Title	Magnitude, moment, and measurement: The seismic mechanism controversy and its resolution
Author(s)	Miyake, Teru
Citation	Miyake, T. (2017). Magnitude, moment, and measurement: The seismic mechanism controversy and its resolution. Studies in History and Philosophy of Science Part A, 65-66, 112-120.
Date	2017
URL	http://hdl.handle.net/10220/44522
Rights	© 2017 Elsevier Ltd. This is the author created version of a work that has been peer reviewed and accepted for publication by Studies in History and Philosophy of Science Part A, Elsevier Ltd. It incorporates referee's comments but changes resulting from the publishing process, such as copyediting, structural formatting, may not be reflected in this document. The published version is available at: [http://dx.doi.org/10.1016/j.shpsa.2017.02.002].

TITLE

Magnitude, Moment, and Measurement: The Seismic Mechanism Controversy and Its Resolution

AUTHOR [FAMILY NAME, GIVEN NAME]

Miyake, Teru

AFFILIATION

Nanyang Technological University, Singapore

ADDRESS

14 Nanyang Drive, School of HSS #06-05, Singapore, 637332, tmiyake@ntu.edu.sg

ABSTRACT

This paper examines the history of two related problems concerning earthquakes, and the way in which a theoretical advance was involved in their resolution. The first problem is the development of a physical, as opposed to empirical, scale for measuring the size of earthquakes. The second problem is that of understanding what happens at the source of an earthquake. There was a controversy about what the proper model for the seismic source mechanism is, which was finally resolved through advances in the theory of elastic dislocations. These two problems are linked, because the development of a physically-based magnitude scale requires an understanding of what goes on at the seismic source. I will show how the theoretical advances allowed seismologists to re-frame the questions they were trying to answer, so that the data they gathered could be brought to bear on the problem of seismic sources in new ways.

HIGHLIGHTS

- First philosophically informed history of an important controversy in seismology
- Examination of the development of a measurement scale in seismology
- Discusses acquisition of knowledge about inaccessible but complicated systems
- Discusses epistemic relation between abstract models and real systems

KEYWORDS

Seismology; Geophysics; Measurement; Observation; Evidence; Idealization

1 Introduction

What is an earthquake? An event happens at a fault, usually deep beneath the earth's surface, where it cannot be seen directly. This event generates elastic waves that propagate through the earth and reach its surface. An earthquake, in the ordinary sense of the word, occurs when these elastic waves reach us, and they are experienced by us as ground motion, or recorded by seismometers. Because of its salience for us, this ground motion is the main thing that comes to mind when we think of an earthquake. But if we want to understand how earthquakes occur, or perhaps even to predict them, we need to understand the event that happens at the origin of the earthquake, at the fault, deep within the earth. This event is called the *seismic source*. The main way in which we obtain knowledge about the source is by taking observations made at the earth's surface by seismometers, and trying to extract information about the source from these observations.

The acquisition of this knowledge is made difficult by two circumstances. First, seismic sources are usually buried deep within the earth, so we do not have direct access to them. All of the processes that happen at the seismic source are hidden from us. Second, what happens at the seismic source is a complicated process, presumably involving rock fracture at a fault which itself has a complicated structure. How, then, do we obtain knowledge about the seismic source from observations at the earth's surface? One way to try to do it is hypothetico-deductively—make hypotheses about what goes on at the seismic source, deduce predictions from these hypotheses about the observations that we ought to see at the earth's surface, and then compare these predictions with observations. This method, however, is subject to well-known problems such as underdetermination, which I will not go over here. The other way to try to obtain knowledge is through measurement. The problem, in our case, is that in order to make a measurement, we need to have a model of the seismic source—but, of course, how can we have such a model without some prior knowledge about what the seismic source is like?

This paper will cover a period during which seismologists moved from a more or less hypothetico-deductive way of trying to obtain knowledge about the seismic source, to being able to measure some of its physically meaningful parameters. This story will advance through two threads that will come together at the end. The first thread involves the development of a scale for measuring the size of earthquakes. For a long time, such scales were entirely empirical—there was no direct connection between these scales and actual physical parameters of seismic sources. It took a theoretical breakthrough in the 1960s for measurement of such physical parameters to become possible, and a physically-based magnitude to be developed. The second problem concerns our understanding of what happens at the seismic source. There were two basic models for the mechanism of the seismic source, and during a period from the 1930s through the 1960s, there was a controversy concerning which of these models best described its target. The controversy was finally resolved in the 1960s via the same theoretical breakthrough already mentioned. These two problems are linked, because the development of a physically-based magnitude scale requires an understanding of what goes on at the seismic source. The theoretical breakthrough has even greater significance, however, because for the first time it allowed seismologists to start bringing seismic wave observations directly to bear on their determinations of the physical processes at work in seismic sources.

The rest of the paper consists of four sections. Section 2 examines the history of earthquake intensity and magnitude scales up to the 1960s. All of these scales were empirical, in the sense that they were defined in terms of what happens locally, near seismic instruments, not in terms of physical processes that occur at the seismic source. Section 3 concerns studies of the seismic source. Here, I focus on the concepts of the single couple and the double couple, and the seismic mechanism controversy, a decades-long dispute over whether the single couple or the double couple is the proper way to represent the seismic source. A significant issue that I discuss here is the role of intuition in this controversy. Section 4 details theoretical advances in the 1960s that finally allowed the controversy to be resolved, and the development of a physically-based magnitude scale. Finally, Section 5 concludes the paper by discussing how the theoretical advances enabled new ways of bringing data to bear on understanding the mechanisms of seismic sources.

