PHY 411 Advanced Classical Mechanics
(Chaos)
U. of Rochester Spring 2002

S. G. Rajeev

March 4, 2002

Contents

Preface to the Course

Introduction

Only simple, exceptional, mechanical systems admit an explicit solution in terms of analytic functions. This course will be mainly about systems that cannot be solved in this way so that approximation methods are necessary. In recent times this field has acquired the name 'Chaos theory', which has grown to include the study of all nonlinear systems. I will restrict myself mostly to examples arising from classical physics. The first half of the course will be accessible to undergraduates and experimentalists.

Pre-requisites

I will assume that all the students are familiar with Classical mechanics at the level of PHY 235, our undergraduate course (i.e., at the level of the book by Marion and Thornton.) Knowledge of differential equations and analysis at the level of our math department sophomore level courses will also be assumed.

Books

I recommend the books *Mechanics*, by Landau and Lifshitz and *Mathematical Methods of Classical Mechanics* by V. I. Arnold as a general reference although there is no required textbook for the course. The course will be slightly more advanced than the first book but will not go as far as Arnold's book. The more adventurous students should study the papers by Siegel, Kolmogorov, Arnold and Moser on invariant torii.

Homeworks,Exams and Grades

I will assign 2-3 homeworks every other week. Some will involve simple numerical calculations. There will be no exams. The course will be graded Pass-Fail.

Syllabus

- Newton's equations of motion.
- The Lagrangian formalism; generalized co-ordinates.

•Hamiltonian formalism; canonical transformations; Poisson brackets.
•Two body problem of Celestial mechanics. Integrability of the equations of motion.
•Perturbation theory; application to the three body problem.
•Restricted three body problem; Lagrange points.
•Normal co-ordinates; Birkhoff's expansion; small denominators and resonances.
•Invariant torii; Kolmogorov-Arnold-Moser theorems.
•Finding roots of functions by iteration: topological dynamics; onset of chaos by bifurcation.

Special Topics for Advanced Students

•Ergodic systems. Sinai billiard table; geodesics of a Riemann surface.
•Quantum Chaos: Gutzwiller's trace formula.
•Chaos in number theory: zeros of the Riemann zeta function.
•Spectrum of random matrices.

Chapter 1

Introduction

•Physics is the oldest and most fundamental of all the sciences; mechanics is the oldest and most fundamental branch of physics. All of physics is modelled after mechanics.

•The historical roots of mechanics are in astronomy-the discovery of reguarity in the motion of the planets, the sun and the moon by ancient astrologers in every civilization is the beginning of mechanics.

•The first regularity to be noticed is periodicity- but often there are several such periodic motions superposed on each other. The motion of the Sun has at least three such periods: with a period of one day, one year and 25,000 years (precession of the equinoxes). We know now that the first of these is due to the rotation of the Earth relative to an inertial reference frame, the second is due to the revolution of the Earth and the last is due to the precession of the Earth's axis of rotations.

•The first deep idea was to regard all motion as the superposition of such periodic motion- a form of harmonic analysis for quasi-periodic functions. This gave a quite good *description* of the motion.

•The precision at around 500 AD is already quite astonishing. The *Aryabhatiyam* gives the ratio of the length of the day to the year to an accuracy of better than one part in ten million; adopts the heliocentric view when convenient; gives the length of each month to four decimal places; even suggests that the unequal lengths of the months is due to the orbit of the Earth being elliptical rather than circular. There were similar parallel developments in China, the Arab world and elsewhere at that time.

•Astronomy upto and including the time of Kepler was mixed in with Astrology and many mystic beliefs formed the motivation for ancient astronomers.

•Kepler marks the transition from this early period to the modern era. The explanation of his three laws by Newtonian mechanics is the first deep result of the modern scientific method.

•Mechanics as we think of it today is mainly the creation of one man: Isaac Newton. The laws of mechanics whch he formulated by analogy with the axioms of Euclidean geometry form the basis of mechanics to this day, although the beginning of the last century saw two basic changes to the fundamental laws of physics.

•These are the theory of relativity and quantum mechanics. We still do not have a theory that combines these into a single unified science. In any case we will largely ignore these developments in these lectures.

Chapter 2

The Kepler Problem

•Much of mechanics was developed in order to understand the motion of planets. Long before Copernicus, many astronomers knew that the apparently erratic motion of the planets can be simply explained as circular motion around the Sun. For example, the *Aryabhateeyam* written in 499 AD gives many calculations based on this model. But various religious taboos and superstitions prevented this simple picture from being universally accepted. It is ironic that the same superstitions (e.g., astrology) were the prime cultural motivation for studying planetary motion.
•Kepler used Tycho Brahe's accurate measurements of planetary positions to find a set of important refinements of the heliocentric model. The three laws of planetary motion he discovered started the scientific revolution which is still continuing.

1 The first law of Kepler is: *Planets move along elliptical orbits with the Sun at a focus*.

2 An *ellipse* is a curve on the plane defined by the equation, in polar co-ordinates (r, θ)

$$\frac{\rho}{r} = 1 + \epsilon \cos\theta.$$

2.1 The parameter ϵ must be between 0 and 1 and is called the *eccentricity*. It measures the deviation of an ellipse from a circle: if $\epsilon = 0$ the curve is a circle of radius ρ . In the opposite limit $\epsilon \to 1$ (keeping ρ

fixed) it approaches a parabola. The parameter $\rho > 0$ measures the size of the ellipse.

•A more geometrical description of the ellipse is this: Choose a pair of points on the plane F_1, F_2 , the *Focii*. If we let a point move on the plane such that the sum of its distances to F_1 and F_2 is a constant, it will trace out an ellipse.
•Derive the equation for the ellipse above from this geometrical description. (Choose the origin of the polar co-ordinate system to be F_2). What is the position of the other focus F_1 ?
•The line connecting the two farthest points on an ellipse is called its *major axis*; this axis passes through the focii. The perpendicular bisector to the major axis is the *minor axis*. If these are equal in length, the ellipse is a circle; in this case the focii coincide. The length of the major axis is called $2a$ usually. Similarly, the semi-minor-axis is called b .
•Show that the major axis is $\frac{2\rho}{1-\epsilon^2}$ and that the eccentricity is $\epsilon = \sqrt{\left[1 - \frac{b^2}{a^2}\right]}$.
•The eccentricity of planetary orbits is quite small: a few percent. Comets and some asteroids and planetary probes have very eccentric orbits.
•If the eccentricity is greater then one, the equation describes a curve that is not closed, called a *hyperbola*.
•The second law of Kepler concerns the angular velocity of the planet:

3 *The line connecting the planet to the Sun sweeps equal areas in equal times* .

3.1 Since the rate of change of this area is $\frac{1}{2}r^2\frac{d\theta}{dt}$, this is the statement that

$$\frac{1}{2}r^2\frac{d\theta}{dt} = \text{constant}.$$

•The third law of Kepler gives a relation between the size of the orbit and its period.

4 *The ratio of the cube of the semi-major axis to the square of the period is the same for all planets*.

•Newton's derivation of these laws from the laws of mechanics was the first triumph of modern science.

5 In the approximation in which the orbits are circular, Kepler's laws imply the force on a planet varies inversely as the square of the distance from the Sun.

5.1 The eccentricities are small (a few percent); so this is a good approximation. The first law then states that the orbits are circular with the Sun at the center; the second that the angular velocity is a constant. This constant is $\dot{\theta} = \frac{2\pi}{T}$, where T is the period. So the acceleration of the planet is pointed towards the Sun and has magnitude $r\dot{\theta}^2 = 4\pi^2 \frac{r}{T^2}$. The third law says in this approximation that $\frac{r^3}{T^2} = K$, the same constant for all planets. Thus the acceleration of a planet at distance r is $\frac{K}{4\pi^2}\frac{1}{r^2}$. Thus the force on a planet must be proportional to its mass and inversely proportional to the square of its distance from the Sun.

•Extrapolating from this Newton arrived at the Universal Law of Gravity:

6 The gravitational force on a body due to another is pointed along the line connecting the bodies; it has magnitude proportional to the product of masses and inversely to the square of the distance.

6.1 If the positions are $\mathbf{r}_1, \mathbf{r}_2$ and masses m_1, m_2 , the forces are respectively

$$\mathbf{F}_1 = [\widehat{\mathbf{r}_2 - \mathbf{r}_1}]\frac{Gm_1m_2}{|\mathbf{r}_1 - \mathbf{r}_2|}, \quad \mathbf{F}_2 = [\widehat{\mathbf{r}_1 - \mathbf{r}_2}]\frac{Gm_1m_2}{|\mathbf{r}_2 - \mathbf{r}_1|}$$

7 Newton's second law gives the equations of motion of the two bodies:

$$m_1\frac{d^2\mathbf{r}_1}{dt^2} = \mathbf{F}_1, \quad m_2\frac{d^2\mathbf{r}_2}{dt^2} = \mathbf{F}_2.$$

8 The key to solving the equations of motion is the set of conserved quantities.

9 Since $\mathbf{F}_1 + \mathbf{F}_2 = 0$ for any isolated system (Newton's third law) the total momentum is always conserved:

$$m_1 \frac{d\mathbf{r}_1}{dt} + m_2 \frac{d\mathbf{r}_2}{dt} = \mathbf{P}, \quad \frac{d\mathbf{P}}{dt} = 0.$$

10 We can change variables from $\mathbf{r}_1, \mathbf{r}_2$ to the *center of mass* and *relative* co-ordinates: $\mathbf{R} = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{m_1+m_2}, \mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$ to get

$$\frac{d^2\mathbf{R}}{dt^2} = 0, \quad m\frac{d^2\mathbf{r}}{dt^2} = -\hat{r}\frac{Gm_1m_2}{|\mathbf{r}|^2}$$

where $m = \frac{m_1m_2}{m_1+m_2}$ is the *reduced mass*.

10.1 The first equation is trivial; the second is the same as that of a single body at position $\mathbf{r}$ moving in a central force field: we have reduced the two body problem to the one body problem.

11 A *central force field* is pointed along the radial vector and has magnitude depending only on the radial distance. Angular momentum $\mathbf{L} = m\mathbf{r} \times \frac{d\mathbf{r}}{dt}$ is conserved in any central force field.

12 Hence the vector $\mathbf{r}$ always lies in the constant plane orthogonal to $\mathbf{L}$.

13 In plane polar co-ordinates, the angular momentum is

$$L = mr^2\frac{d\theta}{dt}.$$

14 We have jsut derived Kepler's second law: since m is a constant, conservation of angular momentum implies that the areal velocity $\frac{1}{2}r^2\frac{d\theta}{dt}$ is a constant.

15 A central force field is *conservative*: it is the gradient of a scalar function: $\mathbf{F} = -\nabla U$ where U is a function only of the radial distance.

15.1 For the gravitational force, $U(r) = -\frac{Gm_1m_2}{r^2}$.

16 In any conservative force field, the total energy is conserved:

$$\frac{1}{2}m\left[\dot{r}^2 + r^2\dot{\theta}^2\right] + U(r) = E.$$

17 To determine the shape of the orbit we need r as a fucntion of θ. We can eliminate t in favor of θ in teh above equation using conservation of angular momentum:

$$\frac{d\theta}{dt} = \frac{L}{mr^2} \Rightarrow \frac{dr}{dt} = \frac{dr}{d\theta}\frac{L}{mr^2} = -\frac{L}{m}\frac{du}{d\theta}$$

where $u = \frac{1}{r}$.

18 Thus we get

$$\frac{L^2}{2m}[u'^2 + u^2] = E - U(u), \Rightarrow \theta - \theta_0 = \int_{u_0}^{u}\left[\frac{2m}{L^2}[E - U(w)] - w^2\right]^{-\frac{1}{2}} dw$$

19 Exercise: Solve this ODE and obtain the equation for the ellipse for the case of the Kepler problem.

20 The rate of change of area is a constant, $\dot{A} = \frac{1}{2}r^2\dot{\theta} = \frac{L}{2m}$; the total area of the ellipse is $A = \pi ab$. These two facts combine to give us the period of the motion, $T = \frac{A}{\dot{A}}$.

21 Exercise: Derive Kepler's third law by simplifying the expression for the period.

Chapter 3

The Action

•Newton formulated mechanics in terms of differential equations. Even at that time there was a parallel view (eg., Leibnitz) that the laws of mechanics could be formulated as a variational principle.

•For example the position of a body of in stable equilibrium is the minimum of potential energy. More generally any extremum of the potential is an equilibrium. Is there a similar principle which determines the path of a particle in motion?

•Let Q be the space of all instantaneous positions of the system. For a point particle it is the Euclidean space R^3 ; for a system of n particles it is R^{3n} ; for a pendulum of length L it is S^1 , the circle of radius L . This space is called the *configuration space.*

•The path is described by a map from the real line (the set of values of *time*) to the configuration space. This curve must be differentiable so that we can speak of velocities: so the configuration space must be a differentiable manifold and the curve must be differentiable.

•The *dimension* of the configuration space is the number of real numbers we must specify to fix the position of the particle at any instant. It is also called the *number of degrees of freedom*.

•Given the initial time, initial position and the final time and final position, we expect a unique curve to connect them that satisfies the laws of motion. This is reasonable if these laws are second order differential equations.

•Among all curves with prescribed endpoints, the curve that satisfies the laws of motion are the extrema of a quantity called the *action*.

•This action is the integral of a function of the position, velocity and time:

$$S[q] = \int_{t_1}^{t_2} L(q, \dot{q}, t) dt$$

If we vary q slightly $q \to q + \delta q$

$$\delta S = \int \left[\delta q^i \frac{\partial L}{\partial q^i} - \delta \dot{q}^i \frac{\partial L}{\partial \dot{q}^i} \right] dt$$

•The variation must vanish at the endpoints: $\delta q^i(t_1) = 0 = \delta q^i(t_1)$. This gives (after an integration by parts)

$$\frac{\partial L}{\partial q^i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} = 0.$$

These differential equations are called the *Euler-Lagrange* equations.

