The Seismogenic Zone of Subduction Thrust Faults

WHAT WE KNOW AND DON'T KNOW

R. D. Hyndman

Abstract

There have been great advances recently in characterizing and understanding earthquakes on subduction thrust faults; this paper discusses some of the many questions that remain. Important seismic characteristics of subduction thrust faults and their physical associations include the following: (1) The maximum thrust earthquake magnitude, $M_x$, is highly variable among subduction zones; $M_x$ may be related to the downdip seismogenic width, i.e., updip and downdip rupture limits, or to the physical characteristics and stress on the fault. (2) The term "seismic coupling," i.e., fraction of relative motion that is accommodated seismically, needs careful definition. Meaningful use of the term requires specification of the downdip seismogenic width. Some subduction zones appear to be completely locked, with no aseismic slip between megathrust events; others have mostly aseismic slip. (3) The term "seismic asperity" also needs careful definition; it commonly describes fault regions that had especially large slip in a great earthquake. However, inferences that such areas always have larger earthquake displacement and that they are associated with fault physical characteristics are not yet firmly established. (4) Subduction thrust faults are concluded to be weak. The commonly favored explanation is regionally elevated fluid pressures, but weak fault zone materials and dynamic rupture processes also have been proposed. (5) Most subduction thrusts have consistent updip and downdip seismogenic limits, i.e., an updip aseismic zone tens of kilometers wide commonly limited by a temperature of $100^\circ$-$150^\circ$C. There is not yet agreement on the mechanism responsible. The downdip limit is frequently the intersection of the thrust with the fore-arc Moho, i.e., $\sim40$ km for continent subduction, less for island arcs. However, deeper thrust events have been observed in some regions. For very hot subduction zones a critical seismogenic temperature limit of $\sim350^\circ$C is reached at a shallower depth.
The reflection character of subduction thrust faults appears to change from a usually strong negative reflection in the updip aseismic zone, to a thin, sharp but weaker interface in the seismic portion, to a broad shear zone for the deeper aseismic zone. (7) Displacements on subduction thrust faults occur over a range of speeds, from earthquake rupture (seconds), to rates that generate tsunamis (minutes), to slower slip seen only in geodetic data. The speed controls are still unclear. (8) Immediately downdip of the seismogenic zone, slip on the aseismic zone in some areas occurs in slow slip events lasting a few weeks to months with intervals of a year to a few years. There are associated seismic tremors with no clear onset.

Introduction

Most of the world’s great earthquakes (M ≥ 8), many intermediate magnitude events, and most large tsunamis are generated by rupture on the “seismogenic zone” of subduction thrust faults (fig. 2.1). In this discussion, I outline some of the important seismic characteristics of subduction thrust faults and their physical associations; what I think we know and what we don’t know. Most of what we know has come from remote observation. There are no boreholes as yet into the seismogenic portion of subduction thrusts, although there has been drilling through the updip aseismic portion by the Ocean Drilling Program [e.g., Moore et al., 2001, and reference therein]. There also has been limited drilling through active land faults, and there is important information from exhumed ancient subduction thrust faults [e.g., Heermance et al., 2003; Hashimoto et al., 2002, and references therein]. In the next few years we hope to have much new data from regional and even global networks.

Some studies have found that earthquakes of magnitude M 9.0 or greater [e.g., Sato et al., 2004] are generated as the result of ruptures that propagate over a longer distance than M 8, and that the rupture谙on increases with magnitude. In contrast, other studies suggest that the rupture zone stays relatively constant as a function of magnitude. For example, the 1960 Chile earthquake [e.g., Sykes, 1969] was M 9.5 and traveled around 1000 km.

Since the 1960 Chile earthquake, studies of the earthquake rupture process have increased in magnitude and sophistication. A number of recent, large earthquakes have provided new insights into the rupture process. The 2011 Tohoku-Oki earthquake [e.g., Satake et al., 2012] was M 9 and provided a unique opportunity to study the rupture process in detail.
new data from deep drilling by the Japanese research ship CHIKYU. Seafloor precision geodetic systems now being developed also should make a very important contribution. What we know about the behavior of subduction thrust seismic zones, and associations of that behavior with the physical composition and state of the thrust, comes mainly from (1) great and smaller thrust earthquakes, (2) land geodetic data, (3) ocean drilling and exhumed subduction thrusts, (4) seismic reflection and wide angle data, (5) thermal data and models, (6) electrical sounding and other geophysical data, and (7) stress indicators near the plate boundary.

The seismic behavior of subduction thrust faults is highly variable, both regionally and locally. Some subduction zones generate events with a maximum magnitude of \( M \sim 7 \); others have great earthquakes of magnitude over \( M 9 \). Some subduction thrusts are largely aseismic between infrequent great earthquakes and are inferred to be fully locked. Some have very frequent but only small to intermediate magnitude earthquakes. Some subduction zones have very large seismic moment release; others have very little. Below, I discuss some of the subduction thrust observations and their physical explanations, along with some of the more important questions about the seismic behavior of subduction faults. I do not discuss the seismological characteristics of great earthquakes in any detail; there are a number of excellent reviews of that subject [cf. Kanamori, 1986, 1983; Nishenko, 1991; Pacheco and Sykes, 1992; Ruff, 1996] and the nature of the great earthquake cycle and associated elastic and viscoelastic deformation [Wang, this volume]. I focus on the physical associations on the thrust with the characteristics and variations of great and smaller thrust earthquakes.

