The Astronomical Journal, 128:2501–2509, 2004 November

TILTING SATURN. I. ANALYTIC MODEL

William R. Ward
Department of Space Studies, Southwest Research Institute, Suite 400, 1050 Walnut Street, Boulder, CO 80302; ward@boulder.swri.edu

and

Douglas P. Hamilton
Department of Astronomy, University of Maryland, College Park, MD 20742-2421; hamilton@astro.umd.edu
Received 2003 December 30; accepted 2004 July 15

ABSTRACT

The tilt of Saturn's spin axis to its orbit plane is 26.°7, while that of Jupiter is only 3.°1. We offer an explanation for this puzzling difference owing to gravitational perturbations of Saturn by the planet Neptune. A similarity between the precession period of Saturn's spin axis and the 1.87×10^6 yr precession period of Neptune's slightly inclined orbit plane implicates a resonant interaction between these planets as responsible for tilting Saturn from an initially more upright state. We make a case that Saturn was captured into this resonance during the erosion of the Kuiper belt, which decreased the rate of regression of Neptune's orbit plane. Penetrating the resonance pumped up Saturn's obliquity to its current value. The spin axis may also be librating in the resonance with an amplitude $\psi \gtrsim 31^\circ$, and we discuss possible causes of this and the implied constraint on Saturn's moment of inertia. Matching the current pole position to the predicted outcome could place constraints on early solar system processes.

Key words: planets and satellites: individual (Neptune, Saturn) — solar system: formation — solar system: general

1. INTRODUCTION

Jupiter and Saturn consist predominantly of hydrogen and helium acquired from the primordial solar nebula during the planet-building epoch. The formation of such gas giants is believed to commence with the collisional accretion of a several–Earth-mass core from solid ice-and-rock planetesimals, followed by the accretion of the gaseous component once the core reaches a critical size (e.g., Pollack et al. 1976; Wuchterl et al. 2000 and references therein). Seemingly at odds with this picture, however, are the very dissimilar obliquities, θ, of these planets to their orbit planes, namely, 26.°7 for Saturn versus only 3.°1 for Jupiter.

The obliquities of the other planets in the solar system are likely due to the stochastic nature of their accumulation from solid planetesimals (Lissauer & Safronov 1991; Dones & Tremaine 1993; Chambers & Wetherill 1998; Agnor et al. 1999), and the rock/ice cores of Jupiter and Saturn probably had nonzero obliquities as well. However, their massive gas components derived from the nebular disk would have added angular momentum nearly perpendicular to their orbit planes, overwhelming that of the cores and ultimately resulting in small obliquities for both planets. With 95 Earth masses and a 10.7 hour rotation period, Saturn has considerable spin angular momentum, making it problematic that an impact could have sufficiently changed its pole direction after its formation. Why then is Saturn's obliquity so large?

We suggest the answer lies in solar system events following the formation of the planets that caused an initially upright Saturn to suffer a tilt. We are not the first to seek such a mechanism. It has been proposed that the obliquities of the outer planets may result from a "twist" of the total angular momentum of the solar system during the collapse of the molecular cloud core that led to its formation (Tremaine 1991). It is possible to choose a timescale for this event that would affect the planetary obliquities from Saturn on out but have only a minor influence on Jupiter. Although this cannot be ruled out, this paper presents what we believe is a more compelling mechanism for generating Saturn's obliquity, predicated on a similarity between Saturn's spin-axis precession period and the regression period of Neptune's orbit plane (Harris & Ward 1982), which seems too close to be a coincidence. We propose that this period match and the obliquity of Saturn are cause and effect through the operation of a secular spin-orbit resonance between these bodies. This type of interaction is already known to cause large-scale oscillations of the obliquity of Mars (Ward 1973, 1974, 1979; Laskar & Robutel 1993; Touma & Wisdom 1993). Here and in a companion paper (Hamilton & Ward 2004, hereafter Paper II), we detail how this mechanism could also account for the spin-axis orientation of Saturn, as well as provide a sensitive constraint on its moment of inertia.

2. PRECESSIONAL MOTIONS

2.1. *Spin Axis*

The equation of motion for a planet's unit spin-axis vector, $\boldsymbol{s}$, is $d\boldsymbol{s}/dt = \alpha(\boldsymbol{s} \cdot \boldsymbol{n})(\boldsymbol{s} \times \boldsymbol{n})$, where $\boldsymbol{n}$ is the unit vector normal to the planet's orbit plane. The precessional constant depends on the strength of the torque exerted on the planet and its spin angular momentum. For Saturn, most of the solar torque is exerted on its satellites instead of directly on the planet, Titan ($M_{\rm Titan} = 1.34 \times 10^{26}$ g) being by far the dominant one (Ward 1975). Saturn's oblate figure gravitationally locks the satellites to its equator plane so that the system precesses as a unit (Goldreich 1965). The precessional constant can be written

$$\alpha = \frac{3}{2}\frac{n^2}{\omega}\frac{J_2 + q}{\lambda + l} \tag{1}$$

(Ward 1975; French et al. 1993), where ω is the spin frequency of Saturn, n is its heliocentric mean motion, $J_2 = 1.6297 \times 10^{-2}$ is the coefficient of the quadrupole moment of its gravity

field, and λ is the moment of inertia of Saturn normalized to $M_S R^2$ with M_S and R being the mass and radius of the planet. The quantity

$$q \equiv \frac{1}{2}\sum_j \frac{(m_j/M_S)(a_j/R)^2 \sin(\theta - i_S)}{\sin\theta} \quad (2)$$

is the effective quadrupole coefficient of the satellite system with q/J_2 being the ratio of the solar torque on the satellites to that directly exerted on the planet, and

$$l \equiv \sum_j (m_j/M_S)(a_j/R)^2(n_j/\omega) \quad (3)$$

is the angular momentum of the satellite system normalized to $M_S R^2\omega$, where M_j, a_j, and n_j are the masses, orbital radii, and mean motions of its satellites, with i_j being the inclination of a satellite orbit to the equator of Saturn.[1] French et al. (1993) give $q = 0.05164$ and $l = 0.00278$; Hubbard & Marley (1989) find $\lambda = 0.2199$ from their interior models, but this value is uncertain by as much as 10% (French et al. 1993; M. Marley 2003, private communication). With these numbers, equation (1) yields $\alpha = 0\farcs8306\ \mathrm{yr}^{-1}$ and $\alpha\cos\theta = 0\farcs7427\ \mathrm{yr}^{-1}$. If the orbit plane of Saturn were fixed in inertial space, its spin-axis precession would occur at constant obliquity θ with a period $P = 2\pi/(\alpha\cos\theta) = 1.745\times10^6$ yr.

Nicholson & French (1997) have analyzed 22 reported ring-plane crossings spanning a period of 280 yr to estimate Saturn's pole precession frequency as $0\farcs51 \pm 0\farcs14\ \mathrm{yr}^{-1}$, but this is low, largely because of a $\sim$700 yr modulation due to Titan's $0\fdg32$ proper inclination (Nicholson et al. 1999). Using the nutation model of Vienne & Duriez (1992), the current rate can be predicted to be 68% of the long-term value, implying $\alpha\cos\theta = 0\farcs75 \pm 0\farcs21\ \mathrm{yr}^{-1}$.

