries, has never been resolved. Quotations from the U.S. Constitution, which empowered the Congress to establish patent and copyright laws "to promote the progress of science and the useful arts," cannot settle the question whether such laws actually serve this purpose. That patent protection may induce investment in further development and commercial application of new inventions is more readily conceded than that it is effective in inducing inventive activity where the corporate form of industry prevails. Doubts concerning this function of patents have been expressed with increasing frequency since the government contribution to research and development came to exceed half the total outlay. In 1959 private industry paid for less than 38 percent of total research and development in the United States. How important, then, can patent protection be in inducing inventive activity? Let us examine the arguments and sift the evidence that have been presented to answer this question.

The author is professor of economics at Princeton University, Princeton, N.J. This article is adapted from a section of a book on the production and distribution of knowledge in the United States, which is to be published this summer.

Large Corporations and Employed Inventors

Patent protection is supposed to serve as an incentive to invest in inventive work or to invest in development and plant construction or to disclose inventions that have been made. No matter which of these purposes are stressed, it is widely held nowadays that patents are not really important as incentives for large corporations, but only for independent inventors or for small firms competing with large ones. This view, strangely enough, is most emphatically stated by representatives of large corporations. Statements of this sort can be found in almost all Congressional hearings on patent legislation of the last 25 years.

If this contention is true, and we have no reason to doubt it, we are faced with the odd situation that patents as incentives for socially desirable activities are unnecessary for those who own the bulk of all patents. In the United States about 60 percent of all patents are assigned to corporations before issuance, which ordinarily indicates that the patented inventions were made by inventors employed by these corporations. Of all patents owned by corporations conducting research and development in 1953, 51 percent were owned by firms with more than 5000 employees, 30 percent by firms with between 1000 and 5000 employees, and only 19 percent by firms with less than 1000 employees. Thus it appears that those who hold most of the patents, the large corporations, testify that the patent system is not necessary for them, but only for those who hold the smallest number of patents.

In reply to the question whether patents are essential to the continuance of large expenditures for research and development, an officer of a large company stated that he might cut down these expenditures to perhaps one-half of the amount spent at that time if patent protection were removed. It happened, however, that approximately one-half of the research and development budget of that company was then devoted to the tasks of securing patents and enforcing the exclusive rights which they were supposed to confer. Hence, if the company were suddenly relieved of the necessity of spending money on obtaining patent rights and litigating about them, the remaining half of its budget would still buy the same amount of genuine research and development work. Most officers of large patent-holding corporations—except those in the chemical industry—do not think that their research expenditures depend on patent protection. For example, Robert E. Wilson, petroleum researcher and oil company executive, speculating about the possibly adverse consequences of a "weakening of the patent system," contended that this would *least* affect the research policies of *large* companies (*1*).

This judgment can be supported by deduction from the theory of oligopolistic competition: no firm in competition with a few others can afford to let its rivals steal a march upon it as far as the technological base of its competitive position is concerned. The research and development work is essential for the maintenance of its position. It cannot allow itself to fall seriously behind in the technological race, regardless of whether inventions promise it a 17-year patent protection, which in fact as a result of obsolescence means usually