• *Example*: Free particle. For a free particle in R^3 , all points are the same, and all directions are the same and all instants of time are the same. These imply that the Lagrangian is independent of the Cartesian co-ordinates q^i , and of time t . It is a function only of $\dot{\mathbf{q}}^2$. Galelean invaraince implies in fact that it is proportional to tehe square of the velocity: $\frac{1}{2} m \dot{\mathbf{q}}^2$. The constant m is the *mass*.

•A basic principle of mechanics is that the motion of a system reduce to that of a free particle for short enough time intervals. We can interpret this to mean that the equations of motion are second order quasi-linear differential equations (quasilinear means linear in the velocities: all the nonlinearities are in the coordinates). The Lagrangian for a wide class of systems is of the form

$$L(q, \dot{q}, t) = \frac{1}{2} g_{ij}(q, t) \dot{q}^i \dot{q}^j + A_i(q, t) \dot{q}^i - V(q, t)$$

where g_{ij} is a positive symmetric tensor.

•The simple example of a particle moving in R^3 in the presence of a potential has

$$L(q, \dot{q}, t) = \frac{1}{2} m \dot{q}^i \dot{q}^i - V(q, t)$$

Chapter 4

Symmetry and Conservation Laws

•It is not possible to improve on the discussion in the book by Landau and Lifshitz, so I just refer you to it.

Chapter 5

Systems with one degree of Freedom

•Consider a system with one degree of freedom with Lagrangian

$$L = \frac{1}{2}m\dot{x}^2 - V(x).$$

Any conservative system with one degree of freedom can be brought to this form by a choice of co-ordinates. For example the more general Lagrangian

$$L = \frac{1}{2}g(q)\dot{q}^2 - U(q).$$

can be brought to our form by the change of variables

$$x(q) = \int_{q_0}^{q} \sqrt{g(q)}dq, \quad V(x) = U(q(x)).$$

•The equations of motion can be integrated once to get a first order ODE. The simplest way to understand this is to not that the total energy is conserved:

$$\frac{1}{2}\dot{x}^2 + V(x) = E$$

Thus we can reduce the solution of the equation to quadrature:

$$t(x) - t_0 = \sqrt{2}\int_{x_0}^{x} \frac{dx}{\sqrt{[E - V(x)]}}.$$

Inverting this function gives position as a function of time.

• *Exercise* For the simple harmonic oscillator:

$$V(x) = \frac{1}{2}\omega x^2,$$

so that

$$t(x) - t_0 = \sqrt{2}\int_{x_0}^{x} \frac{dx}{\sqrt{[E - \frac{1}{2}\omega x^2]}}.$$

Evaluate the integral by trigonometric substitutions.

•The *simple pendulum* is a ball of mass m hung from a fixed point by a rigid rod of length l . The Lagrangian is

$$L = \frac{1}{2}ml^2\dot{\theta}^2 + mgl\cos\theta.$$

Conservation of energy gives

$$\frac{1}{2}ml^2\dot{\theta}^2 - mgl\cos\theta = E.$$

The energy thus has minimum value $-mgl$ when the ball is at rest at $\theta = 0$. If $E < mgl$ the ball will oscillate back and forth; when $E > mgl$ the ball will rotate around the point of support.

•The integral above will give the answer for time as a function of angle in terms of an *elliptic integral*. There are many excellent books on elliptic integrals eg. *Higher Transcendental Functions* ed. by Erdelyi. We will look instead at the angle as a function of time which involve single functions called *elliptic functions*.

•The Jacobi elliptic functions $\mathrm{sn}\,(u,k)$, $\mathrm{cn}\,(u,k)$, $\mathrm{dn}\,(u,k)$ are single valued meromorphic functions of two complex variables u and k. They are generalizations of the trigonometric functions. They can be defined by the system of differential equations

$$\frac{d\,\mathrm{sn}\,(u,k)}{du} = \mathrm{cn}\,(u,k)\,\mathrm{dn}\,(u,k), \quad \frac{d\,\mathrm{cn}\,(u,k)}{du} = -\,\mathrm{sn}\,(u,k)\,\mathrm{dn}\,(u,k)$$

$$\frac{d\,\mathrm{dn}\,(u,k)}{du} = -k^2\,\mathrm{sn}\,(u,k)\,\mathrm{cn}\,(u,k).$$

with the boundary conditions

$$\text{sn}\,(0,k) = 0,\ \text{cn}\,(0,k) = 1,\ \text{dn}\,(0,k) = 1.$$

•It follows that (writing s for c for cn etc.)

$$s^2 + c^2 = 1, \quad k^2 s^2 + d^2 = 1, \quad d^2 - k^2 c^2 = 1 - k^2$$

for all (u,k) (they are like conservation laws for the ODE). (Prove this by explicit differentiation!) The differential equation can thus be written as

$$s'^2 = (1 - s^2)(1 - k^2 s^2).$$

•From the symmetry of the equation and boundary condition under reflection $u \to -u$, we see also that $\text{sn}\,(-u,k) = -\,\text{sn}\,(u,k),\ \text{cn}\,(-u,k) = \text{cn}\,(u,k),\ \text{dn}\,(-u,k) = \text{dn}\,(u,k)$.
•Clearly when $k = 0$ the solution is $\text{sn}\,(u,0) = \sin u$, $\text{cn}\,(u,0) = \cos u$ and $\text{dn}\,(u,0) = 1$. These are the limiting values of the Jacobi functions.
•The differential equation for the angle of the pendulum can be solved in terms of sn by some change of variables. Once we notice the similarity between the two equations, we put in the ansatz

$$u = \omega t, \sin\frac{1}{2}\theta(t) = A\ \text{sn}\,(u,k).$$

With these definitions the ODE for the pendulum reduces to

$$2ml^2\omega^2 A^2 \left[\frac{d\ \text{sn}\,(u,k)}{du}\right]^2 = [E + mgl - 2mglA^2\ \text{sn}^{\,2}(u,k)][1 - A^2\ \text{sn}^{\,2}(u,k)]$$

•Comparing with the ODE for sn we get

$$E + mgl = 2mglA^2 = 2ml^2A^2\omega^2, A = k$$

as one way to reduce our ODE to Jacobi's. (There are is another which corresponds to a symmetry of $\text{sn}\,(u,k)$ under $k \to \frac{1}{k}$.)
•We get

$$\omega = \left[\frac{g}{l}\right]^{\frac{1}{2}}, \quad k^2 = \frac{E + mgl}{2mgl}.$$

The solution is thus

$$\sin\frac{1}{2}\theta(t) = k\ \mathrm{sn}\ (\omega(t-t_0), k).$$

•In the limit $E \to -mgl$,we get $k \to 0$, so that the solution reduces to small oscillations around the point $\theta = 0$

$$\theta(t) \sim 2k\sin(\omega(t-t_0))$$

This is a familiar result . Indeed we recognize ω as the angular frequency of this simple harmonic motion.

•When $E < mgl$ we expect oscillatory motion around the equilibrium point: there isnt enough energy to go all the way around. Thus physically we expect $\mathrm{sn}\ (u,k)$ to be a periodic function. We can even get a formula for the period from the earlier formula for time as a function of position: it is just four times the time it takes to go from $\theta = 0$ to the maximum value of the angle, which is the root of $E + mgl\cos\theta = 0$. The period can be expressed in terms of the *complete elliptic integral*

$$K(k) = \int_0^{\frac{\pi}{2}} \frac{d\phi}{\sqrt{[1-k^2\sin^2\phi]}}.$$

In fact,

$$\mathrm{sn}\ (u + 4K(k), k) = \mathrm{sn}\ (u,k).$$

•We can analytically continue the function to the complex plane in the u variable. Surprsingly, sn is also periodic with an imaginary period: it is a *doubly periodic function* of u . There is a simple physical way to understand this. Replacing $t \to it$ in the equation of motion amonts to reversing the sign of the potential and of energy. Thus in imaginary time we get the same problem! Changing $E \to -E$ is the same as $k \to \sqrt{[1-k^2]}$. So we have a formula for the other period:

$$\mathrm{sn}\ (u + 4iK'(k), k) = \mathrm{sn}\ (u,k)$$

with

$$K'(k) = K(\sqrt{[1-k^2]}).$$

(Here the prime does *not* stand for the derivative.)

•Since $\mathrm{sn}\,(0,k) = 0$ we see that it has an infinite number if zeros, at the points of a lattice in the complex plane. A doubly periodic analytic function must have some singularities: otherwise it would be constant by Lioville's theorem. $\mathrm{sn}\,(u,k)$ has simple poles on a lattice as well.

• *Exercise* Show that $\mathrm{sn}\,(u,1) = \tanh u$. What kind of motion of the pendulum does this correspond to?. Show that as $E \to \infty$ the motion reduces to uniform circular motion with a large frequency.

• *Exercise* Use the differential equation to relate $\mathrm{sn}\,(u,\frac{1}{k})$ to $\mathrm{sn}\,(u,k)$. Show then that as $E \to \infty$ the solution tends to uniform circular motion.

•There is much more to the story; but you have to study the rest of this fascinating story from other books. I recommend the book *Elliptic Functions* by K. Chandrashekaran.

Chapter 6

Nonlinear Oscillations in One Dimension

22 An oscillation whose amplitude is not infinitesimal is described by a nonlinear equation.

23 The simplest case is a one dimensional system with a potential that has a unique extremum, which is a minimum (one stable equilibrium).

23.1 For example, $V(x) = \frac{1}{2}kx^2 + \lambda x^4$ with $k, \lambda > 0$ has a unique minimum at the origin.

23.2 Qualitatively such oscillations are similar to the harmonic oscillator: all orbits are periodic. The main difference is that the period might depend on the energy (which is a measure of how much the oscillation departs from the equilibrium.)

23.3 The Lagrangian

$$L = \frac{1}{2}m\dot{x}^2 - V(x)$$

leads to the equation of motion

$$m\ddot{x} = -V'(x).$$

23.4 It is often convenient to regard the equations as a pair of first order equations for position and momentum

$$\frac{dx}{dt} = \frac{p}{m}, \quad \frac{dp}{dt} = -V'(x).$$

23.5 Given the value of the position and momentum at any time, these equations determine them for all future times. There are many reliable numerical methods (for example, the Runge–Kutta method, built into Mathematica) that compute this solution with sufficient accuracy.

23.6 Thus, it is convenient to regard the solution as a curve in the plane (x, p) ; if we give initial conditions (where the system is at any given time as a point in this plane) the equations determine where it will be for all future times.

24 The motion of a conservative system is described as a curve in the *phase space*; each point in phase space is a set of values of position and momentum.

24.1 In our case of oscillations in one position variable, the phase space is two dimensional.

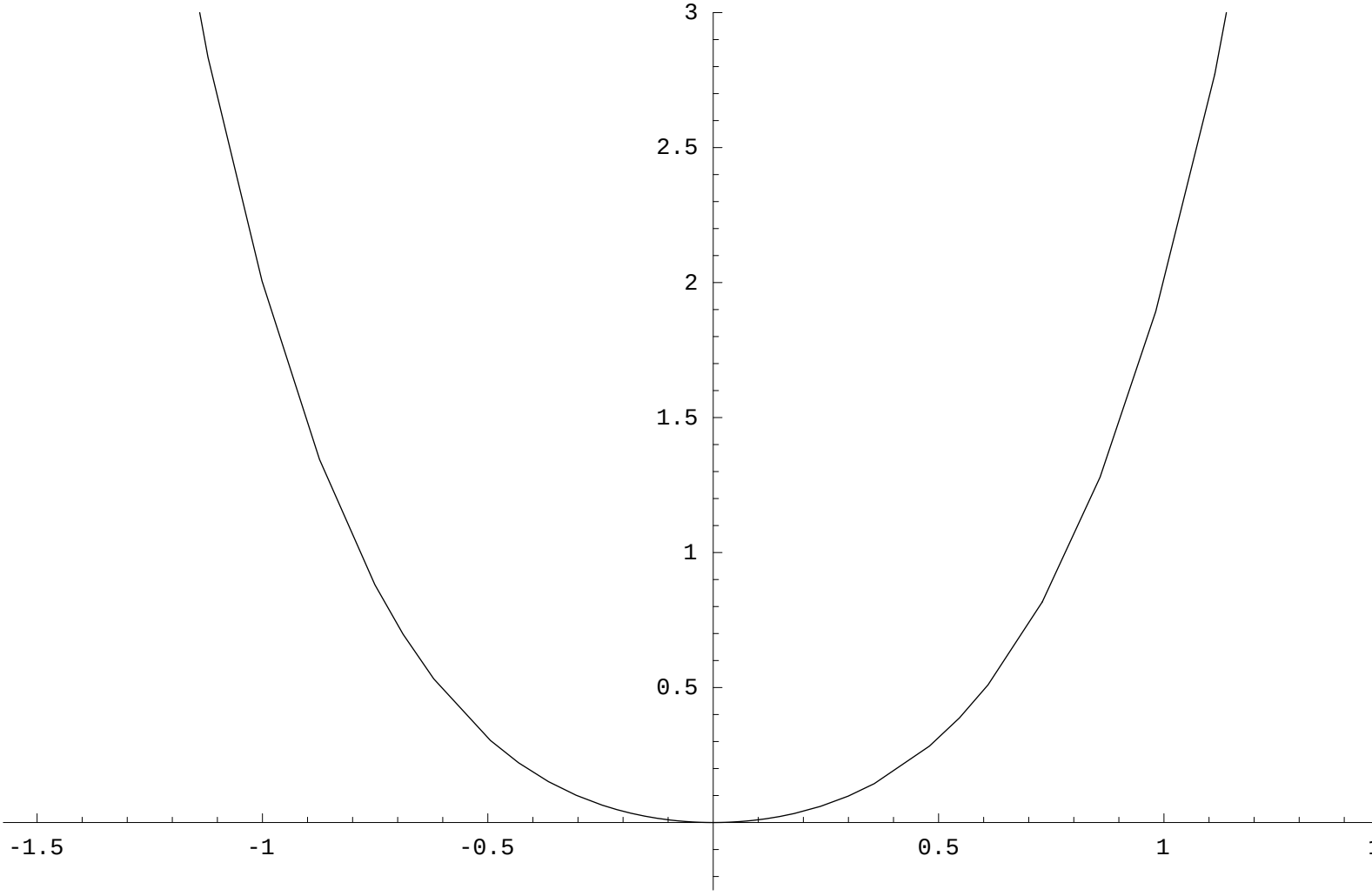
25 In the case of a system of one degree of freedom, the conservation of energy determines the shape of these curves.