2.2. *Orbit Plane*

All of the planetary orbits have small inclinations to the invariable plane and undergo nonuniform regressions due to their mutual gravitational perturbations. The inclination I and ascending node Ω of a given planet are then found from a superposition,

$$\sin\frac{I}{2}\sin\Omega = \sum_j \frac{I_j}{2}\sin(g_j t + \delta_j),$$
$$\sin\frac{I}{2}\cos\Omega = \sum_j \frac{I_j}{2}\cos(g_j t + \delta_j) \quad (4)$$

(e.g., Brouwer & van Woerkom 1950; Bretagnon 1974; Applegate et al. 1986; Bretagnon & Francou 1992), comprising many terms of amplitudes $\{I_j\}$ and frequencies $\{g_j\}$. Nevertheless, most of them are of only minor importance, and the largest-amplitude terms for Saturn's orbit are listed in Table 1; they represent contributions from three of the eight fundamental modes of a Laplace-Lagrange solution of the secular evolution of the solar system as given by Bretagnon (1974). The first of these is due to a strong 4.92×10^4 yr mutual orbital precession of Jupiter and Saturn; the next two are perturbations to Saturn's orbit plane due to the nodal regressions of Uranus (4.33×10^5 yr) and Neptune (1.87×10^6 yr), respectively.

[1] The Laplace plane at the distance of Iapetus is inclined by $14\fdg8$ to Saturn's equator, and Iapetus precesses about its normal in $\sim3\times10^3$ yr. Thus its average i is $14\fdg8$ as well.

TABLE 1
LARGEST-AMPLITUDE TERMS FOR SATURN'S ORBIT

j	g_j (arcsec yr^{-1})	I_j (deg)
16	−26.34	0.910
17	−2.99	0.045
18	−0.692	0.064

The variations in inclination and precession rate of the orbit cause a complicated time dependence for the orbit normal $\boldsymbol{n}(t)$ in the equation of motion for $\boldsymbol{s}$. This can lead to oscillations of the planet's obliquity as the spin axis attempts to precess about the moving orbit normal. In a linearized solution (e.g., Ward 1974; eq. [5] below), conspicuous oscillations occur because there is a near-match between the spin-axis precession rate $\alpha\cos\theta$ and $-g_{18}$, but again, the relative closeness of these frequencies for Saturn is due in part to its current obliquity. On the other hand, there are very good reasons to believe that both α and g_{18} were different in the past. For example, the spin-axis precession rate would have varied during the early contraction of Saturn as it cooled soon after formation (Pollack et al. 1976; Bodenheimer & Pollack 1986), while the frequency g_{18} would have been faster in the early solar system as a result of the presence of a larger population of objects in the Kuiper belt (Holman & Wisdom 1993; Duncan et al. 1995; Malhotra et al. 2000), whose gravitational influence would have increased Neptune's regression rate. Thus, if the similar values of $\alpha\cos\theta$ and $-g_{18}$ are not coincidental, something must have maintained this relationship during these changes. This paper applies the theory of secular spin-orbit resonance to the Saturn-Neptune interaction and demonstrates the ability of the resonance to drive up Saturn's obliquity from an initially near-zero value. Our companion paper presents numerical experiments that further support the efficacy of this mechanism.

3. SECULAR SPIN-ORBIT RESONANCE

3.1. *Spin-Axis Trajectories*

Since there are many terms in equation (4), neither the inclination nor the regression rate $\dot\Omega$ of the orbit is constant. In the case of small angles, the equation of motion for the spin axis can be linearized and solved analytically to give an expression for obliquity variations of the form

$$\theta \approx \bar\theta - \sum_j \frac{g_j I_j}{\alpha\cos\bar\theta + g_j}\sin(\alpha t\cos\bar\theta + g_j t + \Delta_j) \quad (5)$$

(Ward 1974), where $\bar\theta$ is a long-term average obliquity and the Δ_j are phase constants that depend on the observed planetary orbits. In general, the various sinusoidal terms cause rapid oscillations compared with any change in $\bar\theta$. However, if $\alpha\cos\bar\theta \to -g_J$ for some $j = J$, that term's small denominator will cause its amplitude to become very large while its frequency becomes very slow. The combination of this J-term plus $\bar\theta$ can then be replaced by a slowly moving nonlinear guiding center $\theta_{gc}(t)$ about which the other terms cause a high-frequency circulation of the spin axis (Ward 1992). It turns out that the high-frequency terms do not interfere much with the motion of the guiding center even if $\alpha\cos\theta_{gc}$ passes through $-g_J$, although in this case the linearized version of $\theta_{gc}(t)$ is no longer valid. One can show both analytically (Ward et al. 1979) and numerically (Ward 1992; Paper II) that the motion of the

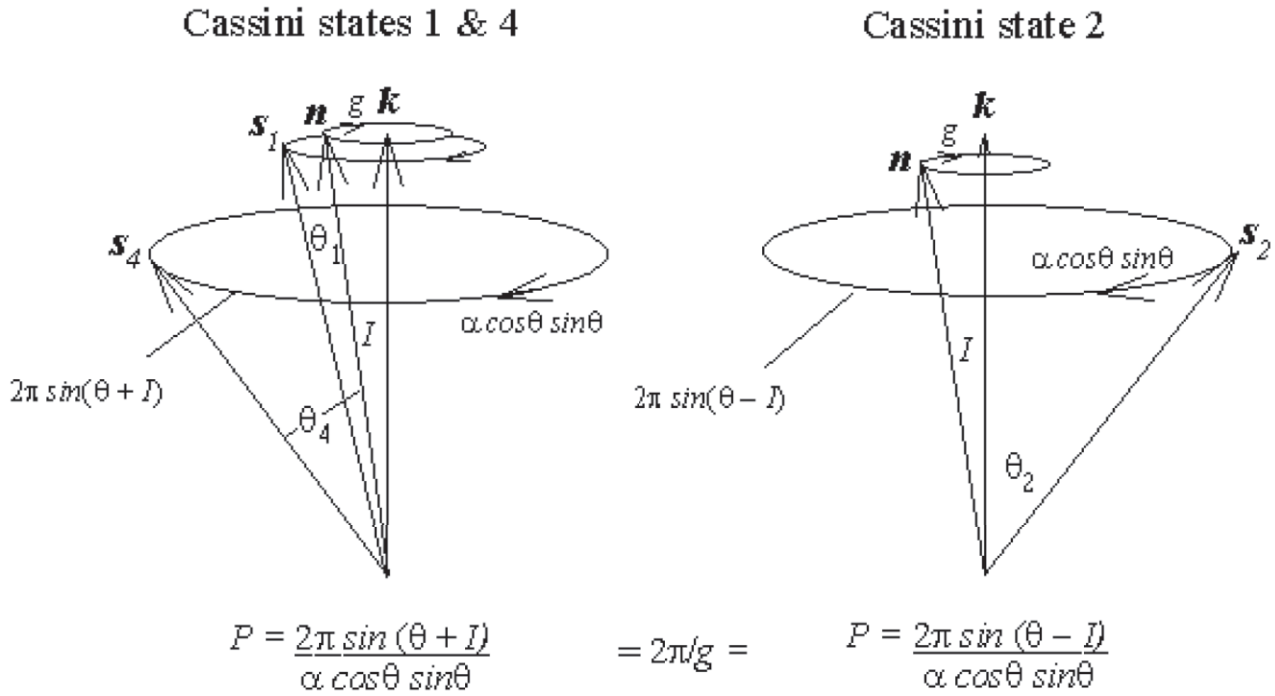

FIG. 1.—Coprecession of orbit normal $\boldsymbol{n}$ and Cassini state position of spin axis $\boldsymbol{s}_i$ about the normal $\boldsymbol{k}$ to the invariable plane.

guiding center is quite similar to the spin-axis motion in the case of uniform orbital precession provided we set $I = I_j$ and $\dot{\Omega} = g_J$. We turn to that case now.

