Delayed formation of the equatorial ridge on Iapetus from a subsatellite created in a giant impact

Andrew J. Dombard,1 Andrew F. Cheng,2 William B. McKinnon,3 and Jonathan P. Kay1

Received 20 October 2011; revised 5 January 2012; accepted 6 January 2012; published 7 March 2012.

[1] The great equatorial ridge on Saturn’s moon Iapetus is arguably the most perplexing landform in the solar system. The ridge is a mountain range up to 20 km tall and sitting on the equator of Iapetus, and explaining its creation is an unresolved challenge. Models of its formation must satisfy three critical observations: why the ridge (1) sits exactly on the equator, (2) is found only on the equator, and (3) is thus far found only on Iapetus. We argue that all previously proposed models fail to satisfy these observations, and we expand upon our previous proposal that the ridge ultimately formed from an ancient giant impact that produced a subsatellite around Iapetus. The orbit of this subsatellite would then decay, once Iapetus itself had despun due to tides raised by Saturn, until tidal forces from Iapetus tore the subsatellite apart. The resultant debris formed a transient ring around Iapetus, the material of which rained down on the surface to build the ridge. By sequestering the material in a subsatellite with a tidally evolving orbit, formation of the ridge is delayed, which increases the likelihood of preservation against the high-impact flux early in the solar system’s history and allows the ridge to form on thick, stiff lithosphere (heat flow likely <1 mW m⁻²) required to support this massive load without apparent flexure. This mechanism thus explains the three critical observations.


1. Introduction

[2] Iapetus has proven to be one of the most peculiar bodies in the solar system. Up until recently, this third largest moon of Saturn (mean radius of 734.3 km [Thomas, 2010]) was most notable for the large semimajor axis of its orbit around Saturn (~59 Saturn radii) and its hemispheric albedo dichotomy, now thought to be a product of spatially variable ice mobilization on its surface [Spencer and Denk, 2010]. Beginning in 2004, however, NASA’s Cassini spacecraft obtained the first high-resolution images of Iapetus, revealing a world even more peculiar than initially thought [e.g., Porco et al., 2005; Denk et al., 2010].

[3] Iapetus possesses a distinctly pronounced oblate spheroid shape, with an equatorial radius greater than the polar radius by 33.6 ± 2.8 km [Thomas, 2010]. This difference translates into a flattening of ~4.5%, compared with Earth’s 0.3% rotational flattening. For a body in hydrostatic equilibrium, an equatorial bulge of that size would indicate that Iapetus should be spinning once every ~16.5 h [Castillo-Rogez et al., 2007]; this moon, however, is synchronously locked to Saturn, rotating (and revolving) once every ~79 days. Models for Iapetus’s flattened shape include a fossilized rotational bulge [Castillo-Rogez et al., 2007; Robuchon et al., 2010] and long-wavelength, axisymmetric deformation of the lithosphere [Sandwell and Schubert, 2010; Kay and Dombard, 2011].

[4] Even more bizarre than the bulge is Iapetus’s peerless equatorial ridge [Porco et al., 2005]. Arguably one of the most astonishing features in the solar system, it is so big that it is clearly visible in global views of the moon (Figure 1). This mountain range is up to 20 km high, 200 km wide (translating into a mass of order 0.1% the total mass of Iapetus), and sits perfectly straight, exactly on the equator. The ridge runs >75% of the circumference of the satellite [e.g., Singer and McKinnon, 2011], though not continuously, and has been modified by subsequent impacts and mass wasting (i.e., landslides [Singer et al., 2009]). The cross-sectional shape is in places trapezoidal, with a flat top and sometimes a central trough, and with slopes of ~15° [Giese et al., 2008]. Notably, the ridge appears to be supported by the lithosphere without an obvious flexural signal [Giese et al., 2008; Dombard and Cheng, 2008] (see also below), and the rest of the surface of Iapetus is dominated by impact craters, with only a few examples of other geomorphic features that are far less impressive in scale than the ridge [Singer and McKinnon, 2011]. The ridge is heavily cratered and thus appears ancient [Denk et al., 2010].

[5] Clearly, the formation of the ridge was one of the key events in the evolution of Iapetus. In this paper, we explore...
models of its origin. Any model must explain why the ridge (1) sits exactly on the equator, (2) is found only on the equator, and (3) is thus far only found on Iapetus. First, we explore the flexural support of the ridge. Next we review past models or scenarios, arguing that they do not explain the three critical observations. Then, we propose a new model for the formation of the ridge [see Dombard et al., 2010] and discuss the implications.