2 Earthquake Intensity and Magnitude Scales[1]

At the seismic source, the occurrence of a certain kind of sudden motion along a fault results in the generation of seismic waves. I will first briefly discuss these waves

---

[1] For the historical details in this section, I have drawn on Howell (1990; esp. ch. 6, pp. 97-118).

before turning to intensity and magnitude scales, which are the main subject of this section. Seismic waves propagate outwards from the source through the earth's interior and along its surface, and they are detected by seismometers located at various points on the earth's surface. In 1889, Ernst von Rebeur-Paschwitz recorded seismic waves in Potsdam that he later proposed emanated from an earthquake that had happened minutes earlier in Tokyo. This was the first recording of seismic waves that had traveled deep through the earth's interior. By the end of the nineteenth century, a well-developed theory already existed for the treatment of waves that travel through an elastic solid medium, owing to the fact that physicists earlier in the century had taken the luminiferous ether to be such an elastic solid. In 1828, Simeon Poisson showed that the interior of a homogeneous elastic medium would support two kinds of waves; dilatational waves (known as P waves in seismology), and shear waves (S waves). Lord Rayleigh and A. E. H. Love later showed that in addition to these, a third and fourth kind of wave, respectively called Rayleigh waves and Love waves, can be created at the surface of such a medium. P waves and S waves are collectively referred to as "body waves," while Rayleigh waves and Love waves are referred to as "surface waves". Richard Oldham identified what he took to be P, S, and Rayleigh waves on seismographic observations of a large Indian earthquake in 1897, and Love waves were identified on seismograms after Love proposed their existence in 1911. The recording of seismic waves from large numbers of earthquakes became routine in the first couple of decades of the twentieth century.

The intensity of the shaking we feel from an earthquake can vary widely, from barely perceptible tremors to the kind of strong shaking that can bring down buildings and highway overpasses. We might thus think of recording the size of an earthquake in terms of the kind of effects that we observe on the ground. Early work on such scales was carried out in the earthquake-prone country of Italy. The first widely used scale for recording earthquake intensity, the joint work of Michele Stefano de Rossi and Francois A. Forel, was published in 1883. It was a ten-degree scale, ranging from degree I, which is a shock that can only be felt by an experienced observer, through degree V, which can be felt by everyone and causes some disturbance to furniture, to degree X, which involves "great disaster, ruins, disturbance of strata, fissures in the earth's crust, rock falls from mountains" (Howell, 1990, p. 100). Giuseppe Mercalli invented a similar scale with twelve degrees in 1887, a scale which, with minor alterations, was adopted by Harry

Wood and Frank Neumann in 1931 for their investigations of earthquakes in Southern California. The Wood-Neumann version of the Mercalli scale is still in use today in the United States.

Intensity scales are based on subjective judgments of felt motion and phenomena such as damage to buildings. They tell us how much the ground shakes at a particular place on the surface of the earth, so they are quite useful for the general public, who naturally want to know how much the ground shakes where they actually are. Seismic intensity is not, however, a direct indicator of the size of the seismic source. As one might expect, the farther one is from the source, the weaker the shaking will be. If what one wants to measure is not the amount of shaking at some particular place on the surface of the earth, but the size of the event at the seismic source, one needs a different scale.

The kind of scale that is taken to correspond in some way to the size of the seismic source itself is called a *magnitude scale*. When an earthquake occurs, the shaking caused locally at various places on the earth's surface is recorded by seismometers. Because the shaking from an earthquake is stronger the closer you are to the source, the maximum amplitude recorded on a seismometer will generally be larger as one gets nearer to the source. In 1931, the Japanese seismologist Kiyoo Wadati first noted that there is a roughly logarithmic relation between this maximum amplitude and the distance from the seismic source to the seismometer on which the shaking is recorded. Charles Richter (1935) later proposed a scale in which the logarithm of the maximum amplitude recorded on a seismometer at 100 km from the source of an earthquake would be used as the measure of the size of the earthquake. This scale is now known popularly as the Richter scale. Seismologists refer to earthquake magnitude as measured on this scale as *local magnitude*, or M_L. The amplitude here is defined in terms of the amplitude recorded by a particular kind of seismometer, the Wood-Anderson torsion seismometer, which was the instrument being used by Richter and his colleagues in the 1930s. The main reason that Richter developed this scale was to provide a convenient way of comparing the relative sizes of earthquakes in Southern California. The aim was not to measure any physical parameter of an earthquake, and in fact the only substantive difference between magnitude and maximum intensity was in its being defined in terms of instrumentation rather than subjective reports of observed effects at the epicenter (Howell, 1990, p. 104). That the local magnitude is defined in terms of a particular kind of seismometer results in various shortcomings. Seismometers differ in their frequency response characteristics, so it is not a straightforward matter to convert magnitudes measured using other types of seismometers to local magnitude (Howell, 1990, p. 106). Further, the Wood-Anderson seismometer responds best to ground motion with a period of less than one second, but wave amplitudes for larger earthquakes with a magnitude beyond 7.0 begin to saturate in this range, and this can make local magnitude inaccurate for larger earthquakes (Beroza and Kanamori, 2009, p. 4).

Magnitude scales were subsequently developed which were no longer defined in terms of readings of a particular seismometer, but in terms of the amplitudes of particular kinds of seismic waves. For example, *body-wave magnitude*, or m_b, is defined in terms of the amplitudes of initial P waves that arrive from distant earthquakes. *Surface-wave magnitude*, or M_S, is defined in terms of amplitudes of surface waves at a period of 20 seconds, measured very far away from the source. M_S is what usually is quoted by the press as being the magnitude of an earthquake. These magnitude scales solve some of the problems with local magnitude, but they are still empirical scales. That is, they are defined entirely in terms of what happens at or near seismometers, not in terms of what happens physically at the seismic source.

Naturally, one might wonder if indeed magnitude correlates with some real physical quantity attributable to the source. Magnitude is sometimes said to correlate roughly to the energy released in an earthquake. Gutenberg and Richter (1942, 1956) investigated connections between surface wave magnitude and the total amount of seismic energy that gets released in the form of seismic waves. Because surface wave magnitude is defined in terms of surface wave amplitude at a period of 20 seconds, the assumption must be made that the amount of seismic wave energy radiated into surface waves in this frequency range is proportional to the total energy radiated into all waves at all frequencies. However, this assumption does not hold for very large earthquakes with a magnitude of 8 or greater (Beroza and Kanamori, 2009, p. 5).

