Great Earthquakes on Plate Boundaries

Thorne Lay

Summary and Keywords

Earthquakes involve sudden shear sliding motion between large rock masses across internal contact surfaces called faults. The slip on the fault releases strain energy previously stored in the surrounding rock that accumulated due to frictional resistance to sliding. Most earthquakes are directly caused by plate tectonics, and locate in the cool, brittle rock near Earth’s surface. Events with seismic magnitude measured 8.0 or greater are called great earthquakes and involve slip of from several to tens of meters across faults with lengths from 100 to more than 1,000 kilometers. These huge ruptures tend to occur on or near plate boundaries; the largest are on shallow-dipping plate boundary faults (megathrusts) found in compressional regions called subduction zones, where one tectonic plate is thrusting under another. Some great earthquakes occur within bending or detaching plates as they deform seaward of or below a subduction zone. Yet others occur on plate boundary strike-slip faults where two plates are shearing horizontally past one another, or within deforming plate interiors. Elastic wave energy released during the fault sliding is recorded and studied by seismologists to determine the fault location, orientation and sense of sliding motion, amount of radiated elastic wave energy, and distribution of slip on the fault during the event (co-seismic slip). Geodetic methods measure elastic strain accumulation prior to an earthquake, co-seismic slip, and afterslip on the fault that occurs without earthquakes, along with viscous deformation of the mantle as it responds to the fault offset. Great earthquakes commonly locate under the ocean, and the sudden motion of the seafloor generates tsunami—gravitational water waves that can be recorded with ocean floor pressure sensors (these waves are also used to determine co-seismic slip). As seismic, geodetic, and tsunami modeling methods have progressed over the past 50 years, our understanding of great earthquake rupture processes and earthquake interactions has advanced steadily in the context of plate tectonics and improved understanding of rock friction. All faults have heterogeneous frictional properties inferred from non-uniform sliding during each event, with areas of large slip instabilities called asperities having slip-velocity weakening friction and other
areas having slip-velocity strengthening friction that results in stable sliding. The seismic wave shaking and tsunami waves can cause great devastation for humanity, so efforts are made to anticipate future earthquake hazards. As plate tectonics steadily move Earth’s plates, elastic strain around plate boundary faults accumulates and releases in a repeated stick-slip sliding process that causes a limited degree of regularity of faulting. Given the history of prior earthquakes on a given fault, we can identify seismic gaps where future slip events are likely to occur. With geodesy we can also now measure locations of accumulating slip deficit relative to plate motions, as well as variation in seismic coupling, which characterizes the fraction of plate motion accounted for by earthquake failure.

Keywords: asperities, elastic rebound theory, fault friction, geodesy, great earthquakes, megathrust faults, plate tectonics, seismic coupling, seismic gaps, seismology, slip deficit, stick-slip sliding, tsunami

Earthquakes are one of Earth’s most significant natural hazards, responsible for extensive loss of life and destruction in many places around the globe throughout history. They are manifestations of on-going large-scale motions driven by planetary cooling and will continue indefinitely. Humanity can, however, mitigate the effects of earthquakes by understanding their nature, by building structures that can withstand their effects, and by using technology to give warnings of imminent strong shaking or to evacuate coastal regions when large tsunami are generated. During the past 50 years, we have developed a good understanding of the fundamental causes of the largest events, great earthquakes.

This article outlines many of the key ideas underpinning current understanding of great earthquakes, but it gives little attention to the individuals who played major roles in developing those ideas. As scientific understanding involves progressive elaboration and modification of all key ideas, it quickly becomes unwieldy, and too lengthy, to give full credit to all contributors to a mature and diverse topic such as great earthquakes. Nonetheless a personal listing of major contributors and the great earthquake-related ideas they are associated with is provided here, with no intention of being comprehensive.
Table 1. Seminal Contributors to Understanding of Great Earthquakes

<table>
<thead>
<tr>
<th>Key Player(s)</th>
<th>Critical Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>B. Atwater &amp; C. Goldfinger.</td>
<td>Geological field studies revealing great earthquakes in Cascadia.</td>
</tr>
<tr>
<td>A. Dziewonski.</td>
<td>Systematic moment tensor inversion for moderate to great earthquakes.</td>
</tr>
<tr>
<td>H. Kanamori.</td>
<td>Time domain modeling of long-period surface waves for great ruptures, moment magnitude scale, asperity model, rupture models, rapid moment tensor inversion for tsunami warning.</td>
</tr>
<tr>
<td>H. F. Reid.</td>
<td>Post-1906 San Francisco earthquake articulation of the elastic-rebound theory.</td>
</tr>
<tr>
<td>K. Sieh.</td>
<td>Paleoseismic investigations of great earthquake in California and Sumatra.</td>
</tr>
</tbody>
</table>
Table 2. Key Terms

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asperities</td>
<td>Areas of a fault surface that experience large co-seismic slip during an earthquake. These large-slip zones may be associated with variations in frictional or geometrical properties of the fault and are indicators of stress and strain heterogeneity in the fault zone.</td>
</tr>
<tr>
<td>Elastic Rebound Theory</td>
<td>Basic conceptual model for earthquakes as frictional instability events on faults between large moving rock masses. Driven by large-scale motions of the rock masses, elastic strain accumulates in the fault zone due to frictional resistance to sliding. When friction is overcome and sudden fault sliding occurs, the rock releases the volumetric strain energy and rebounds to a lower strain state. Friction and/or lack of available strain energy stops the sliding, and the process repeats.</td>
</tr>
<tr>
<td>Fault Zone</td>
<td>The volume of rock on both sides of a fault that experiences significant strain accumulation and release during stick-slip earthquake faulting. Release of volumetric strain energy when rapid slip occurs on the fault generates seismic P and S waves that propagate outward into the surrounding rock away from the fault zone.</td>
</tr>
<tr>
<td>Interplate</td>
<td>Faults and earthquakes on plate boundaries.</td>
</tr>
<tr>
<td>Intraplate</td>
<td>Faults and earthquakes within plates.</td>
</tr>
<tr>
<td>Magnitude</td>
<td>A measure of earthquake size based on the amplitude of recorded ground motions for an earthquake. Different magnitude scales are used for different seismic wave arrivals and different periods. All use logarithmic measures of the amplitude and correct for distance from the source to the observing location.</td>
</tr>
<tr>
<td>Megathrust Fault</td>
<td>The primary rock contact surface at the plate boundary between converging plates in a subduction zone. Earthquakes on megathrusts involve shallow dipping thrust faulting.</td>
</tr>
<tr>
<td>Term</td>
<td>Description</td>
</tr>
<tr>
<td>--------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Moment Magnitude</td>
<td>$M_W$, the preferred seismic magnitude measure for low frequency strength of an earthquake. $M_W = (\log_{10} M_0 - 9.1)/1.5$, where $M_0$ is the seismic moment expressed in units of Newton-meters.</td>
</tr>
<tr>
<td>Seismic Coupling</td>
<td>The fraction of total slip on a fault that occurs during earthquakes relative to total relative plate motion across a fault over a given interval of time.</td>
</tr>
<tr>
<td>Seismic Gap</td>
<td>Region where it is known that large earthquakes have ruptured a fault in the past, and it has been a significant fraction of the average large earthquake recurrence interval since that last large event.</td>
</tr>
<tr>
<td>Seismic Moment</td>
<td>$M_0$, a measure of earthquake size given by average shear slip motion over the fault during an earthquake. $M_0 = \mu AD$, where $\mu$ is fault zone rigidity, $A$ is total area that slipped during the event, and $D$ is average slip over the ruptured area.</td>
</tr>
<tr>
<td>Seismic Waves</td>
<td>Elastic waves traveling in a solid involve two basic types of disturbances, P-waves and S-waves. P-waves involve compressional and dilational motions in the direction the waves are traveling, and are analogous to sound waves in a fluid. S-waves involve shearing motions perpendicular to the direction the waves are traveling and only exist in solids. P-waves travel faster and are the first (Primary) wave, while S-waves travel slower and are the second (Secondary) wave to arrive. Near the Earth’s surface above an earthquake P-waves and S-waves interact to establish a Rayleigh wave, which is an interference pattern that travels along the surface, and S-waves reverberate in shallow low velocity layers giving another surface wave called a Love wave.</td>
</tr>
<tr>
<td>Slip Deficit</td>
<td>A calculation of the difference between total relative motion of large rock masses in a given time minus the amount of slip on their boundary fault observed to have occurred in that time.</td>
</tr>
<tr>
<td>Stick-Slip Sliding</td>
<td>Frictional slip instability involving change from static friction with no relative motion across a fault to lower dynamic friction with offset slip occurring on the fault, followed by return to static friction.</td>
</tr>
</tbody>
</table>
Great Earthquakes on Plate Boundaries