2. The Ridge and Lithospheric Flexure

[6] Giese et al. [2008] and Dombard and Cheng [2008] noted the lack of an obvious flexural signal, indicating strong lithospheric support of the ridge. We expand on this notion here. From a near global topographic model of Iapetus derived from stereo imagery [Schenk, 2010], we extract pole-to-pole topographic profiles every 5° of longitude from 140° to 170° W longitude (Figure 2), which is the most continuous, best developed portion of the ridge and thus should best show any flexure. From the individual profiles and from the average, no flexural signal is obvious. Low areas are seen on either side of the ridge, but their relation to lithospheric flexure is not apparent, as the troughs could simply be due to high-standing topography peripheral to the ridge. In any event, lithospheric flexure under the ridge would be expected to be symmetric across the equator, which clearly this topography is not. On the other hand, a flexural signal with a magnitude of ~1 km could be hidden on Iapetus, because large-amplitude, long-wavelength topography is prevalent on this satellite [Schenk, 2010], as is evident in Figure 2.

[7] We simulate the deformation of the lithosphere of Iapetus under a ridge load, using the commercially available MSC.Marc finite element package, which we have used many times in the study of icy satellite geodynamics [e.g., Dombard and McKinnon, 2000, 2006a, 2006b; Dombard et al., 2007; Kay and Dombard, 2011; Dampsz and Dombard, 2011; A. J. Dombard et al., Flanking fractures and the formation of double ridges on Europa, submitted to Icarus, 2011]. We simulate a plane-strain system (which neglects membrane support; justified below) of water ice, 400 km deep (roughly the depth to a rocky core if Iapetus is differentiated, estimated by considering the mass of Iapetus and the densities of the component rocky and icy materials) and 1200 km wide, subdivided into 6000 quadrilateral elements (120 evenly spaced elements horizontally and

Figure 1. This global view of Iapetus from NASA’s Cassini spacecraft shows the great equatorial ridge, a mountain range up to 20 km tall and 200 km wide that sits perfectly on the equator. Image PIA06166 from the NASA Planetary Photojournal (http://photojournal.jpl.nasa.gov), courtesy NASA/JPL-Caltech.

Figure 2. The topography of Iapetus [see Schenk, 2010]; elevations are given with respect to the global biaxial figure. (a) A near global model shown in a cylindrical projection and centered on 180°W longitude. The blue lines mark the zone from which profiles are taken. (b) Pole-to-pole (positive north) elevation profiles, extracted every 5° from 140° to 170°W longitude. (c) Pole-to-pole elevation profiles. The gray lines are the individual profiles, and the black line is the average. No signal of lithospheric flexure (i.e., symmetric flanking troughs) is obvious.
1m Wm

Figure 2) suggests the heat flow has always been less than 0.22 m s

that scales with the surface gravitational acceleration of

and bottom boundaries. Gravity is applied as a body force

1 mm [see sensitive to the ice grain size; we assume a grain size of

grain-boundary sliding and grain-boundary diffusion, are

of Gammon et al.

brittle faulting). We assume the elastic properties of water ice

do not exist in a state of collapse. A zone, 80 km deep. Our results demonstrate that the ridge would
effectively exist in a state of collapse. A zone, 80–160 km
distant from the equator, of high plastic strains (up to
5.5%) marks in essence a hinge fold where the lithosphere
is breaking. Consequently, almost 90% of the initial height
is lost as the thin lithosphere founders beneath this massive
load. Producing a final profile with the ridge over 15 km tall
would require a larger load, exacerbating the state of col-
lapse. Thus, models that appeal to epochs of high heat flow
are not consistent with the lithospherically supported state
of the ridge.

[13] Conversely, the ridge can be supported when the heat
flow is lower. For a heat flow of 3 mW m\(^{-2}\), we predict a
flexural trough comparable in depth to the topographic low
on the south side of the ridge (compare the low at 400 km
distance from the ridge in Figure 2), but this topographic
low is farther from the ridge than we predict and is not
symmetric across the equator, which would be expected if
the observed flanking lows were flexural in origin. To sup-
port the ridge yet limit the depth of the flexural trough to
<1 km would require lower heat flows <1 mW m\(^{-2}\), condi-
tions that would have a long horizontal scale where the
deforation would reach to the poles, violating our planar
strain assumption.