## Subduction Thrust Earthquakes: Maximum Magnitude and "Seismic Coupling"

Some subduction zones generate thrust earthquakes of magnitude greater than \( M 9 \), whereas others have maximum earthquake magnitudes of \( M \sim 7 \) (fig. 2.2). The largest historical earthquakes, \( M \sim 9 \), have occurred in the subduction zones of southern Chile, Cascadia, southern Alaska, the Kuril trench, and north Sumatra [e.g., Pfauffer, 1969, 1972; Abe and Kanamori, 1980]. In contrast, only earthquakes of magnitude less than 7 or 7.5 have been recorded for most of the southwest Pacific island arcs; the smallest maximum, \( M \), appears to be for most of the Mariana-Izu-Bonin subduction zones [e.g., Pacheco and Sykes, 1992; Pacheco et al., 1993].

Since the magnitude of earthquakes increases systematically with the fault rupture area [e.g., Wells and Coppersmith, 1994], this difference means that subduction thrusts producing \( M < 7.5 \) earthquakes probably have seismic behavior only in small patches, at most a few tens of kilometers across, whereas \( M 9 \) earthquakes have seismic rupture that may be over areas of \( \sim 100 \) km.
Some subduction thrusts produce $M \approx 9$ earthquakes, others only less than $M \approx 7$; what are the controls: maximum fault area; seismic coupling; seismic versus aseismic slip?

down-dip and ~1000 km along strike, such as Alaska 1964 [e.g., Plafker, 1969, 1971], Chile 1960 [e.g., Plafker and Savage, 1970], and northern Sumatra 2004 events. The conclusion that subduction thrusts with only $M_x \approx 7$ earthquakes have only small patches that are seismic is supported by their generally small seismic moment release rate; their seismic efficiency ("seismic coupling," see below) is very small [e.g., Pacheco et al., 1993]. Most of the plate convergence is inferred to be accommodated aseismically. Although they may be very infrequent, the large events represent a much greater seismic moment release rate than the many $M < 7$ earthquakes. The subduction zones that have exhibited great $M \approx 9$ earthquakes are found to have seismic efficiency close to 1 [e.g., Pacheco et al., 1993]. Plate convergence is accommodated mainly seismically over a defined down-dip width. For most of these great earthquake regions, land geodetic data also require almost complete thrust locking be-
tween earthquake events; Cascadia, southwest Japan, and Chile are examples [e.g., Hyndman and Wang, 1995; Hyndman et al., 1995; Brooks et al., 2003].

"Seismic coupling" is an expression that needs careful definition [Wang and Dixon, 2004]. It usually refers to the fraction of the plate convergence rate at a subduction zone that is accommodated in thrust earthquakes [e.g., Ruff and Kanamori, 1983; Peterson and Seno, 1984; Jarrard, 1986; Pacheco et al., 1993; McCaffrey, 1997]. The remainder is inferred to be accommodated by some form of aseismic slip. In a kinematic description, a locked or fully coupled fault has no or low slip between great earthquakes. It is important to recognize that this definition does not involve any inferences of stress condition or fault properties [Wang and Dixon, 2004]. Partial coupling usually refers to the fraction of plate convergence that is accommodated seismically. The aseismic motion may occur in post seismic transient slip, in steady creep motion, or in slow slip events (discussed below). In estimating the seismic component of plate convergence, it is important to recognize that some component of plate convergence may be accommodated by crustal shortening in the back arc and possibly the fore arc [e.g., Hindle et al., 2002; Mazzotti and Hyndman, 2002; Norabuena et al., 2004]. Also, transient slip must be considered, including tsunami earthquakes and post seismic slip of great events that are not included in the seismic moment [Wang, this volume].

A critical parameter in the calculation of seismic coupling that often is not emphasized is the down dip seismogenic width. The calculated coupling is inversely proportional to the assumed seismogenic width. A less ambiguous expression may be "seismic efficiency," the fraction of relative convergence motion that is accommodated seismically over a defined downdip fault width. However, in the section below I will follow the common use of "seismic coupling." It probably is preferable to use the less ambiguous quantity, the average seismic moment release rate per unit length of subduction zone, because the downdip seismogenic width is often poorly known. It is also important to recognize that subduction zones with only infrequent great earthquakes have poor statistical sampling. There may have been no such events in the historical record.

To calculate the seismic coupling, the seismic slip from the average seismic moment release rate is compared to the convergence rate to give a seismic coupling efficiency "$\alpha$" [e.g., Pacheco et al., 1993, and references therein]. The coupling $\alpha = 1$ if all convergence is accommodated in earthquakes over the defined downdip seismogenic width. However, it is important to recognize that the calculation of $\alpha$ involves a number of poorly know parameters, especially the critical variable, the downdip width of the seismogenic zone "$W$." The computed seismic coupling is inversely proportional to the choice of $W$. The width $W$ has often been fixed, taking the updip limit near the trench and the downdip limit at depth 40–50 km (i.e., width of ~100 km) based on a common maximum rupture depth of great earthquakes [e.g., Pacheco et al., 1992; Tichelaar and Ruff, 1993]. In the discussion below I will use the expression
"apparent seismic coupling" for use where the downdip seismogenic width is unconstrained and is assumed.

