1. The averaging principle

Consider motion in a near-integrable Hamiltonian with n degrees of freedom,

$$H(\mathbf{J},\boldsymbol{\theta}) = H_0(\mathbf{J}) + \epsilon H_1(\mathbf{J},\boldsymbol{\theta}). \tag{1}$$

where $(\mathbf{J},\boldsymbol{\theta})$ are action-angle variables in the unperturbed Hamiltonian and $\epsilon \ll 1$. Equations of motion are

$$\dot{\mathbf{J}} = -\frac{\partial H}{\partial \boldsymbol{\theta}}, \quad \dot{\boldsymbol{\theta}} = \frac{\partial H}{\partial \mathbf{J}}. \tag{2}$$

To lowest order in ϵ

$$\mathbf{J} = \text{const}, \quad \boldsymbol{\theta} = \boldsymbol{\theta}_0 + \boldsymbol{\Omega} t, \quad \boldsymbol{\Omega} = \frac{\partial H_0}{\partial J}. \tag{3}$$

To first order the actions and angles have oscillatory terms of relative amplitude ϵ and frequency of order Ω.

Suppose H_0 is independent of k of the n actions ("slow" actions); thus $H_0 = H_0(J_1, \ldots, J_{n-k})$ (the "fast" actions). The averaging principle consists of the replacement of H_1 by

$$\overline{H}_1(\mathbf{J}, \theta_{n-k+1}, \ldots, \theta_n) = \frac{1}{(2\pi)^k} \int d\theta_1 \cdots d\theta_{n-k} H_1(\mathbf{J},\boldsymbol{\theta}). \tag{4}$$

In words, $\overline{H}_1$ is the average of H_1 over the fast angles. Since $\overline{H}_1$ is independent of the fast angles, the fast actions are all conserved and the system is reduced from n degrees of freedom to k degrees of freedom.

Arnold (1989, p. 292) has commented: "this principle is neither a theorem, an axiom, nor a definition, but rather a physical proposition, i.e., a vaguely formulated and, strictly speaking, untrue assertion."

2. The averaging principle in quasi-Keplerian systems

The averaging principle is relevant to planetary systems, triple star systems, and to nuclear star clusters surrounding supermassive black holes (see Fouvry's talk on vector resonant relaxation, which is an example of secular effects in such systems). The difference is that nuclear star clusters have far more bodies (10^6 instead of ~ 10) and the bodies are on much more eccentric and inclined orbits.

The Hamiltonian is the standard N-body Hamiltonian

$$H = \sum_{i=1}^{N} \left(\tfrac{1}{2} v_i^2 - \frac{GM}{r_i} \right) - \sum_{i=1}^{N} \sum_{j<i} \frac{G m_i m_j}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{5}$$

where $m_i \ll M$. Although this has $3N$ degrees of freedom we will often focus on the case where we follow only one body subject to fixed perturbations.

$$H = \tfrac{1}{2}v^2 - \frac{GM}{r} + \epsilon\Phi(\mathbf{r}, t). \tag{6}$$

In the Kepler problem each planet has three actions

$$J_1 = \sqrt{GMa}, \quad J_2 = J_1(1 - \sqrt{1-e^2}), \quad J_3 = J_1\sqrt{1-e^2}(1 - \cos I) \tag{7}$$

where a, e, I are the semimajor axis, eccentricity, and inclination. For planets e and I are small, so

$$J_2 \simeq \tfrac{1}{2}J_1 e^2, \quad J_3 \simeq \tfrac{1}{2}J_1 I^2. \tag{8}$$

The unperturbed or Kepler Hamiltonian is

$$H_0 = -\frac{GM}{2a} = -\frac{G^2M^2}{2J_1^2} \tag{9}$$

Thus H_0 depends only on J_1 so $\Omega_2 = \Omega_3 = 0$. Thus J_2 and J_3 are slow actions.

This means we have to treat terms in H_1 that depend on θ_1 differently from terms that do not. Let

$$\overline{H}_1 = \frac{1}{2\pi}\int d\theta_1 H_1(\mathbf{J}, \boldsymbol{\theta}). \tag{10}$$

Since $\overline{H}_1$ is independent of θ_1,

$$\dot{J}_1 = -\frac{\partial \overline{H}_1}{\partial \theta_1} = 0. \tag{11}$$

In the solar system this implies that the semimajor axes of the planets are constant (fractional change in Earth's semimajor axis is only a few parts in 10^{-5}).

On the other hand

$$\dot{J}_{2,3} = -\frac{\partial \overline{H}_1}{\partial \theta_{2,3}} = \text{const}. \tag{12}$$

In the solar system this implies that the eccentricities and inclinations grow without limit.