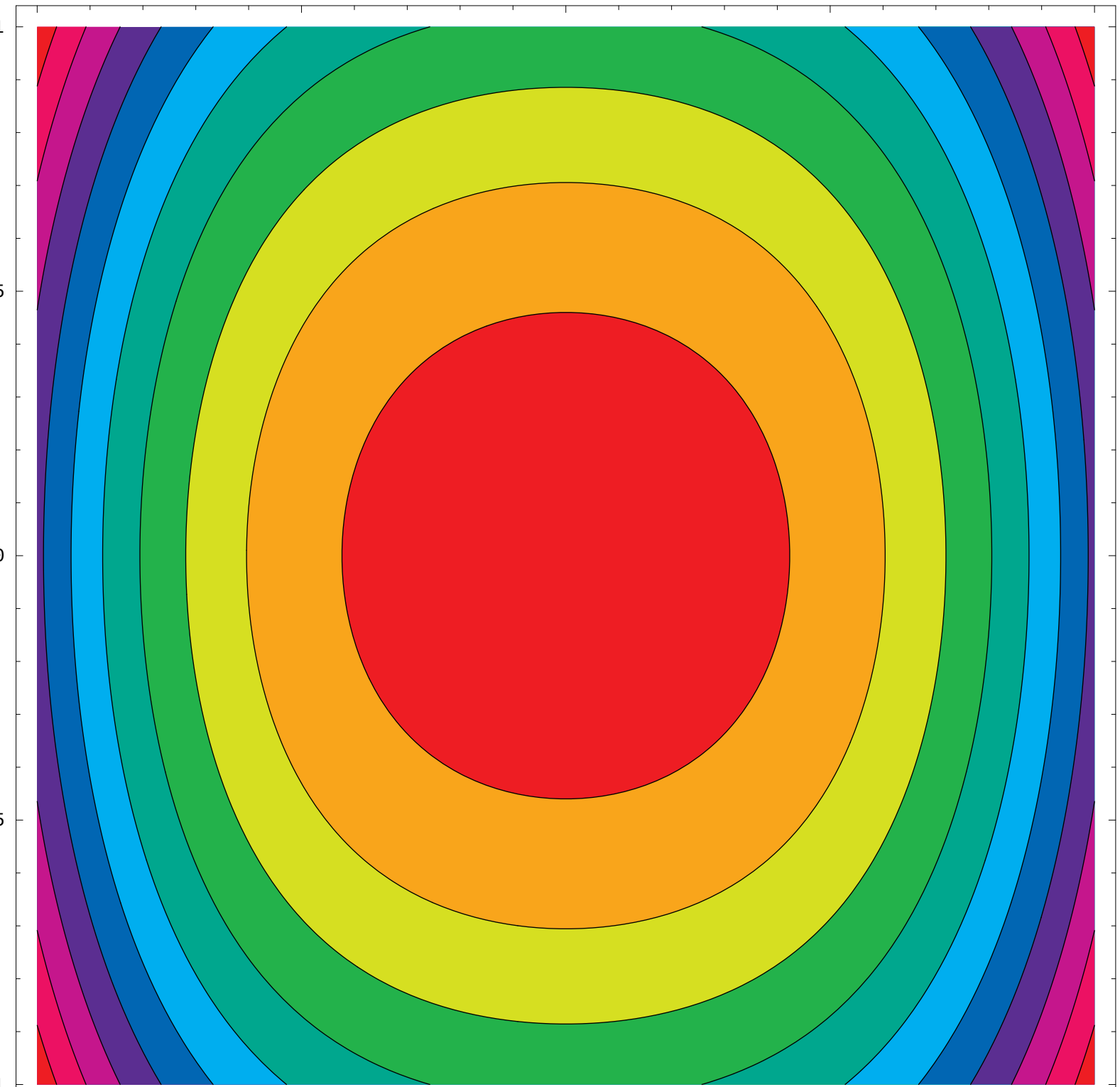
25.1 The energy expressed as a function of position and momenta is called the *hamitonian*:

$$H(x, p) = \frac{p^2}{2m} + V(x);$$

We can imagine the hamiltonian as a function on the plane (x, p) , plotted as a sort of 'hill'. The bottom of the valleys of this plot describe stable equilibrium points. The system evolves along contour plots,i.e., curves along which the energy is a constant.

25.2 If $V(x)$ has a single stable equilibrium point, these curve are ellipses for small values of energy; as energy grows they will still be closed curves, but no longer ellipses.


3
2.5
2
1.5
1
0.5
-1.5
-1
-0.5
0.5
1



25.3 The period of the orbit is

$$T(E) = 2\sqrt{\left[\frac{2}{m}\right]} \int_{x_1(E)}^{x_2(E)} \frac{dx}{\sqrt{[E - V(x)]}}$$

where $x_1(E), x_2(E)$ are solutions of the equations $V(x) = E$. (They are called *turning points*). In this case there will be exactly two solutions, as long as the energy is greater than the minimum value. The system oscillates between these two points.

25.4 As we approach the minimum value of energy the turning points approach merge with each other and we get a static solution.

26 The area enclosed by the curve of constant energy $W(E)$, has many interesting properties; for example the period of the orbit is $T(E) = \frac{dW(E)}{dE}$.

26.1 Prove this!.

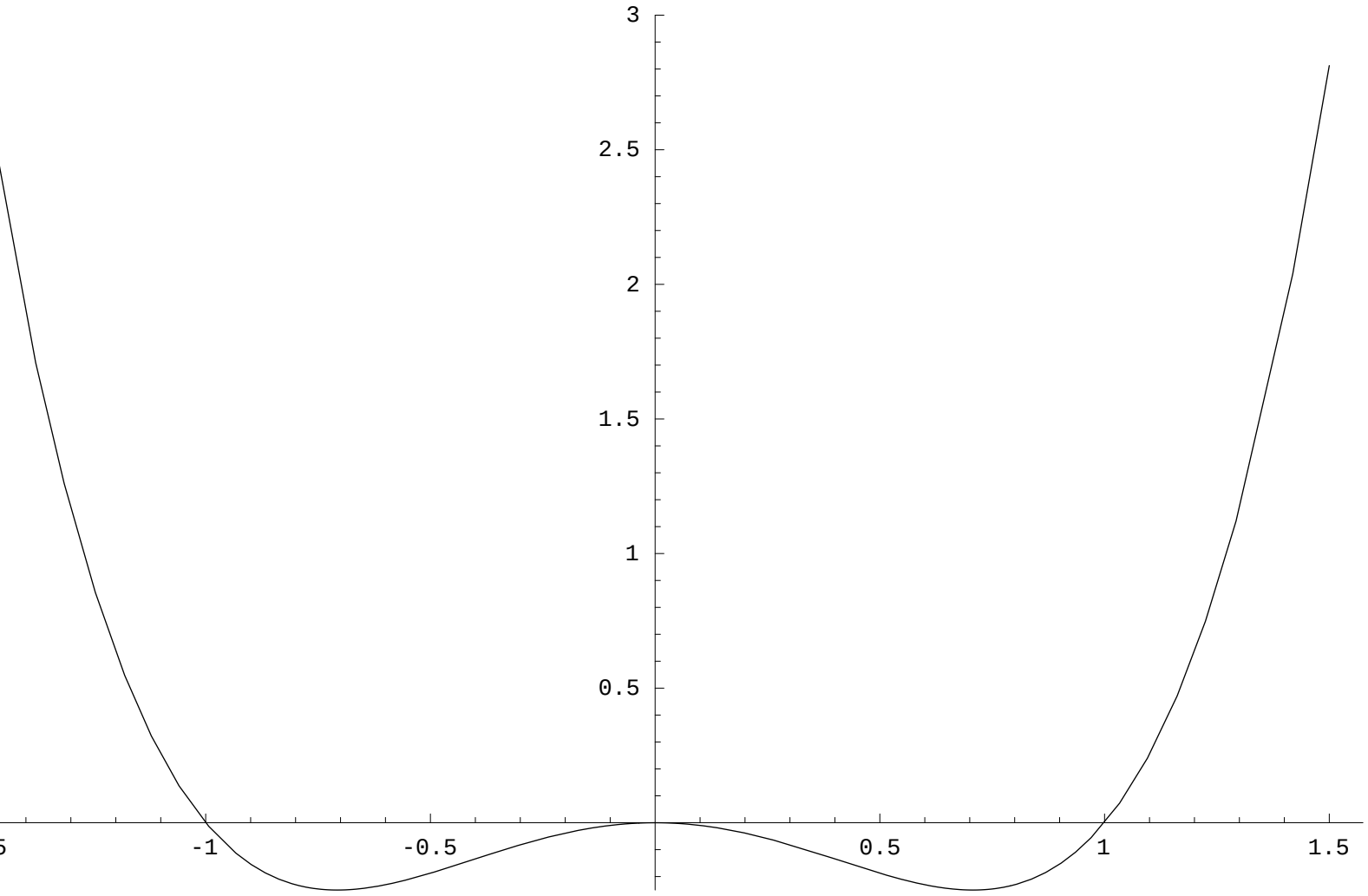
26.2 Thus the area of an orbit is a monotonically increasing function of the energy.

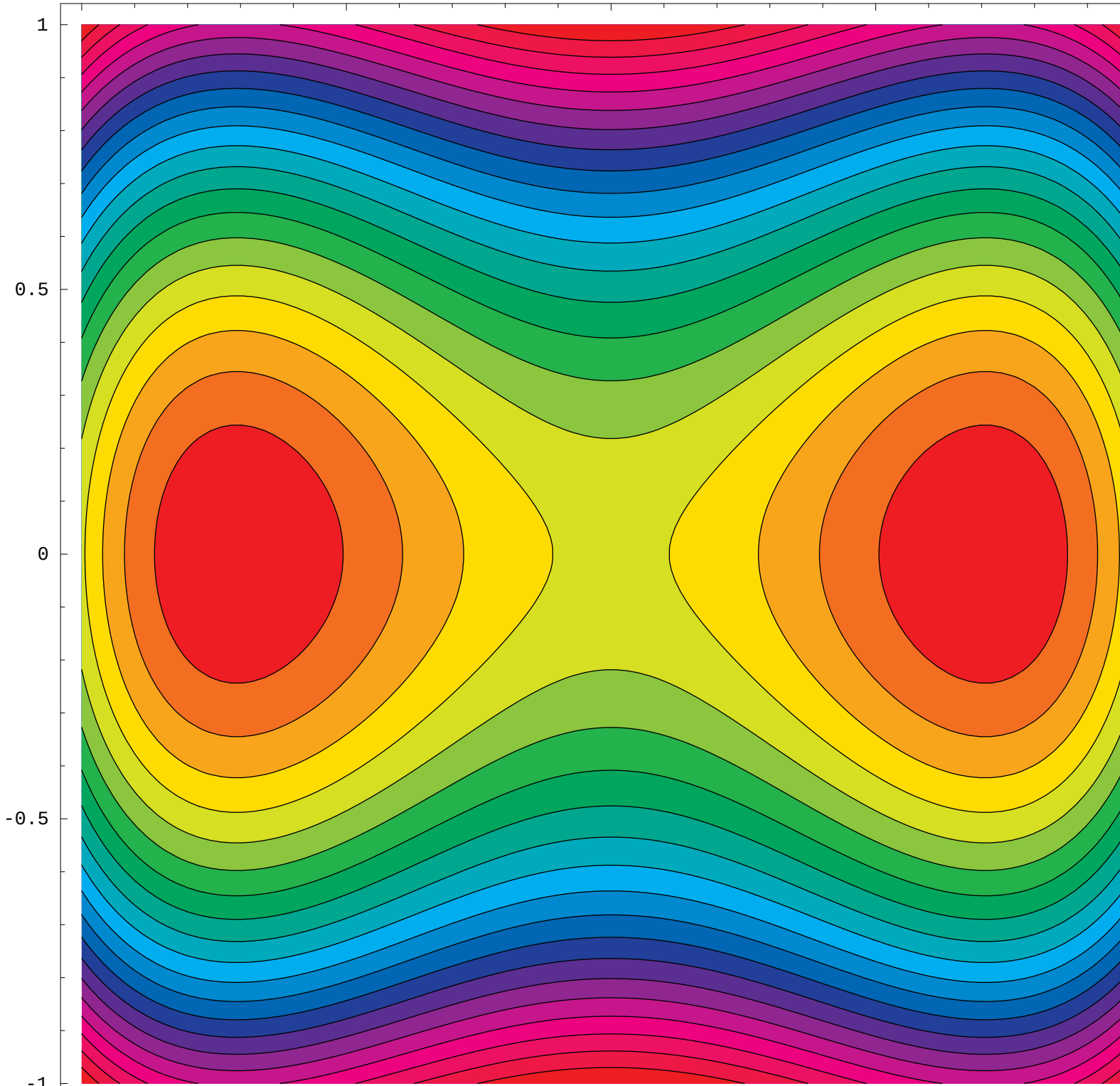
27 In the semiclassical approximation of quantum mechanics, the energy levels are given by the *Bohr-Sommerfeld* condition:

$$W(E) = n\hbar, n = 0, 1, 2, \cdots$$

27.1 The number of energy levels of energy less than E is the area enclosed by the curve $H(x, p) = E$ in units of Plank's constant.