Consider an orbital precession obtained by retaining only a single term J in the precession equation (eq. [4]). In this case, $\boldsymbol{n}$ maintains a constant inclination $I = I_J$ to the fixed normal to the invariable plane $\boldsymbol{k}$ and precesses at a constant rate $\dot{\Omega} = g_J = g$ (Fig. 1). If a coordinate frame rotating with angular frequency g is adopted, the orbit normal $\boldsymbol{n}$ will appear fixed. The equation of motion for the unit spin vector of a planet now takes the form

$$\frac{d\boldsymbol{s}}{dt} = \alpha(\boldsymbol{s}\cdot\boldsymbol{n})(\boldsymbol{s}\times\boldsymbol{n}) + g(\boldsymbol{s}\times\boldsymbol{k}). \tag{6}$$

This problem is well studied (e.g., Colombo 1966; Peale 1969, 1974; Ward et al. 1979; Henrard & Murigande 1987), and an exact integral of the motion can be found, which is also the relevant portion of the Hamiltonian of the system, $H = -(\alpha/2)(\boldsymbol{n}\cdot\boldsymbol{s})^2 - g(\boldsymbol{k}\cdot\boldsymbol{s})$ (e.g., Ward 1975). In the next section, we use H together with the elegant Cassini state theory as developed by Colombo, Peale, and others.

3.2. *Cassini States*

Colombo (1966) showed that the unit spin axis $\boldsymbol{s} \equiv (x, y, z) = (\sin\theta\cos\phi,\ \sin\theta\sin\phi,\ \cos\theta)$ traces out a closed curve on the unit sphere, $x^2 + y^2 + z^2 = 1$, given by its intersection with a cylindrical parabola,

$$[z + (g/\alpha)\cos I]^2 = -2(g/\alpha)\sin I\,(y - K), \tag{7}$$

as shown in Figure 2. This describes a family of parabolae with latus rectum $p = -(g/\alpha)\sin I$ and axis $z_0 = -(g/\alpha)\cos I$ but various vertices K. Depending on the choice of g/α and the resulting location of the axis z_0, there are either two or four locations (called Cassini states and denoted by the vectors $\boldsymbol{s}_1$ through $\boldsymbol{s}_4$ in Fig. 2) where a parabola is tangent to the unit sphere for some value of the vertex and the trajectory degenerates to a point. Here the spin axis $\boldsymbol{s}$ remains coplanar with $\boldsymbol{n}$ and $\boldsymbol{k}$ and stationary in the rotating frame, which means that in inertial space, these vectors coprecess at the same rate g, as depicted in Figure 1. Two of the states (labeled 1 and 4) are on the same side of $\boldsymbol{k}$ as $\boldsymbol{n}$, while state 2 is on the opposite side (Cassini state 3, which is retrograde, will not further concern us here). If the convention introduced by Peale (1974) of measuring θ_i clockwise from $\boldsymbol{n}$ is used, the state obliquities can be found from the single relationship

$$(\alpha/g)\cos\theta_i \sin\theta_i + \sin(\theta_i - I) = 0, \tag{8}$$

obtained by setting $\dot{\theta}$ and $\dot{\phi}$ to zero in equation (6). This is equivalent to a quartic equation, which could be solved explicitly for its four roots. Figure 3 shows a plot of the Cassini state obliquities θ_1, θ_2, and θ_4 as functions of α/g, where a value of $I = I_{18}$ has been adopted. It can be seen that for $|\alpha/g|$ less than some critical value, states 1 and 4 do not exist. This value can be determined by differentiating equation (8) and setting $d\theta/d(\alpha/g) = \infty$ at $(\alpha/g)_{\rm crit}$ to find the condition $(\alpha/g)_{\rm crit}\cos 2\theta + \cos(\theta - I) = 0$. Combining with equation (8), one can solve for the critical values of $\theta_1 = \theta_4 = \theta_{\rm crit}$ and $(\alpha/g)_{\rm crit}$, namely,

$$\tan\theta_{\rm crit} = -\tan^{1/3} I, \quad (\alpha/g)_{\rm crit} = -(\sin^{2/3} I + \cos^{2/3} I)^{3/2}. \tag{9}$$

For I_{18}, $\theta_{\rm crit} = -5\overset{\circ}{.}92$ and $(\alpha/g)_{\rm crit} = -1.016$.

Rearranging equation (8) into the form $[(\alpha/g)\cos\theta + \cos I]\tan\theta = \sin I \ll 1$, it is clear that the left-hand side can be made small either by a small value of $\tan\theta$ for which $\cos\theta \sim O(\pm 1)$ or by making the square-bracketed term small. These conditions yield the following approximate formulae:

$$\theta \approx \tan^{-1}\left(\frac{\sin I}{1 \pm \alpha/g}\right), \quad \theta \approx \pm\cos^{-1}\left(-\frac{g\cos I}{\alpha}\right). \tag{10}$$

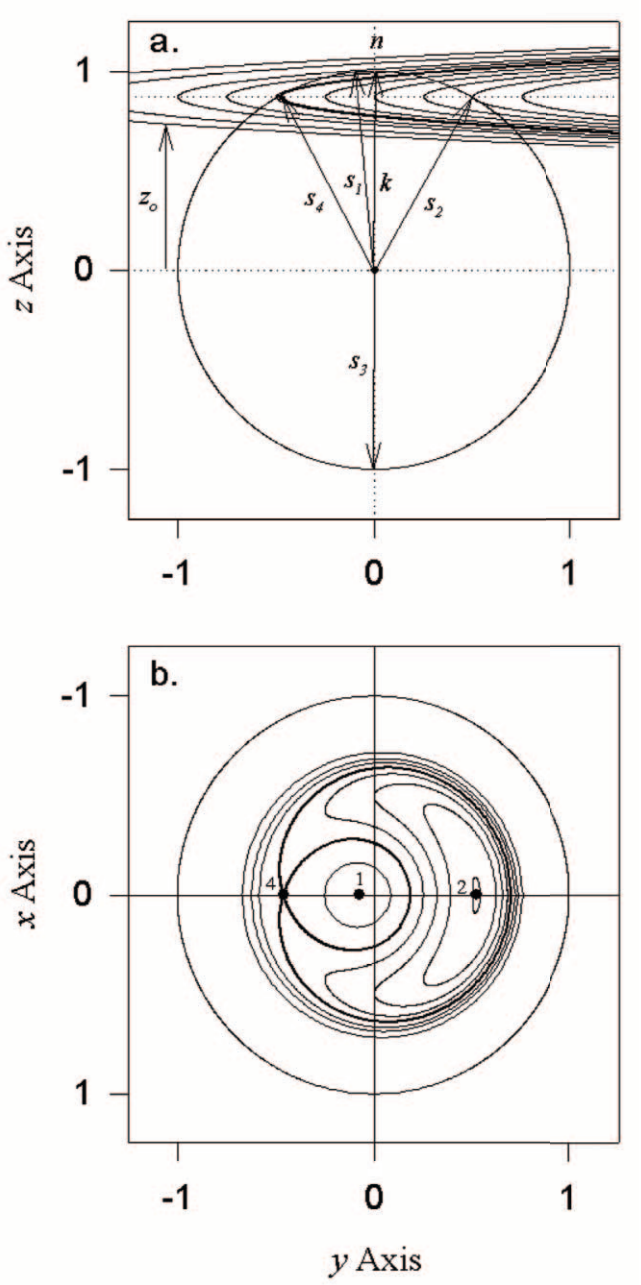

FIG. 2.—(*a*) Spin-axis trajectories and Cassini states (shown by vectors $\boldsymbol{s}_1$ through $\boldsymbol{s}_4$) traced on the unit sphere for a coordinate system rotating about the normal to the invariable plane, $\boldsymbol{k}$, with nodal regression frequency g. The Cartesian coordinate system has its z-axis in the direction of the orbit normal $\boldsymbol{n}$ and its x-axis along the line of the orbit's ascending node, so that the vector $\boldsymbol{k}$ lies in the y-z plane, inclined by angle I to $\boldsymbol{n}$. The z-coordinate of the spin-axis position is the cosine of the obliquity, $z = \cos\theta$. (*b*) Polar view of the same unit sphere. The trajectory passing through state 4 is the separatrix, which partitions the unit sphere into the three domains. The inclination employed this example to more clearly illustrate the morphology of the trajectories is an order of magnitude larger than the actual $j = 18$ term used in Fig. 4.