Three explanations have been proposed for the apparent variation in seismic coupling, one related to seismogenic fault area, one to stress state, and one to frictional conditions. The first is that the downdip width of the seismogenic or coupled zone is wide where there are great earthquakes and very narrow where there are only smaller thrust earthquakes [e.g., Hyndman et al., 1997]. Within the defined downdip width \( W \), the thrust then is fully or almost fully seismically coupled in both cases (seismic efficiency of 1). The rupture length of great earthquakes along the margin is commonly 2 to 3 times the downdip width [e.g., Tichelaar and Ruff, 1993] (there are important exceptions with very long ruptures along strike, such as Cascadia, southern Chile, and Sumatra) so the maximum magnitude decreases strongly with decreasing downdip width. The maximum downdip depth may be the most important variable (see below). This explanation is supported by a maximum earthquake depth of \( \sim 10 \) km for the subduction zones such as Mariana with \( M_x \sim 7 \), compared to 40–50 km for the subduction zones with great earthquakes such as southern Chile and Alaska, \( M_x \sim 9 \) [e.g., Pacheco et al., 1993; Zhang and Schwartz, 1992; Tichelaar and Ruff, 1993]. If, for subduction zones like Mariana, the updip and downdip seismogenic limits approximately coincide, only a few patches will be seismic. This explains, at least in part, the very small seismic moment release rate and is one limiting case explanation. The other limiting case is that thrust earthquakes are distributed over a broad width to a depth limit of \( \sim 40 \) km but that only a few patches are seismic, producing only intermediate maximum magnitude earthquakes (see discussion of Wang, this volume). The updip and downdip seismogenic limits generally are nearly constant along strike for each subduction zone segment where there are no significant changes in plate age, plate convergence rate, and plate dip. The limited local variation in these limits suggests that they are related to thrust physical conditions or state common to significant subduction zone lengths along strike. These limits are discussed in the next section.

The second explanation for variations in apparent seismic coupling and for subduction zone maximum earthquake magnitude focuses on differences in the stress regimes, for example between southwest Pacific arcs and east Pacific continental subduction zones [e.g., Scholz and Campos, 1995]. The former are inferred to be in extensional regimes with the arcs receding from the trenches and subducting slabs, and the latter in compression with the continents overriding the subducting slab [Hyndman, 1972; Uyeda and Kanamori, 1979]. Scholz and Campos [1995] argue that extensional regimes have only small earthquakes; in contrast, compressional regimes have large megathrust earthquakes. However, the west Pacific subduction zones of Kuril, northeast Japan, and southwest Japan all have had great earthquakes. Also, the apparent seismic coupling factor is high in the southwest Japan subduction zone and lower in the northeastern Japan subduction zone [Astiz et al., 1988], whereas the force interaction as reflected by upper plate stress is opposite [Wang and Suyehiro,
The Seismogenic Zone of Subduction Thrust Faults

The seismogenic width is

1999; Wang and Dixon, 2004). For the areas with large earthquakes, subduction occurs beneath thick continental crust. A more consistent association seems to be that small maximum magnitudes are associated mainly with island arcs whereas great earthquakes are associated with subduction beneath continental crust (i.e., “Andean” type subduction). The latter association is at least partly explained above; the fore-arc crust in island arcs is thin so the fore-arc Moho and aseismic fore-arc mantle behavior are reached at a shallow depth.

The third explanation for variations in apparent seismic coupling is variations in fault properties, i.e., frictional state, or force required for rupture. An association that has not been well quantified is that very smooth incoming oceanic plates have infrequent but very large events (up to M 9), whereas rough incoming plates with seamounts, fracture zones, etc. that cause stress concentrations usually have smaller more frequent thrust earthquakes (M < 8) [e.g., Bilek et al., 2003]. Thrust irregularities such as seamount chains and aseismic ridges may especially limit the along-strike lengths of subduction thrust earthquakes and their character [e.g., Kodaira et al., 2002]. A related association is that great earthquakes commonly occur where there are large accretionary sedimentary prisms [e.g., Ruff, 1989]; the latter may smooth the subduction thrust such that there are few stress concentrations and strain can build up to very large ruptures. An example is the larger apparent coupling for the southwest Japan, Nankai trough, compared to the northeast Japan trench [e.g., Astiz et al., 1988; Pacheco et al., 1993]. Other examples of subduction zones with great thrust earthquakes and large accretionary prisms are southeastern Alaska, Cascadia, and southern Chile [e.g., Oleskevich et al., 1999]. However, large accretionary prisms only occur for continental subduction zones where there is a large source of sediment, so the main association of large earthquakes may be with subduction beneath continents rather than with accretionary prisms.

Subduction Thrust Asperities

“Seismic asperity” is another expression that is used in a number of different ways and needs careful definition (see also Wang [this volume]). The simple and unambiguous use is for patches or regions within the overall area of rupture in a specific great earthquake that has especially large displacement [e.g., Lay and Kanamori, 1981; Lay et al., 1982]. The earthquake slip distribution may be mapped through seismic waveform modeling, and the updip region in a few cases by tsunami modeling [e.g., Johnson, 1999; Tsunitaka et al., 2002]. Confusion comes with the extension of this definition to associations with spatial variations in the physical properties of the thrust interface (fig. 2.3). Asperities are often inferred to be “stronger” than the surrounding region of the thrust, and therefore they accommodate most of the plate convergence seismically, whereas adjacent areas have more aseismic slip (cf. reviews by Scholz [1990] and Ruff [1992]). Although a number of clear associations have been suggested between the patterns of subduction thrust seismic behavior and physical
features on the incoming plate such as seamounts, aseismic ridges, and fracture zones [e.g., Bilek et al., 2003; Kodaira et al., 2002], only few associations of local variations in rupture displacement in great earthquakes with localized physical features on thrusts have been conclusively established [e.g., Igarashi et al., 2003; Zweck et al., 2002].

If there is a physical connection with areas of large rupture, we expect that there will be more slip in these asperity regions for all great earthquake ruptures that include them and that adjacent areas will always have more aseismic slip. This behavior is opposite to the concept of “seismic gaps” in which areas of little or no recent rupture displacement are expected to have greater rupture in future events [McCann et al., 1979; Wyss and Wiemer, 1999, and references therein]. In the latter case, the long-term average earthquake slip may be nearly constant along a subduction zone. In one model, there may be fully locked patches surrounded by regions that are freely slipping. A number of repetitions of great earthquakes in the same region are required to establish which view is more correct. There have been only a few reported repeat ruptures of the same area by historical great earthquakes [e.g., Wyss and Wiemer, 1999; Schwartz, 1999], and the first of the two repeated events is usually too old for the high-quality seismic data required to model the rupture distribution. New geodetic data may allow connection between areas on the thrust that are locked and large earthquake rupture asperities.