<table>
<thead>
<tr>
<th>Tsunami</th>
<th>A long-wavelength ocean gravity wave generated by displacement of a large water volume by seafloor motion from earthquake faulting, landslides into or beneath the water, asteroid or meteorite impacts, or volcanic eruptions. A tsunami increases in amplitude as the wave encounters shallower water.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavefront</td>
<td>Surface within the rock medium for which the phase (or travel time) of P and S waves from the source is uniform. Surface waves and tsunami waves have wavefronts that expand along the Earth’s surface.</td>
</tr>
</tbody>
</table>

### Earthquakes and Plate Tectonics

Earthquakes involve sudden shear sliding motion between rock masses across internal rock contact surfaces called faults, along with the associated ground vibrations from passage of radiated elastic waves and static ground deformation. The ground vibrations and deformation, and induced tsunami, landslides, ground liquefaction, and disruption of built infrastructure, comprise the major earthquake hazards. Most earthquakes are a direct consequence of plate tectonics, which involves relative motions between a spherical mosaic of large cool, stiff rock masses (lithospheric plates) that subdivide Earth’s surface and translate coherently over warmer, viscously deforming rocks at greater depth. Cooling of the planet’s interior by very slow viscous mantle convection powers the plate motions. This has persisted for billions of years, slowly reorganizing the surface configuration of continents (continental drift) and creating (sea-floor spreading) and recycling (subducting) oceanic plates. The theory of plate tectonics was developed in the 1960s, and it is now the fundamental paradigm of the earth sciences.

Earthquakes of all sizes occur within the cool (< 600 °C), brittle rock in the upper portion (< ~50 km depth) of the lithospheric plates. Great earthquakes, with seismic magnitudes ≥8.0, involve sliding displacements, from several to tens of meters across faults with lengths from ~100 to ~1300 kilometers, with the fault slip lasting from ~60 to ~450 s. The energy release can set up vibrations of the entire planet that persist for days. These huge ruptures tend to occur on or near plate boundaries (Figure 1); the largest are on shallow-dipping plate boundary faults (megathrusts) found in regions of plate convergence called subduction zones, where one tectonic plate is thrusting under another. Subduction zones are found on most circum-Pacific margins, as well as in the Mediterranean, Caribbean, Sumatra-Java, Scotia, and Himalaya regions. Some great
earthquakes occur within bending or detaching oceanic plates as they deform seaward of or below a subduction zone. Yet others occur on plate boundary strike-slip faults where two plates are shearing horizontally past one another. A handful of great earthquakes have been recorded within deforming plate interiors, including several earthquakes as deep as 700 km, within cold slabs of oceanic lithosphere that are subducting into the upper mantle. Deep earthquakes are fewer in number and less damaging than shallow earthquakes.

As plate tectonics inexorably moves Earth’s surface plates relative to one another, elastic strain in the rock volume around many faults (the fault zone) accumulates and releases in a repeated stick-slip process that causes a certain degree of regularity of faulting. We can think of earthquakes as sudden slip events that happen in arrears, with the frictionally delayed, shear-strained rock in the fault zone releasing strain episodically to keep up with steady large scale offsets between surrounding undeformed rock masses far from the fault. If the episodic sliding occurs too slowly (time scales much longer than one hour), the event is called a slow slip event, and there is no seismic wave generation. Some faults slip continuously without episodic build-up and release of strain in a process called aseismic creep. Plate motions, with typical relative plate velocities on the order of 0.1 m/yr, thus provide the total slip budget that plate boundary fault systems must match over time by cumulative earthquake slip, slow slip, and/or stable sliding. Rock deformation experiments and earthquake ground deformation observations indicate limiting strain levels of $\sim 10^{-5}$ to $10^{-4}$ ($\sim$1–10 cm across a length scale of 1 km) prior to either brittle fracture of a rock or slip on pre-existing faults in the rock at cool lithospheric conditions.
Thus, there are intrinsic limits to how much strain can accumulate and how much slip can occur in earthquakes.

Even huge earthquakes are only small local steps in the long-term global plate tectonics process. Plate boundary faults can experience thousands or millions of large slip events during their existence, and while individual events can have devastating impact on humanity, we know that the geological process will continue to produce future earthquakes repeatedly. It is often hundreds or thousands of years between large earthquakes that rupture a given portion of a fault, as it takes that long to build up many meters of slip deficit at annual loading rates of ~0.1 m/yr. Only a few fault segments have experienced repeated failure in seismically recorded great earthquakes.

If we know the history of prior earthquakes on a given fault, we can identify portions of a fault where past ruptures have occurred, enough time has passed for substantial strain to have accumulated in the fault zone, and future slip events are thus likely to occur. With measurements of ground position over time, we can now directly quantify locations of accumulating slip deficit relative to surrounding plate motions. Essentially, we can detect the elastic strain building up in the fault zone as the rock flexes toward the limiting stress and strain that will overcome static friction and initiate an earthquake slip event that results in stress and strain reductions in the fault zone. While we have a relatively short observation interval for plate boundary interplate earthquakes relative to the geologic time scales involved in plate tectonics, estimates can be made of variation in seismic coupling of interplate faults, which characterizes the fraction of the total plate motion accounted for by cumulative earthquake slip events rather than by slow sliding that does not produce earthquakes. This information can guide long-term hazard assessments. Intraplate earthquakes on faults within plates tend to lack well-defined tectonic loading rates, but historical activity and crustal strain accumulation measurements can still guide hazard assessments.

In almost all environments, our inability to directly measure absolute stress (and cumulative strain since the last event, if it preceded modern deformation rate measurements) and stress needed for failure in fault zones precludes us from knowing how close to failure a fault is. Thus, even if we know the loading rate, we do not know when earthquakes will occur or how big they will be. This uncertainty is compounded by interactions with stress changes produced by nearby earthquakes on the same or a separate fault that may advance or delay reaching the failure limit, as well as by dynamic loading by seismic waves from remote earthquakes that can trigger failure earlier than expected given the local loading rate and cumulative deformation.
Recording Earthquake Motions, Historical Earthquakes, Paleoseismology

The sudden frictional instability of earthquake slip on a fault releases strain energy stored in the fault zone volume that accumulated prior to the rupture due to frictional resistance to sliding. A small portion of the energy released during the fault sliding expands outward in all directions through the surrounding rock as elastic \( P \) and \( S \) waves that propagate with velocities determined by the rock properties (density, incompressibility, and rigidity). Elastic waves involve ground accelerations produced by rapid time-varying spatial gradients of stresses in the rock as the fault slips and stress in the fault zone reduces; the expanding wavefronts of motions in the rock associated with the two types of elastic waves are fully characterized by the elastodynamic theory of seismology. When these wavefronts reach the earth’s surface, the ground shaking can be recorded by seismometers with accurate clocks, giving ground motion time series called seismograms. The arrival times, amplitudes, and directions of seismic wave ground motions are studied by seismologists to determine the source location and amount of radiated elastic wave energy, the fault orientation and direction of sliding motion across the fault, the distribution of slip over the portion of the fault that ruptured during the event (co-seismic slip), and the change in shear stress acting on the fault (stress drop).

Even modest size earthquakes (down to seismic magnitudes of about 4.5) produce seismic wave motions that can be detected everywhere on the surface using modern instruments, enabling seismological analysis of global faulting processes.

Satellite-based geodetic methods use the Global Positioning System (GPS) and the expanded Global Navigation Satellite System (GNSS), or Interferometric Synthetic Aperture Radar (INSAR) to precisely measure absolute land surface ground position changes over time. These are used to model elastic strain accumulation prior to an earthquake, co-seismic slip distribution, and afterslip on the fault that occurs with or without aftershocks, along with viscous deformation of the mantle as it responds to the earthquake faulting offset. Ground deformations are largest in the fault zone, so geodesy is most useful for events under land areas, although progress is being made in developing seafloor geodesy. Great earthquakes commonly locate under the ocean, and the sudden motion of the seafloor displaces the overlying water layer, generating tsunami, relatively slowly propagating gravitational water waves. Tsunami can be measured with continuously recording ocean floor pressure sensors, and these recordings can be used to determine offshore co-seismic slip.