When $|\alpha/g| \gg |\alpha/g|_{\rm crit}$, the first expression approximates states 1 and 3, corresponding to $\pm\alpha$, respectively, while the second gives states 2 and 4. When $|\alpha/g| \ll |\alpha/g|_{\rm crit}$, the first expression gives states 2 and 3, while the aforementioned square-bracketed quantity cannot approach zero and the two corresponding roots of the quartic equation are complex. In this case, states 1 and 4 do not exist. States 1 through 3 are stable in the sense that if the spin axis is slightly displaced from them, it will tend to circulate the state; state 4 is unstable in this regard

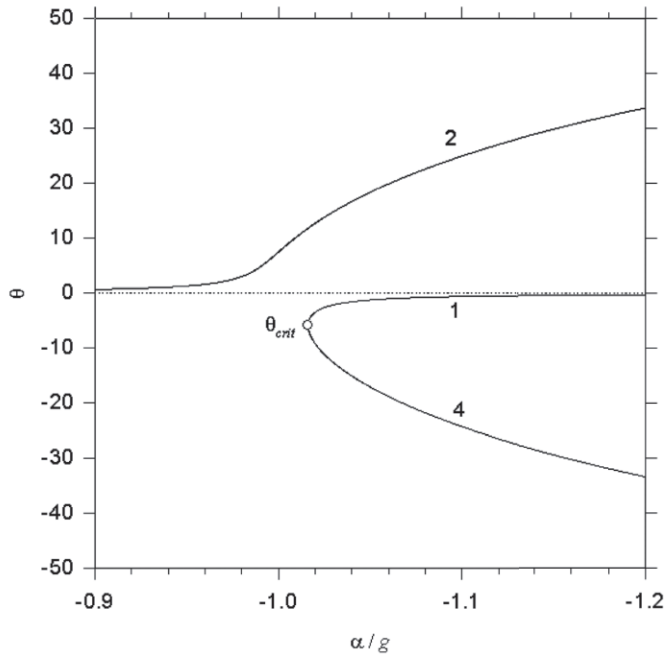

FIG. 3.—Obliquities of Cassini states 1, 2, and 4 as a function of the frequency ratio α/g. Spin-axis trajectories circulate about stable states 1 and 2; state 4 is unstable and lies on a separatrix. States 1 and 4 merge and disappear at the critical frequency ratio.

and lies on a separatrix (Fig. 2*b*) that partitions the unit sphere into three domains, each containing a stable state.

3.3. *Spin-Axis Position*

To evaluate whether Saturn could be in the resonance, its spin axis must be located with respect to the $j = 18$ reference frame defined by a z-axis that lies along the pole of the term and an x-axis along its ascending node on the invariable plane of the solar system. The right ascension and declination of $\boldsymbol{s}$ with respect to the equator and equinox at epoch J2000.0 are $40\overset{\circ}{.}595$ and $83\overset{\circ}{.}538$, respectively (Yoder 1995). Rotating about the vernal equinox by Earth's obliquity, $23\overset{\circ}{.}439$ (Yoder 1995), gives $\boldsymbol{s}$ with respect to the ecliptic and equinox as shown in Table 2. The normal $\boldsymbol{k}$ to the invariable plane from Allen (1976) is listed as well. Also included in the table are the x- and y-components of each vector in each system. Since the inclination of the invariable plane is very small, to first-order accuracy the coordinate system can be transformed to that plane by subtracting the components of $\boldsymbol{k}$ from $\boldsymbol{s}$. We now introduce the $j = 18$ pole from Applegate et al. (1986), who give[2] $\{p, q\} = \sin(I_{18}/2) \times \{\sin, \cos\}\,\Omega_{18} = N_8\,\{\sin, \cos\}\,\delta_8$, where $N_8 = -10^{-3.25}$ and $\delta_8 = 203\overset{\circ}{.}518$. Setting $\sin(I_{18}/2) \approx (\sin I_{18})/2$, and recalling that the longitude of the pole is 90° behind its ascending node Ω_{18}, we find the components of $\boldsymbol{n}$ listed in Table 2. We can transform again to a system with $\boldsymbol{n}$ at the origin by subtracting its components from the other vectors. This gives the vectors in an intermediate system. A final counterclockwise rotation of the coordinate system by $23\overset{\circ}{.}524$ puts $\boldsymbol{k}$ in the y-axis. The spin axis lies $\psi = 90^\circ - 59\overset{\circ}{.}12 = 30\overset{\circ}{.}88$ from $\boldsymbol{k}$. Figure 4 shows polar views of the $j = 18$ system for $\alpha/g_{18} = -1.16$ along with the separatrix and $\boldsymbol{s} = (x_s, y_s, x_s) = (0.236, 0.394, 0.888)$. The separatrix is more narrow than the example of Figure 2

[2] In the Applegate et al. (1986) notation, our $j = 18$ is their $j = 8$.

TABLE 2
COORDINATES OF SATURN SPIN AXIS, $j = 18$ POLE, AND NORMAL TO INVARIABLE PLANE

Reference Frame	Vector	Colatitude[a] (deg)	Longitude (deg)	x	y
Ecliptic/equinox	$\boldsymbol{s}$	28.049	79.529	8.546×10^{-2}	4.624×10^{-1}
	$\boldsymbol{k}$	1.579	17.584	2.626×10^{-2}	8.322×10^{-3}
Invariable plane	$\boldsymbol{k}$	0	...	0	0
	$\boldsymbol{s}$	27.253	82.572	5.920×10^{-2}	4.541×10^{-1}
	$\boldsymbol{n}$	0.0644	−66.476	4.488×10^{-4}	-1.031×10^{-3}
Intermediate	$\boldsymbol{k}$	0.0644	113.524	-4.488×10^{-4}	1.031×10^{-3}
	$\boldsymbol{s}$	27.317	82.644	5.875×10^{-2}	4.551×10^{-1}
	$\boldsymbol{n}$	0	...	0	0
$j = 18$ system	$\boldsymbol{k}$	0.0644	90	0	1.124×10^{-3}
	$\boldsymbol{s}$	27.317	59.120	2.355×10^{-1}	3.939×10^{-1}
	$\boldsymbol{n}$	0	...	0	0

[a] For $\boldsymbol{s}$, the colatitude is the obliquity; for $\boldsymbol{k}$ and $\boldsymbol{n}$ it is the inclination.

because the inclination is an order of magnitude smaller. Since the trajectory does not enclose $\boldsymbol{n}$ at the origin, such motion produces the longitude libration that is diagnostic of resonance trapping.

3.4. *System Evolution*

If the frequency ratio α/g changes for some reason, the Cassini states migrate along the unit sphere in accordance equation (8) and Figure 3. If changes occur slowly enough, it can be shown that the area enclosed by the spin-axis trajectory about the local Cassini state remains nearly invariant (see, e.g., Peale 1974; Ward et al. 1979). "Slowly enough" means that the state migration rate is much less than the rate of spin-axis motion: the so-called adiabatic limit. In particular, if the spin axis starts near state 2 with $|\alpha/g| \ll 1$, it will remain so as $|\alpha/g|$ increases and the state migrates away from $\boldsymbol{n}$ (at $\theta = 0$). As the frequency ratio passes through the critical value, the obliquity rises steeply and can become quite large; this is resonance *capture*.