Weak Subduction Thrust Faults

It has been recognized for many years that shallow angle thrust faults must be weak to accommodate large subhorizontal displacements [Hubbert and Rubey,
1959; Raleigh and Evernden, 1981; Davis et al., 1983). Other arguments based on regional stress and force balance [e.g., Wang and He, 1999], and on the lack of thermal anomalies due to frictional heating, have been developed subsequently [e.g., Wang et al., 1995, and references therein] (fig. 2.4). On some margins the maximum principal stress is margin parallel rather than in the direction of convergence [e.g., Wang et al., 1995]. Margin-normal stress on several margins is determined to be less than or equal to the vertical stress [e.g., Wang and Suwahiro, 1999; Wang et al., 1995]. The conclusion of weak large faults also applies to continental strike-slip faults [e.g., Zoback et al., 1987; Hickman, 1991; Lachenbruch and McGarr, 1990]. The arguments for weak strike-slip faults comes first from the maximum horizontal principal stress approaching orthogonal (65°–85°) to the fault as the fault is approached, suggesting that the fault is moving at very low levels of shear stress. Second, weak strike-slip faults are concluded from the lack of the thermal anomaly expected for motion on strong faults. The inferred strength of large faults is <20 MPa, similar to the stress relieved in large earthquakes (stress drop). In contrast, the overall strength of the brittle upper lithosphere is estimated to be much greater, >50–100 MPa, as predicted by friction laws [e.g., Hickman, 1991; Zoback and Townend, 2002].

The reasons for the weakness of large faults have been much discussed. The three main possibilities are (1) elevated fluid pressures [e.g., Magee and Zoback, 1993]; (2) low coefficients of friction due to the fault zone material, i.e., clay, serpentine, silica gel [e.g., Vrolijk, 1990; Toro et al., 2004]; and (3) dynamic weakening, i.e., slip shear heating, propagation of dilation waves along the fault, and fluidization of fault-zone material [e.g., Brune, 1993; Shi et al., 1998; Brodsky and Kanamori, 2001; Ma et al., 2003; Melosh, 1996; Lachenbruch, 1980]. The elevated fluid-pressure explanation appears favored by most authors [see Davis et al.,

![Image](image_url)

**Figure 2.4** Subduction thrust faults are very weak based on regional stress, earthquake, and thermal data.
especially for subduction thrust faults, following the well-established ac-
ceptance of this explanation for the low angle of thrust faults in sedimentary
fold and thrust belts [Hubbert and Rubey, 1959]. For subduction thrusts, there is
an ample supply of fluid from the consolidation of underthrust material and
from dehydration reactions in the downgoing slab. The fluid supply for contin-
ental transcurrent faults is less obvious, but some suggestions have been made
[e.g., Kirby et al., 2002]. Secure confirmation awaits drilling of the seismogenic
portion of subduction thrust faults by the International Ocean Drilling Pro-
gram (IODP) drill ship CHIKYU and of active continental faults such as the San
Andreas fault by the San Andreas Fault Observatory at Depth (SAFOD) project.

Updip Limit of Subduction Thrust
Earthquakes

Both great subduction earthquakes and smaller thrust events usually do not
extend to the trench (fig. 2.5); there is an updip aseismic zone commonly tens
of kilometers wide (fig. 2.4) [e.g., Byrne et al., 1888; Byrne and Fisher, 1990]. This
 updip limit, or upper stability transition depth, is defined by (see summary by
Oleskevich et al. [1999]) (1) the updip rupture limit in great earthquakes, as de-
termined by waveform modeling; (2) the updip limit from modeling of the tsu-
namis generated by great earthquakes; (3) the updip limit of great earthquake
aftershocks; and (4) the updip limit of small thrust earthquakes on the subduc-

Figure 2.5  Great earthquakes usually do not rupture to the trench; there is an updip
aseismic zone. This limit is important for tsunami generation.
tion thrust between great events. In several subduction zones the coast is close enough for land geodetic data to provide some constraint to the updip limit of the locked zones, although the resolution is low [e.g., Lundgren et al., 1999; Norabuena et al., 2004]. Seafloor geodetic measurements that are in progress by United States and Japanese groups should soon give additional information [e.g., Spiess et al., 1998]. All of these definitions are based on different measures and sometimes give different locations [e.g., Norabuena et al., 2004], but usually they are consistent within the common horizontal resolution of 10–20 km [e.g., Oleskerich et al., 1999]. Correlation of seismic properties with composition and state changes is very dependent on how well this updip rupture or seismogenic limit can be defined. The accuracy varies greatly for different subduction zones, with horizontal definition relative to the trench ranging from about ±5 km for well-studied subduction zones with recent great earthquakes to only very generally (±30 km) for others. Where the limit has been determined, it ranges from ~30 km from the trench for the Nankai, southwest Japan subduction zone [e.g., Ando, 1975; Toshioka et al., 2002; Obana et al., 2001], which is young and hot, to ~100 km for the cool shallow dipping southeast Alaska subduction zone [e.g., Alpern et al., 1969; Johnson et al., 1996].