The seismic wave shaking, ground deformation, and tsunami waves that are recorded by geophysical instruments provide our primary information about the nature of otherwise inaccessible earthquake sources, which are deeply embedded in the rock and often under oceans. Of course, these phenomena can cause great devastation for humanity, and prior
to quantitative seismological and geodetic scientific measurements, earthquakes were mysterious events that spawned many mythologies.

Seismometers with accurate timing and calibrated ground motion recording were first invented in the 1870s, and only after ~1900 have large events been recorded by seismograms from enough global seismic stations to provide reliable quantitative measures of the source processes, such as seismic magnitudes that indicate the ground motion amplitude at specified wave periods (Figure 2). Early seismometers were unable to record the complete strong ground shaking produced by great earthquakes, and it was not until the 1970s that broadband digital seismic instrumentation was developed capable of on-scale recording of all ground motions. The causal connection between fault sliding and earthquake shaking was not accepted until the 1890s, and the plate tectonics context of great earthquakes was not recognized until the 1960s. These considerations indicate that our information about great earthquake rupture processes is still very data-limited, and our understanding of their large-scale plate tectonics context extends back only about 50 years.

Of course, humans have been exposed to earthquakes for far longer than we have had instruments to record and study them. The history of where and when large earthquakes have occurred has been documented to varying degrees by a few civilizations, with the longest accurate lists of large events spanning several thousand years in Japan, China, and the Middle East. Damage caused by large historical events provides some information about their locations and sizes, mainly by calibrating seismic shaking and damage intensity measures relative to recent observations for events of known seismic magnitude. These lists of historic events are strongly biased toward continental regions and subduction zones near population centers like Japan. Earthquake activity elsewhere under the oceans, in remote regions of the world, and as deep as 700 km below the surface in some places, was not recognized prior to deployment of a global distribution of seismometers that allowed such events to first be well located in the 1920s to 1950s. Seismological determination of faulting motions for such events played a major role in
revealing the global distribution of plate boundaries and relative motions of plates in the 1960s, but for most earlier events, we have little seismological information about the faulting that was involved if there was no surface rupture.

Much longer records of earthquake ruptures can be deduced from geological features analyzed in the field of paleoseismology. Trenches excavated along faults can reveal repeated disruptions of the near-surface sediments indicating ruptures that can be dated using radiocarbon-dating methods for organic debris. Offset streams, uplifted scarps, coastal terraces, coral reefs, and other surface features can reveal multiple and cumulative earthquake slip contributions that constrain a fault’s rupture history over several centuries to a few tens of thousand years. Sequences of uplift and inundation of coastlines, with exhumed beaches or drowned forests, and inland tsunami deposits can reveal cycles of strain accumulation and release in subduction zones. Offshore turbidites, or submarine avalanches, induced by large earthquakes, leave layered deposits in channels in the continental slope and on the abyssal plain that can be dated to reveal past events. These methods have emerged and matured over the past few decades and provide important information about the timing and frequency of great earthquakes that preceded the seismological record. The global record remains very non-uniform, particularly for remote subduction zones along sparsely inhabited island arcs.

Megathrust Earthquakes—First Recognition and Early Ideas

By the 1950s, seismological observations had established the distribution of global circuiting bands of earthquake activity produced by faulting concentrated near now recognized plate boundaries, and catalogs of earthquakes with source strengths from different types of seismic magnitude measurements had been developed. Determinations of radial variations of P and S velocity with depth from the earth’s surface to the innermost core had improved to a point where major earthquakes (magnitude 7.0–7.9) anywhere in the planet could be located to within several tens of kilometers accuracy. Developments in seismological theory and analysis of the directions of recorded P and S wave ground motions from an expanding global network of well-calibrated seismometers enabled systematic determinations of shear faulting orientations for abundant moderate to strong earthquakes (magnitudes 5.0–6.9) in the 1960s. Displays of the inferred shear faulting geometries called focal mechanisms became important for resolving deformation patterns as plate tectonics was being elucidated.

Early in this process, two huge earthquakes occurred that launched modern studies of great earthquakes. The first was the Valdivia, Chile earthquake on May 22, 1960, which is the largest earthquake to be seismologically recorded, with a moment magnitude, $M_W$
of ~9.5. $M_W$ is an energy based magnitude determined from the seismic moment, $M_0$; $M_W = (\log_{10} M_0 - 9.1)/1.5$. $M_0 = \mu AD$, where $\mu$ is fault zone rigidity, $A$ is total rupture area, and $D$ is average slip over the rupture. $M_0$ can be estimated from very long period seismic waves that sense the entire energy release; for the 1960 Chile event, $M_0$ is $\sim 2.4 \times 10^{23}$ N-m (some estimates are twice as large as this, and there may have been more than one fault involved). The 1960 rupture was about 900 km long, extending along the southern Chile coast from 37° S to 45° S, with a width of about 200 km extending from the offshore Chile trench to under the coastline, and slip averaging ~30 m. The offshore slip generated huge tsunami that inundated the Chilean coast and spread throughout the Pacific Ocean, causing wave run-up as high as 11 m in Hawaii and 5.5 m in Japan. The second event was the March 27, 1964 Alaska earthquake, the second largest seismically recorded earthquake, with $M_W = 9.2$ ($M_0 \sim 8.2 \times 10^{22}$ N-m), which ruptured a stretch of fault ~750 km long and ~250 km wide offshore of southern Alaska, with ~10 m average slip. The fault sliding in 1964 lasted about 4 minutes.

These huge events occurred offshore and were so large that most seismic recording systems went off-scale for the early arriving seismic waves, so at the time there was substantial uncertainty in exactly what the fault geometry was. It took a combination of systematic field measurements of deformation on land in the overriding plates and advances in analysis of long-period seismic surface waves that orbited the planet multiple times before they became small enough to be recorded on-scale, to robustly establish these events as being on shallowly dipping fault planes with oceanic plates underthrusting the continental margins, rather than involving vertical motions on steeply dipping faults. The thrust fault solutions ruled out large-scale counterclockwise rotation of the Pacific basin, as had been speculated previously based on horizontal shearing along the San Andreas fault in California observed for the great 1906 San Francisco earthquake. This recognition of the faulting orientation for the two huge events in Alaska and Chile validated the emerging idea that plate convergence in subduction zones was a major tectonic process by which old oceanic lithosphere sinks into the mantle, thereby accommodating creation of new ocean lithosphere by seafloor spreading far away at mid-ocean ridges.

Deployment of the World Wide Standardized Seismographic Network (WWSSN), primarily for the purpose of monitoring underground nuclear testing, was nearly complete in 1964. While most early seismograms were off-scale, enough records were obtained to indicate that the slip on the fault for the 1964 Alaska earthquake was not uniform, but had large patches of very high co-seismic slip surrounded by areas of lower slip. Azimuthal patterns in the long-period ground motions for both events (once the waves had traveled far enough to be recorded on-scale) established that the ruptures had spread over the fault at a speed of about 3 km/s, piling up energy in the direction of
rupture expansion (southward for the 1960 Chile event and westward for the 1964 Alaska event). These events launched the systematic seismological and geodetic study of great earthquake rupture processes that has ensued for 50 years. The seismograms for the 1960 and 1964 earthquakes also provided the first extensive measurements of standing wave oscillations of the planet, called normal modes, and these were used in determining models of the seismic velocity and density variations with depth through the mantle and core.

Seismic Gaps

In the light of the newly introduced plate tectonics framework, combined with rapidly advancing seismic and geodetic techniques for analysis of the faulting process of great earthquakes, the historical earthquake activity near circum-Pacific margins began to be reassessed in terms of three categories: interplate thrust faulting, intraplate intermediate and deep focus earthquakes in subducted oceanic lithosphere, and plate bending and tearing processes (Figure 3). A sequence of great earthquake ruptures along the Alaska-Aleutian megathrust in 1938 ($M_S$ 8.3), 1946 ($M$ 8.1), 1957 ($M_w$ 8.6), 1964 ($M_w$ 9.2), and 1965 ($M_w$ 8.7) demonstrated that the entire plate boundary had failed in a suite of temporally clustered, but non-overlapping shallowly dipping thrust events that correspond to the latest strain release steps in the long-term underthrusting of the Pacific plate beneath the North American plate. Similar recognition that the largest recent earthquakes in island arc and continental arc subduction zones usually involve shallow thrust faulting on the interplate megathrust faults driven by plate convergence was firmly established by the late 1970s. Most historical events cannot be fully quantified in terms of faulting geometry, but the very largest shallow events in most subduction zones can usually be assumed to be caused by interplate thrust-faulting. There are a few exceptions; regions such as the Marianas, Tonga, and Java have few recorded large interplate thrust events, with the largest well-characterized events having involved plate bending and tearing processes.
Great Earthquakes on Plate Boundaries

Figure 3. Schematic vertical cross-section through a subduction zone showing the locations where great earthquakes tend to occur. Most are on the megathrust plate boundary fault and involve shallow-dipping thrust faulting. Some are below the outer trench slope and involve slab-bending and slab-pull. Others are intraslab events at intermediate depths below the volcanic arc in the overlying plate.