28 Thus the density of energy levels at energy E is equal to the classical period divided by Plank's constant: $\frac{T(E)}{\hbar}$.


3
2.5
2
1.5
1
0.5
-1
-0.5
0.5
1
1.5


1
0.5
0
-0.5
-1

29 Now consider the case where $V(x)$ has two stable equilibrium points.

29.1 If there are two stable equilibrium points, there must be one unstable equilibrium in between them. (Prove this!).

29.2 Let us (for simplicity) assume that the potential is an even function: $V(x) = V(-x)$; so that the two stable points are symmetrically placed around the origin, and the origin itself is an unstable equilibrium point.

29.3 Exercise: Generalize the analysis below to the case where $V(x)$ is not symmetric.

29.4 With two stable equilibria, the orbits for small energy will be approximately ellipses around each stable equilibrium point: the set of points with a given value of energy is *disconnected*. For large energy they will be curves that surround both stable equilibria: it is a connected curve.

29.5 Let E_{min} be the minimum of the potential, and E_{cr} its value at the unstable equilibrium point . For $E_{min} < E < E_{cr}$, there are four real solutions to the equation $V(x) = E$, $x_1(E) < x_2(E) < 0 < x_3(E) < x_4(E)$. The system oscillates between either $x_1(E)$ and $x_2(E)$ or between $x_3(E)$ and $x_4(E)$; it never reaches the unstable equilibrium point at the origin.

29.6 As $E \to E_{cr}$, the two middle turning points $x_3(E)$ and $x_4(E)$ approach each other and eventually dissappear (i.e., become complex); for $E > E_{cr}$, there are only two real solutions $x_1(E), x_4(E)$ for the equation $V(x) = E$. The system oscillates between these two points, which means it passes through all three equilibrium points.

29.7 Thus at the *critical* value of energy E_{cr} the nature of the orbits change drastically. If the system starts out at this point with an infinitesimally small but small postive value of energy it will fall into the positive minimum and return to the original point after a long time; with an infinitesimally small but negative momentum it will fall into the other stable equilibrium point and return. The period of this motion is infinite (we will see this soon).

29.8 Thus we discover the significance of unstable equilibria: they describe orbits that are the boundary between those that circle one of the equilibrium points and those that surround both.These are called *critical orbits*.

29.9 The critical orbit has infinite period; more precisely, the period $T(E)$ diverges as $E \to E_{cr}$. The nature of this divergence is independent of the details of the potential. As long as the unstable equilibrium point is non-degenerate, (this means that V'' is non-zero there) $T(E) \sim \log |E - E_{cr}|$ as $E \to E_{cr}$.

30 The area of orbits is a subtle quantity when $E_{min} < E < E_{cr}$.

30.1 For, the curve encloses two separate (disconnected) regions; the derivative of the are of each connected component is the period of that piece.

30.2 This function has a singularity at the critical values of energy: two disconnected areas merge into one so that suddenly the area doubles. The area of each piece will have a jump discontinuity, but the sum of the two areas will change continuously.

30.3 But the derivative of even the sum of the areas will diverge logarithmically, as it is proportionalo the period of the critical orbit.

30.4 Thus the density of energy levels has spikes at critical values of the energy . This fact has interesting consequences in the theory of metals, as the conductivity of metals is proportional to the density of states at the Fermi energy.

31 There are even further subtleties in quantum mechanics of double-well potentials: the system can *tunnel* from one of the minima to the other.

Chapter 7

Scattering In One dimension

•Let $V(x) : R \to R$ be a potential on the real line such that $\lim_{|x|\to\infty} V(x) = 0$. A particle moves from $x = -\infty$ with an initial energy E and it will either be reflected back to $x = -\infty$ (if it doesnt have enough energy to overcome some 'hilly' part of the potential) or be transmitted to $x = \infty$. Let us consider for simplicity the case $V(x) \leq 0$ so that all particles are transmitted. Let us set the mass equal to 2 to avoid unpleasant constants.
•Conservation of energy gives

$$\dot{x}^2 + V(x) = E$$

From this we can calculate the time of transit from $x = -L$ to $x = L$:

$$T(E, L) = \int_{-L}^{L} \frac{dx}{\sqrt{[E - V(x)]}}$$

We can compare this to the time it would have taken in the absence of a potential:

$$T_0(E, L) = \int_{-L}^{L} \frac{dx}{\sqrt{E}}.$$

Each of these integrals diverge as $L \to \infty$ but there difference will converge for potentials of finite range: the particle will take a little bit less time to transit in the presence of a potential, as it is moving faster (when $V(x) < 0$). This time difference is

$$T(E) = \int_{-\infty}^{\infty} dx \left\{ [E - V(x)]^{-\frac{1}{2}} - E^{-\frac{1}{2}} \right\}.$$

This quantity is the classical analogue of the scattering phase shift in quantum theory. (Actually an even closer analogue is the difference in the action between the two paths, which is exactly the classical limit of the scattering phase shift).

•Often in physics we are interested in inverse problems. Is it possible to reconstruct the potential $V(x)$ knowing the time delay $T(E)$ of scattering for all energies? That is, can we solve the above integral equation for $V(x)$ given $T(E)$?.

•This is a nonlinear integral equation which is hard to solve. But a little trick will reduce it to a linear integral equation: we try to find the inverse function of $V(x)$. That is we look for $x(V)$. But this is a multiple-valued function: since $V(x)$ goes from 0 at $x = -\infty$ to a negative value to 0 again at $x = \infty$, its inverse is at least double valued. We can avoid this problem by restricting ourselves to potentials that are symmetric $V(-x) = V(x)$ and monotonic in the range $0 < x < \infty$. Then, $x(V)$ will be single valued function with range $0 < x(V) < \infty$ when V is in the range $V_0 < V < 0$. (Here, $V_0 = V(0)$).

$$T(E) = 2\int_{V_0}^{0} \frac{dx(V)}{dV}\left\{[E-V]^{-\frac{1}{2}} - E^{-\frac{1}{2}}\right\} dV.$$

which is a linear integral equation for $\frac{dx(V)}{dV}$. This can be solved using the theory of *fractional integrals*. Thus in this case the potential can be reconstructed from the scattering data.

• *Aside* The *fractional integral of order* μ of a function $f(x)$ is

$$R_\mu f(y) = \frac{1}{\Gamma(\mu)}\int_0^y f(x)(y-x)^{\mu-1}dx.$$

The path of integration in the complex x plane is the straight line connecting the origin to y . It has the following properties (after appropriate analytic continuations):

$$R_\mu R_\nu f = R_{\mu+\nu} f, \quad R_1 f(y) = \int_0^y f(x)dx$$

$$R_\mu \frac{d^n f}{dx^n} = R_{\mu-n} f.$$

This means that R_μ f can be thought of as an analytic continuation of the n th derivative of f to a negative (or even complex) value $-\alpha$.

•In particular,

$$g(y) = \frac{1}{\Gamma(\mu)} \int_0^y f(x)(y-x)^{\mu-1} dx \Rightarrow f(y) = \frac{1}{\Gamma(-\mu)} \int_0^y g(x)(y-x)^{-\mu-1} dx.$$

This is the idea behind solution of integral equations such as the ones above.

Chapter 8

Small Oscillations with Many Degrees of Freedom

See Landau and Lifshitz Chapter 5.

•One of the simplest mechanical systems is the *simple harmonic oscillator* with Lagrangian

$$L = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2.$$

•The equations of motion $m\ddot{x} + kx = 0$ have the well known solutions in terms of trigonometric functions:

$$x(t) = a\cos[\omega_0(t - t_0)]$$

where the *angular frequency* $\omega_0 = \sqrt{(k/m)}$. The constants a and t_0 are constants of integration. They have simple physical meanins: A is the maximum displacement and t_0 is a time at which $x(t)$ has this maximum value. The energy is conserved and has the value $\frac{1}{2}ma^2$.

•Although the solutions $x(t)$ are real for physical reasons, it turns out to be convenient to think of them as the real parts of a complex solution:

$$x(t) = \text{Re } Ae^{i\omega_0 t}.$$

Then $a = |A|$ and $t_0 = -\frac{\arg A}{\omega_0}$.

• *Example* Imagine a particle of mass m moving on a line atttched to two fixed points a, b on either side, by springs of strengths k_1 and k_2 . What is the angular frequency?

$$L = \frac{1}{2}\dot{x}^2 - \frac{1}{2}k_1(x - a)^2 - \frac{1}{2}k_2(x - b)^2$$

which is

$$L = \frac{1}{2}\dot{x}^2 - \frac{1}{2}[k_1 + k_2]\left[x - \left(\frac{k_1 a + k_2 b}{k_1 + k_2}\right)\right]^2 + \text{constant}.$$

Thus we get simple harmonic oscillation around the point $\frac{k_1 a + k_2 b}{k_1 + k_2}$ with angular frequency $\left[\frac{k_1+k_2}{m}\right]^{\frac{1}{2}}$.

•Now consider an external force $F(t)$ acting on a simple harmonic oscillator:

$$m\ddot{x} + kx = F(t).$$

This equation can be solved by the Fourier transform:

$$x(t) = \int \tilde{x}(\omega) e^{i\omega t}\frac{d\omega}{2\pi}, \quad F(t) = \int \tilde{F}(\omega) e^{i\omega t}\frac{d\omega}{2\pi}$$

to get

$$m[\omega_0^2 - \omega^2]\tilde{x}(\omega) = \tilde{F}(\omega).$$

Thus a solution is

$$x(t) = \frac{1}{m}\int \frac{\tilde{F}(\omega)}{\omega_0^2 - \omega^2}\frac{d\omega}{2\pi}.$$

But this integral is singular at two points along the real line: we need to somehow modify the contour of integration to make the integral meaningful. The choice of this rule will depend on two constants, which are determined by the boundary conditions on the equation of motion.

•Let us now consider a more general mechanical system with Lagrangian

$$L = \frac{1}{2}\sum_{ij} g_{ij}(q)\dot{q}^i\dot{q}^j - V(q).$$

Practically any mechanical system (except those with velocity dependent forces) is of this form. The equation of motion is

$$\frac{d}{dt}\sum_j g_{ij}(q)\dot{q}^j + \frac{\partial V}{\partial q^i} = 0.$$

Any point at which $\frac{\partial V}{\partial q^i} = 0$ is an *equilibrium point*: there is a constant solution to the equation of motion at that point.

•We will now study the system in the infinitesimal neighborhood of an equilibrium point. We can choose choose co-ordinates so that this point is at the origin.

$$L = \frac{1}{2} g_{ij} \dot{q}^i \dot{q}^j - \frac{1}{2} K_{ij} q^i q^j - V(0) + O(q^3).$$

Here $K_{ij} = \frac{\partial^2 V}{\partial q^i \partial q^j}(0)$ is the matrix of second derivatives of the potential: also called the *Hessian*.

•The first thing to note is that *stable equilibrium points correspond to minima of the potential*; i.e., K must be positive. An equilibrium point is stable if the linearized equations of motion have bounded solutions for any initial condition. Now we get the equation of motion (using matrix notation)

$$\sum_j g_{ij} \frac{d^2 q^j}{dt^2} + \sum_j K_{ij} q^j = 0.$$

We can look for solutions of the form $q^j = A^j e^{i\omega t}$. The solution is bounded if the frequency is real. Thus for a stable equilibrium point all the frequencies must be real. We get the eigenvalue problem:

$$\sum_j [-\omega^2 g_{ij} + K_{ij}] A^j = 0.$$

•No we already know that g_{ij} is a positive: otherwise kinetic energy would not be a positive function of the velocities. Thus we can regard it as defining an inner product on the n -dimensional real vector space. Thus we can choose a co-ordinate system in which $g_{ij} = 0$ for $i \neq j$ and $g_{ij} = 1$ for $i = j$. There is a procedure (Gramm-Schmidt orthogonalization) that even constructs this orthonormal co-ordinate system explicitly. Then we get a more familiar eigenvalue problem

$$KA = \omega^2 A.$$

We still have some freedom in the co-ordinate system: transformations that are orthogonal with respect to the inner product g will map orthonormal systems to each other. A symmetric matrix such as K can be diagonalized

using such a transformation and the diagonal entries are then the eigenvalues. Now in order that ω be real, the eignvalues of K should be positive: it must be a positive matrix.

•The the eigenvectors of the pair (g, K) are the *normal modes* of the system: they form a co-ordinate system in which each axis has a definite frequency of vibration.

•If the normal frequencies ω_j are all integer multiples of some underlying real number ω all solutions will be periodic. In general, the motion is *quasi-periodic*: each normal mode is periodic with a different period.

•Let us complete the story by studying velocity dependent forces such as magnetic fields. For small oscillations around an extremum of the potential, the Lagrangian has the form

$$L = \frac{1}{2} g_{ij} \dot{q}_i \dot{q}_j + \frac{1}{2} B_{ij} q_i \dot{q}_j - \frac{1}{2} K_{ij} q_i q_j.$$

•The surprise here is that we can have a stable oscillation around an extremum of the potential *even if it not a minimum*. We will derive a more general criterion for the stability of an extremum of the potential.

•The equations of motion are

$$\sum_j g_{ij} \ddot{q} + \sum_j B_{ij} q_j + \sum_j K_{ij} q_j = 0$$

or in matrix notation

$$g\ddot{q} + B\dot{q} + Kq = 0.$$

Here B is antisymmetric, while g and K are symmetric; moreover g is positive.