[14] Alternate thermophysical parameters do not change
this conclusion. Simulations exploring a reasonable range in
ice grain size (0.1–10 mm) show comparable flexure (trough
depths within a few hundred meters), demonstrating that

Figure 3. Results from finite element simulations of the
lithosphere of Iapetus loaded by the equatorial ridge. The
dashed line is the initial shape. The thin solid line is for a heat
flow of 3 mW m\(^{-2}\), while the thick line is for a heat flow of
18 mW m\(^{-2}\). The lack of an obvious flexural signal (compare
Figure 2) suggests the heat flow has always been less than
~1 mW m\(^{-2}\) since the ridge was created.

50 elements vertically, with a bias to concentrate more ele-
ments near the surface). The ridge is not directly included as
part of the mesh. The thermal state is that of conductive
passage of a basal heat flux to a constant surface temperature of
90 K, employing the thermal conductivity of solid water
ice of Klinger [1981], which is inversely proportional to
temperature. (We assume any bulk heating occurs below the
lithosphere, such that our region of interest is simply passing
heat conductively.)

[8] The rheology generally has three components: elastic
(which ultimately provides the strength of the lithosphere),
viscous (to simulate temperature- and time-dependent ductile
creep), and plastic (a continuum approximation of discrete
brittle faulting). We assume the elastic properties of water ice
from Gammon et al. [1983] and the ductile creep flow laws
of Goldsby and Kohlstedt [2001]. Two creep mechanisms,
grain-boundary sliding and grain-boundary diffusion, are
sensitive to the ice grain size; we assume a grain size of
1 mm [see Dombard and McKinnon, 2006a]. For the plastic
component, we use the results from ice friction experiments
by Beeman et al. [1988].

[9] Free-slip boundary conditions are applied on the side
and bottom boundaries. Gravity is applied as a body force
that scales with the surface gravitational acceleration of
0.22 m s\(^{-2}\) and an assumed density of 950 kg m\(^{-3}\); because
the simulated material is compressible (i.e., elastic Poisson’s
ratio <0.5), the application of gravity requires the initial
stress state be adjusted to lithostatic (vertical and horizontal
stresses equal and growing linearly with depth). The load
imposed by the ridge is simulated as a series of elemental
surface pressures that stepwise approximate a 200 km wide
(100 km half width) and 20 km tall symmetric ridge with a
triangular cross section (see Figure 2). We linearly increase
the magnitude of these surface pressures, to approximate the
growth of the ridge, over a finite time of 1 kyr. This value is
arbitrary; however, the resultant flexure is insensitive to the
growth time [see, e.g., Albert et al., 2000; Dombard et al.,
submitted manuscript]. Our predicted final topography is a
sum of the surface displacement and initial triangular ridge
topography approximated by the surface pressure.

[10] We implement full large-strain deformation, as well
as a formalism that enforces constant dilatation across each
element, which prevents numerical errors that can arise in
the simulation of nearly incompressible behavior (e.g., ductile
creep). Time stepping is automatically controlled to resolve
the minimum viscoelastic Maxwell time in the mesh by a
factor of 4 or greater. To keep the run times reasonable (of
order 10–100 h per simulation), we implement a minimum
viscosity in the mesh of 10\(^{20}\) Pa s (minimum Maxwell
time of 900 years); with simulations using different cutoff
viscosities, we have confirmed our results are not sensitive
to this value.

[11] Our results are shown in Figure 3, which shows the
topography after a simulated time of 10 Myr. The depths to
the brittle-ductile transitions, which we take to be representa-
tive of an effective elastic lithospheric thickness, are ~50 and
~5 km for the lower and higher heat flow cases, respectively.
Following Turcotte et al. [1981], membrane support is neg-
ligible if the horizontal scale of the deformation is less than
the resolution of a surface spherical harmonic of a degree
approximately given by the square root of the ratio of the
radius to the elastic thickness. For our cases, these degrees
are 4 and 12, or ~1200 and ~400 km for the lower and
higher heat flows. Our simulated deformation fits within
these bounds, justifying our plane-strain assumption.

[12] Most models typically appeal to early formation of
the ridge during epochs of elevated heat flow (and hence thin
lithospheres). For instance, Castillo-Rogez et al. [2007]
suggested that the ridge formed when the depth to the
170 K isotherm was ~15 km deep. We test a somewhat
lower heat flow of 18 mW m\(^{-2}\), which places this isotherm
~20 km deep. Our results demonstrate that the ridge would
effectively exist in a state of collapse. A zone, 80–160 km
distant from the equator, of high plastic strains (up to
~5.5%) marks in essence a hinge fold where the lithosphere
is breaking. Consequently, almost 90% of the initial height
is lost as the thin lithosphere founders beneath this massive
load. Producing a final profile with the ridge over 15 km tall
would require a larger load, exacerbating the state of col-
lapse. Thus, models that appeal to epochs of high heat flow
are not consistent with the lithospherically supported state
of the ridge.