Because tsunami wave analyses show that the updip aseismic zone usually does not contribute significantly to tsunami generation, the motion in this region must be slower than over the few tens of minutes of the tsunami period. There is inadequate seafloor geodetic data as yet to determine if the updip aseismic zone moves mainly as a transient in the hours to days following great events or there is some continuous motion between great events. Wang (this volume) has pointed out that if the adjacent deeper part of the thrust is locked, there is no force to drive motion on the updip portion. The nature and timing of motion on the updip portion of the subduction thrust is an important application of high-precision seafloor geodetic data for the future. In a few events, this updip zone appears to rupture independently of the deeper seismic zone. The slip is over a time interval of a few minutes to tens of minutes such as to generate a large tsunami but to radiate little seismic energy, i.e., a tsunami earthquake. No clear special physical characteristics have as yet been identified for subduction zones having tsunami earthquakes, but they are too infrequent to as yet allow ready correlations [e.g., Okal and Newman, 2001].

The first explanation for the updip limit that was suggested is the reasonable one: the part of the thrust in contact with accreted sediments is aseismic [e.g., Byrne et al., 1988]. Seismic behavior should then start landward of the thrust contact with the crystalline crust of the overlying fore arc, i.e., the “backstop.” At least for smaller prisms, the sediments are generally quite unconsolidated and should have little strength to support elastic strain buildup. However, it is now clear that at some continental subduction zones, great earthquake rupture occurs at or near the base of large accretionary sedimentary prisms. At a few subduction zones that have great earthquakes, most if not all of the seismic rupture zone underlies accreted sedimentary material, i.e., Cascadia, southern
Alaska, and southwest Japan [e.g., Oleskevich et al., 1999]. For other subduction zones where there also are great earthquakes, there is little accreted sediment and substantial undercutting erosion of the fore-arc crust is inferred, i.e., most of South America and northeast Japan [e.g., von Huene and Scholl, 1991]. Thus the updip limit does not appear to be constrained primarily by downdip changes in the bulk physical composition of the overlying fore arc. Variations in the material within the fault zone décollement, however, may be important [e.g., Cloos and Shreve, 1996].

Where there are large accretionary prisms, there is the complication that part of the plate convergence is taken up by shortening of the seaward portion of the prism rather than by slip on the main subduction thrust. The shortening may be on splay thrusts ("out of sequence thrusts") in folding and in bulk shortening. Some earthquake ruptures may rise to the seafloor on faults landward of the trench, giving an updip limit in earthquake and tsunami modeling of that event (e.g., Park et al. [2000] for southwest Japan and Plafker [1972] for southern Alaska). However, most of the long-term thrust motion, including aseismic slip, must extend to near the trench because of the geometric constraint that most of the sediment thickening must occur near the toe of the sedimentary prism.

Two other associations of the updip limit that have been suggested are depth, i.e., pressure-controlled consolidation and temperature. Depth and temperature are roughly correlated for many subduction zones, but as discussed below, there are substantial variations in vertical temperature gradients due to incoming plate age, thickness of insulating sediments on the incoming plate, plate convergence rate, etc. [e.g., Oleskevich et al., 1999]. Thermal models for subduction thrust temperatures are discussed below. For most subduction zones the updip limit for great earthquake rupture appears to correlate best with thrust temperature, with a critical temperature of 100°–150°C [e.g., Vrolijk, 1990; Hyndman et al., 1997] rather than pressure (depth). This conclusion is supported by a recent study to define the updip seismogenic limit and along-strike variations in temperature for the Costa Rica subduction zone by Harris and Wang [2002], Spinelli and Saffer [2004], and Norabuena et al. [2004]. They found a good correlation between variations in the thermal state of the incoming oceanic plate and the depth to the updip seismogenic limit defined by microearthquakes. There is a sharp boundary in the incoming oceanic plate heat flow along the margin that is well correlated with a change in the distance from the trench and the depth of the microearthquake updip limit. Along strike across the boundary, there is a significant change in the seismicity updip limit, but the estimated temperature at the limit on both sides of the boundary is 100°–150°C.

The physical explanation for a temperature-controlled updip limit is still much debated [e.g., Moore and Saffer, 2001; Saffer and Marone, 2003]. The earlier favored explanation was temperature-dependent clay dehydration [Vrolijk, 1990; Hyndman and Wang, 1993] following a proposal that this mechanism was
responsible for the lack of shallow earthquakes in land faults [e.g., Marone and Scholz, 1988]. Laboratory data show that the transition from smectite to illite/chlorite occurs in the appropriate temperature range of 100°–150°C [e.g., Moore and Vrolijk, 1992; Chamley, 1989]. Also, laboratory data show smectite clays to be very weak [e.g., Wang and Mao, 1979]. The slip behavior of smectite and illite/chorite appear to be complex, but recent laboratory studies of their sliding behavior do not confirm the transition from stable-sliding velocity strengthening to stick-slip velocity weakening behavior [Saffer and Marone, 2003]. Other temperature-dependent mechanisms are therefore being examined. This limit is important for tsunami generation and thus the hazard at inland cities. Possible mechanisms include silica and carbonate diagenesis and consolidation changes of permeability that control pore-fluid pressure.

### Downdip Limit of Subduction Thrust Earthquakes

The downdip limit, or lower stability transition depth, on subduction thrust faults (fig. 2.6) is important for the closest approach of the seismic source and thus the hazard at inland cities. This limit is defined by (1) the downdip rupture limit of great earthquakes as determined by waveform modeling, (2) the downdip limit of great earthquake aftershocks, (3) small thrust earthquakes occurring between great events, (4) land geodetic data that define the rupture downdip limit, and (5) geodetic data that provide constraint to the downdip limit of

![Figure 2.6](image)