•Now, we can always choose a basis in which $g = 1$, so that

$$\ddot{q} + B\dot{q} + Kq = 0.$$

Let us now choose an ansatz

$$q(t) = e^{i\omega t} A$$

so that

$$\left[-\omega^2 + iB\omega + K \right] A = 0.$$

•The simplest example is with two degrees of freedom: $B = \begin{pmatrix} 0 & -b \\ b & 0 \end{pmatrix}$ so that

$$\begin{pmatrix} K_{11} - \omega^2 & K_{12} - ib\omega \\ K_{12} + ib\omega & K_{22} - \omega^2 \end{pmatrix} A = 0$$

which has the charateristic equation

$$\omega^4 - [b^2 + \ \mathrm{tr} K]\omega^2 + \det K = 0.$$

•The condition for ω^2 to be real is

$$[b^2 + \ \mathrm{tr} K]^2 - 4 \det K > 0.$$

If the solutions are to be positive, their sum and product must be positive:

$$b^2 + \ \mathrm{tr} K > 0, \det K > 0.$$

•In summary the stability criterion is,

$$\det K > 0, \quad b^2 + \ \mathrm{tr} K > 2\sqrt{\det K}.$$

The eigenvalues are then given by

$$\omega^2 = \frac{1}{2}\left[\ \mathrm{tr} K + b^2 \pm \sqrt{\left([\ \mathrm{tr} K + b^2]^2 - 4 \det K \right)} \right]$$

•Notice that the extremum must be either a minimum or a maximum: a saddle point is not stabilized by this mechanism for two degrees of freedom.
•These formulae establish the stability of the Lagrange points $L4$ and L_5 .
•This is the secret of the stability of the Penning trap. An electrostatic field cannot confine charged particles to a region of space, since it cannot have a minimum. Since $\nabla^2 V = 0$ in a vacuum, at least one of the eigenvalues of $\partial^2 V$ will be negative. But a magnetic field can stabilize system. For example, if the magnetic field is along the third axis,

$$B = \begin{pmatrix} 0 & b & 0 \\ -b & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, B^2 = -b^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

An electrostatic potential which has a minimum along the direction of the magnetic field can now provide a stable equilibrium point for the charged particles, provided that the magnitude of the negative eigenvalues are not too large. For example,

$$K = \begin{pmatrix} k_1 & 0 & 0 \\ 0 & k_2 & 0 \\ 0 & 0 & k_3 \end{pmatrix}$$

is stable if

$$k_3 > 0, \quad k_1 k_2 > 0, \quad b^2 + k_1 + k_2 > 2\sqrt{[k_1 k_2]}.$$

Thus in this particular case, the potential must be a minimum along the magnetic axis and a a maximum in the plane perpendicular to the magnetic field; also, the magnetic field must be large enough.

•The general analysis of stability with several degrees of freedom appears to be complicated.

Chapter 9

The Restricted Three Body Problem

32 We follow the discussion in *Methods of Celestial Mechanics* by D.Brouwer and G. M. Clemence, Academic Press, NY (1961).

32.1 Consider a body of unit mass (e.g., an asteroid) moving in the field of a heavy body of mass M (e.g., the Sun) and a much lighter body (e.g., Jupiter) of mass m . Also assume that the heavy bodies are in circular orbit around each other with angular frequency Ω ; the effect of the light body on them is negligible. All three bodies move in the same plane. This is the *restricted 3 -body problem*.

32.2 Ignoring the effect of the asteroid, we can solve the two body problem to get

$$\frac{Mm}{M+m}R\Omega^2 = \frac{GMm}{R^2}, \Rightarrow \Omega^2 = \frac{G(M+m)}{R^3}.$$

32.3 Choose the center of mass of the heavy bodies as the origin of a polar co-ordinate system. Then the position of the Sun is at $(\nu R, \pi - \Omega t)$ and the Moon is at $((1-\nu)R, \Omega t)$, where R is the distance between them, and $\nu = \frac{m}{M+m}$. The distance from the asteroid to the Sun is

$$\rho_1(t) = \sqrt{\left[r^2 + \nu^2 R^2 + 2\nu r R\cos[\theta - \Omega t]\right]}$$

and to Juptiter is

$$\rho_2(t) = \sqrt{\left[r^2 + (1-\nu)^2 R^2 - 2(1-\nu) r R \cos[\theta - \Omega t]\right]}\,.$$

The lagrangian for the motion of the asteroid is

$$L = \frac{1}{2}\dot{r}^2 + \frac{1}{2} r^2 \dot{\theta}^2 + G(M+m)\left[\frac{1-\nu}{\rho_1(t)} + \frac{\nu}{\rho_2(t)}\right].$$

32.4 In this co-ordinate system the Lagrangian has an explicit time dependence: the energy is not conserved. Change variables to $\chi = \theta - \Omega t$ to get

$$L = \frac{1}{2}\dot{r}^2 + \frac{1}{2} r^2 [\dot{\chi} + \Omega]^2 + G[M+m]\left[\frac{1-\nu}{r_1} + \frac{\nu}{r_2}\right]$$

where

$$r_1 = \sqrt{\left[r^2 + \nu^2 R^2 + 2\nu r R \cos\chi\right]}\,, \quad r_2 = \sqrt{\left[r^2 + (1-\nu)^2 R^2 - 2(1-\nu) r R \cos\chi\right]}$$

are now independent of time.

32.5 Now the hamiltonian in the rotating frame,

$$H = \dot{r}\frac{\partial L}{\partial \dot{r}} + \dot{\chi}\frac{\partial L}{\partial \dot{\chi}} - L = \frac{1}{2}\dot{r}^2 + \frac{1}{2} r^2 \dot{\chi}^2 - G[M+m]\left[\frac{r^2}{2R^3} + \frac{1-\nu}{r_1} + \frac{\nu}{r_2}\right]$$

is a constant of the motion. This is called the 'Jacobi integral' in classical literature.

32.6 The hamiltonian is of the form $H = T + V$ where T is the kinetic energy and V is an *effective potential energy*:

$$V(r,\chi) = -G[M+m]\left[\frac{r^2}{2R^3} + \frac{1-\nu}{r_1} + \frac{\nu}{r_2}\right]$$

It consists of the gravitational potential energy plus a term due to the centrifugal barrier, since we are in a rotating co-ordinate system.

32.7 The effective potential $V(r,\chi)$ is conveniently expressed in terms of the distances to the massive bodies,

$$V(r_1,r_2) = -G\left[M\left\{\frac{r_1^2}{2R^3}+\frac{1}{r_1}\right\}+m\left\{\frac{r_2^2}{2R^3}+\frac{1}{r_2}\right\}\right]$$

using the identity

$$\frac{1}{\nu}r_1^2+\frac{1}{1-\nu}r_2^2 = \frac{1}{\nu(1-\nu)}r^2+R^2.$$

(We have removed an irrelevant constant from the potential.)

32.8 Sometimes it is convenient to use cartesian co-ordinates, in which the hamiltonian and lagrangian are,

$$H=\frac{1}{2}\dot{x}^2+\frac{1}{2}\dot{y}^2+V(x,y),\quad L=\frac{1}{2}\dot{x}^2+\frac{1}{2}\dot{y}^2+\Omega\left[x\dot{y}-y\dot{x}\right]-V(x,y).$$

32.9 It is obvious from the above formula for the potential as a function of r_1 and r_2 that $r_1=r_2=R$ is an extremum of the potential.There are two ways this can happen: the asteroid can form an equilateral triangle with the Sun and Jupiter on either side of the line joining them. These are the Lagrange points L_4 and L_5 . These are actually maxima of the potential. In spite of this fact, they correspond to stable equilibrium points because of the effect of the velocity dependent forces.

32.10 Let us find the frequencies for small oscillations around these points. Explicit calculation of the second derivatives w.r.t. x,y give

$$\partial^2 V := K\Omega^2 = -\begin{pmatrix} \frac{3}{4} & \pm\frac{3\sqrt{3}}{4}(1-2\nu) \\ \pm\frac{3\sqrt{3}}{4}(1-2\nu) & \frac{9}{4}\end{pmatrix}\Omega^2,\quad B=2\Omega\begin{pmatrix}0&-1\\1&0\end{pmatrix}.$$

for L_4 and L_5 respectively. Thus $\mathrm{tr}K=-3, \det K=\frac{27}{16}\nu(1-\nu)$. The condition for stability is

$$\det K>0,\quad 4+\ \mathrm{tr}K>2\sqrt{\det K}\Rightarrow \nu(1-\nu)>0,\quad 27\nu(1-\nu)<1.$$

Since $0<\nu<1$ from its defintion, the first condition is trivially satisfied. Thus the stability condition reduces to

$$\nu<\frac{1}{2}\left[1-\sqrt{\left(1-\frac{4}{27}\right)}\right]\sim 0.03852.$$

The frequencies are given by

$$\omega^2 = \frac{1}{2} \pm \frac{1}{2}\left[1 - 27\nu(1-\nu)\right]^{\frac{1}{2}} .$$

32.11 The libration periods as multiples of the period T are

$$T_{1,2} = T\frac{1}{\sqrt{2}}\left[1 \pm \sqrt{\{1 - 27\nu(1-\nu)\}}\right]^{-\frac{1}{2}} .$$

32.12 For the Sun-Jupiter system $\nu = 0.00095388$, orbital period= 11.86 years so that the periods are 147.54 years or 11.90 years. For the Earth-Moon system $\nu = \frac{1}{81}$, orbital period= 27.32 days so that the Libration periods are 90.8 days and 28.6 days.

32.13 Lagrange thought that these special solutions were artificial and that they would never be realized in nature. But we now know that there are asteroids (*Trojan asteroids*) that form an equilateral triangle with Sun and Jupiter.

32.14 Lagrange discovered something even more astonishing: the equilateral triangle is an exact solution for the *full three body problem*, not assuming one of the bodies to be infinitesimally small.

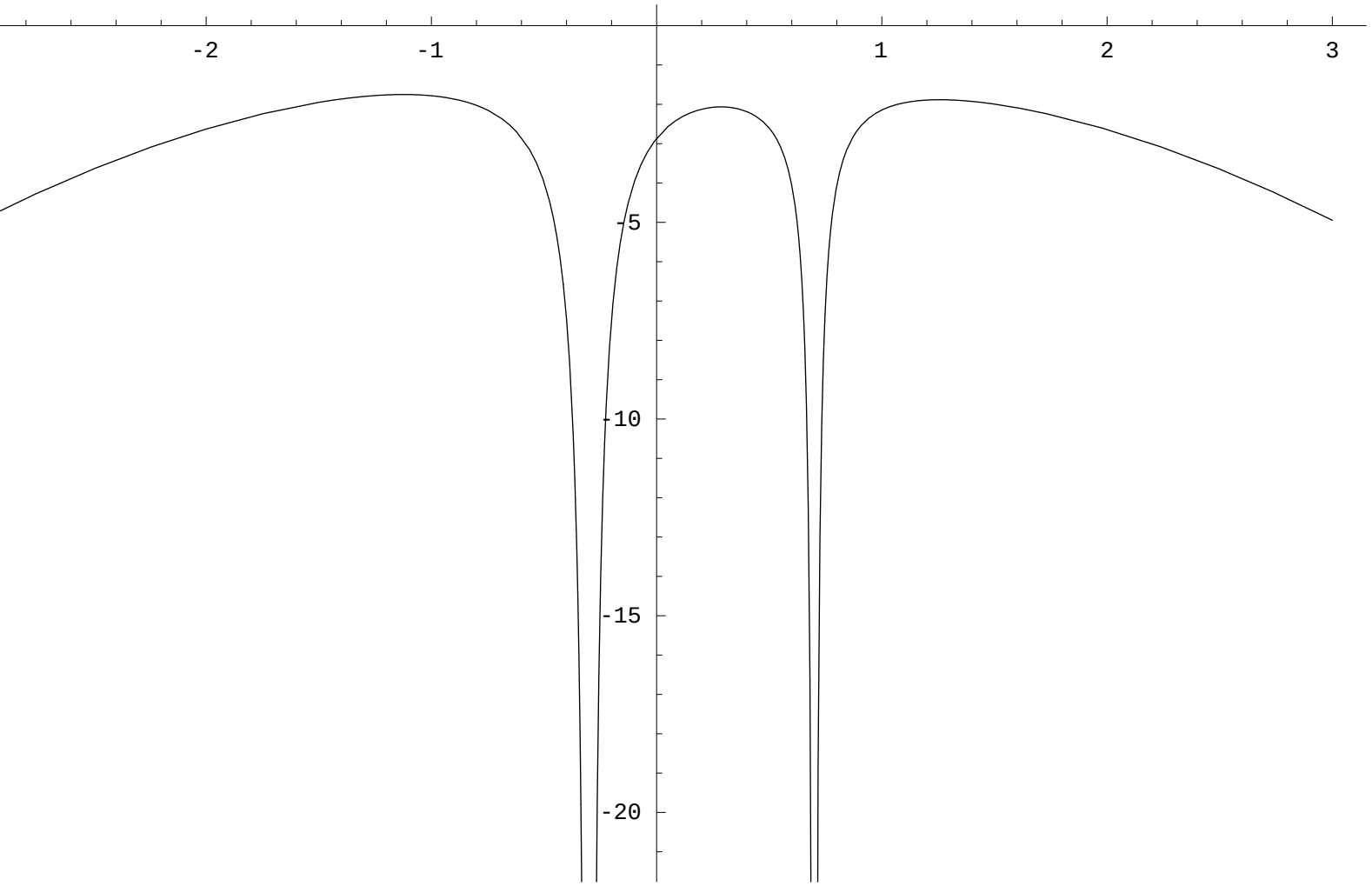
32.15 If the change of variables from x, y to r_1, r_2 were invertible everywhere, there would have been no other extrema. However, if all three bodies lie on the same line, the variables r_1 and r_2 are not independent: $r_2 = R - r_1, r_2 = -R + r_1$ or $r_2 = r_1 + R$ respectively if the asteroid, the Jupiter or Sun is in the middle on this line. Then there is the possibility of additional extrema of $V(x, y)$.