[15] Conversely, the ridge can be supported when the heat
flow is lower. For a heat flow of 3 mW m\(^{-2}\), we predict a
flexural trough comparable in depth to the topographic low
on the south side of the ridge (compare the low at ~400 km
distance from the ridge in Figure 2), but this topographic
low is farther from the ridge than we predict and is not
symmetric across the equator, which would be expected if
the observed flanking lows were flexural in origin. To sup-
port the ridge yet limit the depth of the flexural trough to
<1 km would require lower heat flows <1 mW m\(^{-2}\), condi-
tions that would have a long horizontal scale where the
deforation would reach to the poles, violating our planar
strain assumption.
temperature structure has the stronger control on lithospheric thickness. This is consistent with characteristic (MPa) stress levels under the ridge, too high for grain size–sensitive creep mechanisms to have much influence. Situations with a near surface porous layer [e.g., Castillo-Rogez et al., 2007] would predict even more deformation, because the porous layer would weaken the material and would provide a thermal blanket on the surface, permitting higher temperatures at depth for the same heat flow and therefore thinning the lithosphere.

[15] Thus, the observed state of the ridge, up to 20 km tall with no obvious signal of lithospheric flexure, suggests that the ridge formed after Iapetus had cooled and the surface heat flow was <3 mW m$^{-2}$. For comparison, such low heat flows were not realized in the thermal models of Castillo-Rogez et al. [2007] until after at least 1 Gyr of evolution. Thus, any model for the formation of the ridge has to explain its lithospheric support.

3. Past Models

[16] Models for the formation of the ridge can be divided into two broad classes: endogenic and exogenic. Endogenic models appeal to regional to global stresses and resultant tectonics. The most commonly cited stress source, acting either alone or in concert with other stresses or mitigating factors, is due to despinning of the satellite [Porco et al., 2005; Castillo-Rogez et al., 2007; Robuchon et al., 2010]. The highly oblate shape of Iapetus suggests an initial spin period far shorter than the current 79 days (likely of order 10 h). The change in shape from a more to less oblate spheroid would have resulted in shortening of the surface centered on the equator [e.g., Melosh, 1977]. A problem was recognized early on, however [Porco et al., 2005]: while compressive stresses may peak at the equator, E-W stresses dominate N-S stresses, which would produce N-S trending thrust faults, which are inconsistent with the E-W trend and symmetric structure of the ridge.

[17] Some authors have proposed that deformation could have been localized at the equator. Sandwell and Schubert [2010] did not specify a localization mechanism, while Melosh and Nimmo [2009] and Beuthe [2010] appealed to a thinner equatorial lithosphere, increasing in thickness at higher latitudes. These later models demonstrated that such thickness variations can change the character of elastic shell models of planetary lithospheres subjected to despinning, planetary contraction, or other stress sources (alone or in combination), yielding in some circumstances conditions consistent with the formation of the ridge. On the other hand, these models do not explain why such a ridge is only found on Iapetus (other satellites undoubtedly have similar lithospheric thickness variations and were despun, albeit much more quickly [Peale, 1977]). Nor do these models explain the lone nature of the ridge. While stresses (of the correct orientation) can peak at the equator, they decrease fairly slowly with increasing latitude and can still be quite large at the poles. The prediction should be, then, that the ridge should be the most pronounced, not only, feature on Iapetus.

[18] A last class of endogenic models postulates upwarping of the lithosphere from below, via an unspecified mechanism [Giese et al., 2008] or solid-state convection within the icy interior of Iapetus [Czechowski and Leliwa-Kopytynski, 2008; Roberts and Nimmo, 2009]. Regardless of the difficulties of generating such equatorially symmetric convection, such models fail because they implicitly require very thin lithospheres that can deform on the scale of the width of the ridge (100–200 km), meaning the ridge would have collapsed after loss of this dynamic support. At the very least, a signal of lithosphere flexure should be quite pronounced, but no topographic signal is apparent (see above).