**Figure 2.6**  Great earthquakes have a variable maximum depth of rupture ~10–50 km, with deeper aseismic motion. This limit is important for closest approach and thus seismic hazard at inland cities.
the interseismic locked zone. Correlation of earthquake characteristics with downdip composition and state changes is very dependent on how well this limit can be defined. Moreover, there appears to be a transition zone of substantial width between fully seismic and fully aseismic zones [e.g., Wang et al., 2003]. The accuracy of the downdip limit varies greatly for different subduction zones, depths resolved to approximately ±5 km (10–20 km in horizontal distance) for well-studied margins, to ±10 or 20 km (~50 km horizontal) for others (see summary by Oleskevich et al., [1999]). The downdip limits range from ~10 km for the Mariana subduction zone to ~50 km for some continental subduction zones. The consistent downdip depth limit along the length of a number of subduction zones gives encouragement that the limit is controlled by some physical composition or state change downdip that varies regionally rather than random variations in stress or state. A good example is the South America subduction zone where the maximum depth of rupture in great earthquakes and of the interseismic locked zone is consistently between 40 and 50 km [Tichelaar and Ruff, 1991; Oleskevich et al., 1999]. Two downdip state and composition changes have been suggested for the limit. The downdip seismic limit for most, but not all, subduction zones appears to agree with either a maximum temperature of 350°C or the thrust intersection with the fore-arc serpentinitized mantle [e.g., Hyndman et al., 1997]. Further study is needed to determine if the apparent exceptions are due to inadequate knowledge, for example, of the depth of the fore-arc mantle intersection, or if other processes control the limit in these cases [e.g., Seno, 2005].

**Temperature Limit**

The downdip rupture limit for hot subduction zones subducting young oceanic crust appears to be temperature controlled at ~350°C. Detailed numerical thermal models specific to each subduction zone are required for accurate temperature estimates on the subduction thrusts. Examples where the temperature limit appears to apply are Cascadia [Savage et al., 1991; Hyndman and Wang, 1993, 1995], Mexico [Currie et al., 2002], southern Chile [Tichelaar and Ruff, 1993; Oleskevich et al., 1999], and southwest Japan [Hyndman et al., 1995]. The maximum temperature of 350°C agrees with the seismic temperature limits estimated for continental strike-slip fault zones. This temperature corresponds well to the transition from velocity-weakening (seismic) to velocity-strengthening (aseismic) behavior based on laboratory data for crustal composition rocks [e.g., Tse and Rice, 1986; Blanpied et al., 1991, 1995]. There may be a tapered transition zone downdip from full great earthquake rupture to no motion that extends to ~450°C. That temperature corresponds to a rapid increase in fault instantaneous shear stress in laboratory data for quartz-feldspathic rocks [Tse and Rice, 1986]. At higher temperatures the shear stress increases rapidly with increasing shear velocity, so rapid seismic slip should not occur. It ap-
Hydrated Fore-Arc Mantle Limit

In many areas the maximum seismogenic depth from the thermal limit is at great depths, much greater than those observed. In these cases the maximum depth usually appears to be close to the intersection of the subduction thrust with the fore-arc Moho [e.g., Ruff and Tichelaar, 1996; Hyndman et al., 1997; Oleskevich et al., 1999; ANCORP Working Group, 2003]. Accurate verification of the fore-arc mantle limit is subject to the problem of defining the depth of the intersection of the thrust with the fore-arc Moho. First, the fore-arc Moho itself is usually difficult to define because the Moho velocity contrast is small. This is interpreted to be because the P wave velocity of the fore-arc mantle is reduced by serpenitization to velocities similar to the crust [e.g., Christensen, 2004; Hyndman and Peacock, 2003, and references therein]. Second, there is evidence in a few places of the fore-arc Moho turning upward or turning downward close to its interaction with the subduction thrust fault. Therefore, even a well-determined regional depth to the fore-arc Moho does not provide assurance of defining the intersection accurately. Receiver function studies giving S-wave structure are providing much improved constraints across the thrust-Moho intersection [e.g., Bostock et al., 2002], and these structure questions should soon be better resolved. Another problem is the potential of sediment or of fore-arc crustal material being carried downward in a channel along the décollement. This process could result in the thrust contact being within felsic crustal composition material (i.e., seismic behavior) in places well below the fore-arc Moho contact.

An explanation for the fore-arc Moho seismic limit may be that the fore-arc mantle is serpenitized, providing an aseismic thrust plane [Hyndman et al., 1997; Peacock and Hyndman, 1999]. There is considerable data supporting the hypothesis that the fore-arc mantle contains substantial serpenitine [e.g., Hyndman and Peacock, 2003, and references therein]. The laboratory data, however, do not yet give a clear story of its seismic/aseismic behavior [Moore et al., 1997; Reinen et al., 1991], but the weak layered structure of serpenitine is such that seismic behavior is not expected. An additional factor that may make the thrust below the fore-arc Moho aseismic is that it may contain significant amounts of talc owing to rising silica-rich fluids from the dehydrating underlying slab [Peacock and Hyndman, 1999]. Talc has a very weak layered mineral structure, and it is unlikely to behave seismically.
Seismic and Aseismic Reflection Character of the Subduction Thrust

The subduction thrust exhibits substantial variations in multichannel seismic reflection character downdip that may allow mapping the updip and downdip seismogenic limits. Good-quality deep multichannel seismic reflection is available for only a few subduction zones, and the change from seismogenic to deeper aseismic behavior often occurs near the coast so reflection data must be combined. However, there is evidence for changes in thrust reflection character at both the updip and downdip seismogenic limits (fig. 2.7). Ocean Drilling Program (ODP) has shown that the thrust in the updip aseismic region is quite thin and, at least in sedimentary prisms, has high fluid pressure [e.g., Brown et al., 2003]. The multichannel reflection character in this updip region is variable from inferred positive to negative impedance contrasts that may relate to lateral pore-pressure variations [e.g., Shipley et al., 1992, 1994]. A clear change in décollement reflection amplitude has been observed for the southwest Japan subduction thrust from strong negative polarity to very weak at 30–40 km from the deformation front; this is close to the updip limit from other constraints [Bangs et al., 2004]. This is also the position of an increase in accretionary wedge taper, indicating an increase in fault frictional strength.