Is there any way of using seismometric observations to measure physical parameters of the seismic source? Before seismologists could make such measurements, they needed a set of physical parameters that could be taken to represent relevant physical details of the seismic source. But the seismic source is the site of a complicated process that occurs along a fault, usually buried deep within the earth. Because of the inaccessibility of the seismic source, however, the exact nature of this process was largely unknown until the latter half of the twentieth century. So seismologists would need three things before such measurements of physical parameters would become possible. First, they required an understanding of the relevant physical processes involved at the seismic source. Second, they had to identify any parameters that could be taken to represent relevant physical features of the seismic source. Third, they needed to find a way to determine the values of those parameters from seismic wave observations. The way in which all these pieces were put into place is a complicated story, but it principally involves a decades-long controversy over the proper way to represent the seismic source mechanism, to which I will now turn.

3 Seismic Source Mechanisms

The modern theory of the seismic source began with Harry Reid's investigations of the 1906 San Francisco earthquake that devastated the city. Reid proposed what is now known as the *elastic rebound theory*. The idea is quite simple. The material making up the earth's interior is elastic to a good approximation at the scales that are dealt with in seismology. Now, imagine a solid block made out of a homogeneous elastic material, say a block of jello. If there are no forces on the surface or within the body of the block, it is in a state of equilibrium. Imagine gently pressing on the block of jello with your finger until a depression forms. If you remove your finger quickly, the jello will spring back into its original equilibrium position. Thus, when any part of such an elastic material is displaced away from its equilibrium position, a force (*stress*) forms such that the part has a tendency to return to the equilibrium position. Suppose we now take this block of jello and make a horizontal cut through the middle of the entire block. Next, imagine gently pushing the entire top surface of the block to the left, while also pushing the entire bottom surface of the block to the right. What happens to the block? Suppose the surfaces on either side of the cut are extremely slippery. Then the entire upper half of the block will simply move to the left, and the entire bottom half of the block will move to the right, while both the top and bottom half remain at equilibrium. But now suppose the surfaces on either side of the cut are "sticky", so there is a tendency for the cut to remain stuck together as the top and bottom surfaces of the entire block move. In this case, as we move the upper and lower surface of the entire block bit by bit, the stress at the cut builds and builds until it becomes great enough to overcome the stickiness. The top and bottom surfaces of the cut will then suddenly slip past each other towards an equilibrium position.

This slip will generate elastic waves that travel throughout the block of jello.

The mechanism of seismic sources, according to the elastic rebound theory, is very much like this sudden slip along a discontinuity within an elastic medium. Within the earth, there are forces on either side of a fault that work in opposite directions, but the fault remains stuck together until the stress builds up to the point at which the slip occurs. Seismic waves are the elastic waves that are generated by such a slip. Now, most seismic sources are located deep underground, where we have no direct access to the faults. Although Reid's theory seems intuitively plausible, we had, at the beginning of the twentieth century, very little direct evidence about the seismic mechanism. In fact, with regard to the deepest earthquakes, there is even now some question as to whether they occur along faults at all—some seismologists have suggested that deep earthquakes could be caused by explosive or implosive sources due to sudden phase changes of rock deep within the earth (see Frohlich, 2006, Chapter 7). If we wanted to investigate seismic source mechanisms, one way in which we can envision doing so would be to examine the seismic waves that are radiated outward from the source and are picked up by seismometers at the earth's surface. From the patterns of seismic waves picked up by seismometers, we might hope to be able to reconstruct at least some features of the seismic source mechanism. As more and more seismometers were deployed throughout the world in the first few decades of the twentieth century, seismologists began thinking about how to do this.

Starting in the 1930s, several groups of seismologists carried out studies to obtain information about the seismic source from seismic wave observations. Most of these studies involved observations of the initial motions of P and S waves recorded at seismometers distributed over a wide area of the earth's surface. If the mechanism of a seismic source involves the sudden displacement of a large amount of material along a roughly planar fault, then it seems reasonable to think that initial motions of P and S waves at these seismometers would display a certain symmetry around the plane of the fault. So, if you have seismometers distributed over a wide area of the globe, certain symmetries and patterns in the seismometer readings of initial motions of P and S waves ought to give you information about, for example, the orientation of the fault plane. Of course, there are complications. For example, P and S waves do not travel in straight lines through the earth—they follow curved paths, and they get reflected and refracted at various boundaries, and so on. It is thus quite complicated, but still in principle possible to gain information about the seismic source. Looking for such symmetries presupposes that the seismic mechanism involves displacement along a roughly planar fault, although some studies actually tried to see if other seismic mechanisms, such as conically shaped sources or explosive or implosive sources, would also fit the patterns of initial motions of P and S waves.

These studies took place roughly in the period from the 1930s through the 1960s (Honda, 1962; Stauder, 1960, 1962). They had two main aims. The first aim was to catalogue the orientations of the fault planes of a large number of earthquakes. This work played an important role in the history of seismology, because it was the determination of fault plane orientations that provided one of the major sources of support for the existence of transform faults, which in turn was a key element of the new theory of plate tectonics that was emerging in the 1960s.[2]

I will, however, focus more on the second, for it is linked to the problem mentioned earlier, of coming up with a way to measure physical parameters of the seismic source. The aim was to answer the question: *What is the proper representation of the seismic source, a single couple or a double couple?* The debate over this question, which I will call the *seismic mechanism controversy*,[3] lasted for several decades, from the 1930s, through the war years and beyond, being finally resolved during the 1960s. The controversy is interesting philosophically, because, for a long time, seismologists entertained mistaken intuitions about this question and its implications for their understanding of the physical processes at work in the seismic source. A theoretical breakthrough in the 1960s showed that these intuitions were wrong, and it was this breakthrough that led quickly to the development of a way to characterize and measure relevant physical parameters of the seismic source.

Before we can discuss that, however, I must first explain what a single couple and double couple are, and why they are significant. As I have mentioned, from the mid-nineteenth century through the first decade of the twentieth, mathematical physicists developed the theory of waves in elastic media. Of particular importance for seismic mechanism studies are results by George Stokes in 1849 and A. E. H. Love in 1904 in which they solved the equations for the elastic waves that would radiate outward from a

---

[2] Sykes (1967) is the classic paper that is widely taken to have verified the existence of transform faults from the determination of fault plane orientations.