[19] The other broad class of models for the formation of the ridge is exogenic. Taking a cue from the prominent equatorial ridge on the small (≈1.2 km across), rapidly spinning (period of ≈2.75 h) asteroid (66391) 1999 KW₄, Kreslavsky and Nimmo [2010] suggested that ancient Iapetus was spinning far faster than initially believed, so fast in fact that it was near its rotational stability limit. At the equator then, centrifugal forces were comparable to gravity, and the ridge was raised by the extreme rotational potential. This hypothesis, as proposed, can be discounted, however, because the whole satellite would have been horribly flattened, with extreme shortening of the polar regions. That is, the surface should be massively tectonized. Furthermore, unlike this small asteroid, a larger body like Iapetus would, near its rotational stability limit, deform into a triaxial, Jacobi ellipsoid (similar to the large Kuiper belt object Haumea [Rabinowitz et al., 2006]) and not maintain axial symmetry [Weidenschilling, 1981].

[20] Another exogenic model is due to Ip [2006], who proposed that the ridge originated as a primordial ring of debris, drawn out of the Saturnian subnebula, in orbit around an accreting Iapetus. The material in the ring collisionally evolved, dissipating orbital energy, which had two main consequences. First, the debris settled into a thin ring centered over the equator of Iapetus (i.e., no inclination), and second, the energy loss allowed the orbits to spread such that many ring particles hit the equator of Iapetus in multiple, small, grazing (very shallow angles) impacts at sub-hypervelocity speeds of a few hundred meters per second, slowly building up the ridge. At first blush, this proposal seems like it best satisfies the three critical observations of the ridge. The location on and only on the equator is a direct consequence of the penultimate step as a ring. Ip [2006] also contended that the hypothesis explains the fact why the ridge is apparently unique. He recognized that Iapetus has the largest Hill sphere (or gravitational zone of dominance), relative to body size, of any major satellite of the outer solar system. Despite the assertion of Ip [2006], Iapetus’s largest Hill sphere only means that Iapetus should have the largest, not only, equatorial ridge. The surface density of the purported ring can be estimated by smearing the mass of the ridge (~0.1% the mass of Iapetus) within the Hill sphere of Iapetus; this surface density can then be applied within the Hill spheres of the other major satellites of the outer solar system. This predicted mass can be used to estimate crudely the potential heights of equatorial ridges on these other satellites. Typically, these predicted ridges are hundreds of meters to kilometers tall, with ridges >6 km on Callisto and Titan. Clearly, a corollary of Ip’s proposal is that kilometer-scale tall equatorial ridges should be commonplace, yet nothing like Iapetus’s ridge has been observed. Another issue is the uncertainty that a gravitationally modest body such as Iapetus could draw a subsubnebula or, more specifically, a particle disk out of the Saturnian subnebula.
An additional exogenic model invokes an ancient giant impact [Levison et al., 2011] (as do we). A primary objective of this work is the creation of a relatively large (0.5%/1.5% of the mass of Iapetus), impact-generated “subsatellite” whose orbital evolution results in the despinnings of Iapetus before (usually) reimpacting this moon. The authors, however, also speculate that remnant debris from the impact that remained within the Roche limit of Iapetus (and hence could not coalesce into a subsatellite) formed the ridge. The final stage of this proposal is similar to that of Ip [2006]: evolution of orbital debris to a ring over the equator and deorbiting of this material to form the ridge, although the models differ in that in the Levison scenario, the material is more tightly bound to Iapetus (i.e., within the Roche limit, or ∼2.5–3 the radius of Iapetus). Their hypothesis, however, also suffers from the same shortcoming as the Ip [2006] proposal. As we discuss more thoroughly below, the Roche limits of most major outer satellites are well within the orbital stability zones of these moons, and all these satellites were undoubtedly pounded by large impacts toward the end of planetary accretion. Nominally, then, equatorial ridges should be common.

4. Formation via “Giant” Impact

[22] Giant impacts undoubtedly occurred in the waning stages of planetary accretion, as the largest member of an orbital zone swept up the second, third, etc., next largest members. Under proper circumstances, such large, planet-scale impacts could eject a substantial amount of material into orbit around the target body, some of which may have coalesced into a satellite. Our moon is thought to have formed this way [e.g., Canup and Righter, 2000]. Because of lower encounter velocities, Pluto’s moon Charon is believed to have formed via a process coined “intact capture,” where a giant impact on Pluto dissipated orbital energy, allowing the impacting body to go into orbit around Pluto [Canup, 2005, 2011]. Presumably, relatively large or “giant” impacts also occurred on the forming satellites of the outer solar system [Canup and Ward, 2006]. Here, we propose that an impact-generated subsatellite formed around Iapetus, the material in which eventually formed the equatorial ridge [cf. Dombard et al., 2010].