32.16 Since r_1, r_2 are even functions of y , their derivatives w.r.t. y are odd; so the derivatives vanish if $y = 0$. Thus the Jacobian determinant for change of variables from x, y to r_1, r_2 vanishes along the line connecting the two massive bodies.

32.17 Since $V(x,y)$ depends on y only through r_1, r_2 , its derivative w.r.t. y also vanishes on the line $y=0$; also $\frac{\partial^2 V}{\partial x \partial y} = 0$ on this line. In fact $\frac{\partial^2 V}{\partial y^2} > 0$ since the potential should increase as we go farther from the massive bodies.

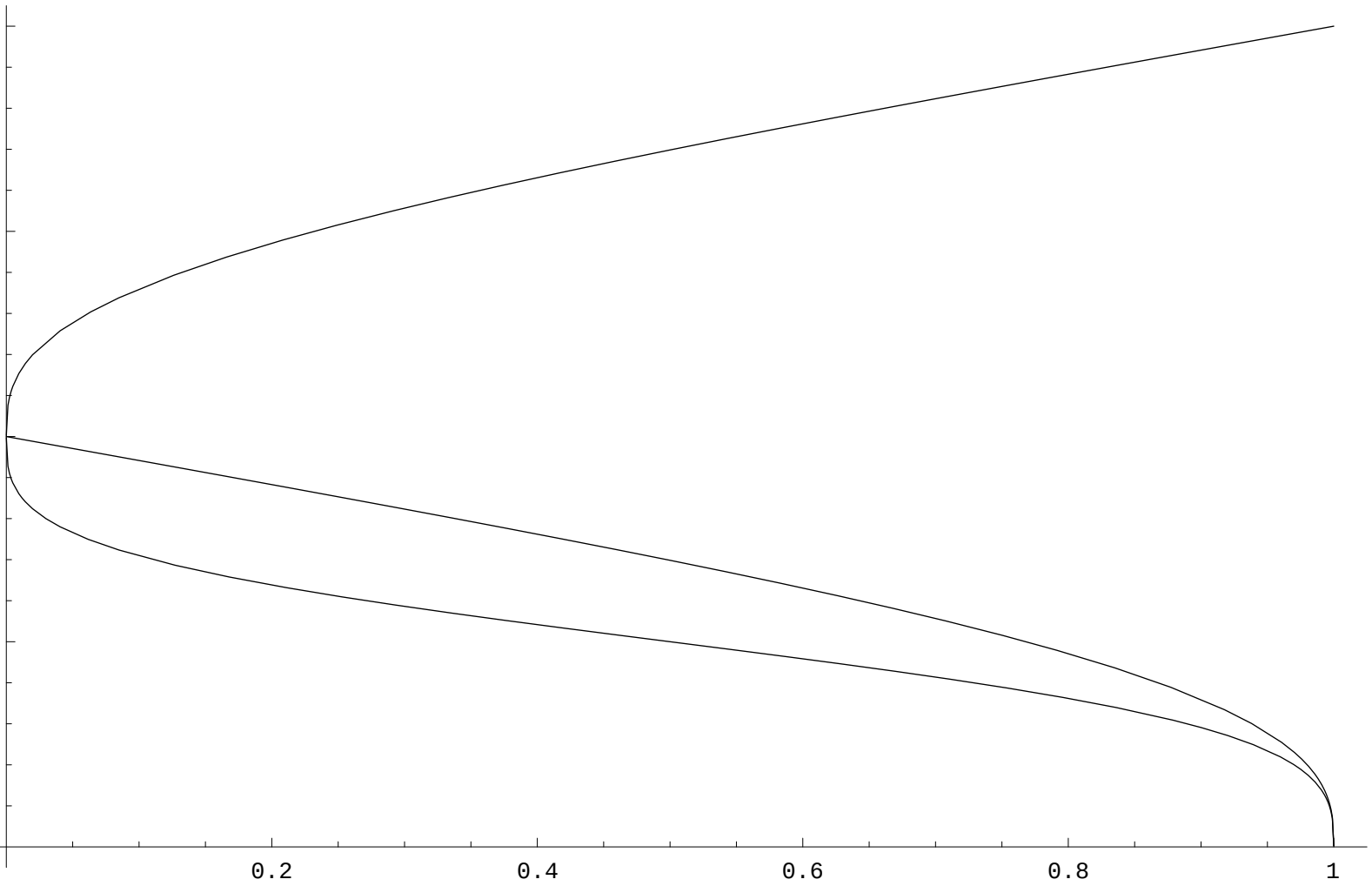
32.18 Thus, it is enough to look for for an extremum of the effective potential as a function of x . In addition to the minima at the position of the Sun and Jupiter, we will have have three maxima: one in between the two primaries L_1 and close to Jupiter, another close to Jupiter but on the opposite side from the Sun L_2 , a third one opposite from the Jupiter (L_3). These three points are saddle points of the potentials, since $\frac{\partial^2 V}{\partial x^2} < 0, \frac{\partial^2 V}{\partial x \partial y} = 0, \frac{\partial^2 V}{\partial y^2} > 0$

32.19 We plot here the potential along the line connecting the primaries, in units where $M + m = 1, R = 1$ and for the case $\nu = 0.3$.


-2
-1
1
2
3
-5
-10
-15
-20

32.20 The same method as before shows that these three solutions are unstable: they correspond to saddle points of the potential. As a function of x they are maxima and as a function of y , L_1, L_2, L_3 are minima.

32.21 Here is a plot of the r_1 -co-ordinate (in units of R) of the Lagrange points L_1, L_2, L_3 as a function of ν :


0.2
0.4
0.6
0.8
1

32.22 For the Earth-Moon system $\mu \sim \frac{1}{81}$ and $R \sim 3 \times 10^5$ km; hence the points $L_{1,2}$ are at a distance of 5×10^4 km from the Moon.

Chapter 10

Time Dependent Linear Systems

•Often we are interested in small corrections to a given orbit: a generalization of the perturbations around static solutions that we discussed in the last chapter. In particular we are interested in the staibility of periodic orbits.

•These problems arise for example in celestial mechanics. The orbit of the moon around the earth is periodic with a period of about 28 days. The correction due to the gravitational field of the Sun can be thought of as a small perturbation, which will lead to a linear differential equation with periodic coefficients (Hill's equation).

•The equations of motion of the Lagrangian $L(q,\dot{q}) = \frac{1}{2}\dot{q}_i\dot{q}_i - V(q)$ can be written as a first order system

$$\frac{d}{dt}q_i = p_i, \quad \frac{d}{dt}p_i = -\frac{\partial V}{\partial q_i}.$$

Here, p_i is the momentum conjugate to q_i . Let $(q_{0i}(t), p_{0i}(t))$ be a solution and $q_i(t) = q_{0i}(t) + \xi_i(t), p_i(t) = p_{0i} + \eta_i(t)$ be another solution close to it. If $(\xi(t), \eta(t))$ is infinitesimally small, we can linearize the equations of motion:

$$\frac{d}{dt}\xi^i = \eta_j, \quad \frac{d}{dt}\eta_i = -\frac{\partial^2 V}{\partial q^i \partial q^j}(t)\xi_j.$$

•If $z = (\xi, \eta)$ we can write this as a linear system of differential equations

with time dependent coefficients:

$$\frac{d}{dt}z = A(t)z, \quad A = \begin{pmatrix} 0 & 1 \\ -V'' & 0 \end{pmatrix}.$$

•If A is a constant matrix the solution is

$$z(t) = e^{At}z(0) = \left[\sum_{n=0}^{\infty} \frac{(At)^n}{n!}\right] z(0).$$

The series converges because it can be dominated term by term by the exponential series

$$\sum_{n=0}^{\infty} \frac{(|A|t)^n}{n!}$$

where $|A| = \sup_u \frac{||Au||}{||u||}$ is the norm of the matrix. ($|A|$ is basically the largest magnitude of the eigenvalues of A .)

•There is a similar convergent power series expansion for the solution of the time dependent linear system. This is in fact a way to prove the existence of solutions of such a syste,. It also proves the real analyticity of the solution as a function of time, provided $A(t)$ itself is real analytic.

•The trick is to write the ODE as a system of integral equations:

$$z(t) = z(0) + \int_0^t A(t_1)z(t_1)dt_1$$

and iterate it to get an infinite series:

$$\begin{aligned} z(t) \;=\; & z(0) + \int_0^t A(t_1)z(0)dt_1 + \int_0^t dt_2 \int_0^{t_2} dt_1 A(t_2)A(t_1)z(0) + \cdots \\ & \int_0^t dt_{n-1} \int_0^{t_{n-1}} dt_{n-2} \cdots \int_0^{t_2} dt_1 A(t_{n-1}) \cdots A(t_2)A(t_1)z(0) \\ & + \int_0^t dt_n \int_0^{t_n} dt_{n-1} \cdots \int_0^{t_2} dt_1 A(t_n) \cdots A(t_2)A(t_1)z(0) + \cdots. \end{aligned}$$

Now, suppose $|A(t_1)| < K(t)$ for all $0 < t_1 < t$. Then the n th term is smaller than

$$\frac{[K(t)t]^n}{n!}.$$

This proves the convergence of the infinite series, again by dominating it term by term by the exponential series.

•The *path ordered exponential* of A is defined by the infinite series

$$\hat{e}^{\int_0^t A(t_1)dt_1} = \sum_{n=0}^{\infty} \int_{t>t_n>t_{n-1}\cdots t_1>0} dt_n \cdots dt_1 A(t_n)A(t_{n-1}) \cdots A(t_2)A(t_1).$$

If $A(t)$ is constant, this becomes just the usual exponential series for e^{At} . The times in the integration are ordered; if $A(t)$ is a constant we can replace this by the integral over $[0,t]$ except we divide by the number of permutations $n!$. This is how you get the exponential series.

•If $A(t)$ is of zero trace for all time, its path ordered exponential is of determinant one. For constant matrices it follows from the identity

$$\det e^A = e^{\,\mathrm{tr}A}.$$

More generally, for any matrix valued function $g(t)$

$$\frac{d}{dt} \log \det g(t) = \mathrm{tr}\left[\frac{d}{dt} g(t)\right] g^{-1}(t).$$

(Prove this by first order perturbation theory!). If $g(t) = \hat{e}^{\int_0^t A(t_1)dt_1}$, we have

$$\left[\frac{d}{dt} g(t)\right] g^{-1}(t) = A(t), \quad g(0) = 1.$$

From this it follows that $\det g(t) = 1$ if $\mathrm{tr}A(t) = 0$.

•Now suppose that $A(t)$ is periodic: $A(t+T) = A(t)$. It doesnt follow that its path ordered exponential $g(t)$ is periodic. But it will be periodic upto a factor:

$$g(t+T) = Mg(t)$$

where

$$M = g(T).$$

The point is that $g_1(t) = g(t+T)$ and $g(t)$ satisfy the same differential equations; the only difference is in the initial conditions: $g_1(0) = g(T)$ whereas $g(t0) = 1$. So the solutions are the same upto a constant multiple.

•A solution $q_0(t)$ is stable if the equation for perturbations around it has always bounded solutions. Now, after each period the solution changes by a factor of M :

$$z(t + nT) = M^n z(t).$$

The question is whether the sequence M^n is bounded. This is true whenever the eigenvalues of M are distinct and are all of magnitude less than or equal to one.

Chapter 11

Hamiltonian Mechanics

•We start with the notion of a Legendre transform in calculus. Let $f : R^n \to R$ be a *convex* function; i.e., the second derivative f'' is a positive matrix everywhere. Then its Legendre transform $\hat{f}(p)$ is defined by

$$\hat{f}(p) = \mathrm{extr}_x \left[p \cdot x - f(x) \right].$$

Here, extr_x denotes the extremum in the variable x . There is a unique extremum for the quantity in the brackets because $f''(x)$ is invertible everywhere. The extremum is the solution of the non-linear equations:

$$f'(x) = p$$

which has a solution (*the inverse function theorem of calculus*) in a domain where $f''(x)$ is non-zero.

•The Legendre transform is itself a convex function. The transform is its own inverse:

$$\hat{\hat{f}}(x) = f(x).$$

•A useful example is a quadratic function,

$$f(x) = \frac{1}{2} x^T A x$$

for some positive matrix x . Then the extremum is at $x_0 = A^{-1} p$ and

$$\hat{f}(p) = \frac{1}{2} p^T A^{-1} p.$$

The above properties are obvious here.

•More generally the extrema of f and $\hat{f}$ are given by the pair of equations

$$p = f'(x), \quad x = \hat{f}'(p).$$

The solutions of these equations are inverse functions of each other. (*Prove this!*). Thus the extremization problem has two equivalent formulations, either in terms of f or $\hat{f}$.

•The Legendre transform is an approximation to the Fourier transform; it is the leading order of the stationary phase approximation. But we wont puruse this matter here.

•The Legendre transform arises in many physical situations where we an extremization problem has an equivalent formulation in two 'dual' descriptions. (If we had regarded f as a concex function ona vector space V , the Legendre transform would have been a convex function on the dual space.)

•For example, in thermodynamics, pressure and volume are dual variables. The internal energy can be viewed either as a convex function of pressure or of volume; the two versions of the energy are Legendre transforms of each other.

•In mechanics, we are interested in an extremum of a function of an infinite number of variables: the action viewed as a function of the position and velocities. There is a 'dual' formulation of this problem where the velocities are traded for conjgate variables called momenta. This gives a new formulation of the action principle, the *Hamitonian formulation*.