[3] After Frohlich (2006, pp. 321-324). See also Pujol (2003).

point force in a homogeneous elastic medium. Here, a point force is a finite force acting in a particular direction at a given point in the medium. Now, of course, an actual seismic source does not occur at a point. The mechanism presumably involves some kind of motion along an extended fault, which may be many kilometers long, and is bound to have an extremely complicated structure. But just as in optics extended light sources are often approximated as point sources when observed from far away, we might expect that extended sources of elastic radiation can be approximated as forces that work at a point in an elastic medium.

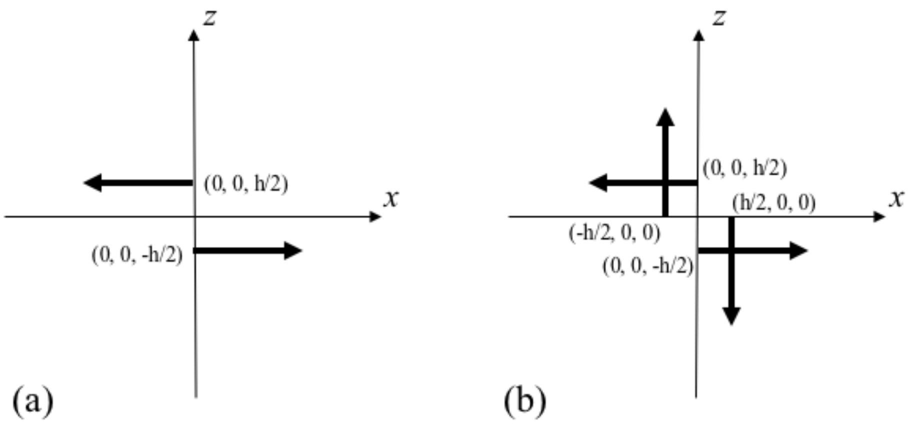

Figure 1. Graphical representations of (a) a single couple and (b) a double couple. Based on Honda (1962), figures 2-4 and 2-5.

Now, consider two such unidirectional point forces, one that acts at the point (0, 0, $h/2$), with a magnitude of $1/h$ in the negative direction of the x-axis, and the other that acts at the point (0, 0, -$h/2$), with a magnitude of $1/h$ in the positive direction of the x-axis (that is, it has the same magnitude as the first point, but it is oriented in exactly the opposite direction). Now take the limit as h goes to zero. The two point forces will converge on one point, but the moment will stay constant, so we end up with a pair of forces that act in opposite directions at a single point, with finite moment (see Figure 1a). This kind of point source is called a *single couple*. Next, consider two single couples, one whose forces are oriented along the x-axis, and another whose forces are oriented along the z-axis, where the moments of these two single couples are equal and opposite. This kind of point source is called a *double couple*, and it has zero moment, because the moments of its component single couples cancel (see Figure 1b).

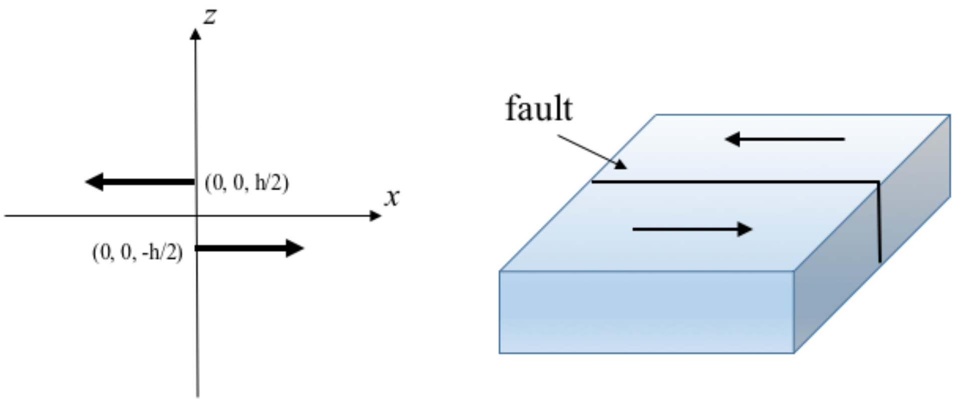

Figure 2. Comparison of a single couple (left) and elastic rebound along a fault (right).

Now, why would the single couple and double couple be of interest to a seismologist? Consider the problem of trying to extract information about a seismic source from observations of seismic waves far away from the source. One place to start would be to examine the elastic wave radiation patterns from sources in a homogeneous elastic medium. A unidirectional point force is the simplest such source, but it seems clear that the forces that will be generated through elastic rebound at a fault will be more complicated than that. If you think of the stresses that would be induced in a fault, it seems, intuitively, that you would have forces that are pointed in opposite directions on either side of the fault, and this suggests the single couple as the simplest possible model of the seismic source (see Figure 2). But there is a problem with the single couple. A single couple has positive moment, which means that if the seismic mechanism can indeed be approximated by a single couple, there would be an unbalanced non-zero moment generated within the earth, which would violate conservation of angular momentum. The double couple appears to be the simplest alternative that would generate zero net force and moment. In spite of this problem with the single couple, however, it appears, as will become clear, that most seismologists during the controversy favored the single couple over the double couple as a model of the seismic source.

In the 1920s a Japanese seismologist, Hiroshi Nakano, calculated and investigated the elastic wave radiation patterns that would be generated by sets of forces including single couples and double couples in a homogeneous elastic medium.[4] Because the ways in which P and S waves propagate through such a medium differ, he calculated two sets of radiation patterns: one for P waves, and one for S waves. We can think about these radiation patterns in the following way. Imagine a single or double couple buried at the center of a huge homogeneous elastic sphere. Now imagine being at a point on the surface of this sphere, and examining the first arrivals of P waves and S waves from this source. The amplitudes of these first arrivals will depend on how the single or double couple is oriented relative to your position on the sphere. Some locations on the sphere, for example, will be nodes, where the amplitudes will be minimal, while in other locations the amplitudes will be maximal. Because the patterns of initial motions differ for single and double couples, an examination of these patterns would, it was thought, allow seismologists to determine whether the seismic source is better approximated by a single couple or a double couple.