[23] Because of the low encounter velocities of roughly coorbital bodies sharing an orbital zone with a young Iapetus about Saturn (collision speeds would be dominated by Iapetus’s escape velocity of ∼0.5 km s⁻¹), we envision intact capture of a subsatellite, although we do not discount the possibility of coalescence from a debris disk. The resultant subsatellite does not need to be large; an icy, porous body less than ∼100 km in radius supplies enough mass to account for the ridge. The size of the subsatellite relative to Iapetus (less than ∼14%) is smaller than the equivalent in the Earth-Moon (∼27%) and Pluto-Charon (∼52%) systems.

[24] The initial orbit of the subsatellite was likely fairly prograde or retrograde and on the spin period of Iapetus. The orbit of a retrograde subsatellite would always decay, but a subsatellite in a prograde orbit would either retreat or decay if the period of revolution was less than the primary’s spin period (like the Earth-Moon system) or greater, respectively. Goldreich and Soter [1966] derived the time rate of change of the semimajor axis of a secondary body in orbit around a primary; under the assumption of a time invariant, dissipative elastic body, this equation can be integrated to yield

\[
\dot{a} = \left( \frac{a_0}{a} \pm \frac{39}{2} \frac{k_2}{M} \frac{GM}{R^3} \right) \frac{1}{2} \left( \frac{a}{R} \right)^{2/3},
\]

where \( \dot{a} \) is the semimajor axis normalized to the radius of the primary (the subscripted version denotes the initial value), \( k_2 \) and \( Q \) are the second degree tidal Love number and tidal dissipation quality factor of the primary body (e.g., Iapetus), \( m \) and \( M \) are the masses of the secondary and primary bodies, \( G \) is the universal gravitational constant, \( R \) is the radius of the primary, and \( t \) is time. The orbital period is thus \( 2\pi \sqrt{(\frac{aR}{GM})} \). The \( k_2 \) of a uniform, elastic sphere is given by

\[
k_2 = \frac{3/2}{1 + \frac{19}{12} \frac{\mu}{\rho gR^3}},
\]

where \( \mu \) is the elastic shear modulus (∼3.5 GPa for ice), \( \rho \) is density, and \( g \) is the surface gravitational acceleration; for an elastic Iapetus, \( k_2 \) is ∼8 × 10⁻³.

[25] The plus or minus in equation (1) denotes cases of orbital retreat (+) or decay (−). Because we are considering an event occurring during planetary accretion (effectively time zero for the solar system), the spin period of Iapetus was undoubtedly much shorter than the current 79 days, likely of order 10 h [see, e.g., Castillo-Rogez et al., 2007], so any subsatellite in a prograde orbit would initially retreat until Iapetus was despun, presumably by Saturn tides. (At 0.1% the mass of Iapetus, the subsatellite is too small to affect appreciably the spin of Iapetus [cf. Levison et al., 2011].) Robuchon et al. [2010, equation 4] describes the time rate of change of the angular rotational frequency \( \omega \) of a satellite, which again assuming a time invariant, dissipative elastic secondary is trivial to integrate [cf. Goldreich and Soter, 1966]:

\[
\omega = \omega_0 - \frac{3}{2} \frac{k_2 M_{GG}}{Q CM} R^3 C \frac{R}{\alpha} \wedge \frac{\dot{\alpha}}{\alpha},
\]

where \( M_{GG} \) is the mass of the gas giant planet and \( \alpha \) is the semimajor axis of the satellite’s orbit around the gas giant. \( C \) is the normalized moment of inertia for the satellite (0.4 for an uniform sphere). For large icy satellites, this value ranges from ∼0.31 (Ganymede) to ∼0.35 (e.g., Callisto). We assume a value of 0.35.

[26] Figure 4 shows the possible configurations of our proposed Iapetus/subsatellite system, assuming a subsatellite/Iapetus mass fraction of 0.1% (comparable to the mass of the ridge), a \( Q \) for Iapetus of 100 (consistent with general
of Figure 4 is placed on a nonlinear sliding bar. Thus, the salient points of Figure 4 are that (1) initially close (less than \( \sim 10 \, R_I \) (Iapetus radii)) retrograde orbits would decay rather quickly and (2) all prograde orbits should expand to a semimajor axis greater than \( \sim 18–19 \, R_I \) before reversing direction. As we will discuss in section 5, this distance is still within Iapetus’s gravitational zone of dominance.