The 1960 Chile earthquake occurred along a plate boundary region where historical huge earthquakes were documented by European colonists in 1837, 1737, and 1575. We lack detailed knowledge of these older events other than the length along the coast that experienced strong shaking and tsunami run-up. But if they are assumed to have involved interplate thrust-faulting, we can infer that to first-order the plate boundary in southern Chile experiences stick-slip failure with ~130 years between \( M \sim 9 \) great events. The 1964 Alaska rupture zone does not have an overlapping previously recorded event, but uplifted terraces indicate a prior rupture ~900 years ago, although the western portion of the rupture may have failed in an earthquake in 1788. With the Pacific plate underthrusting Alaska at about 5.8 cm/yr, strain comparable to that released in the 1964 event should build up in ~175–330 years, so the earthquake history may be incomplete, or there may be a significant component of aseismic slip.

The dual recognition that (1) large interplate thrust events tend to fill-in the plate boundary to accommodate the convergent plate motions in subduction zones, and (2) that historical earthquakes indicate that each stretch of the plate boundary re-ruptures over time, provided a major break-through in understanding why and where Earth’s greatest earthquakes occur. The 1,300-year history of great earthquakes along the Nankai Trough in southwest Japan, with the most recent involving adjacent ruptures in 1946 (\( M_w 8.1 \)) and 1944 (\( M_w 8.0 \)), preceded by either single events with \( M \sim 8.4\text{–}8.6 \) or paired \( M \sim 8.0\text{–}8.4 \) ruptures of the corresponding region in 1854(2), 1707, 1605–1498, 1361–1360, 1099–1096, 887, and 684, provided the first long sequence of up to eight ruptures of the same megathrust fault, with average recurrence intervals of ~176–180 years. It was also
recognized that the two most recent events did not rupture the entire fault length that failed in 1854, leaving a region of potential high strain accumulation along the subduction zone, close to Tokyo, that is now called the *Tokai seismic gap*. This region has been the focus of Japanese earthquake prediction efforts for several decades.

The notion of *seismic gaps* as sections of faults that have experienced prior earthquakes (so it is known that they are seismogenic rather than aseismic) and have presumably accumulated significant strain since the last event (i.e., the time since the last event is a large fraction of the characteristic recurrence time for that fault section), was well-developed by the late 1970s. Circum-Pacific subduction zones and large plate boundary strike-slip faults like the San Andreas were considered in identifying regions where near-future large earthquakes might fill a seismic gap. Many areas that were identified as seismic gaps by 1980 have ruptured during the past 35 years, notably along Sumatra, Hokkaido, Mexico, Costa Rica, Chile, and Peru. Some great events have struck in areas that were gaps with unknown potential for rupturing given a lack of prior activity, including the central Solomon Islands, the Santa Cruz Islands region, and the central Kuril subduction zone. Some other seismic gaps, such as the southern San Andreas fault in California, where the last large rupture was in 1857, and prior events have happened about every 150 years on average, are still considered likely to host a future major earthquake within the next few decades. Very few regions that have experienced a great earthquake in the past 50 years or so involved re-rupture of a region much sooner than might have been expected given prior history; the notion of a required strain-accumulation time appears to be valid to first order. This behavior is best observed only for great earthquakes; smaller events tend to have more chaotic appearance, but when the actual slip distribution is known pretty well they too can have some regularity in localized regions, as for the sequence of $M \sim 5.0–6.4$ Parkfield earthquakes in 1857, 1881, 1901, 1922, 1934, 1966, and 2004.

This conceptual model of plate boundary stick-slip behavior keeping up with plate loading rates is intellectually appealing and sensible. However, efforts to develop rigorous statistical frameworks for assigning probabilities to timing of future large earthquakes have been less successful, in part due to lack of long-term observations for characterizing the statistical parameters, and in part due to the non-linear nature of frictional instabilities and influence of static and dynamic stress transfer from nearby or even remote earthquakes. It is also clear that earthquakes have varying slip and varying stress drop over the fault during each event, and we lack information about slip distributions for historical ruptures so that apparent re-rupture may be a misinterpretation of complementary slip distributions. In addition, earthquakes tend to cluster in space and time, with it being difficult to identify exactly what areas ruptured during historic events for which we have limited information. Finally, some regions have ruptured either in
single huge earthquakes or in overlapping sets of smaller earthquakes, so the behavior of a region may vary with time. This is true for the Nankai Trough sequence mentioned above, and for the subduction zone along Colombia-Ecuador (the January 31, 1906 $M_W$ 8.8 earthquake rupture zone subsequently failed in three smaller events in 1942 ($M_S$ 7.9), 1958 ($M_S$ 7.8), and 1979 ($M_W$ 8.1)). For Colombia, the historical record is limited, with no record of an overlapping event as large as in 1906 back as far as 1575. This leaves it uncertain whether the repeat times for $M \sim 8$ events are characterized by intervals of 36–73 years, with synchronization leading to a much bigger event every \sim 210 years, or whether triggering effects dominate and only produce huge events infrequently when along-fault stress conditions favor rupture growth.

Continental collision can produce megathrusts with one continental plate underthrusting another continental plate. The primary example of this is along the Himalayan front, where the Indian lithosphere underthrusts the Eurasian lithosphere. Major and great earthquakes have been seismically recorded along the Himalayan megathrust (1905, $M_W$ 7.8; 1934, $M_W$ 8.2; 1950, $M_W$ 8.5; 2015, $M_W$ 7.8), and historical events with $M \sim 8.1$ to 8.6 are documented in 1842, 1803, 1555, 1505, 1400, and 1344. The recurrence time for events along this plate boundary is poorly known, but the vast increase in population in northern India results in tremendous present-day seismic hazard in the region. Given long inter-seismic strain accumulation times in several places along the margin, the possibility of several previously adjacent rupture zones failing together in a single huge event must be considered.

While a seismic gap may be clearly identified, whether it will be filled by a single huge event or by several smaller events is not known in advance, so we really do not know what to expect, although we have a basis for evaluating what is viable. A good example is provided by the northern Chile subduction zone, where a great earthquake with $M \sim 8.5$–8.8 struck in 1877, and the recognition of the associated seismic gap prompted recent seismic and geodetic instrument deployments in the area. On April 1, 2014, an $M_W$ 8.0 event ruptured only about 20% of the rupture area of the 1877 event, analogous to the behavior in Colombia. Prior activity in the region is poorly documented, but claims were made that the event was not the expected earthquake in the gap, and that the remainder of the gap should fail in a single big event. This misrepresents our actual understanding of how variable even great earthquake behavior can be. We can expect that one or more additional great events will occur to complete strain release of the northern Chile gap.

Along Sumatra, there is a stretch of the plate boundary located between great ruptures in 2005 and 2007, where the last great rupture was in 1797. This region is called the Mentawai gap, and it potentially could fail in a magnitude \sim 8.3 earthquake if the entire region ruptures. As the last recently unbroken stretch of plate boundary along Sumatra,
and one straddled by recent large ruptures, this gap is now a focus of instrumentation deployment and preparation for earthquake response in the nearby city of Padang and along the Mentawai Islands offshore of Sumatra.

At present, the seismic gap notion is primarily useful for evaluating the possible region of near-future great earthquakes, as it provides some guidance on what fault segments have plausibly accumulated significant strain and slip deficit relative to plate motion since prior great earthquake ruptures. Direct measurement of current slip deficit accumulation with GPS can validate the prospect of a seismic gap rupturing in the future, but the timing and size of the future rupture is not currently predictable.