Starting in the 1930s, several groups of seismologists worked on the problem of determining, from the patterns of initial motions of P and S waves detected on the earth's surface, whether the seismic source mechanism is a single couple or a double couple. There were several factors that made the issue difficult to settle. First, as I have already mentioned, because P and S waves within the real earth are not actually traveling through a homogeneous medium, they propagate along curved paths, and are reflected and refracted at various discontinuities. P waves can get converted to S waves at some of these discontinuities, and vice versa. Sophisticated methods for accounting for these complications had to be developed. Second, the P wave radiation pattern is exactly the same for a single couple and a double couple, so observations of P waves alone cannot distinguish between a single and a double couple. In order to make the distinction, seismologists needed to examine the initial motions of S waves. Unfortunately, P waves travel faster than S waves, so on seismograms the initial motions of S waves tend to get obscured by the tail of the P wave. Third, these studies depend on observations carried out at seismometers distributed over a wide area. These seismometers often were located in remote areas, and run by inexperienced staff.[5] Finally, we should keep in mind that

---

4 Nakano, and most authors until the 1960s, referred to the single couple as a "Type I source" and the double couple as a "Type II source".

5 See the discussion of such observations in Byerly (1960, p. 1496).

there was no guarantee that all seismic sources must exhibit the same mechanism, nor was there a guarantee that all seismic sources could be approximated by either a single couple or a double couple.

At this point, I want to re-examine the single couple / double couple issue in sharper focus. I stated that the question seismologists were interested in is: *What is the proper representation of the seismic source, a single couple or a double couple?* But there are actually two questions here that were not clearly separated out by seismologists until the 1960s. The first question is: *Does the elastic wave radiation pattern that emanates from a seismic source look like the radiation pattern of a single couple, or that of a double couple?* The second question is: *Supposing that the radiation pattern from a seismic source looks like that of a double couple (or single couple), what does that tell us about the nature of the actual seismic source?* The first question may be answered through empirical studies, by simply comparing seismic wave observations with theoretically derived elastic wave radiation patterns, although of course the empirical work here is by no means straightforward, as I have just explained.

Now, suppose we have shown that the radiation pattern that emanates from a seismic source looks like the radiation pattern of a double couple (or single couple). In order to make a further inference about the nature of the seismic source, or to actually measure physical parameters of the seismic source, one needs to find a way of connecting that double couple (or single couple) with physical features of the actual seismic source. One must bridge the gap between the theory of seismic sources, on the one hand, and the theory of single couples and double couples, on the other. In hindsight, this seems quite clear, but in actual fact the very existence of this inferential gap was not seen clearly until the 1960s. It was probably obscured by certain very natural intuitions. According to the elastic rebound theory, the seismic source mechanism involves two surfaces moving suddenly in opposite directions along a fault. The intuition is that one may approximate such a mechanism by a set of two forces working in opposite directions (see Figure 2; see also Figure 3 in Benioff, 1964), and a single couple is exactly that—a set of two force components working in opposite directions. Moreover, the double couple, with four force components, two of which are working at right angles to the other two, looks very different. Thus, an intuitive picture emerged that strongly linked the elastic rebound theory to the single couple. According to this picture, if it turns out that the radiation pattern from a seismic source looks like that of a single couple, this would be evidence that the seismic source involves elastic rebound. On the other hand, if it looks more like a double couple, this would be evidence against the elastic rebound theory, and seismologists would have to start looking for another theory of the seismic mechanism. This intuitive link between the single couple and the elastic rebound theory, however, turned out to be mistaken.

Let us now return to the history of the controversy. In the 1930s, a Japanese group led by Hirokichi Honda, concentrating its studies on earthquakes in the vicinity of Japan, concluded that the patterns of S wave initial motions seemed to fit a double couple model for the seismic source. At around the same time, an American group led by Perry Byerly performed a number of studies aiming to determine the orientation of the fault plane of seismic sources. In order to perform these studies, Byerly needed to model the seismic source mechanism, and relying on the intuitions I have described, he modeled it as a single couple. Other groups, including a Dutch group and a Soviet group, also carried out studies under the assumption that the source is a single couple. Because of the complications in the observations of initial motions that I have mentioned, however, there was not enough empirical evidence to decide conclusively whether the elastic wave radiation patterns fit a single couple or a double couple more closely. These initial motions studies were interrupted for a while due to World War II, and then resumed after the war, continuing into the 1960s.

During this entire period, the single couple was strongly associated with the elastic rebound theory. In fact, when the initial motions of some earthquakes seemed to indicate a double couple instead of a single couple, some seismologists started questioning the elastic rebound theory, rather than the purported link between the single couple and the elastic rebound theory (Benioff, 1964, p. 1401). Indeed, the major proponent of the double couple, Honda, proposed a mechanism based not on a fault, but a catastrophic process in which a volume of earth originally under shear stress suddenly collapses due to the loss of rigidity (Aki, 1979, p. 48). This is presumably because he thought at the time that the double couple was incompatible with the elastic rebound theory. As I have mentioned, a single couple source would violate conservation of angular momentum, but this apparently was not enough to overcome the intuitive appeal of the single couple for seismologists at the time. There are ways in which the excess angular momentum can be accounted for. Stauder (1962), for example, suggested a source mechanism that involves slip along a fault followed by a rotation of the fault plane in order to restore equilibrium.

The controversy was ultimately resolved in the 1960s, when it was shown through a series of theoretical results that the radiation patterns from elastic rebound along a fault are equivalent to those from a double couple. The significance of these results was that they effectively reversed the answer to the second question that I gave above: *Supposing that the radiation pattern from the seismic source approximates a double couple (or single couple), what does that tell us about the nature of the actual seismic source?* Prior to the 1960s, and throughout the period of the controversy, seismologists had thought that the observation of a single couple radiation pattern would be evidence that the source mechanism is elastic rebound, while the observation of a double couple radiation pattern would be evidence that some other source mechanism is at work. On the contrary, the 1960s results showed that a double couple radiation pattern indicates that the source mechanism is elastic rebound.