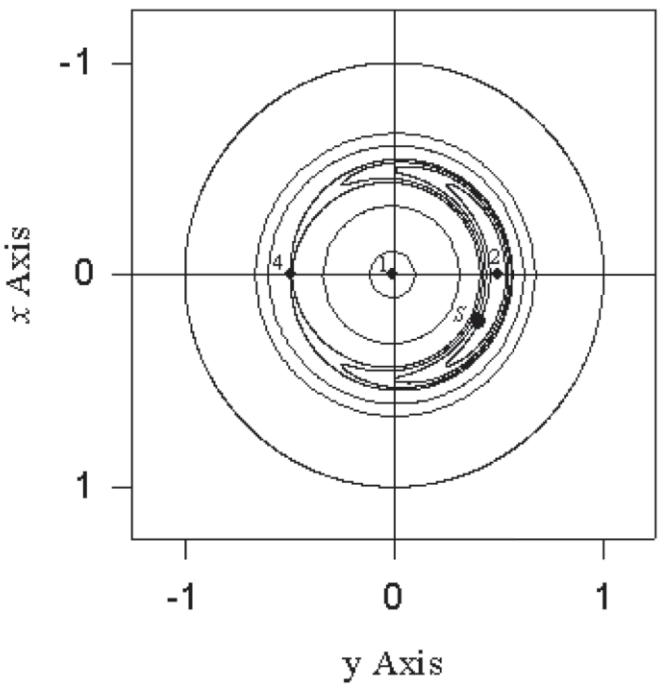

FIG. 4.—Illustration of the polar view of the unit sphere for the $j = 18$ frame of reference for $\alpha/g_{18} = -1.16$. The amplitude of the $j = 18$ term is $I_{18} = 0.^{\circ}064$. In addition to the Cassini states and separatrix, the current spin-axis position of Saturn is indicated. The spin axis lies inside the separatrix and circulates about state 2 on elongated trajectories that produce libration.

By contrast, with $|\alpha/g| \gg 1$, state 1 is near $\boldsymbol{n}$ but rotates away as $|\alpha/g|$ decreases, while state 4 rotates toward it. The two states eventually merge at $(\alpha/g)_{\rm crit}$; past this, state 2 is the only prograde state. A spin axis initially close to state 1 will track its motion until its merger with state 4. At this point it is left stranded and must establish a new trajectory about state 2, enclosing an area that may no longer be small. This sequence is resonance passage in the *noncapture* direction and results in a "kick" to the obliquity. Consequently, passage through the resonance is not a reversible process, with the outcome depending on direction. Employing elegant analytical expressions derived by Henrard & Murigande (1987) for the areas inside each of the domains, the above arguments are easily quantified. The area inside the separatrix containing state 2 can be written as

$$A_2 = 8\rho + 4\tan^{-1}T - 8z_0\tan^{-1}(1/\chi), \tag{11}$$

where

$$\chi \equiv \sqrt{-\tan^3\theta_4/\tan I - 1}, \tag{12}$$

which starts at $\chi = 0$ for $\theta_4 = \theta_{\rm crit}$ and diverges as $\theta_4 \to -\pi/2$. The remaining functions can be written

$$\rho \equiv \frac{\chi \sin^2\theta_4 \cos\theta_4}{\chi^2\cos^2\theta_4 + 1}, \quad T \equiv \frac{2\chi\cos\theta_4}{\chi^2\cos^2\theta_4 - 1}, \tag{13}$$

where the second-quadrant value of $\tan^{-1} T$ is to be used when $T < 0$. The other two domain areas can now be written in terms of equation (11),

$$A_1 = 2\pi(1 - z_0) - \tfrac{1}{2}A_2, \quad A_3 = 2\pi(1 + z_0) - \tfrac{1}{2}A_2. \tag{14}$$

Figure 6 below displays the domain areas as a function of $\alpha/g = \sin I/\sin\theta_4 - \cos I/\cos\theta_4$. Note that when $\chi = 0$ we have $\rho = 0$, $\tan^{-1}T = \pi$, and $\tan^{-1}(1/\chi) = \pi/2$. Substituting these values into equation (11) yields the critical value of $A_2 = A_{\rm crit} = 4\pi(1 - z_{\rm crit})$ for which A_1 vanishes, indicating the

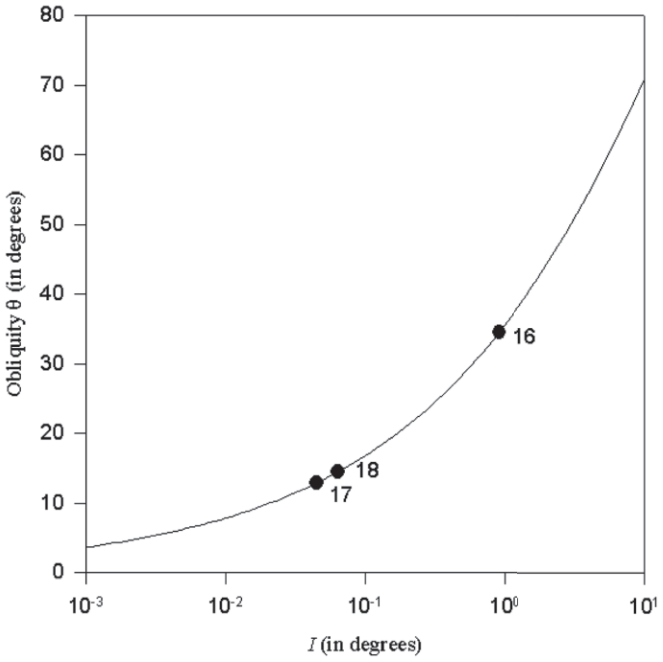

FIG. 5.—Adiabatic values of the obliquity excited by an arbitrarily slow resonance passage in the noncapture direction as a function of the inclination amplitude I. The values corresponding to the $j = 16$, 17, and 18 terms for Saturn's orbit are indicated.

merger of states 1 and 4. Finally, setting $z_{\rm crit} = -(g/\alpha)_{\rm crit} \times \cos I$, the area surrounding state 2 at merger becomes

$$A_{\rm crit} = 4\pi[1 - (1 + \tan^{2/3} I)^{-3/2}]. \quad (15)$$

If state 2 were to then migrate near the orbit normal (i.e., $|\alpha/g| \ll 1$), the precession would become almost uniform, with an obliquity given by

$$\cos\theta = 1 - \frac{A_{\rm crit}}{2\pi} = \frac{2}{(1 + \tan^{2/3} I)^{3/2}} - 1. \quad (16)$$

Figure 5 shows this obliquity as a function of inclination amplitude. Again, these are adiabatic values corresponding to arbitrarily slow passage. The obliquities for amplitudes in Table 1 are indicated on the curve; for $I_{18} = 0\overset{\circ}{.}064$, $\theta = 14\overset{\circ}{.}5$.

4. TUNING MECHANISMS

If Saturn was captured into a secular spin-orbit resonance with Neptune, how and when did this occur? For the present frequency ratio, state 1 lies very close to the pole position of the $j = 18$ term (Fig. 4). It is only because of Saturn's large obliquity that the ratio of Neptune's orbital precession rate to Saturn's pole precession rate could be near unity. To account for this as a result of resonance capture, the case must be made for either an increase in α or a decrease in $|g_{18}|$ to "tune" the system through the critical frequency ratio in the proper direction. Below we discuss possible adjustments in the early solar system that could account for this, although they may not be unique.