**Comparative Subductology**

One of the other aspects of great subduction zone earthquakes that began to receive attention around 1980 is the variation in maximum observed earthquake size between different subduction zones. In some regions, such as the Marianas, Vanuatu, and Tonga, the largest interplate thrusting earthquakes in the seismological record have $M_W \leq 8.0$, whereas huge ruptures with $M_W \geq 9.0$ have struck in Chile, Alaska-Aleutians, Cascadia, Kamchatka, Japan, and Sumatra. Regions with documented huge earthquakes tend to have 150 to 250 km wide, shallow-dipping megathrusts with relatively uniform geometry for long portions of the subduction zone. This allows ruptures to spread over large areas as needed to give large seismic moment. Geometrically complex, heterogeneous environments may tend to delimit rupture expansion, precluding huge events. Regions like the Marianas and Tonga have relatively steep plate boundary dip and, because they are island arcs, the upper oceanic plate has relatively thin crust. The along-dip seismogenic fault width is thus smaller, ~100 to 150 km, so very long single ruptures would be required to give a huge seismic moment. Thus, while a 1,000-km-long rupture of the entire Tonga arc could conceivably produce a magnitude ~9.0 event, the probability of the plate interface being stress state uniform enough to allow such a total-length rupture in a single event is considered intrinsically very low; thus extraordinarily infrequent. This accounts for why no such event has been documented historically in that region. Worst-case, or as we call them, *black-swan*, events with very low probability are of special difficulty for earthquake hazard preparations. As such, there is usually a greater focus on higher probability events that have been experienced historically. Extreme events can happen and will almost always catch us underprepared and surprised, even when such scenarios are recognized as intrinsically possible.

There is no question that the ~115 year seismological history, and even the several-century long lists of historic great earthquake occurrence in many regions, is short
relative to the recurrence time of great earthquakes. As a result, the largest viable events in most regions may not have been experienced historically. The subduction zone along Cascadia has not experienced a great earthquake since 1700, and while that event produced a tsunami recorded in Japan, it was not until paleoseismic analysis of drowned and emerged coastlines and correlation of offshore turbidite sequences along the entire boundary that the regional hazard of great earthquakes was recognized. Nevertheless, at least 18 past great earthquakes along the boundary have now been inferred, with an average recurrence interval of 500 years along the Washington to Vancouver region and perhaps 400 years along the Oregon coast. Ongoing strain accumulation is now being measured using GPS along the Cascadia margin, confirming that the offshore fault is frictionally stuck and accumulating strain, even while the deeper portion of the plate boundary experiences repeated slow slip events with seismic tremor. But this region clearly demonstrates the danger of excessive reliance on the recorded seismological history of large events (Figure 1) for hazard assessment, given the long recurrence times for some regions.

The ratio of cumulative earthquake seismic moment from historical events divided by the maximum potential seismic moment across the entire frictionally controlled interface for a given time interval, or the seismic coupling, is about 30% when all subduction zones are considered together. This indicates that either the historical record greatly underrepresents long-term earthquake occurrence, or more likely, that the majority of plate convergence is aseismic. Assuming that frictional behavior does not allow both fault creep and stick-slip events on the same portion of the fault, about 30% of the frictionally controlled plate contact actually experiences earthquake faulting while adjacent areas are creeping. Regions such as Northern Chile, Southern Chile, Cascadia, Nankai Trough, Central Aleutians, Mexico, Southern Peru, and Sumatra have seismic coupling coefficients close to 1 (100% of the plate motion is accounted for by earthquake slip over the megathrust), whereas Izu-Bonin, Northern New Zealand, Southern Tonga, Central America, Java, Southern Kermadec, and the Marianas have seismic coupling coefficients of 0.1 or less (Figure 4). A few regions have intermediate seismic coupling coefficients, namely Honshu, Kamchatka, Hokkaido, Alaska, Central Chile, and Colombia. While new occurrence of a huge event, or discovery of a past huge event in any of the regions with inferred low seismic coupling coefficient can completely reset the estimate for that region, it does appear that great earthquake behavior varies from region to region in a significant way.
Figure 4. Estimates of seismic coupling coefficient plotted versus ΔF_n, the reduction of normal force on the subduction interface relative to a reference state. The latter were calculated from plate tectonic forces. The curves delineating gray areas were obtained, independently of the data shown, from calibration with the Izu-Bonin/Marianas system. From Scholz and Campos (2012), which presents updated results from an earlier study.

Efforts to understand why seismic coupling may vary between subduction zones from the end-member cases of Mariana-type decoupling to Chilean-type very strong coupling have considered large scale plate tectonic characteristics of each subduction zone, including the convergence rate at the trench, lithospheric age of the subducting plate, geometry of the plate interface (dip, width, etc.), presence of back-arc spreading, presence of sediments on the seafloor, topographic structure of the underthrusting plate (presence of seamounts, spreading ridges, fracture zone, etc.), and other attributes. This analysis of seismogenic behavior in the context of large-scale plate features is called comparative subductology. In the early 1980s, the most promising predictors of maximum earthquake size in different subduction zones appeared to be a combination of convergence rate and lithospheric age. Relatively young, thermally buoyant oceanic lithosphere that resists subduction tends to be associated with large maximum size event interplate seismicity, while very old, negatively buoyant oceanic lithosphere that is readily subducting tends to have smaller maximum size interplate events, especially if there is back-arc spreading in the upper plate and progressive seaward roll-back of the trench position over time. High convergence rate tends to increase maximum earthquake size, but there are counter-examples such as the Tonga subduction zone, which has the highest trench convergence rate globally of as much as 0.25 m/yr (largely due to rapid back-arc spreading), but no recorded interplate thrust event larger than M_w 8.0. Regions with large amounts of sediment on the subducting plate, such as along the 1964 Alaska rupture zone, may have larger ruptures due to the sediments producing uniform frictional state over large
regions, but huge earthquakes have also struck regions with little sediment cover on the subducting plate.

The short history of recorded great earthquakes thus makes inference of tectonic controls on their generation quite uncertain, and some very recent great earthquakes have struck in regions that had been inferred to have relatively low coupling. Subduction of rough bathymetric structure and mid-ocean ridges appears to weaken or disrupt the plate boundary frictional resistance, such that the historic record has no history of great earthquakes in regions such as where the Tehuantepec, Cocos, Nazca, and Chile rises subduct; however, there are exceptions, such as for the April 1, 2007 Solomon Islands ($M_W 8.1$) earthquake that ruptured the megathrust right across a triple junction where the obliquely spreading rise between the Solomon and Australian plates is subducting. The November 16, 2006 Kuril Islands ($M_W 8.4$) earthquake ruptured a region that had not been confidently known to have experienced great thrusting in recorded history. The nature of seismic coupling in this region had been questioned because the upper plate structure is very distinct from adjacent regions where great earthquakes have occurred. The volcanic island chain is disrupted, and the forearc region has an unusual basin; yet these features did not prevent shallow rupture of the megathrust with up to 14 m of co-seismic slip. The 2004 Sumatra-Andaman earthquake ($M_W 9.15$) ruptured a 1,300 km long fault along the plate boundary between the India and Burma/Sunda plates. For the northern portion of this rupture along the Nicobar and Andaman Islands, the relative plate motion is primarily arc-parallel and accommodated by back-arc faulting, with minor (<0.02 m/yr) present-day convergence, yet arc-perpendicular thrusting occurred, indicating strong partitioning of the relative plate motion onto different fault systems.

The current situation is that limited confidence can be placed in using simply parameterized large-scale tectonic attributes to gauge the maximum size earthquake potential of a region. However, the balance of regional plate tectonic forces acting on the normal stress load on the megathrust fault, involving contributions from convergence, back-arc spreading, and slab sinking does appear to account for much of the variability in inferred seismic coupling (Figure 4), indicating that conditions for great earthquake occurrence are influenced by the tectonic setting.

**Asperities and Fault Zone Heterogeneity**

While the overall distribution of great earthquakes appears to be somewhat influenced by plate tectonic forces, it is generally hard to relate details of individual ruptures to specific environmental conditions. This is not too surprising given the non-linear instability of stick-slip failure and the long prior history of previous overlapping and adjacent ruptures
that have left residual stress distributions on each fault segment that affects later ruptures. Seismological methods have been applied to analyze the complex P and S wave ground vibrations produced by great earthquakes since about 1980, revealing that all ruptures are influenced by heterogeneous frictional properties inferred from non-uniform sliding during each event. Localized regions within a rupture that have particularly large slip are called asperities. In the rock friction literature, asperities are tiny point contacts on a frictional surface; seismological usage in the asperity model is more generic, and the large co-seismic slip may be due to variations in frictional failure strength, uniformity of strain release over a large area, and/or geometric obstructions on the fault surface that allow large stress accumulation prior to failure. The recognition of variable slip in individual events complicates the cumulative slip budget approach for great earthquakes. Even though the entire fault surface must ultimately slip enough to keep up with relative plate motions, that process may be accommodated by a mix of inter-seismic, co-seismic and post-seismic sliding of adjacent regions with either stick-slip or stable sliding behavior. Asperities may or may not be persistent features; geometric structures may endure for multiple earthquakes, and frictional heterogeneities may vary from event to event if there are pore fluid variations or other changeable properties involved.