I will discuss those results below, but before I do so, I want to give some indication of how surprising these results were. The seismologist Keiiti Aki, who was one of the first seismologists to apply these results to the study of actual seismic mechanisms, describes them as follows:

> This conclusion was unexpected, in view of arguments used in a long-standing debate on the question of whether earthquakes should be modeled by a single couple or by a double couple. Those who advocated the single-couple theory did believe that earthquakes were due to slip on a fault, but they intuitively thought for many years that such slip was equivalent to a single couple (composed of two forces corresponding to the motions on opposite sides of the fault). An intuitive approach is often dangerous in elastodynamics. On the other hand, some of those who advocated the double-couple theory thought that an earthquake must be voluminal collapse under pre-existing shear stress. The fault theory of earthquake sources (now recognized as the equivalent of a double couple) has gained strong support from increasing amounts of data obtained very close to the source region, as well as support from the radiation patterns observed at great distances. (Aki and Richards, 2002, p. 42)

Although a convincing qualitative argument can be given for the double couple being the correct representation of an elastic rebound source (Benioff, 1964), the most conclusive factor in deciding the issue was a set of theoretical results that bridged the theory of elastic dislocations with the theory of point sources.[6] These theoretical results led, moreover, to the development of methods that would enable the measurement of seismic source parameters from seismometric observations, as I will discuss in the next section.

4 Body Force Equivalents and Seismic Moment

By the 1960s, the theory of elastic radiation from point sources such as single couples and double couples had been well-studied, and the controversy over which type of couple best represented the seismic mechanism had been going on for decades. As I have mentioned, however, there are actually two questions here. Even if it was established that the radiation patterns emanating from a seismic source are identical to those that emanate from a single couple or a double couple, the question remains as to what that tells us about the nature of the actual seismic source. Throughout the controversy, seismologists had been led intuitively to think that there is a link between the elastic rebound theory and the single couple. But this intuition turned out to be wrong. What was needed was a theoretical means of bridging the gap between the single couple / double couple and the elastic rebound theory.

The gap was more difficult to bridge than might be thought. The theories of the single couple and double couple concern body forces—that is, forces that act on points within a volume—in elastic media. The elastic rebound theory involves considering surface forces, that is, the elastic forces that act on two surfaces—the upper and lower surfaces of the fault. This requires a different set of tools, those of dislocation theory, the origins of which lie in the work of Vito Volterra in 1907, further developed by Love in 1927. As I have already described, we can think of a fault as a surface buried within an elastic medium in which there are certain stresses on either side of the surface. When an earthquake occurs, the points on one side of the fault are suddenly displaced relative to the points on the other side. This sudden displacement results in a *displacement discontinuity*. The surface within an elastic body across which the displacement discontinuity is created is called a "dislocation"—this is Love's translation of an Italian

[6] It's not entirely clear which arguments were the most convincing to seismologists at the time, but contemporary seismologists such as Pujol (2003) and Frohlich (2006) take these theoretical results to have ultimately resolved the issue.

term used by Volterra.

In the late 1950s, a Russian geophysicist, A. V. Vvedenskaya, and an American, J. A. Steketee, started using Volterra's theory of dislocations to attempt to model seismic sources. In the early 1960s, a series of papers (Maruyama, 1963; Burridge and Knopoff, 1964; Haskell, 1964; see also Pujol, 2003), building on the work of Vvedenskaya and Steketee, established an equivalence between the seismic radiation patterns that would emanate from a seismic dislocation and those that would emanate from a double couple. What is meant here by "equivalence"? Suppose there is a fault buried within an elastic medium, along which a displacement discontinuity is suddenly created. This will generate a certain pattern of elastic waves that will travel through the medium. Now, suppose we take the same elastic medium, but remove the fault. In its place we put a point source such as a single couple or a double couple. Note that a single couple or a double couple is simply a system of forces within the medium, and such a system would give rise to elastic displacements, but there will be no discontinuity. We might now ask the following questions: Is there a point source that will generate exactly the same pattern of elastic waves, far away from the source, that the displacement discontinuity generates? If so, what kind of point source is it? The results of Maruyama, Haskell, and Burridge and Knopoff showed that, for any given seismic dislocation, one can always find an equivalent double couple, and there will never be an equivalent single couple.[7]

These results are widely taken to have resolved the seismic mechanism controversy in favor of the double couple. To be clear, what was actually established here was a link between the elastic rebound theory and the radiation patterns due to double couples. The results, by themselves, are purely theoretical, and do not provide any evidence about what actual seismic sources are like. Perhaps the biggest immediate consequence for seismology was the severing of the strong intuitive link between the single couple and the elastic rebound theory, as I have mentioned. But many seismologists still thought in the mid-1960s that actually observed seismic radiation patterns seemed to indicate a single couple. It was open to such seismologists to either deny that actual earthquake mechanisms involve elastic rebound, or to accept the elastic rebound theory but posit that there is some additional and unaccounted for physical aspect of the seismic

---

7 The double couple is not unique—there could be other equivalent sources, but the important thing in this case is that there are no equivalent single couples. See Aki and Richards (2002, ch. 3).

mechanism.

Over the long term, the most important consequence of these equivalence results was that they led to new ways of bringing seismographic observations to bear on the seismic source. The most direct application was in a paper by Keiiti Aki (1966), where a set of equations from Burridge and Knopoff (1964) were used to derive a now-famous equation: $M_0 = \mu\bar{u}S$. On the righthand side are parameters which describe a seismic dislocation over a fault: S is the area of the fault surface, $\bar{u}$ is the average dislocation over the fault surface, and μ is a constant. On the lefthand side is a parameter, M_0, the scalar seismic moment, which represents the size of the equivalent double couple.[8] The significance of this result is that M_0 can be estimated from amplitudes of seismic waves at the earth's surface. Aki (1966) measured M_0 for an earthquake that occurred in Niigata, Japan in 1964, and then, using estimates of μ and S made independently, estimated the average dislocation across the fault of this earthquake (the calculated value was around 4 meters).