4.1. *The Kuiper Belt*

The regression of Neptune's nodal line is caused by the orbit-averaged gravity of the planets interior to it. If there were planets exterior to Neptune, they would each contribute to g_{18} by an amount

$$\delta g \approx -\frac{n_{\rm N}}{4}\left(\frac{M_p}{1\ M_\odot}\right)\left(\frac{a_{\rm N}}{a_p}\right)^2 b^{(1)}_{3/2}(a_{\rm N}/a_p), \quad (17)$$

where M_p and a_p are the planet's mass and semimajor axis, $a_{\rm N} = 30.1$ AU and $n_{\rm N} = 7\overset{''}{.}85 \times 10^3$ yr^{-1} are the semimajor axis and mean motion of Neptune, $M_\odot$ is the solar mass, and the quantity $b^{(1)}_{3/2}(\gamma)$ is a Laplace coefficient (see, e.g., Brouwer & Clemence 1961). Pluto does this, but its mass is so small ($M_{\rm P} = 2.2 \times 10^{-3}$ Earth masses, $M_\oplus$; $a_{\rm P} = 39.5$ AU) that its fractional contribution is only $\delta g/g_{18} \approx 3 \times 10^{-5}$. However, Pluto is generally regarded as a remnant of a larger Kuiper belt population that was eroded away over time (see, e.g., Holman & Wisdom 1993; Duncan et al. 1995). The contribution of a primordial Kuiper belt of surface density σ to Neptune's precession can be estimated by replacing M_p with $2\pi\sigma r\, dr$ in equation (17) and integrating over the width of the belt. Starting at Pluto's distance, $r_{\rm K} \sim 40$ AU, and integrating to $\sim$50 AU, where recent observations indicate an outer edge to the belt (Allen et al. 2001; Trujillo & Brown 2001), yields a fractional contribution of order $\delta g/g_{18} \approx 10^{-2} M_{\rm K}/(1\ M_\oplus)$, or about 1% for each Earth mass of material. The present mass $M_{\rm K}$ of the Kuiper belt is on the order of a few times $10^{-1}\ M_\oplus$, but its primordial mass, estimated by extrapolating the planetesimal disk from Neptune into the region, could have been as high as a few times $10\ M_\oplus$ (e.g., Stern & Colwell 1997; Farinella et al. 2000; Malhotra et al. 2000). This is sufficient to place Saturn to the left of $(\alpha/g)_{\rm crit}$ in Figure 3, implying that it passed through the resonance in the capture direction as the mass of the belt diminished. We should also point out that a concomitant outward migration of Neptune (see, e.g., Hahn & Malhotra 1999) would result in a decreasing $|g_{18}|$ as well. Numerical experiments of the erosion of the belt indicate a timescale in excess of $O(10^8)$ yr for the portion of the belt beyond $\sim$40 AU. The final obliquity of Saturn would then simply be the limiting value it acquired by the time the Kuiper belt's mass was exhausted.

4.2. *Spin-Axis Librations*

If Saturn is currently trapped in the resonance, Figure 4 shows that it is librating about state 2. The inferred libration amplitude $\psi_{\rm max}$ is sensitive to where we put the separatrix or, equivalently, to the exact value of α/g_{18}. There is some uncertainty in g_{18}, but it is probably small; the planetary theory constructed by Bretagnon (1974) including fourth-order long-period terms with short-period term corrections gives $0\overset{''}{.}691$ yr^{-1}, while Applegate et al. (1986) Fourier-transform the orbital elements of a 100 Myr numerical integration of the outer five planets to find $0\overset{''}{.}692$ yr^{-1}. The greatest uncertainty in α is through the moment of inertia of Saturn, but resonance occupancy places a constraint on λ.

The smallest area enclosed by a librating trajectory is found by making the current spin-axis position the amplitude $\psi_{\rm max}$, while the largest area is found by putting Saturn's pole on the separatrix itself, for z_0 either greater or less than z_s so that $\psi_{\rm max} = \pi$. The area inside the current trajectory can be found as a function of α/g,

$$A = \int_{\rm traj} \sin\theta\, d\theta\, d\phi = \int dz\, d\phi = 2\int_{z_1}^{z_2} [\pi/2 - \phi(z)]dz, \quad (18)$$

where $\sin\phi(z) = [K + (z - z_0)^2/2p]/(1 - z^2)^{1/2}$ and the integration limits are the values of z for which $\sin\phi(z) = 1$

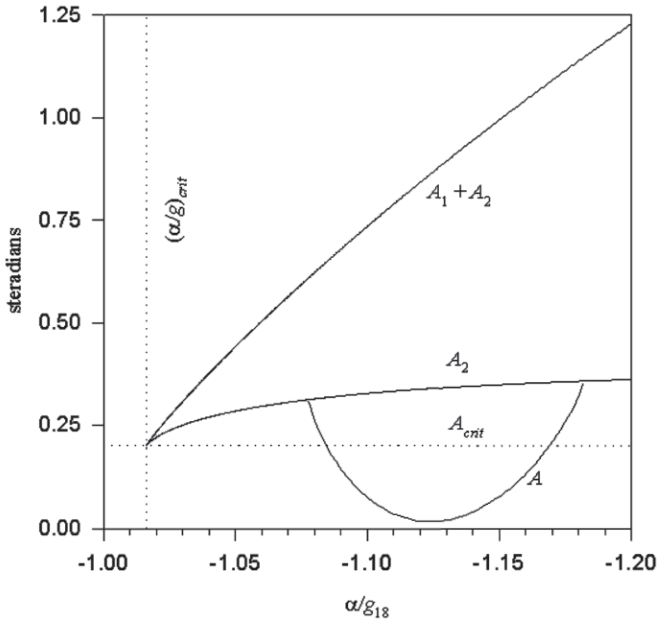

FIG. 6.—Domain areas as a function of frequency ratio. The top line is the sum of domains 1 and 2; the second curve is domain 2 only. The lowest curve shows possible loci of area A enclosed by Saturn's current spin-axis trajectory. Intersections with A_2 limit the frequency ratio for trapping; intersections with $A_{\rm crit}$ (*horizontal dotted line*) limit the curve to precapture obliquities less than 14°.5, for which capture is certain.

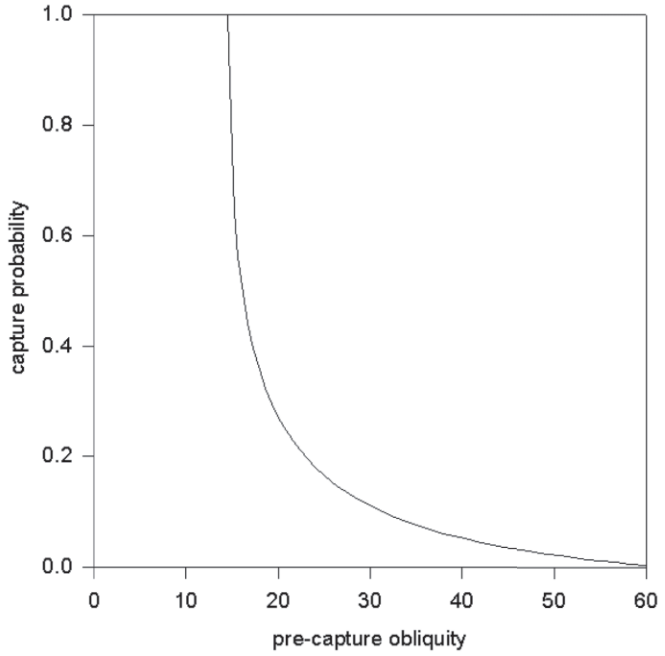

FIG. 7.—Capture probabilities into domain 2 as a function of the precapture obliquity.