Present day investigations of great earthquake ruptures use a combination of seismic, geodetic and tsunami observations from global instrumentation to determine the slip-distribution in space and time over the ruptured portion of a fault. Examples of the final slip distribution and time-varying energy release for finite-source models for the great earthquakes of February 27, 2010 Chile (MW 8.8), and March 11, 2011 Japan (MW 9.0) are displayed in Figure 5, although visualizing the actual spreading of the rupture and slip at each instant of time requires a time-varying map animation.
Great Earthquakes on Plate Boundaries

Figure 5. Maps summarizing rupture characteristics for (a) the March 11, 2011 Tohoku, Japan ($M_w$ 9.0), and (b) the 27 February 2010 Maule, Chile ($M_w$ 8.8) earthquakes. The white stars indicate the epicentral locations used for each rupture model. The co-seismic slip distributions are those determined from teleseismic body wave recordings for the Tohoku event and for the Chile event. The vectors indicate the variable slip direction for subfaults, with the contoured color scale indicating the total slip at each position. The position and timing of sources of coherent short-period teleseismic P wave radiation in the bandpass indicated in each panel imaged by back-projection of recordings at North American seismic stations, mainly from the EarthScope Transportable Array, are shown by the colored circles, with radius scaled proportional to relative beam power. The rectangles in (a) indicate estimated source locations of high frequency (5–20 Hz) strong ground motions. Note that the regions with large slip locate up-dip, toward the trench (dashed line) in each case, whereas the coherent short-period radiation is from down-dip, near the coastline. From Lay (2015).

Many factors likely contribute to the non-uniform slip during any given rupture. The February 27, 2010 Chile ($M_w$ 8.8) earthquake nucleated in a region that had last experienced a great earthquake in 1835 (an event chronicled by Charles Darwin who was in Valdivia, Chile at the time on his famous voyage on the Beagle). The seismic gap had been instrumented prior to the 2010 event, but the actual slip on the fault during the earthquake was low in the region of the 1835 earthquake, while there was as much as 20 m of slip near the rupture zone of a more recent event in 1928 ($M_w$ 8.0). The rupture was delimited on the north by the rupture zone of a major earthquake in 1985 ($M_w$ 7.8) and on the south by the great 1960 Chile earthquake. Given the plate convergence rate of 0.066 m/yr, only about 5.4 m of slip deficit should have accumulated near the 1928 rupture, whereas 12 m of slip deficit could have accumulated near the 1835 zone, but that is not what was released during the 2010 rupture. Many large aftershocks struck within the low slip region near the 1835 rupture, indicating that it is not still locked, thereby implying a high component of aseismic slip in that region. Lacking details of the slip distribution in the 1928 zone makes it difficult to infer whether there was overlap in slip with the 2010 rupture, but clearly the earlier event did not release all of the regional slip deficit that had accumulated prior to 1928. This illustrates the typical challenge raised by the non-uniform slip and variable stress drop in sequences of great events.
When the asperity model was first introduced around 1980, attempts were made to relate the slip heterogeneity inferred from seismic waves for great earthquakes to tectonic forcing (Figure 6). Huge, continuous slip patches with ~100 km scale-lengths found for the largest earthquakes, like 1960 Chile ($M_W 9.5$), 1964 Alaska ($M_W 9.2$), 2004 Sumatra ($M_W 9.2$), and 2011 Japan ($M_W 9.0$), represent relatively uniform large-strain accumulation in the largest asperities typical of Chilean-type subduction. Groups of moderate size slip patches with 30-50 km scales are found for complex ruptures in the $M_W \sim 8.2-8.6$ events in the Kuriles, Aleutians, and Southern Peru, with doublet and triplet sequences of $M_W \sim 8.0$ events being found in Vanuatu and the Solomon Islands. Small asperities are found for the largest events in the Marianas and Tonga regions. In part, these variations can be related to geometry of the subduction zones, but specific variations in frictional strength and strain accumulation remain the focus of active investigations.

![Figure 6](image)

**Figure 6.** Left: Source time functions (moment rate) for great earthquakes along with long period P wave motions from which they are determined. Right: Asperity model for different subduction zones, indicating areas with unstable stick-slip sliding (hatched) zones on plate boundary faults. The unstable friction zones have large slip during big earthquakes and are called asperities. Variations in the area fraction of unstable to stable sliding frictional behavior varies from region to region. Strong seismic coupling is associated with large asperities and weak coupling is associated with small or no asperities.

For many historic events our primary information about their rupture extent comes from the aftershock distribution, although sometimes tsunami recordings at tide gauges provide useful constraints on the uplifted seafloor area. Aftershock zones have been used to estimate the rupture area and seismic moment of large historic events. But modern studies of co-seismic slip distributions indicate that aftershocks on the megathrust fault tend to be located outside of the regions of major co-seismic slip; this can lead to overestimates of the rupture area. This is compounded by the recognition that many aftershocks occur off the main thrust plane, activated by changes in stress and fluid distribution induced by the main faulting.
Areas that experience large slip in great events tend to have few overlapping events in the inter-seismic period, suggesting strong locking of the fault within the asperities. For example, there has been almost no recorded interplate thrust faulting other than tiny microearthquakes on the Cascadia subduction zone in the seismological record. This was often invoked in the false interpretation that the region must be slipping aseismically, before paleoseismic observations and GPS slip deficit measurements demonstrated otherwise. The southern San Andreas Fault along the 1857 rupture has very little on-fault seismicity; it appears that the fault is rather uniformly locked over large regions, and strain accumulation in the fault zone is clearly detected by GPS estimates of ground deformation on traverses across the fault.

Portions of megathrusts outside of major asperities sometimes have regularly repeating small earthquakes that appear to be repeated ruptures of isolated small asperities surrounded by stably sliding portions of the fault. These events are increasingly used to identify regions where large slip will not occur in great earthquakes, as well as to estimate the plate interface deformation rate in stably sliding regions. After nearby great events occur, the creeping portions of the fault sometimes appear to slide at accelerated rates for several years, and this post-seismic slip-rate increase causes a reduction in the time between repeating earthquakes. As much of the post-seismic deformation is aseismic, geodetic measurements are essential for detecting this behavior. For some large earthquakes off of the coast of Japan, the amount of post-seismic deformation has been found to exceed the amount of co-seismic slip, indicating how the total sliding budget of the interplate motion is partitioned.

Deployments of GPS instruments along many subduction zones are now monitoring upper plate deformation rates, typically involving landward flexing of the plate as shear strain in the wedge accumulates. This has been steadily improving for about 20 years as instrumentation was developed, GPS and GNSS satellite constellations were expanded, and ability to process and monitor the deformation improved. Prior to the March 11, 2011 Japan earthquake, the extensive GPS network in Honshu showed westward ground velocities of the coastline that indicated coupling on the plate boundary, but the on-shore measurements could not reliably resolve how the slip deficit was distributed over the full width of the megathrust. Such geodetic measurements can resolve lateral variations in slip deficit accumulation along the subduction zone, and this can partially delimit asperity dimensions, but how far an earthquake will expand depends on non-linear effects that cannot be anticipated. GPS measurements indicated slip deficit accumulation along the 2010 Chile earthquake zone that has some correspondence to the actual co-seismic slip distribution, but these observations did not define how far north the rupture was able to extend. Future geodetic measurements of this type are essential for all subduction zones, and development of offshore geodetic measurement capability would help immensely for long-term detection of strain build-up prior to future great earthquakes.
Great Earthquake Rupture Variations With Depth

Lateral variations in great earthquake occurrence have received much attention, but in the past 15 years there has also been interest in characterizing the variation in rupture properties with depth along the megathrust. The down-dip edge of the seismogenic portion of the megathrust involves a transition from frictional stick-slip behavior to ductile deformation between the plates that occurs without brittle earthquake faulting. This transition takes place by about 35 km depth in oceanic arcs, and by about 50 km in continental arcs, with some influence of the upper plate crustal thickness (Figure 3). In regions with relatively young subducting lithosphere, like southwest Japan and Cascadia, there is a transition zone with slow slip and seismic tremor down-dip of the stick-slip regime. The up-dip edge of the seismogenic zone may extend all the way to the trench or it may be at a depth of about 10 km along a seismic front, with no earthquakes along the shallowest toe of an accretionary prism. This up-dip edge may be influenced by presence of weak, water-saturated sediments with slip-velocity strengthening material that undergoes stable sliding, or there may be rough horst and graben structure or seamounts on the underthrusting plate that produces patchy asperities with slip-velocity weakening friction that causes shallow stick-slip faulting. The shallow frictional environment may also involve conditional stability, with stable sliding occurring at normal strain rates, but rapid earthquake slip occurring when high strain rates from deeper earthquake displacements drive rupture up to the trench. In between the up-dip and down-dip limits of interplate earthquake behavior, the megathrust undergoes increasing pressure, temperature, expulsion of fluids, development of impermeable shear zones, smoothing of bathymetric roughness, induration of sediments, phase changes in sedimentary clay minerals, and other effects that may produce depth-varying changes in friction and seismic radiation during rupture.