The equivalence results thus finally made it possible to utilize seismic wave observations in determining the values of parameters that are taken to represent actual physical features of the seismic source. The results showed that, if an actual seismic source can be approximated as an elastic dislocation, then there will be an equivalent double couple. This equivalence can then be applied to the problem that I discussed in Section 2—that of coming up with a magnitude scale that is based on a real physical parameter of the seismic source. The size of an earthquake can, quite simply, be defined in terms of the size of its equivalent double couple, or the scalar seismic moment. Thus, in the late 1970s, a new magnitude scale based on the seismic moment was developed: the *moment magnitude*, M_W. It was shown that this scale can be defined so that it agrees with the surface wave magnitude for smaller magnitudes, but also extends past the point where the surface wave magnitude saturates (Hanks and Kanamori 1979, p. 2348).

5 Measuring the Earth's Interior

In the introduction, I mentioned that there are two features of seismic sources

[8] More specifically, M_0 is a scalar representing the size of the force components of the double couple. A double couple has zero total force and zero total moment because the components cancel, but the size of the components is non-zero. See Aki (1966, p. 83), and Pujol (2003, p. 164).

that make it extremely difficult for us to obtain knowledge about them. First, seismic sources are usually buried deep within the earth, so we do not have direct access to them. All of the processes that happen at the seismic source are hidden from us. Second, what happens at the seismic source is a complicated process, presumably involving rock fracture at a fault which itself has a complicated structure. One way to try to deal with these problems is to proceed hypothetico-deductively. But the problem is that mere approximate agreement between predictions and observations is a very thin basis on which to form a detailed picture of what is going on at the seismic source. The other way to proceed is to try to manipulate seismic wave observations so that they tell you things about the seismic source—to try to measure its properties. But this is difficult in our case, because we have very little prior knowledge about the seismic source. Recent work on measurement has focused on the important role of models in measurement (Morgan and Morrison, 1999; Tal, 2012). If the thing we want to measure properties of is not directly accessible, as is the case with the seismic source, there is a very difficult question of whether the model of measurement really is such that, when we make a measurement, we are able to estimate, to some degree of approximation, the actual values of real properties of that thing. And this question becomes even more difficult if the thing we want to measure properties of is thought to be highly complex, as is the case with the seismic source.

One way to proceed is to start by measuring properties that do not depend on inaccessible, complicated physical details of the thing we want to measure. Consider, for example, the way in which we treat systems that are actually extremely complex, such as the earth and other planets, as if they were point masses. We do not think twice about idealizing planets as such imaginary points, measuring their masses, and then taking these masses to represent something real about such systems. But if we go all the way back to the *Principia*, we see that Newton did a lot of work in Sections 12 and 13 of Book 1 (Newton 1999 [1687], pp. 590-621) to show that the attractive force due to a spherically symmetric body with density varying along its radius is exactly equivalent to the attractive force due to a point mass whose mass is equivalent to the total mass of that body, given inverse square gravity working between all particles of matter. The results in Section 12 and 13 underwrite the physical meaningfulness of the measured mass of this imaginary point, and thus allow this measurement to be made without having to worry about the complicated details of the earth's mass structure. In a similar way, the seismic moment is an abstract quantity that represents a physical fact about the seismic source as a whole, which is independent of the complicated physical details of the seismic source. The physical meaningfulness of this quantity is underwritten by the results from the 1960s showing the equivalence between the radiation pattern from a double couple and that from an elastic dislocation, just as the results in Sections 12 and 13 of the *Principia* underwrite the physical meaningfulness of point masses.

But this is not the only thing that the 1960s equivalence results enabled. While, on the one hand, they underwrote this method of abstracting away from physical details of the source, they also led to ways of starting to deal with such complications. Seismologists knew all along that actual faults are bound to be much more complex than the simple elastic rebound model, but they had no resources for investigating such complications. The equivalence results showed that a seismic mechanism that purely involves elastic rebound at a fault surface would have a seismic radiation pattern that is exactly the same as that of a double couple. So the equivalence results allowed a new way of bringing seismic data to bear on the problem of the seismic mechanism. Departures from a pure double couple radiation pattern could now be interpreted as showing that there is some further physically relevant factor at work, in addition to elastic rebound. Investigations of the seismic radiation patterns, which previously were formulated around the question of whether the seismic mechanism is a single couple or a double couple, could now be re-formulated around the questions: Are there earthquakes whose radiation patterns have significant departures from a double couple radiation pattern? If so, what are these departures telling us about the seismic source?

Indeed, the mathematical work leading to these equivalence results enabled new ways of representing seismic sources other than slip along a fault. It was suggested that there could, for example, be explosive or implosive mechanisms that might be caused by sudden phase changes deep within the earth. There could also be fault systems that are not planar, or are more complicated. Such other seismic mechanisms would have force equivalents[9] that are different from single couples or double couples. Leon Knopoff and M. J. Randall (1970) published a series of papers in which they considered these other

---

[9] That is, force systems whose radiation patterns would be equivalent to those of the seismic sources in question. These systems are systems of body forces (forces that act on points within a volume) as opposed to surface forces (forces that act on surfaces), so they are usually called "body force equivalents", after Burridge and Knopoff (1964).

force systems, particularly one that they called a *compensated linear vector dipole*, or CLVD. They showed that a linear combination of double couples and CLVDs would suffice to capture the force equivalents of all possible seismic mechanisms. Further work in the 1970s led to the development of the *seismic moment tensor*, which can represent any particular seismic source as a linear combination of various kinds of force systems such as double couples and CLVDs. The components of these tensors are calculated automatically for a large number of earthquakes every year from observations of normal mode frequencies.[10] Most earthquakes have a strong double couple component, but if other components are significantly large, this can be taken to be an indication of other physical mechanisms at work. In a discussion of deep earthquakes, for example, Frohlich (2006) explores the possibility of the existence of isotropic or CLVD components for deep earthquakes. The study of such components, if they indeed exist, can reveal things about other possible mechanisms at work in deep earthquakes.