[27] The subsatellite, whether initially prograde or retrograde, would have ultimately encountered Iapetus’s Roche limit of at \( \sim 2.5–3 \, R_I \). We anticipate that the subsatellite would have accreted as a rubble pile, and even if not, have been pervasively fractured during its impact formation/capture. Thus inward of \( \sim 2.5–3 \, R_I \), tidal forces should ultimately overwhelm the subsatellite, tearing it apart. The resultant debris would then collisionally evolve, which would (1) comminute the material, in principle; (2) dissipate any remaining inclination in the subsatellite’s orbit, placing the debris into a ring over the equator; and (3) dissipate orbital energy. For a ring mass of 0.1% that of Iapetus, the surface mass density and viscosity would be far higher than that of Saturn’s main rings, which would lead to very rapid settling of debris into Iapetus’s equatorial plane and viscous spreading within that plane [e.g., de Pater and Lissauer, 2010, section 11.2]. The endgame of our scenario is very similar to that proposed by Ip [2006], with a debris ring mostly deorbiting and impacting the surface of Iapetus at very shallow angles and subsonic speeds of a few hundred meters per second. Over time, the repeated impacts added material to the surface in a narrow zone, building the ridge.

5. Discussion

[28] Our hypothesis for the formation of the equatorial ridge explains the three critical observations. The first two (why on and only on the equator) are explained via the hypothesis’s penultimate step as a debris ring. The last observation (why only Iapetus) is a natural consequence of the satellite’s unique orbital position far from its gas giant parent. Figure 5 shows the zones of prograde orbital stability for all the major satellites of the outer solar system (i.e., those large enough to be roughly spherical), calculated by determining the semimajor axis at which is reached a minimum nondimensional Jacobi constant [see Hamilton and Burns, 1991, equation 5]. Above a critical value (\( \sim 9.4 \)), orbital trajectories are circular and stable, and at somewhat smaller values (between 9 and \( \sim 9.4 \)), trajectories can become chaotic. For values <9, the Hill curves no longer fully enclose the satellite, meaning any subsatellite may (though not necessarily will) be lost to an orbit around the gas giant planet. The Jacobi constant also depends on orbit inclination; thus the ranges in Figure 5 are found by considering the most restrictive (critical Jacobi constant of 9 and inclination of 90°) and forgiving (9.4 and 0°) cases. Also plotted are the Roche limit for each satellite and the range in Charon’s semimajor axis in the Canup [2005] simulations (in satellite radii), which we take to be representative of intact capture. Although all the satellites in Figure 5 likely suffered giant impacts, Iapetus clearly has more stable phase space than any other satellite; many satellites only have stable orbital zones within their Roche limits. At best they could only retain a postimpact debris disk, and it is uncertain that the amount of retained debris could yield a ridge (we consider it unlikely, or at least
ridge survival unlikely (see below), because otherwise ridges would be common). A handful of other satellites do have orbital stability zones that could retain a subsatellite, though not nearly as extensive a stability zone as Iapetus. The Jacobi constant–derived ranges in Figure 5 are equivalent to \(~34\%-45\%\) of the size of the Hill sphere \(r_H = \left[ M/(3M_{\odot}) \right]^{1/3} \), and retrograde orbits are stable in a wider zone, potentially the size of the Hill sphere [Hamilton and Burns, 1991]. Thus by Figure 5, the general retention of an impact-generated subsatellite is not precluded for satellites other than Iapetus, but the parameter space is far more restrictive.

[29] In the event that a subsatellite was formed within the orbital stability zone, a likely fate is our ridge-forming scenario. Equations (1) and (3) can be inverted to estimate timescales (again, enhanced dissipation, via viscoelasticity for example, would decrease these timescales). In general, the time for the orbit of a prograde subsatellite to expand to the edge of its respective stability zone is greater than the despinning time of the satellite, meaning the expansion halts before orbit destabilization. We can further use the inverted equation (1) to estimate maximum timescales for the orbital evolution (Figure 6), assuming the starting and ending semimajor axes are known; we use 2.5 times the orbital evolution (Figure 6), assuming the starting and ending semimajor axes are known; we use 2.5 times the Roche limit, as shown in Figure 5 for prograde orbits (solid circles), every subsatellite would evolve and form a ridge within 10 Myr, except for Iapetus (cf. Figure 4). Because we are considering a process occurring during planetary accretion, these short timescales mean that any equatorial ridge on outer satellites produced from a prograde orbit, other than on Iapetus, would have to survive over 4.5 Gyr of impact bombardment, including the intense bombardment in the waning stages of accretion and the possibility of a Late Heavy Bombardment in the outer solar system [e.g., Dones et al., 2009]. We doubt that any such feature would not have been erased.