(Ward et al. 1979). Saturn's current pole position constrains allowable values for the vertex, $K = y_s + (z_s - z_0)^2/2p$. The minimum $K_{\rm min} = y_s$ occurs when $z_0 = z_s$ and $\alpha/g_{18} = -(\cos I)/z_s = -1.126$. The associated value of $A = 0.0194$ sets a minimum precapture obliquity of $\theta_{\rm min} = 4^\circ.5$. A curve of A-values compatible with the current $\boldsymbol{s}$ is included in Figure 6. The inferred range of uncertainty in the frequency ratio, $1.078 < |\alpha/g| < 1.182$, can be related to Saturn's moment of inertia through equation (1), that is, $0.2233 < \lambda < 0.2452$ gives the range of values for which Saturn could be currently trapped in the resonance. The minimum is ~1% larger than the Hubbard & Marley value but well within its uncertainty.

For I_{18}, $A_{\rm crit} = 4\pi(1 - z_{\rm crit}) = 0.2003$; projecting this case back to $|\alpha/g| \ll |\alpha/g|_{\rm crit}$ gives the maximum obliquity, $\theta = \cos^{-1}(2z_{\rm crit} - 1)$, for certain capture, which turns out to recover equation (16), that is, $\theta = 14^\circ.5$. From Figure 6, $A = A_{\rm crit}$ for α/g_{18} of -1.084 and -1.169. The corresponding inertia range is $0.2257 < \lambda < 0.2438$. Capture is possible for larger θ, but at a decreasing probability given by

$$P = |\dot{A}_2/\dot{A}_3| = 2/(1 - 4\pi\dot{z}_0/\dot{A}_2). \quad (19)$$

The rate of change of A_2 is found from $\dot{A}_2 = \dot{z}_0\,\partial A_2/\partial z_0 + \dot{p}\,\partial A_2/\partial p$, where the partial derivatives are given by Henrard & Murigande (1987):

$$\partial A_2/\partial z_0 = -8\tan^{-1}(1/\chi), \quad \partial A_2/\partial p = 8\rho/p. \quad (20)$$

Substitution into equation (19) yields

$$P = \frac{2}{1 + \frac{1}{2}\pi[\tan^{-1}(1/\chi) - \rho/z_0]^{-1}}, \quad (21)$$

which is to be evaluated at the moment of separatrix crossing. If, when $|\alpha/g| \ll |\alpha/g|_{\rm crit}$, the original obliquity is $\theta_0 > 14^\circ.5$, the area $A_0 = 2\pi(1 + \cos\theta_0)$ of the unit sphere below the trajectory is less than value of A_3 when the separatrix first appears, that is, the spin axis is in domain 3 outside of the separatrix. As $|\alpha/g|$ increases, both A_1 and A_2 increase at the expense of A_3, as shown in Figure 6. The spin axis crosses the separatrix when $A_0 = A_3$, or $A_2 = 4\pi(z_0 - \cos\theta_0)$. Combining with equation (11), this condition reads

$$z_0\left(1 + \frac{2}{\pi}\tan^{-1}\frac{1}{\chi}\right) - \frac{1}{\pi}(2\rho + \tan^{-1}T) = \cos\theta_0, \quad (22)$$

which can be used to determine the transition values of χ and θ_4. Using these in equation (21) yields the probability. Figure 7 shows the capture probabilities as a function of the precapture obliquity.

5. DISCUSSION

What would be a likely origin of Saturn's libration? One clear possibility is a late impact. A fortuitous grazing impact near Saturn's pole at its escape velocity will shift the spin axis by $\delta\theta \sim (m/\lambda M_{\rm S})(GM_{\rm S}/\omega^2R^3)^{1/2} \sim 7^\circ[m/(1\ M_\oplus)]$, where m is the mass of the projectile; more probable impact parameters and angles would require several Earth masses. On the other hand, if the impact postdated resonance capture, the elongated nature of the trajectories decreases the required obliquity change by a factor $(\tan I/\tan\theta_0)^{1/4} \approx 0.22$ (see Paper II). Indeed, if the impactor shifts the axis more than the half-width of the separatrix itself, $\Delta\theta_s = 2(\tan I/\tan\theta_0)^{1/2} = 5^\circ.4$, it would knock Saturn out of the resonance. Another way to generate libration is by a somewhat nonadiabatic passage through the $j = 18$ resonance in the capture direction on a timescale comparable to the libration time of $\sim 8 \times 10^7$ yr. We note that this is not too different from the erosion timescale of the Kuiper belt, and we numerically assess this possibility in Paper II. Nonadiabaticity may also be introduced if Neptune migrates in a stochastic manner (Hahn & Malhotra 2000). This could cause diffusion of the spin axis inside the separatrix. If the g_{18} splits into a cluster

of similar terms, this could introduce chaos in a manner similar to that found for Mars (Touma & Wisdom 1993; Laskar & Robutel 1993). However, current orbit theory does not yet show much evidence for this.

An intriguing alternative is to evoke *two* passes through g_{18} starting with a nonadiabatic passage in the noncapture direction during Saturn's Kelvin-Helmholtz contraction. During contraction, the spin angular momentum of Saturn, $L = \lambda M_S R^2 \omega$, remains constant, so that $\omega \propto R^{-2}$. Equation (1) then indicates that α increases with the quantity $J_2 R^2$, while the rotationally induced value of J_2 is $\omega^2 R^3 / G M_S$ (e.g., Kaula 1968). Consequently, $J_2 R^2 \propto \omega^2 R^5 \propto R$, and $\alpha = \alpha_0 (R/R_0 + q)/(1+q)$ was larger when Saturn was more distended, where α_0 and R_0 denote the current precessional constant and planetary radius. Accordingly, an increase of $\delta R/R_0 \gtrsim (1+q)\delta g/\alpha_0$ would more than compensate for a primordial increase in g_{18}, reinstating Saturn to the right of the resonance in Figure 3. Contraction then drives Saturn through the resonance in the noncapture direction with a maximum induced obliquity of 14°.5. If the passage is fast enough to break the adiabatic invariant at some point, the induced obliquity, $\theta \approx gI(2\pi/\dot{a})^{1/2}$, will be less than the adiabatic value (see Appendix). Consequently, the minimum precapture obliquity $\theta_{\min}$ could be used to put a limit on how fast Saturn could have contracted during its first resonance passage. The minimum characteristic timescale, $\tau \equiv \alpha/|\dot{\alpha}|$, for resonance passage is

$$\tau_{\min} \approx \frac{2\pi}{g_{18}} \left(\frac{\theta_{\min}}{2\pi I_{18}} \right)^2 \approx 2 \times 10^8 \text{ yr}, \tag{23}$$

which in turn implies a minimum characteristic contraction timescale $\tau_R \equiv R/\dot{R} = \tau_{\min}/(1 + q/J_2) \approx 5 \times 10^7$ yr. This is consistent with models of the contraction of the gas giant planets using modified stellar evolution codes (e.g., Pollack et al. 1976; Bodenheimer & Pollack 1986).