•Recall that the action is

$$S[q] = \int L(q(t), \dot{q}(t), t)dt.$$

By analogy to the above, let us define the Legendre transform of the lagrangian with respect to the velocities:

$$H(q(t), p(t), t) = \sum_i p_i(t)\dot{q}^i(t) - L(q(t), \dot{q}(t), t)$$

where the $p(t)$ and $\dot{q}(t)$ are related by the condition that the r.h.s be an extremum:

$$p_i = \frac{\partial L}{\partial \dot{q}^i(t)}.$$

(We are supposed to solve thse equation to eliminate $\dot{q}$ in terms of p in the formula for H .) This function $H(q,p,t)$ is the *hamiltonian*.

•Conversely the lagrangian is the Legendre transform of the hamiltonian. The action is then

$$S[q,p] = \int \left[\sum_i p_i \dot{q}^i(t) - H(q(t),p(t),t)\right] dt$$

The equations of motion in this alternative picture are obtained by varying q and p as independent variables:

$$\frac{dq^i}{dt} = \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q^i}.$$

These are *Hamilton's equations*.

• *Example* Let $L = \frac{1}{2}m\dot{q}^2 - V(q)$, the Lagrangian of a partcle of mass m in a potential $V(x)$. The variable p is just the momentum $p = m\dot{q}$ and H the energy:

$$H(q,p,t) = \frac{p^2}{2m} + V(q).$$

Whenever the Lagrangian does not depend explicitly on time, the hamiltonian is the energy.

•Hamilton's equations are a system of $2n$ first order differential equations for a system with n degrees of freedom.The set of initial conditions is a manifold of dimension $2n$, called the phase space. Through each point in this space is a unique curve which is the solution of the equations of motion.

• *Example*The system with hamiltonian $H = \frac{p^2}{2m} + \frac{1}{2}kx^2$ is the simple harmonic oscillator. The equations of motion are linear

$$\frac{dq}{dt} = \frac{p}{m}, \quad \frac{dp}{dt} = -kq.$$

The solutions are ellipses in the plane (q,p) .

•Any observable quantity of a mechanical system can be expressed as a function of co-ordinates and momenta; i.e., as functions on the phase space. (We exclude for now any explicit dependence on time for f as well as the hamiltonian.) It is interesting to ask how such an observable changes with time:

$$\frac{df}{dt} = \sum_i \frac{\partial f}{\partial q^i}\frac{dq^i}{dt} + \sum_i \frac{\partial f}{\partial p_i}\frac{dp_i}{dt}.$$

By Hamilton's equations,

$$\frac{df}{dt} = \sum_i \left(\frac{\partial f}{\partial q^i}\frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial H}{\partial q^i} \right).$$

•The quantity in the brackets is of fundamental significance in classical mechanics. It is called the *Poisson bracket*. In general the Poisson bracket of a pair of observables is defined by

$$\{f, g\} = \sum_i \left(\frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i} \right).$$

•Thus Hamilton's equations are equivalent to the statement

$$\frac{df}{dt} = \{f, H\}$$

for any observable. In particular a *conserved quantity* is a function tat has no explicit time dependence and also has zero total time derivative: in other words a conserved quantity has zero Poisson brackets with the hamiltonian:

$$\{f, H\} = 0$$

•The Poisson bracket satisfies the identities

$$\{f, g\} = -\{g, f\}, \quad \{\{f, g\}, h\} + \{\{g, h\}, f\} + \{\{h, f\}, g\} = 0, \quad \{f, gh\} = \{f, g\}h + g\{f, h\}.$$

The first identity is obvious. The second (*the Jacobi identity*) is hard, but straightforward to prove by a long calculation. The last is just the statement that the Poisson brackets is linear in the first derivatives of the functions.

•It is more important to know why these are interesting identities; indeed why this whole way of thinking of mechanics is useful. The understanding of any mechanical system starts with that of its conserved quantities. Some of them will be obvious but others can be quite obscure at first. The above identities show for example that the product (*easy!*) and the Poisson brackets (*use the Jacobi identity!*) of two conserved quantities are also conserved.

•But that is just the beginning. Imagine that we could find a complete set of observables which are either constant in time, or depend linearly on time.

•Then using these observables as co-ordinates the solution of the system would be as trivial as that for the free particle. Thus instead of trying to solve differential equations directly we shift to looking for appropriate changes of co-ordinates.

•But what are the allowed changes of co-ordinates that preserve Hamilton's equations;i.e., so that the new co-ordinates (Q, P) also satisfy

$$\frac{dQ^i}{dt} = \frac{\partial H}{\partial P_i}, \quad \frac{dP_i}{dt} = -\frac{\partial H}{\partial Q^i}.$$

A straightforward calculation will show that the conditions are

$$\{Q^i, Q^j\} = 0 = \{P_i, P_j\}, \quad \{Q^i, P_j\} = \delta^i_j.$$

(Here δ^i_j is the *Kronecker* symbol, which is 1 for $i = j$ and 0 otherwise.) Variables that satisfy these relations are said to be *canonical*: they are the "correct" co-ordinates to use in classical mechanics.

•There is an enormous freedom within the world of canonical co-ordinates: we can even mix momenta and position with each other. If we can find canonical co-ordinates (θ^i, I_i) such that the hamiltonian only depends on the 'momenta' I_i we have solved the equations of motion. For in these co-ordinates,

$$\frac{dI_i}{dt} = 0, \quad \frac{d\theta^i}{dt} = \frac{\partial H}{\partial I_i}.$$

•Discovering such *normal co-ordinates* is the holy grail of classical mechanics: once we find such a co-ordinate system the equations are solved in a trivial way. The whole question is then whether such co-ordinates exist and if so how to find them.

•The key to understanding these issues is, surprisingly, geometry. Let us first look at the hamrmoic oscillator:

$$H = \frac{p^2}{2m} + \frac{1}{2}kq^2.$$

We saw that the orbits are ellipses in the phase plane: these are just the curves of constant energy:

$$\frac{p^2}{2m} + \frac{1}{2}kq^2 = E.$$

The particle moves along this ellipse as time evolves. In this case we know explicit answer:

$$q(t) = \left[\frac{2E}{k}\right]^{\frac{1}{2}} \cos[\omega(t-t_0)], \quad p(t) = \sqrt{[2mE]}\sin[\omega(t-t_0)]$$

Thus the change of variables $q = A\cos\theta, p = B\sin\theta$ suggests itself, where A, B are the axes of the ellipse. What is the conjugate I of the angle variable θ ? It turns out to be (upto a constant) the *area of the ellipse!*. To see this, let us think of A, B as functions of I . Then the canonical relations among the pairs (q, p) and (θ, I) give

$$\{q, p\} = \frac{\partial q}{\partial I}\frac{\partial p}{\partial \theta} - \frac{\partial q}{\partial \theta}\frac{\partial p}{\partial I} = 1.$$

which gives

$$A(I)B'(I) + A'(I)B(I) = 1, A(I)B'(I) - A'(I)B(I) = 0.$$

In other words $AB = I$ or $I = 2E\omega$. The area of the ellipse is the product of the semi-axes times π . Thus the variable I is just the area of the ellipse upto a factor of π .

•Why do we get such a geometrical answer? The point is that

$$\{\theta, I\} = \frac{\partial(\theta, I)}{\partial(q, p)}$$

is just the Jacobian of the transformation $(q, p) \mapsto (\theta, I)$. Thus canonical transformations preserve the area. Indeed, the area of a curve can be written either as $\int_\gamma p dq$ or $\int_\gamma I d\theta$ as long as ther two co-ordinates are canonically related.

•Thus we can generalize the above ideas to an arbitrary conservative system with one degree of freedom. The curves $H(p, q) = E$ enclose an area $I(E) = \int_{\gamma_E} p dq$. This area can be chosen as one of the co-ordinates in phase space. It is obviously a conserved quantity since it depends only on energy. Conversely energy can be thought of as a function of I : $H = f(I)$. The conjugate variable θ will depend on time linearly since it satisfies Hamilton's equations:

$$\frac{d\theta}{dt} = \frac{\partial f(I)}{\partial I}.$$

Indeed $\theta = \omega(I)[t - t_0]$ where the frequency is $\frac{\partial f(I)}{\partial I}$. In the case of the harmonic oscillator this frequency is even independent of energy or I ; but in general this is not true.

Chapter 12

Hamilton-Jacobi Theory

•Suppose we are presented a system with hamiltonian $H(q,p)$ in some canonical co-ordinate system (q^i, p_i) . If we can find a new canonical co-ordinate system (θ^i, I_i) such that the hamiltonian depends only on the momentum variables I_i , then we have solved the equations of motion. For, in this system, the Hamiltons equations are

$$\frac{dI_i}{dt} = 0, \quad \frac{d\theta^i}{dt} = \omega^i(I)$$

where the frequencies are $\omega^i(I) = \frac{\partial H(I)}{\partial I_i}$. Such a system will be called *normal co-ordinate system*; they are also called *action-angle* variables.

•The motion is on surfaces of constant I_i , on which the n variables θ^i are co-ordinates; they depend linearly on time. Often this is a surface of finite volume and the co-ordinates θ^i are some sort of angles: the surface is a torus. Indeed, there is a vector field (the velocity) which is everywhere non-zero and has constant components in these co-ordinates. The only compact surface on which that is possible is a torus. It is conventional to choose the period of each angle to be 2π . These surfaces of constant I_i are called *invariant torii* (they are invariant under time evolution.)

•If there is a real number ω and non-zero integers n_i such that $\omega_i = n_i\omega$, the motion is periodic. But generically, the frequencies are not related: there is no set of integers such that $\sum_i \omega_i = 0$. In this case there is a period $\frac{2\pi}{\omega_i}$ in each variable separately, but no common period for the motion as a whole: it is *quasi-periodic*.

•A hamiltonian for which such a normal co-ordinate system exists is said to be *integrable*. It is found that such hamiltonians are 'rare': under small

perturbations we will lose this property. Still much of what we know about mechanics comes from such examples.

•A co-ordinate independent characetrization of integrablity is that the whole phase space can be split up ('foliated') into invariant torii.

•Thus much of mechanics is about the search for the normal co-ordinates or equivalently the invariant torii. Even when they dont exist the attempt to construct them (by perturbing around an integrable limiting case) will reveal deeper structure of the system and even about mechanisc in general.

•Under small perturbations of an integrable system, some of the torii remain invariant although they suffer some deformation: the phase space is divided up into regions bounded by these surviving torii. (Of course by time evolution one can never cross an invariant torus.) To any compact subset of the phase space, we can associate the ratio of the volume of the points that lie on invariant torii to the total volume. This ratio can be made to approach one by choosing the strength of the perturbation small enough. Thus small perturbations dont destroy totally the integrability of a system: 'most' initial data still lie on invariant torii. This is believed to 'explain' the stablity of the solar system in spite of the fact that it is not an integrable system: it is a small perturbation of an integrable system so most orbits are still either on invariant torii or are bounded by them.

•What I gave above is a paraphrasing of one of the most profound results in all of mechanics, may be even physics: the Kolmogorov-Arnold-Moser theorem. We will say more about it later.

•However the structure of the invariant torii is enormously compicated even with a small perturbation: inside each invariant torus there can be another and so on ad infinitum: there is a self-similar structure for the typical system. Even this case of small perturbations around an integrable system is poorly understood still.

•There are mechanical systems which as far from being integrable of course. An extreme example would be an *ergodic system* for which each orbit is dense on the energy surface. Clearly then there are no constants of the motion other than energy and no invariant torii either. These systems are believed to be described by statitical mechanics.

•There is a wide variety of possiblities in between. Of special interest are periodic orbits. For generic systems there are only a countable number of these (periodic orbits are 'rare'). The actions of these periodic orbits form a kind of *classical spectrum* for a mechanical system.

•Indeed Gutzwiller has shown that this classical spectrum of periodic orbits determines the eigenvalues of the quamtum mechanical hamiltonian, in the limit as $\hbar \to 0$.

•So how does one look for normal co-ordinates? Suppose we are given the hamiltonian as a function of some canonical co-ordinates (q, p) . We seek new canonical co-ordinates (θ, I) such that the hamiltonian is independent of the θ ('angle') variables. The canonical transformation from (q, p) to (θ, I) must have a generating function $W(q, I)$, such that

$$p_i = \frac{\partial W(q, I)}{\partial q^i}, \quad \theta^i = \frac{\partial W}{\partial I_i}.$$

•Putting this into the hamiltonian, we get a first order partial differential equation for W :

$$H(\frac{\partial W}{q^i}, q) = E.$$

This is called the *Hamiltonian-Jacobi equation*. If we can solve this equation we have found the canonical transformation to the normal co-ordinates, which is the same as solving the equations of motion.

•The constants of integration in solving this equation are the 'action' variables I_i .

•Thus we have transformed the original system of oridnary differential equations (Hamilton's equations) to a single first order non-linear partial differential equation. In fact this is a reversible process: any single first order partial differential equation can be turned to a system of ordinary differential equations. The curves that solve this system are the 'characteristics'.

•Let us look at an examples from outside of mechanics, the propagation of waves in an inhomogenous medium. We will think of them as light waves but the story is the same for sound waves or water waves. The wave equation is then

$$-\nabla^2 \phi + \frac{n^2(x)}{c^2} \frac{\partial^2}{\partial t^2} \phi = 0.$$

(We have ignored the polarization of light; also some additional approximations about the gradient of n have been made. See Born and Wolf for the complete story.) Here n is the refractive index and c is the velocity of light.