![Image](image.png)

**Figure 2.7** The subduction thrust seismic reflection image appears to be sharp where seismic and a broad band deeper where aseismic.
Bangs et al. [2004] interpret these changes as due to loss of fluids and loss of excess fluid-pressure downdip.

Within the locked seismogenic zone, the multichannel seismic reflection images generally show a thin although weak seismic reflector and inferred sharp thrust contact. At greater depth beneath the downdip seismogenic limit, where resolved, the multichannel reflection images in several areas show a ~5 km thick reflective band that may be a shear zone [Nedimovic et al., 2003; ANCORP Working Group, 2003] above the downdropping plate. This layer may have high Poisson’s ratio indicative of high pore pressure [Cassidy and Ellis, 1991]. There is support for this shear zone thickening with depth in exhumed faults [e.g., Sibson, 1992]. The thrust reflection thickness change thus may allow mapping the downdip seismic-aseismic transition [Nedimovic et al., 2003].

Speed of Slip on Subduction Thrust Faults

Displacements on subduction thrust faults occur over a range of speeds, from earthquake rupture (seconds), to rates that generate tsunamis (minutes), to slower slip seen only in geodetic data. The speed controls are still unclear. Some slip events are too slow to generate much seismic energy (fig. 2.8) but fast enough to generate tsunami (“tsunami earthquakes”), i.e., slip duration of minutes [e.g., Kanamori, 1972; Pelno and Wiens, 1992; Okal and Newman, 2001; Bilek and Lay, 2002]. Tsunami events are primarily defined for the shallower updip portions of subduction thrusts where slip can generate tsunamis. A few slip events are still slower and are seen only in geodetic data [e.g., Sato et al., 2004]. In this case, slower slip events are defined mainly on the deeper

---

Figure 2.8 Some thrust events are fast (seismic), some are intermediate (tsunami earthquakes), and some are slow (seen only by geodetic measurements).
part of subduction thrusts that affect land geodetic data. Models for fault slip behavior must accommodate this range of event slip rates. Such models also must accommodate the evidence noted above that subduction thrust faults are very weak, from thermal data and models, from regional stress estimates, and from earthquake data. High fluid pressures that reduce the effective fault normal stress are suggested, as discussed above.

**Downdip Slow Slip Events and Seismic Tremor**

Recent GPS geodetic data have indicated slow slip events downdip of the thrust seismogenic zone, not associated with normal abrupt onset earthquakes (fig. 2.9) [Dragert et al., 2001; Lowrey et al., 2001; Miller et al., 2002; Ozawa et al., 2002]. The subduction thrust just downdip of the seismogenic zone also has been found to exhibit seismic tremor [e.g., Obara, 2002]. On one subduction zone, the seismic tremor has now been associated with the slow slip events, i.e., Episodic Tremor and Slip (ETS) [Rogers and Dragert, 2003]. The mechanism of slow slip is beginning to be examined [Shibazaki and Yoshihisa, 2003; Yoshida and Naoyuki, 2003]. The location of the slip and tremor events may provide a new method for defining the downdip end of the locked seismogenic zone. Of special significance is that each of these slow slip events loads the locked part of the thrust farther updip. Therefore the time interval of their occurrence represents a period of increased probability of great earthquake rupture. This time-dependent risk is now being examined [Mazzotti and Adams, 2004].

![Diagram of tremor and slow slip](Figure 2.9) Slow slip and seismic tremor on the thrust downdip of the seismogenic zone.
Conclusions

Until recently, most of what was known about the seismic behavior of subduction thrust faults came from teleseismic studies of the earthquakes that they generate. However, there have been great advances recently in characterizing and understanding earthquakes on subduction thrust faults based on seismic reflection and wide-angle structure studies, ocean bottom seismograph recording of microearthquakes, geodetic data, thermal data, stress indicators, drilling the updip portion of the thrust, exhumed ancient subduction faults, and a wide variety of modeling work. Important seismic characteristics of subduction thrust faults and their physical associations include the following:

1. The subduction thrust maximum earthquake magnitude $M_x$ is highly variable among subduction zones; $M_x$ is at least partly related to the down-dip seismogenic width, i.e., up dip and down dip rupture limits, because these limits are a principal control on rupture area. Variations in the fault physical characteristics and stress on the fault also may be important. Continental subduction zones tend to have large maximum thrust earthquake magnitudes and island arcs usually have $M_x$ less than ~7.5.

2. The term "seismic coupling," i.e., fraction of relative motion that is accommodated seismically compared to by aseismic slip, needs careful definition. Meaningful use of the term requires specification of the down dip seismogenic width. Average moment release rate is a less ambiguous quantity. Some subduction zones appear to be completely locked with almost no seismgenic slip between great megathrust events (after post earthquake transients). These faults tend to have very few thrust earthquakes between great events. Other subduction zones have very small seismic moment release rates and are interpreted to have mostly aseismic slip. They usually have small maximum magnitudes and many small thrust earthquakes. Continental subduction zones with large maximum magnitudes usually have large average moment release rates. Island arcs with small maximum magnitudes tend to have only small rates of seismic moment release; the small moment rates may be due in part to very small down dip seismogenic widths.

3. The term "seismic asperity" also needs careful definition. There is no confusion if this term is used to describe fault areas with especially large slip in particular great earthquakes. However, inferences that such areas always have larger displacement in successive great earthquakes and that they are associated with special physical characteristics on the fault are not yet firmly established. There also have been associations of the locations and boundaries of great earthquakes and rupture limits with a number of fault physical characteristics such as seamounts, aseismic ridges, fracture zones etc., but again, in most areas, there have not been enough repeated great events to establish these associations with assurance.