This was the dark matter problem of the eighteenth century. Do the variations in the orbits gradually grow, leading eventually to the catastrophic disruption of the solar system?

This question has fascinated physicists and mathematicians since the time of Newton. Newton apparently believed that the perturbations did grow, stating in his book Opticks in 1730 that the "irregularities" in the solar system arising "from the mutual actions of planets upon one another" would gradually grow until the solar system "wants a reformation," that is, until God intervenes to restore order.

This is the point of view of a theist, i.e.,

Theism: God exists, and continually interacts with humans and the known universe via methods of divine intervention.

Deism: God exists and created the universe but theeafter left it alone and does not interact with humans or the known universe.

Leibniz was a deist. He believed that the perfection of God required the perfection of the solar system, and complained in 1715 that "according to [Newtons] doctrine, God Almighty wants to wind up his watch from time to time . . . he had not, it seems, sufficient foresight to make it a perpetual motion."

The controversy between Newton and Leibniz was influenced by observations of Jupiter and Saturn dating back to Johannes Kepler in 1625, which seemed to show that their semi-major axes were changing linearly in time.

The actions $J_{2.3}$ are nearly zero and using polar coordinates near zero is generally a bad idea. Carry out a canonical transformation from J_2, θ_2 to

$$q_2 = \sqrt{2J_2}\cos\theta_2, \quad p_2 = \sqrt{2J_2}\sin\theta_2, \tag{13}$$

with a similar transformation from J_3, θ_3 to q_3, p_3. The generating function is $S = -\frac{1}{2}p_2^2\cot\theta_2$.

The advantage of these variables is that (a) analogous to the transformation from polar to Cartesian coordinates, they avoid a singularity near the origin. (b) p_2 and q_2 are of order e which is small, and p_3 and q_3 are of order I, which is small. (c) The interaction Hamiltonian $\overline{H}_1$ can be expanded in a Taylor series in the p's and q's.

Therefore expand $\overline{H}_1$ as a power series in the q's and p's and keep only terms up to quadratic order.

$$\overline{H}_1 = \sum_{i,j} \frac{Gm_j}{a_j}[\frac{1}{8}\alpha_{ij}b^1_{3/2}(\alpha_{ij})(q_{2i}^2 + p_{2i}^2 + q_{2j}^2 + p_{2j}^2) - \frac{1}{4}\alpha_{ij}b^2_{3/2}(\alpha_{ij})(q_{2i}q_{2j} + p_{2i}p_{2j})], \tag{14}$$

plus terms in q_{3i}, p_{3i} where

$$b^m_s(\alpha) = \frac{2}{\pi}\int_0^\pi \frac{d\phi\cos m\phi}{(1-2\alpha\cos\phi+\alpha^2)^s}. \tag{15}$$

and $\alpha_{ij} = a_i/a_j$. This is **Lagrange-Laplace theory**.

Any quadratic Hamiltonian can be solved analytically. It is found that (a) All linear terms in the Hamiltonian vanish. (b) The equations of motion are linear and homogeneous; (b) the q's and p's oscillate with a frequency of order $\epsilon\Omega$; (c) the amplitudes of the oscillations are of the same order as the initial eccentricities and inclinations (and therefore not large enough to disrupt the solar system).

This still left unresolved the apparent drift in semimajor axes of Jupiter and Saturn. Then in 1785, Laplace showed that this arises because of a near-resonance between the two planets: their mean motions are related by $2n_{Jupiter} \simeq 5n_{Saturn}$. This near-resonance, sometimes called the Great Inequality, leads to oscillations in the mean motions with a period of about 900 years, and this variation appeared nearly linear over the 150 years between Kepler and Laplace.

Lagrange-Laplace theory provided an early approximate "proof" that the solar system was stable. Many other authors made analytic proofs but these always involved approximations. Poincaré said "Those who are interested in the progress of celestial mechanics...must feel some astonishment at seeing how many times the stability of the Solar System has been demonstrated. Lagrange established it first, Poisson has demonstrated it again, other demonstrations came afterward, others will come again. Were the old demonstrations insufficient, or are the new ones unnecessary?"

The actual "proofs" come from numerical integration of the planetary orbits over timescales of 5–10 Gyr. This is a hard problem for several reasons: (a) it requires $\sim 10^{12}$ timesteps and so is very CPU-intensive; (b) impossible to parallelize because it's an intrinsically serial problem; (c) must manage roundoff error. These problems have now been solved and you can check the stability of the solar system on your laptop in a few days.

The answer is that the system is stable but chaotic, with a short Liapunov time of only 10^7 years.