Rupture of the shallowest portion of megathrusts was generally not thought to occur prior to the recognition of a class of large earthquakes called tsunami earthquakes. These were first recognized, in the early 1970s, as relatively rare events that generate tsunami waves larger than expected for the measured surface wave seismic magnitude. Classic examples are the 1896 Sanriku, Japan $M \sim 8.5$ earthquake, which generated peak tsunami run-up of about 40 m along the northern Honshu coast, and the 1946 $M_w 8.1$ Aleutian earthquake. Both events were found to involve relatively weak generation of short period seismic waves, primarily because the source process was unusually long, due to slow rupture expansion velocity or slow fault sliding velocity. More recent events of this type occurred along the Kuril and Peru subduction zones, but the first to be well recorded by high quality broadband seismic instrumentation was the September 2, 1992 Nicaragua ($M_w 7.6$) earthquake. The data resolved a low average rupture velocity of about 1.2–1.5
km/s, significantly lower than the 2.0–3.0 km/s rupture velocity of typical megathrust events, contributing to an unusually long rupture duration of about 110 s. Tsunami observations indicate that the rupture occurred very close to the trench, and the deeper portion of the megathrust appears to be weakly coupled based on GPS observations on land.

Similar characteristics of very shallow megathrust rupture, slow rupture velocity, and long duration have been found for other recent tsunami earthquakes including the July 17, 2006 Java ($M_W$ 7.8) and October 25, 2010 Mentawai, Indonesia ($M_W$ 7.8) events. The 2010 Mentawai event ruptured up-dip from a great earthquake that struck in 2007 ($M_W$ 8.5) along the Sumatra coast, demonstrating the possibility that tsunami earthquakes can occur in the many places when rupture of deeper events does not extend to the trench. Given that these events occur under relatively low confining pressure near deep trenches, it is generally thought that wet sediments on the megathrust may contribute to the low rupture velocities, but the events clearly demonstrate that large (~10 m) slip deficits can accumulate near the toe of the trench, contrary to the conventional common assumption of stable sliding behavior.

It has only recently become possible to examine depth-varying properties for great earthquake ruptures due to availability of large seismic data sets from global and regional networks. The March 11, 2011 Japan ($M_W$ 9.0) earthquake is the best-recorded great earthquake to date, largely due to extensive seismic and geodetic networks deployed in Japan. This earthquake ruptured the entire 200 km width of the seismogenic zone, with average slip of about 15 m, but peak slip on the shallowest part of the megathrust of 50–60 m. The overall rupture velocity was slow, and the large shallow slip with slow rupture expansion and little short period seismic wave radiation in the up-dip region behaved like a tsunami earthquake. This region probably had not ruptured in a great event since 869. However, co-seismic slip extended all the way down to below the Honshu coast at 50 km depth. The slip from 30–50 km depth was less than 10 m, and involved higher rupture velocity re-rupture of regions that had failed repeatedly in smaller $M < 8.0$ events. Seismic wave radiation from the deep region of slip was found to have relatively enriched short-period wave energy (Figure 5). This was found by nearby seismometers in Japan that record up to 10 Hz shaking and by remote large-scale arrays of short-period stations in Europe and North America that record up to 2 Hz P wave energy.

Similar depth-varying properties of rupture were also found for the February 27, 2010 Chile ($M_W$ 8.8) event (Figure 5) and the December 26, 2004 Sumatra ($M_W$ 9.2) event, both of which also ruptured across the entire megathrust. Analysis of seismic wave radiation from smaller events at different depths on the megathrust have further indicated that there are depth variations in seismic wave energy release, with the upper 15 km of the
megathrust depleted in short-period radiation, and regions below 30 km enriched in short-period radiation on average (Figure 7). This behavior is being studied to understand frictional variations along the fault, but is also of great importance for seismic hazard assessment, as it is the high frequency energy released from the deeper portion of the megathrust that causes most of the shaking damage experienced during great earthquakes.

Figure 7. (a) Schematic cross-section, scaled appropriately for the subduction zone off the northeast coast of Honshu where the great 2011 Tohoku earthquake occurred, indicating four domains of megathrust rupture characteristics: A–near-trench domain where tsunami earthquakes or anelastic deformation and stable sliding occur; B–central megathrust domain where large slip occurs with minor short-period seismic radiation; C–down-dip domain where moderate slip occurs with significant coherent short-period seismic radiation; D–transitional domain, only present in some areas, typically with a young subducting plate, where slow slip events, low frequency earthquakes (LFEs), and seismic tremor can occur. At yet greater depths, the megathrust slides stably or with episodic slow slip or plastic deformation that does not generate earthquakes. (b) Cut-away schematic characterization of the megathrust frictional environment, related to Domains A, B, C, and D defined in (a). Regions of unstable frictional sliding are red regions labeled seismic. Regions of aseismic stable or episodic sliding are white regions labeled aseismic. Orange areas are conditional stability regions, which displace aseismically except when accelerated by failure of adjacent seismic patches. Domain A is at shallow depth where sediments and pore fluids cause very slow rupture expansion even if large displacements occur in tsunami earthquakes. Domain B has large, relatively uniform regions of stable sliding that can have large slip, but generate modest amounts of short-period radiation upon failure. Domain C has patchy, smaller scale regions of stable sliding surrounded by conditionally stable areas. When these areas fail, coherent short-period radiation is produced. Small, isolated patches may behave as repeaters when quasi-static sliding of surrounding regions regularly load them to failure. Domain D is dominated by aseismic sliding, but many small unstable patches can rupture in seismic tremor when slow slip events occur. From Lay (2015).
Intraplate Great Earthquakes Near Subduction Zones

It is rather remarkable that 80 to 100 km thick stiff oceanic lithosphere is able to bend and sink into the mantle, pulled down by its own dense leading edge (slab pull). The bending process places the near-surface environment into horizontal tension in the upward bowed outer rise and outer trench slope regions seaward of subduction zones, and this produces extensional (normal) faulting, with less frequency, deeper compressional faulting (about 50 km deep) in the lower portion of the brittle lithosphere (Figure 3). Great normal faulting intraslab earthquakes have occurred in old subducting lithosphere in regions of moderate interplate seismic coupling (e.g., the 1933 Sanriku-Oki ($M_W \sim 8.3-8.6$) earthquake), and in regions of very low seismic coupling (e.g., the August 19, 1977 Sumba, Indonesia ($M_W 8.3$) and September 29, 2009 Samoa ($M_W 8.1$) earthquakes). Great normal faulting events can also occur seaward of strongly coupled subduction zones following interplate ruptures. The January 13, 2007 Kurile ($M_W 8.1$) normal faulting earthquake ruptured the outer trench slope following the November 16, 2006 Kuril ($M_W 8.4$) underthrusting event. It appears that elastic bending stresses can be modulated by interplate shear stresses. Such large normal faulting earthquakes rupture over a 30–40 km depth range, essentially breaking through the entire brittle lithosphere. Thus, they are sometimes viewed more as slab detachment events than simple bending stress events.

This perspective is supported by great extensional events that have occurred just below the megathrust, apparently dominated by slab pull rather than bending (Figure 3). The June 22, 1977 Tonga ($M_W 8.1$) earthquake ruptured within the sinking Pacific slab at depths of 70 to 90 km, probably rupturing across the entire brittle slab domain. The November 4, 1963 Banda Sea ($M_W 8.3$) event did likewise, near 120 km deep, but with an intraplate compressional mechanism in a region where the slab is strongly contorted. Comparable size great events at these depths have struck beneath Chile (December 9, 1950 $M_W 8.2$; November 11, 1922 $M_W 8.3$), and the Rat Islands, Alaska (August 17, 1906 $M_W 8.3$).