[30] The situation for retrograde subsatellites is more forgiving, because of the potential large stability zones. For most satellites save five (Iapetus, Oberon, Callisto, Titan, and Titania), the maximum times (open circles in Figure 6) are <16 Myr, so again it is doubtful any subsequent ridge would survive to the present day. Of these other five, the maximum timescales are a few hundred Myr for Callisto, Titan, and Titania (not enough to escape erasure by a Late Heavy Bombardment), \(~1\) Gyr for Oberon, and far in excess of the age of the solar system for Iapetus.

[31] Thus, the scenario that we propose does not preclude the formation of a stable subsatellite, or even the retention of any subsequently produced equatorial ridge, for worlds other than Iapetus; it is just far less likely. This does raise the possibility, however, of other equatorial ridges in the solar system. Of the other four more likely possibilities that we consider, Voyager and Galileo images of Callisto did not reveal a ridge. Similarly, Cassini images have not revealed any such feature on Titan, although the geologic and hydrologic processes on Titan would make the long-term retention of a global ridge even less likely. The possibility remains for

**Figure 5.** Orbital stability limits (in satellite radii) for prograde subsatellites around the major satellites of the outer solar system, calculated by determining the semimajor axis at which a minimum nondimensional Jacobi constant is reached. Extrema are found by considering the most restrictive and permissive cases. Also plotted are the Roche limits (hatched zone) and the range of initial semimajor axes from hydrocode simulations of the “intact capture” of Charon around Pluto [Canup, 2005]. Most satellites have orbital stability zones within the Roche limit, while of the remainder, Iapetus by far has the largest phase space to accommodate a subsatellite.

**Figure 6.** Estimates of the maximum orbital decay times for subsatellites around the major satellites of the outer solar system, assuming that the subsatellites begin at the outer edges of the orbital stability limits and have masses of 1% that of the satellites. Solid and open circles denote subsatellites in prograde and retrograde orbits, respectively; differences are due to the fact that orbital stability zones are wider for retrograde orbits. In nearly all cases, these maximum times are <0.5 Gyr, and usually <20 Myr, meaning that any equatorial ridge produced by our proposed scenario would have to survive the intense bombardment in the waning stages of accretion and the possibility of a Late Heavy Bombardment in the outer solar system. We calculate these timescales assuming \(Q = 100\) for each satellite and an elastic \(k_2\) as calculated in the text; decay times scale as \(Q/k_2\) and so can be rescaled accordingly.
the two satellites of Uranus Titania and Oberon, but because of the polar observations of the system by Voyager 2, the equators of the two satellites were not (or barely) imaged [Greeley and Batson, 1997]). Future missions to Uranus should address this issue.

[23] Supporting evidence for our hypothesis may have recently been discovered on Iapetus’s Saturnian sister Rhea. Images by the Cassini spacecraft revealed a splotchy color anomaly, only 10 km wide, on a great circle path nearly aligned with the equator [Schenk et al., 2011]. In fact, this ring is tilted only 1.8° off the equator (~24 km maximum), which is small enough that the “tilt” may be due to uncertainties in the cartographic control network for Rhea. The color anomaly was due to deorbiting of a collisionally evolved, very low mass ring in a retrograde orbit around Rhea, the material of which was likely derived by a modest impact from a modest impactor population today [Tiscareno et al., 2010]). As they noted, their proposal is consistent with the hypothesis bypasses this paradox. By delaying the formation of the ridge (Figure 4), thus allowing Iapetus to cool until the lithosphere thickened, our hypothesis contradicts this view.

[35] Acknowledgments. We thank Paul M. Schenk for sharing his topography and Amandas and Steven A. Hauck for their assistance and acknowledge a NASA PGG grant to W.B.M. and a NASA Early Careers Fellowship (grant NNX10A72G) to A.J. We are grateful for comments from H. Jay Melosh, an anonymous reviewer, and an Associate Editor.

References


A. F. Cheng, Johns Hopkins University Applied Physics Laboratory, 11100 Johns Hopkins Rd., Laurel, MD 20723, USA.

A. J. Dombard and J. P. Kay, Department of Earth and Environmental Sciences, University of Illinois at Chicago, 845 West Taylor St., Chicago, IL 60607, USA. (adombard@uic.edu)

W. B. McKinnon, Department of Earth and Planetary Sciences, Washington University, One Brookings Dr., Box 1169, St. Louis, MO 63130, USA.