3. Milankovich equations

The Lagrange-Laplace theory is only valid for small eccentricity and inclination and we need a theory that works for arbitrary eccentricity and inclination. In principle this can be done by writing $\overline{H}_1(J_1, J_2, J_3, \theta_2, \theta_3, t)$ and solving Hamilton's equations but (a) these equations are complicated and lack any natural structure; (b) they are ill-defined when the eccentricity e or inclination I is zero, or when $e = 1$. These disadvantages can be remedied in a vector-based formalism for secular theory, Since $\mathbf{r}$ and $\mathbf{v}$ form a canonical coordinate-momentum pair, their Poisson brackets are

$$\{r_i, r_j\} = 0, \quad \{v_i, v_j\} = 0, \quad \{r_i, v_j\} = -\{v_i, r_j\} = \delta_{ij}, \tag{16}$$

The angular momentum per unit mass $\mathbf{L} = \mathbf{r} \times \mathbf{v} = \epsilon_{ijk}\hat{\mathbf{n}}_i r_j v_k$, where ϵ_{ijk} is the permutation symbol; throughout this section the summation convention is in force. It is straightforward to show that the Poisson brackets of the components of angular momentum are

$$\{L_i, L_j\} = \epsilon_{ijk} L_k. \tag{17}$$

Since $|L| = [GMa(1-e^2)]^{1/2}$, it proves useful to define a dimensionless angular momentum

$$\mathbf{j} \equiv \frac{\mathbf{L}}{(GMa)^{1/2}} = (1-e^2)^{1/2}\hat{\mathbf{L}}, \tag{18}$$

whose magnitude varies between 0 and 1. It is straightforward to show that equation (17) can be rewritten as

$$\{j_i, j_j\} = \frac{1}{(GMa)^{1/2}}\epsilon_{ijk} j_k. \tag{19}$$

Define the eccentricity vector:

$$\mathbf{e} = \frac{\mathbf{v} \times (\mathbf{r} \times \mathbf{v})}{GM} - \frac{\mathbf{r}}{r}. \tag{20}$$

It is straightforward to show that $\mathbf{e}$ is a fixed vector whose magnitude is the eccentricity and that points towards pericenter.

The orbit-averaged Hamiltonian $\overline{H}_1$ is a function of the size and shape of the orbit, and possibly the time—it's simply the gravitational potential energy of an elliptical wire in the external field. We can specify the orbit by the semimajor axis a and the two vectors $\mathbf{j}$ and $\mathbf{e}$. Thus we can write the orbit-averaged Hamiltonian as $\overline{H}_1(a, \mathbf{j}, \mathbf{e}, t)$. Note that these arguments contain seven phase-space variables (a and the three components of each of the two vectors), but only five are independent, because they are related by the constraints

$$\mathbf{j} \cdot \mathbf{e} = 0, \quad j^2 + e^2 = 1. \tag{21}$$

It is straightforward, though tedious, to show that the Poisson brackets of the components of the eccentricity vector are

$$\{e_i, e_j\} = \frac{1}{(GMa)^{1/2}}\epsilon_{ijk} j_k. \tag{22}$$

Similarly we can show that

$$\{j_i, e_j\} = \{e_i, j_j\} = \frac{1}{(GMa)^{1/2}} \epsilon_{ijk} e_k. \tag{23}$$

The elegant relations (19), (22) and (23) arise because the symmetry group of the Kepler problem is the group of rotations in 4-dimensional space, SO(4).

Let $\mathbf{z} \equiv (\mathbf{r}, \mathbf{v})$. Then Hamilton's equations can be written

$$\dot{\mathbf{z}} = \{\mathbf{z}, H\}. \tag{24}$$

The rate of change of any function $f(\mathbf{z}, t)$ along a trajectory determined by Hamilton's equations is

$$\frac{d}{dt} f[\mathbf{z}(t), t] = \frac{\partial f}{\partial t} + \{f, H\}. \tag{25}$$

The time evolution of j_i under the influence of $\overline{H}_1$ is given by equation (25),

$$\frac{dj_i}{dt} = \{j_i, \overline{H}_1\}. \tag{26}$$

Then from the chain rule

$$\frac{dj_i}{dt} = \{j_i, j_k\} \frac{\partial \overline{H}_1}{\partial j_k} + \{j_i, e_k\} \frac{\partial \overline{H}_1}{\partial e_k} + \{j_i, a\} \frac{\partial \overline{H}_1}{\partial a}. \tag{27}$$