These great ruptures within subducting slabs represent large-scale deformation of the descending plate, but the loading rates are not readily defined, as is the case for interplate deformation. Great intraplate events in the upper 120 km of the slab below the megathrust can produce significant shaking damage because they locate below land on the upper plate. Seismic wave energy from intraplate earthquakes tends to be enriched in high frequency content relative to interplate ruptures, which accentuates the shaking hazard for intraslab and outer rise earthquakes.
Three great earthquakes at very great depth in subducting oceanic slabs have been recorded (Figure 1). The two largest are $M_W 8.3$ events on May 24, 2013, about 610 km deep in the Kuril slab, and June 9, 1994, about 640 km deep under Bolivia. The third is the July 31, 1970 $M_W 8.0$ event 645 km below Colombia. These events, like all deep focus earthquakes at depths greater than 400 km, are rather enigmatic, as they involve extensive shear faulting under immense confining pressures that should inhibit brittle failure. They may be influenced by phase changes in minerals or sudden fluid release from minerals in the descending plates, with rupture possibly involving rapid melting as shear faulting expands. Their great depth makes these events unimportant for seismic and tsunami hazards. Major and great earthquakes at intermediate depths can be very damaging.

Strike-Slip on Plate Boundaries and Continental Deformation Zones

Strike-slip faulting occurs where rock masses are sliding horizontally relative to each other; and it is relatively rare for such events to be great earthquakes because the down-dip fault width tends to be only tens of kilometers, and fault lengths tend to be bounded by structural variations. Large strike-slip earthquakes have struck along the Queen Charlotte Fault (August 22, 1949; $M_S 8.1$) and in 1906 ($M 7.7–7.9$), and January 9, 1857 ($M \sim 7.9$) along the San Andreas Fault. Very large strike slip earthquakes also occur on other plate boundary transform faults such as the Alpine fault in New Zealand and the Anatolian fault in Turkey, but great earthquakes are not common on these boundaries.

The continental collision in southern Asia results in extensive crustal shearing throughout China and large, destructive strike-slip faulting is widespread, resulting in extensive seismic hazard for the large population. The December 16, 1920 Gansu-Ningxia border ($M_W 8.3$) event is among the largest examples. Great events have been also recorded as far to the north as Mongolia (July 9, 1905 and July 19, 1905 $M_W 8.0–8.3$ events). While these events are not on a plate boundary in a strict sense (Figure 1), they appear to be consequences of plate interactions distributed over a broad deformation zone nonetheless, with portions of Asia extruding eastward to accommodate the northward motions of India.

Great strike-slip events have also struck along oceanic fracture zones or within deforming oceanic lithosphere near plate boundaries. These remote events are not particularly hazardous, as they excite only weak tsunami. The largest recorded intraplate strike slip faulting occurred in a complex faulting event on April 11, 2012, with a network of five nearly orthogonal faults rupturing sequentially in an $M_W 8.6$ event within the Indo-Australian plate seaward of the December 26, 2004 Sumatra $M_W 9.2$ megathrust event.
An $M_W$ 8.2 aftershock struck the same day. This deformation appears to be associated with gradual breaking apart of separate Indian and Australian plates, resulting from compressional stress in the plate caused by the Himalayan collision. While a major tectonic process, fortunately the associated seismic hazard is low for the oceanic strike-slip shearing region.

**Great Earthquake Early Warning; Tsunami Warning**

There is as yet no reliable procedure for earthquake prediction in terms of specific location, size, and timing of future events. Identification of seismic gaps and monitoring of regional strain accumulation are valuable activities, but unless there is a still undiscovered precursory phenomena that can be monitored in advance giving robust indications of imminent failure, great earthquakes will continue to catch us by surprise when they occur. Efforts to find reliable precursors continue; some recent great events have been preceded by slow slip events with migrating seismicity sequences that may have some precursory attributes, but other approaches are needed to provide useful hazard mitigation.

Seismic waves and ground deformations recorded in the vicinity of an earthquake can very quickly provide an indication of how much energy has been released, and by immediate analysis of the signals to determine the faulting characteristics, information about the source can be transmitted to regions away from the fault prior to arrival of strong seismic waves (particularly S waves and surface waves) and tsunami waves. This has enabled the development of earthquake early warning systems, now being implemented in many countries, and tsunami warning systems. Smart-response systems that automatically activate hazard mitigation upon detection of first-arriving P wave signals that exceed thresholds for ground shaking have been active in Japan for decades. This has worked well for stopping bullet trains when initial arrivals from earthquakes are detected, prior to the stronger shaking from later arrivals. Ground acceleration detectors can also shut off gas lines, lock elevators, initiate response procedures at critical facilities, and thereby reduce the impact of large earthquakes. With great earthquakes commonly located offshore, rapid determination of source size, location, and faulting mechanism using long-period seismic waves or geodetic deformations (including offshore ocean-bottom pressure sensors) can provide a 15–20 minute warning of regional tsunami waves that prompt evacuation activities. Such procedures are critical for mitigating the effects of great earthquakes near population centers. The value of this was demonstrated by the 2011 Tohoku, Japan $M_W$ 9.0 earthquake for which loss of life was reduced by rapid evacuation despite great tsunami devastation of coastal towns. This event, along with the 2010 Chile $M_W$ 8.8 earthquake, also demonstrated how well seismological
characterization of great events and rapid modeling of deep-water tsunami observations can predict remote tsunami amplitudes around large ocean basins.

**Research Questions and Frontiers**

Recent great earthquakes have been recorded with unprecedented numbers of seismic, geodetic, and tsunami observations, and every event has been the subject of numerous research investigations. This has dramatically increased our observational basis for understanding great earthquakes, and has confirmed both the need for using technology to mitigate impacts of great earthquakes and viability of doing so.

While progress has accelerated, there are persistent challenges that remain in the study of great earthquakes. With new data and new probes of the fault zone environment, increasing diversity of frictional properties of megathrusts and other major faults has been recognized. The mechanisms by which ruptures nucleate, spread in a heterogeneous stress regime, and stop remain unclear for earthquakes of all sizes. Whether asperities are primarily geometrical, controlled by fluids or sediments, and how persistent they are from great event to great event is not well understood. Drops in dynamic friction during rupture have now been demonstrated in the laboratory and may be critical for fault zone heating and earthquake stress drops relative to absolute stress levels.

Stress transfer between earthquakes has been demonstrated by recent great events (Figure 8), with a wide variety of phenomena. Along-fault migrating sequences (Figure 8A) have been documented in the Sumatra subduction zone, from 2004 to 2010, similar to that seen for the Alaska-Aleutian zone from 1938 to 1965. The role of elastic and viscoelastic processes involved in such migrations needs to be better understood. The ability to rupture across tectonic structures such as subducting ridges (Figure 8B) that had been thought to be obstacles to rupture was demonstrated by the 2007 Solomon Islands earthquake. We need to better understand whether this can happen in other regions of “aseismic” ridge subduction. Failure of the shallow megathrust up-dip of great events at greater depth (Figure 8C) has been documented in the Kurile and Sumatra subduction zones, and it is important to develop a capability to tell where this may occur, as it can result in destructive tsunami earthquakes. Some great ruptures manage to spread all the way to the trench, enhancing their tsunami excitation (Figure 8D). How and why this happens for some events and not others is not well understood. Recent great earthquake triggering has occurred with interplate thrusting triggering intraplate extension (Figure 8E; observed in the Kurile islands in 2006–2007) or intraplate extension triggering interplate thrusting (Figure 8F; observed in Samoa-Tonga in 2009).
Understanding this type of dynamic interaction and the compounded seismic and tsunami hazard it entails is a challenge for researchers.

Great earthquakes will continue throughout human existence, and understanding their nature and reducing their impact on humanity is a grand challenge for geophysicists and natural hazard mitigation systems. Long-term observation of the earth system with seismic, geodetic, and tsunami instrumentation is paramount for making future progress in fundamental understanding and effective hazard mitigation. Major new efforts to expand the instrumentation to poorly covered regions and to the offshore environment in subduction zones will need to be undertaken to deal with the global nature of the great earthquake challenge.

**Suggested Readings**


**Thorne Lay**

Department of Earth and Planetary Sciences, University of California Santa Cruz