4. Subduction thrust faults are concluded to be weak, i.e., the maximum shear strength is low, on the basis of stress arguments and the lack of a measur-
able effect in observed heat flow. This low strength is consistent with the small stress drop of large thrust earthquakes. The favored explanation is regionally elevated fluid pressures, but several other processes have been proposed, including weak fault materials and dynamic weakening in the rupture process.

5. Most continental subduction thrusts have consistent updip and down-dip seismogenic limits. There usually is an updip aseismic zone tens of kilometers wide. This limit commonly corresponds to a temperature of 100°–150°C. There is not yet agreement on the mechanism responsible; the dehydration of stable-sliding smectite clays to illite/chlorite occurs at about that temperature, but recent laboratory data have not supported this mechanism. Other chemical alterations that increase the permeability and allow lower pore pressure and seismic behavior at greater depth also have been suggested. The down-dip limit is frequently close to the intersection of the thrust with the fore-arc Moho, i.e., ~40 km for continental subduction, less for island arcs since their fore-arc crusts are usually much thinner. For very hot subduction zones the down-dip limit appears to be shallower and has been found to correspond to the critical seismogenic temperature limit of ~350°C found in laboratory studies of crustal rocks.

6. From a few observations the reflection character of subduction thrust faults appears to change down-dip from the updip aseismic zone, to the seismogenic zone, to the down-dip aseismic zone. A strong reflection, commonly negative, has been observed in the updip aseismic zone that has been associated with high pore pressure. In the seismogenic zone, there usually is a thin sharp but weaker interface. For the deeper aseismic zone a broad shear zone has been observed by deep reflection measurements in a few areas. Such a broad shear zone is observed in field studies of the deeper parts of exhumed faults.

7. Displacements on subduction thrust faults occur over a range of speeds, from earthquake rupture (seconds), to rates that generate tsunamis (minutes), to slower slip seen only in geodetic data. The speed controls are still unclear; variations in fault materials and pore pressure have been suggested.

8. Immediate down-dip of the seismogenic zone, the slip on the aseismic zone in some areas, occurs in slow slip events lasting a few weeks with intervals of a few years. There are associated seismic tremors with no clear onset. The mechanism for these episodic tremor and slip events is not yet well understood, but the times of slow slip may represent a period of higher probability for great thrust earthquakes.

Acknowledgments

I acknowledge the many discussions with colleagues and associates on the subject of the seismogenic zone of subduction thrust faults, including a series of meetings and advisory groups involved in the planning for drilling...
of subduction thrust faults by ODP and IODP. The MARGINS workshop on
the Seismogenic Zone of Subduction Faults was especially helpful. Thorough
reviews of the manuscript by C. Moore and T. Dixon were much appreciated.

References
Aki, K., and H. Kanamori (1980), Magnitudes of great shallow earthquakes from 1953 to 1977,
Algermissen, S. T., W. A. Rinehart, R. W. Sherburne, and W. Dillinger Jr. (1969), Preshocks and
aftershocks of the Prince William Sound earthquake of March 28, 1964, U.S. Coast Geod.
Surv., 211, 23–43.
ANCORP Working Group (2003), Seismic imaging of a convergent continental margin and
plateau in the central Andes (Andean Continental Research Project 1996 (ANCORP96)),
Ando, M. (1975), Source mechanisms and tectonic significance of historical earthquakes along
the Nankai Trough, Japan, Tectonophysics, 27, 119–140.
Astiz, L., T. Lay, and H. Kanamori (1988), Large intermediate-depth earthquakes and the
subduction process, Phys. Earth Planet. Inter., 53, 80–166.
Evolution of the Nankai Trough decollement from the trench into the seismogenic zone:
Inferences from three-dimensional seismic reflection imaging, Geology, 32(4), 273–276.
Bilck, S. L., and T. Lay (2002), Tsunami earthquakes possibly widespread manifestations of
Bilck, S. L., S. Y. Schwartz, and H. R. Deshot (2003), Control of seafloor roughness on
earthquake rupture behavior, Geology, 31(5), 455–458.
Blanpied, M. L., D. A. Lockner, and J. D. Byerlee (1991), Fault stability inferred from granite
Blanpied, M. L., D. A. Lockner, and J. D. Byerlee (1995), Frictional slip of granite at hydrothermal
Bostock, M. G., R. D. Hyndman, S. Rondenay, and S. M. Peacock (2002), An inverted Moho and
Brooks, B. A., M. Bevis, R. Smalley, E. Kendrick, R. Manceda, E. Lauria, R. Maturana, and
M. Araujo (2003), Crustal motion in the southern Andes (26–36 S): Do the Andes behave like
Brown, K., A. Kopf, M. Underwood, and L. Weinberger (2003), Composition and fluid pressure
controls on the state of stress on the Nankai subduction thrust: A weak plate boundary,
Byrne, T., and D. M. Fisher (1990), Evidence for a weak and overpressured decollement beneath
Byrne, D. E., D. M. Davis, and L. R. Sykes (1988), Locs and maximum size of thrust earthquakes
and the mechanics of the shallow region of subduction zones, Tectonics, 7, 833–857.
Cassidy, J. F., and R. M. Ellis (1991), Shear wave constraints on deep crustal reflective zone


Hyndman, R. D., M. Yamano, and D. A. Oleskevich (1997), The seismogenic zone of subduction thrust faults, Island Arc, 6, 244–260.


Sibson, R. H. (1992), Implications of fault valve behavior for rupture nucleation and recurrence, Tectonophysics, 211, 283–293.


Wang, K., Elastic and viscoelastic models of crustal deformation in subduction earthquake cycles, this volume.