6. SUMMARY

Saturn's spin-axis precession period is close to the precession period (1.87×10^6 yr) of Neptune's orbit plane. We propose that these planets are locked into a secular spin-orbit resonance, and that this is the origin of Saturn's relatively large obliquity (26°.7) compared with that of Jupiter (3°.1). We have outlined a sequence of events that could account for the establishment of this resonant state. Initially forming with a small obliquity, Saturn passed through the resonance in the capture direction as the Kuiper belt was depleted, pumping up its obliquity until it eventually acquired its current value. A Saturn currently in resonance places a constraint on that planet's normalized moment of inertia $\lambda > \lambda_{\text{crit}} = 0.2233$ and implies that the spin axis is librating with an amplitude $\geq 31^\circ$. This could be a fossil remnant of a precapture obliquity generated in an earlier resonance pass during the planet's Kelvin-Helmholtz contraction. Alternatively, the librations may have been caused by nonadiabatic conditions during resonance passage (Paper II) or excited by a late impact. Numerical experiments described in our companion paper illustrate resonant capture in detail. The critical precession frequency separating circulating from librating spin-axis trajectories is $\alpha/g_{18} = -1.18$, $\alpha \cos\theta = 0''.730 \text{ yr}^{-1}$, and perhaps further observational data will be able to discriminate between them.

We thank R. Canup, S. Peale, and A. Stern for constructive comments on the manuscript, and P. Goldreich for pointing out that a modest impact could excite Saturn's libration about the resonance. W. R. W. was supported by a research grant from NASA's Planetary Geology and Geophysics Program. D. P. H. acknowledges support from an NSF Career award and from the Southwest Research Institute during his sabbatical stay.

APPENDIX

NONADIABATIC OBLIQUITIES

The equation of motion (eq. [6]) in component form reads

$$\begin{aligned} \dot{x} &= (\alpha z + g \cos I) y - g \sin I z, \\ \dot{y} &= -(\alpha z + g \cos I) x, \\ \dot{z} &= x \sin I. \end{aligned} \tag{A1}$$

Consider again a small-angle approximation: $z \sim \cos I \sim 1$, $\sin I \sim I$, and introduce a new independent variable $\varphi \equiv \int_o^t (\alpha + g)\, dt$, where we take $t = 0$ to be the moment when $\alpha = -g$. The x- and y-equations can be combined to yield a second-order equation $d^2y/d\varphi^2 + y = gI/(\alpha + g)$ with solution

$$y = gI \left(-\cos\varphi \int \sin\varphi\, dt + \sin\varphi \int \cos\varphi\, dt \right). \tag{A2}$$

We now expand $\alpha \approx -g + \dot{\alpha} t$ so that $\varphi \approx \dot{\alpha} t^2/2 \equiv t'^2$. Integration of equation (A2) then gives

$$y = gI\sqrt{\pi/\dot{\alpha}} \left\{ \left[C_1(t') + \tfrac{1}{2} \right] \sin t'^2 - \left[S_1(t') + \tfrac{1}{2} \right] \cos t'^2 \right\} \tag{A3}$$

(Ward et al. 1976), where $C_1(t')$ and $S_1(t')$ are Fresnel integrals and we have required $y \to 0$ as $t' \to -\infty$. As $t' \to +\infty$, C_1 and S_1 approach $\frac{1}{2}$, yielding an obliquity proportional to $\dot{\alpha}^{-1/2}$, namely,

$$\theta \to gI\sqrt{2\pi/\dot{\alpha}}. \tag{A4}$$

REFERENCES

Agnor, C. B., Canup, R. M., & Levison, H. F. 1999, Icarus, 142, 219
Allen, C. W. 1976, Astrophysical Quantities (3rd ed.; London: Athlone)
Allen, R. L., Bernstein, G. M., & Malhotra, R. 2001, ApJ, 549, L241
Applegate, J. H., Douglas, M. R., Gürsel, Y., Sussman, G. J., & Wisdom, J. 1986, AJ, 92, 176
Bodenheimer, P., & Pollack, J. B. 1986, Icarus, 67, 391
Bretagnon, P. 1974, A&A, 30, 141
Bretagnon, P., & Francou, G. 1992, in IAU Symp. 152, Chaos, Resonance and Collective Dynamical Phenomena in the Solar System, ed. S. Ferraz-Mello (Dordrecht: Kluwer), 37
Brouwer, D., & Clemence, G. M. 1961, Methods of Celestial Mechanics (New York: Academic)
Brouwer, D., & van Woerkom, A. J. J. 1950, The Secular Variations of the Orbital Elements of the Principal Planets (Astron. Pap. Am. Ephemeris Naut. Alm., 13, 81) (Washington: GPO)
Chambers, J. E., & Wetherill, G. W. 1998, Icarus, 136, 304
Colombo, G. 1966, AJ, 71, 891
Dones, L., & Tremaine, S. 1993, Icarus, 103, 67
Duncan, M. J., Levison, H. F., & Budd, S. M. 1995, AJ, 110, 3073
Farinella, P., Davis, D. R., & Stern, S. A. 2000, in Protostars and Planets IV, ed. V. Mannings, A. P. Boss, & S. S. Russell (Tucson: Univ. Arizona Press), 1255
French, R. G., et al. 1993, Icarus, 103, 163
Goldreich, P. 1965, AJ, 70, 5

Hahn, J. M., & Malhotra, R. 1999, AJ, 117, 3041
———. 2000, BAAS, 32, 857
Hamilton, D. P., & Ward, W. R. 2004, AJ, 128, 2510 (Paper II)
Harris, A. W., & Ward, W. R. 1982, Annu. Rev. Earth Planet. Sci., 10, 61
Henrard, J., & Murigande, C. 1987, Celest. Mech., 40, 345
Holman, M. J., & Wisdom, J. 1993, AJ, 105, 1987
Hubbard, W. B., & Marley, M. S. 1989, Icarus, 78, 102
Kaula, W. M. 1968, An Introduction to Planetary Physics (New York: Wiley)
Laskar, J., & Robutel, P. 1993, Nature, 361, 608
Lissauer, J. J., & Safronov, V. S. 1991, Icarus, 93, 288
Malhotra, R., Duncan, M. J., & Levison, H. F. 2000, in Protostars and Planets IV, ed. V. Mannings, A. P. Boss, & S. S. Russell (Tucson: Univ. Arizona Press), 1231
Nicholson, P. D., & French, R. G. 1997, BAAS, 28, 1097
Nicholson, P. D., French, R. G., & Bosh, A. S. 1999, BAAS, 31(4), Div. Planet. Sci. abstr. No. 44.01
Peale, S. J. 1969, AJ, 74, 483
———. 1974, AJ, 79, 722
Pollack, J. B., Grossman, A. S., Moore, R., & Graboske, H. C., Jr. 1976, Icarus, 29, 35
Stern, S. A., & Colwell, J. E. 1997, AJ, 114, 841
Touma, J., & Wisdom, J. 1993, Science, 259, 1294
Tremaine, S. 1991, Icarus, 89, 85
Trujillo, C. A., & Brown, M. E. 2001, ApJ, 554, L95
Vienne, A., & Duriez, L. 1992, A&A, 257, 331
Ward, W. R. 1973, Science, 181, 260
———. 1974, J. Geophys. Res., 79, 3375
———. 1975, AJ, 80, 64
———. 1979, J. Geophys. Res., 84, 237
———. 1992, in Mars, ed. H. H. Kieffer, B. M. Jakosky, C. W. Snyder, & M. S. Matthews (Tucson: Univ. Arizona Press), 298
Ward, W. R., Burns, J. A., & Toon, O. B. 1979, J. Geophys. Res., 84, 243
Ward, W. R., Colombo, G., & Franklin, F. A. 1976, Icarus, 28, 441
Wuchterl, G., Guillot, T., & Lissauer, J. J. 2000, in Protostars and Planets IV, ed. V. Mannings, A. P. Boss, & S. S. Russell (Tucson: Univ. Arizona Press), 1081
Yoder, C. F. 1995, in Global Earth Physics, ed. T. J. Ahrens (Washington: Am. Geophys. Union), 1