Using the evaluations of the Poisson brackets in equations (19) and (23), the result simplifies to

$$\frac{dj_i}{dt} = \frac{1}{(GMa)^{1/2}} \epsilon_{ikm} j_m \frac{\partial \overline{H}_1}{\partial j_k} + \frac{1}{(GMa)^{1/2}} \epsilon_{ikm} e_m \frac{\partial \overline{H}_1}{\partial e_k}. \tag{28}$$

This can be rewritten in vector notation as

$$\frac{d\mathbf{j}}{dt} = -\frac{1}{(GMa)^{1/2}} \left(\mathbf{j} \times \frac{\partial}{\partial \mathbf{j}} \overline{H}_1 + \mathbf{e} \times \frac{\partial}{\partial \mathbf{e}} \overline{H}_1 \right), \tag{29}$$

where $\partial f / \partial \mathbf{j}$ is the vector having components $(\partial f/\partial j_1, \partial f/\partial j_2, \partial f/\partial j_3)$ for any function $f(j_1, j_2, j_3)$. Similarly, the time evolution of the eccentricity vector is given by

$$\frac{d\mathbf{e}}{dt} = -\frac{1}{(GMa)^{1/2}} \left(\mathbf{e} \times \frac{\partial}{\partial \mathbf{j}} \overline{H}_1 + \mathbf{j} \times \frac{\partial}{\partial \mathbf{e}} \overline{H}_1 \right). \tag{30}$$

Equations (29) and (30) are the **Milankovich equations**.

It is straightforward to show that the Milankovich equations conserve $\mathbf{j} \cdot \mathbf{e}$ and $j^2 + e^2$. Thus if the constraints (21) are satisfied by the initial conditions, they continue to be satisfied for all time. Because of this property the formula for a given Hamiltonian in terms of $\mathbf{j}$ and $\mathbf{e}$ is not unique—for example, $\overline{H}_1 = j^2$ could also be written $\overline{H}_1 = -e^2$ or $\overline{H}_1 = j^2 + \mathbf{e} \cdot \mathbf{j}$—but the trajectories determined by the Milankovich equations are the same for all of these.

4. ZLK oscillations

Consider a three-body problem consisting of a binary star with a planet orbiting one member of the binary. System must be hierarchical since otherwise it will not be stable.

The planet is a test particle and has semimajor axis a; on a circular orbit. The companion star has semimajor axis $a_c \gg a$. The inclination of the planet orbit to the binary star orbit is I (draw a diagram, including $\hat{\mathbf{n}}_c$). Then orbit-average over both the orbit of the companion and the orbit of the planet. Because $a \ll a_c$ only the quadrupole term is important. The Hamiltonian is

$$\begin{aligned}\overline{H}_1 &= \frac{GM_c a^2}{8a_c^3(1-e_c^2)^{3/2}}[15(\mathbf{e}\cdot\hat{\mathbf{n}}_c)^2 - 6e^2 - 3(\mathbf{j}\cdot\hat{\mathbf{n}}_c)^2] \\ &= \frac{GM_c a^2}{8a_c^3(1-e_c^2)^{3/2}}\left[\frac{6J_2^2}{J_1^2} - \frac{3J_3^2}{J_1^2} + 15\left(1 - \frac{J_3^2}{J_2^2} - \frac{J_2^2}{J_1^2} + \frac{J_3^2}{J_1^2}\right)\sin^2\theta_2\right]. \end{aligned} \tag{31}$$

Result is independent of θ_1 because of averaging, so J_1 and semimajor axis are conserved. Result is independent of θ_3 by accident so J_3 is conserved. Therefore there is only one degree of freedom. If initial eccentricity is zero and initial inclination is I_0 then conservation of the Hamiltonian implies

$$C = 5\left(1 - \frac{\cos^2 I_0}{1-e^2}\right)e^2\sin^2\omega - 2e^2 \tag{32}$$

is a constant. Set $x = e\cos\omega$, $y = e\sin\omega$, then near $e = 0$

$$C = 5\sin^2 I_0 y^2 - 2x^2 - 2y^2. \tag{33}$$

Stable if and only if $|\sin I_0| \le \sqrt{\frac{2}{5}}$ which requires that I_0 is less than 39° or more than 141°.

This instability leads to **von Zeipel-Lidov-Kozai oscillations.** These have at least two remarkable features:

1. The instability is independent of the mass or distance of the companion, so long as (a) there are no other perturbing effects on the planet (e.g., general relativity) and (b) the instability time is short compared to the age of the universe.

2. If the initial inclination is close to 90 degrees then the ZLK oscillation carries the eccentricity to nearly unity and there can be a collision with the star or the planet can come close enough that tides damp the orbit.

ZLK oscillations underpin one of the two main theories for how planets are formed close to their host star (hot Jupiters).