An examination of such further developments will take us too far afield. Let me end with some speculative thoughts about how progress has been made in seismology. Actual seismic sources are complicated, and the seismic waves they generate must pass through a very complexly structured medium before they reach seismological instruments. The raw data that is available to seismologists thus exhibits complexity piled atop complexity. What complicates matters further is that seismologists don't know antecedently all the physical factors that give rise to these complexities. The key to progress in seismology lies thus in the discovery of such physical factors, coming up with ways of quantifying and measuring them, and in characterizing the patterns they elicit in seismological data.[11] Once a physical factor is discovered, quantified, and measured, the patterns it gives rise to can, in effect, be filtered out, so that further physical factors can be discovered. This is the significance of the body force equivalence results. They showed that elastic rebound along a fault will give rise to a certain pattern, equivalent to that

---

[10] There was, in fact, an explosion of research on seismic sources in the 1960s and 70s following the body force equivalence results, and I cannot do justice to it here. This includes, for example, research on finite extended models of the seismic source (including kinematic and dynamic models), and spectral characteristics of the radiated field from seismic sources.

[11] This is similar to the way in which Smith (2014) describes progress in gravity theory. I think that the analogy I made to Newton's *Principia* runs quite deep, but this is not the place to try to make such a case.

produced by a double couple, in the seismological data. The size of that equivalent double couple could then be measured, so that the pattern could then effectively be filtered out (that is, departures from the double couple radiation pattern could then be examined), which then could lead to the discovery of other physical factors at work in the seismic mechanism.

ACKNOWLEDGMENTS

I would like to thank George Smith for constant advice and encouragement, my former NTU colleague Lina Jansson for always helpful discussion, and the organizers of the Cambridge measurement conference: Eran Tal, Hasok Chang, and Daniel Mitchell. Some tough post-talk questions by Nancy Cartwright and comments by two anonymous referees were very helpful in getting me to re-think some of the issues in this paper. This work was supported by Singapore MOE AcRF Tier 1 Grant No. RG167/14. Finally, in a special issue on measurement, I cannot fail to acknowledge my considerable debt to Pat Suppes, who will be missed.

REFERENCES

Aki, K. (1966). Generation and propagation of G waves from the Niigata earthquake of June 16 1964 part 2. Estimation of earthquake moment, released energy, and stress-strain drop from the G wave spectrum. *Bulletin of the Earthquake Research Institute*, 44, 73-88.

Aki, K. (1979). Evolution of quantitative models of earthquakes. *SIAM-AMS Proceedings*, 12, 43-58.

Aki, K., & Richards, P. G. (2002). *Quantitative seismology*. Sausalito: University Science Books.

Benioff, H. (1964). Earthquake source mechanisms. *Science,* New Series, 143, 3613, 1399-1406.

Beroza, G. C., & Kanamori, H. (2009). Comprehensive overview. In H. Kanamori (Ed.), *Treatise on geophysics volume 4: earthquake seismology* (pp. 2-58). Amsterdam: Elsevier.

Burridge, R. & Knopoff, L. (1964). Body force equivalents for seismic dislocations. *Bulletin of the Seismological Society of America*, 54, 6, 1875-88.

Byerly, P. (1960). Earthquake mechanisms. *Science,* New Series, 131, 3412, 1493-96.

Frohlich, C. (2006). *Deep earthquakes*. Cambridge: Cambridge University Press.

Gutenberg, B. & Richter, C. F. (1942). Earthquake magnitude, intensity, energy, and acceleration. *Bulletin of the Seismological Society of America*, 32, 3, 163-91.

Gutenberg, B. & Richter, C. F. (1956). Earthquake magnitude, intensity, energy, and acceleration (second paper). *Bulletin of the Seismological Society of America*, 46, 2, 105-45.

Hanks, T. C. & Kanamori, H. (1979). A moment magnitude scale. *Journal of Geophysical Research*, 84, B5, 2348-50.

Haskell, N. A. (1964). Total energy and energy spectral density of elastic wave radiation from propagating faults. *Bulletin of the Seismological Society of America*, 54, 6, 1811-41.

Honda, H. (1962). Earthquake mechanism and seismic waves. *Geophysical Notes, Tokyo University*, 15, Supplement, 1-97.

Howell, B. F. (1990). *An introduction to seismological research: History and development*. Cambridge: Cambridge University Press.

Knopoff, L. & Randall, M. J. (1970). The compensated linear vector dipole: a possible mechanism for deep earthquakes. *Journal of Geophysical Research*, 75, 26, 4957-63.

Love, A. E. H. (1927). *A treatise on the mathematical theory of elasticity*. Cambridge: Cambridge University Press.

Maruyama, T. (1963). On the force equivalents of dynamical elastic dislocations with reference to the earthquake mechanism. *Bulletin of the Earthquake Research Institute*, 41, 467-86.

Morgan, M., & Morrison, M. (1999). *Models as mediators*. Cambridge: Cambridge University Press.

Newton, I. (1999). *Mathematical principles of natural philosophy*. Trans. (I. B. Cohen & A. Whitman, Trans.). Berkeley: University of California Press. (First published 1687)

Pujol, J. (2003). The body force equivalent to an earthquake: a tutorial. *Seismological Research Letters*, 74, 2, 163-8.

Richter, C. F. (1935). An instrumental earthquake magnitude scale. *Bulletin of the Seismological Society of America*, 25, 1, 1-32.

Smith, G. E. (2014). Closing the loop: testing Newtonian gravity, then and now. In Z. Biener & E. Schliesser (Eds.) *Newton and empiricism* (pp. 262-345). New York: Oxford University Press.

Stauder, W. (1960). S waves and focal mechanisms: the state of the question. *Bulletin of the Seismological Society of America*, 50, 3, 333-46.

Stauder, W. (1962). The focal mechanism of earthquakes. In H. E. Landsberg & J. Van Mieghem (Eds.), *Advances in geophysics* (pp. 1-76). New York: Academic Press.

Sykes, L. R. (1967). Mechanism of earthquakes and the nature of faulting on the mid-oceanic ridges. *Journal of Geophysical Research*, 72, 8, 2131-53.

Tal, E. (2012). The epistemology of measurement: A model-based account. PhD dissertation, University of Toronto.