•For monochromatic waves,

$$\nabla^2\psi + k^2 n^2(x)\psi = 0$$

where k is the wave number. Put $\psi = e^{ikW}$ to get

$$-(\nabla W)^2 + n^2 + \frac{1}{k^2} i\nabla^2 W = 0.$$

If the variational of the refractive index over one wavelength is small, we can ignore the second derivative term of W a first order partial differential equation. The function W (called the *eikonal*) satisfies

$$(\nabla W)^2 = n^2.$$

The analogy to the Hamilton-Jacobi equation is clear.

•It is as if we have a hamiltonian $H(x,p) = \frac{1}{2}[p^2 - n^2]$. The equations for the characteristics are

$$\frac{dx}{d\tau} = p, \quad \frac{dp}{d\tau} = \nabla n^2.$$

These are the equations for the light rays. This formulation is called *Hamiltonian optics*. There is a conjugate description in terms of Langrangians as well. Determining rays is helped by the existence of *integral invariants* which are the analogues of conserved quantities.

•The eikonal equation is the approximation to the wave equation in the limit of short wavelength. Is it possible to think of the Hamilton-Jacobi equation as an approximation to a wave equation? Let $H(p,q) = \frac{p^2}{2m} + V(q)$. We would have by analogy $\psi = e^{\frac{i}{\hbar}W}$ .The constant $\hbar$ has the dimensions of action since that is the dimension of W in the H-J equation.

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V(q)\psi = E\psi.$$

This is just the Schrodinger equation; $\hbar$ is Planck's constant!. Hamilton had some inkling of the existence of such a wave mechanics. But at that time there was no experimental reason to doubt the absolute validity of Newtonian mechanics.

• *Example* Let us solve the Kepler problem this way. In spherical polar co-ordinates

$$H = \frac{1}{2m}\left[p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\phi^2}{r^2\sin^2\theta}\right] + V(r,\theta,\phi).$$

As long as the potential is of the form

$$V = a(r) + \frac{b(\theta)}{r^2} + \frac{c(\phi)}{r^2 \sin^2 \theta}$$

a solution of the form

$$W = R(r) + \Theta(\theta) + \Phi(\phi)$$

exists. This is the method of *separation of variables*. The ordinary differential equations for the functions R, Θ, Φ can be solved in terms of elliptic functions. For details see Landau and Lifshitz.

•If we can solve the H-J equations, and determine the normal co-ordinates, we can solve the equations of motion completely. Such systems are said to be *integrable*. It turns out that not all systems are integrable.

•An integrable system with n degrees of freedom will have n conserved quantities: the 'action' variables. In fact it will have n conserved quantities in *involution*: their mutual Poisson brackets are zero. Conversely whenever there and n independent conserved quantities in involution, a system with n degrees of freedom is integrable.

•The surfaces defined by $I_i =$ constant are torii: they admit vector fields that are non-zero everywhere, $\frac{\partial}{\partial \theta^i}$ .The existence of these *invariant torii* is an abstract characterization of integrability.

•In the normal co-ordinates, the solutions of the equations of motion are real analytic in time. Indeed, this appears to be another characetrization of integrability: if the solutions of the equations of motion are real analytic in time, the system is integrable. But I dont know of a rigorous theorem along these lines.

Chapter 13

Birkhoff Normal Form

•See F. G. Gustavson Astonomical Journal Vol 71, 670 (1966) for a beautiful exposition of these ideas.

•We now start the study of classical systems that are not integrable. Suppose that a system has a stable equilibrium point; Near this point it is well approximated by a harmonic oscillator; there are canonical co-ordinates such that the hamiltonian is

$$H(q,p) = \frac{1}{2}\sum_i \omega_i(p_i^2 + q_i^2) + O(3)$$

upto terms that are cubic in the co-ordinates.

•The higher order terms describes the non-linearities in the equations of motion. We can regard the hamiltonian as a power series

$$H(q,p) = \sum_{r=2}^{\infty} H^r(q,p)$$

where $H^r(q,p) = \sum_{|A|+|B|=r} H_{AB} q^A p^B$ is a homogenous polynomial of order r in (q,p). Here $A = \{A_1, A_2 \cdots A_n\}$ is a multi-index and $q^A = q_1^{a_1} \cdots q_n^{a_n}$ etc. Also, $|A| = a_1 + a_2 \cdots + a_n$, so that the degree of the polynomial is $r = |A| + |B|$. Of course $H^2(q,p) = \frac{1}{2}\sum_i \omega_i(p_i^2 + q_i^2)$.

•Even if the hamiltonian is a smooth function, this series may not converge; it may not be a real analytic function. Still this series makes sense as a 'formal power series'. (See the appendix for a discussion of formal power series.)

•We say that the hamiltonian is *in normal form upto order* s if

$$DH^r = 0 \quad \text{for } r \leq s,$$

where

$$D = \sum_i \omega_i \left(q_i \frac{\partial}{\partial p_i} - p_i \frac{\partial}{\partial q_i} \right).$$

In other words $H_{AB} = 0$ for $|A| + |B| \leq s$ unless

$$\sum_i \omega_i [a_i - b_i] = 0.$$

•To understand what this means, note that the operator D is

$$D = \sum_i \omega_i \frac{\partial}{\partial \theta_i}$$

where θ_i are the angular co-ordinates conjugate to $I_i = q_i^2 + p_i^2$. The simplest case is when the frequencies ω_i are *linearly independent* to order s : there are no integers k_i , with $|k_1| + |k_2| \cdots |k_n| \leq s$ such that

$$\sum k_i \omega_i = 0.$$

Then for a hamiltonian of normal form, the hamiltonian is independent of θ_i and I_i are conserved to order s : they serve as approximate action-angle variables.

•In the case where the frequencies are linearly dependent there may be resonances, and the hamiltonian may depend on additional variables and we wont have enough conserved quantities to 'solve' the system even approximately.

• *Theorem* Given any hamiltonian $H = \frac{1}{2} \sum_i \omega_i [q_i^2 + p_i^2] + O(3)$ and any positive integer s , there exists a canonical transformation to new co-ordinates (Q_i, P_i) such that the new hamiltonian $\Gamma(Q, P)$ is of normal form to order s .

•Thus we can solve any classical mechanical system to arbitrary order if the frequencies are linearly independent.

•The proof proceeds by induction; in fact we will give an algorithm to construct the necessary canonical transformation order by order. Suppose the hamiltonian is already in normal form to some order $s - 1$. (Right at the start we have $s = 3$.) Consider a canonical transformation generated by $\sum_i q_i P_i + w^s(q, P)$ where $w^s(q, P)$ is a homogenous polynomial of order s in the new momenta and old co-ordinates. The old co-ordinates q can be eliminated in terms of the new ones Q using

$$Q_i = q_i + \frac{\partial w^s}{\partial P}.$$

The new hamiltonian is determined by

$$H(q, P + \frac{\partial w^s}{\partial q}) = \Gamma(q + \frac{\partial w^s}{\partial P}, P).$$

If we equate terms of order s of this equation we get

$$Dw^s(q, P) = \Gamma^s(q, P) - H^s(q, P).$$

Now we show that w^s can be chosen such that Γ^s is of normal form; i.e., in the kernel of D. D is a linear operator in the space of homogenous polynomials V^s. It is not invertible, but is antisymmetric so that we have a decomposition $V^s = \ker D + \text{range } D$. In particular we can split $H^s = K^s + R^s$ where $K^s \in \ker D$ and $R^s \in \text{range } D$. Thus if we determine w^s by solving $Dw^s = R^s$, we have $\Gamma^s = R^s$; i.e., Γ^s is in normal form. This completes the proof.

•The solution of the equation $Dw^s = R^s$ is facilitated by diagonalizing D. Being anti-symmetric, it is diagonalized by complex linear combinations $z_j = q_j + ip_j, \bar{z}_j = q_j - ip_j$:

$$Dz^A\bar{z}^B = i\sum_i [a_i - b_i]\omega_i z^A\bar{z}^B.$$

Thus the pseudo-inverse of D may be defined to be

$$D^{-1}z^A\bar{z}^B = \left\{ \begin{array}{cc} 0 & \text{if } \sum_i [a_i - b_i]\omega_i = 0 \\ \frac{1}{i\sum_i [a_i - b_i]\omega_i} z^A\bar{z}^B & \text{otherwise} \end{array} \right\}.$$

Then we just have to express H^s as a polynomial in $z, \bar{z}$ and then set $w^s = -D^{-1}H^s$.

•This will give us the hamiltonian as a function of (q, P). We must now eliminate q, solving

$$q_i = Q_i - \frac{\partial w^s(q, P)}{\partial P_i}$$

as a power series in (Q, P). Substituting into the hamiltonian will give us an infinite series for Γ. Gustavson has worked out the following formula for the term of order r in Γ

$$\Gamma^r = H^r + \sum_{\substack{l - |j| + |j|(s-1) = r \\ 1 \leq |j| \leq l \leq r \\ l \geq 2, s \geq 3}} \frac{1}{j!}\left[\frac{\partial^{|j|}H^l}{\partial P^j}\left(\frac{\partial w^s}{\partial q}\right)^j - \frac{\partial^{|j|}\Gamma^l}{\partial q^j}\left(\frac{\partial w^s}{\partial P}\right)^j\right]$$

Here $j! = j_1!j_2!\cdots j_n!$ etc. To order less than s this says that $\Gamma^r = H^r$. To order s this is the equation that determines w^s . In higher orders this gives the terms of the new hamiltonian in terms of the old one recursively. This formula can be translated into a symbolic computer program.

•By iterating this procedure, the hamiltonian can be reduced to normal form to any order.

•If the frequencies are linearly independent to all orders, we can construct a complete set of conserved quantities $I_i = Q_i^2 + P_i^2$ as a formal power series in the original co-ordinates (q, p) . If this power series were convergent, we would have been able to show that every non-resonant classical mechanical system is integrable. But it turns out that typically, the series are divergent. Still, if the perturbations are small enough this procedure gives a good approximate solution.

Appendix: Formal Power Series

•A convergent power series

$$f(z) = \sum_{n=0}^{\infty} f_n z^n$$

determines an analytic function. The sum, product and derivative of analytic functions translate into operations on the coefficients:

$$[f+g]_n = f_n + g_n, \quad [fg]_n = \sum_{k+l=n} f_k g_l, \quad [f']_n = (n+1) f_{n+1}$$

•These operations involve only a finite number of coefficients. Hence the space of infinite sequences $(f_0, f_1, \cdots)$ is a commutative ring with derivation. If the sequence is rapidly decreasing, this corresponds to an analytic function; otherwise the sequence can be thought of as generalizing the notion of a function, a *formal power series*. In many mathematical situations, this ring of formal power series a good substitute for the ring of functions of one variable. We may still write occasionally

$$f(z) = \sum_{n} f_n z^n$$

but of course the quatity $f(z)$ is not always a function: there may be no meaning to evaluating it at some value of z .

•The idea generalizes to an arbitrary number of variables: the coeffiecients become symmetric tensors:

$$f(z) = \sum f_{i_1 \cdots i_r} z_{i_1} \cdots z_{i_r}.$$

An equivalent description is in terms of multi-exponents $A = (a_1, \cdots a_n)$:

$$f(z) = \sum_{A=0}^{\infty} f_A z_1^{a_1} \cdots z_n^{a_n}.$$

•A divergent series might describe a function that is not analytic. A series $\sum_{n=0}^{\infty} f_n z^n$ is an *asymptotic expansion* of a function, $f(z) \sim \sum_{n=0}^{\infty} f_n z^n$ if there is a wedge $\theta_1 < \arg(z) < \theta_2$ such that

$$\lim_{|z| \to 0} z^{-n} [f(z) - \sum_{0}^{n} f_k z^k] = 0$$

for every n . The limit is taken along some ray within the wedge. The idea is that the difference between $f(z)$ and the sum of the first n terms is of the same order as the $(n+1)$ th term, as long as $\arg(z)$ is within the range.

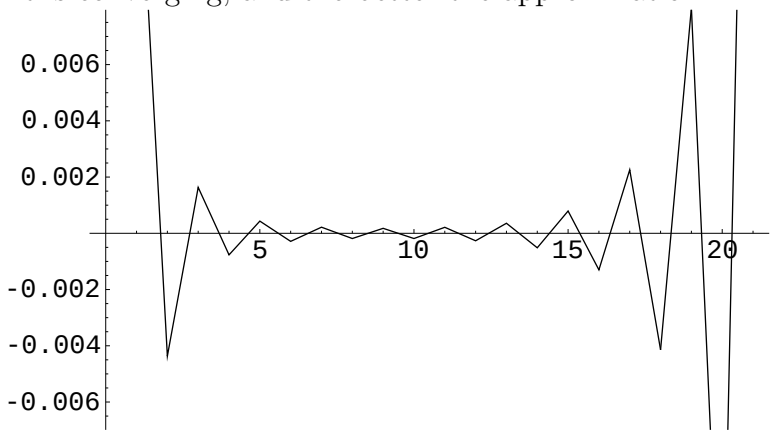
•An asymptotic series, unlike a convergent series does not determine the function uniquely. For example we can always add $e^{-\frac{1}{z}}$ to any function without changing its asymptotic series!. The point is that all the derivatives of $e^{-\frac{1}{z}}$ are zero at the origin and yet it is not analytic.

•If additional information is available about the function we can determine it from its asymtotic series. The ideas to do that involve integral transforms (Borel transform) or Pade approximants.

• *Example* The series $f(z) = \sum_0^\infty n! z^n$ is not convergent. Yet it can be thought of the Taylor series expansion of the function

$$f(z) = \int_0^\infty \frac{e^{-t} dt}{1 - zt}.$$

In fact for moderate values of $|z|$ (say $|z| < 0.1$) we can even find a good approximation to the value of the function by the power series. (Exercise: Write a small program to verify this. It is easy to do in Mathematica.) The point is that the series starts to converge to the value of the function for a finite number of terms before diverging away. The smaller the value of $|z|$, the longer it is converging, and the better the approximation.



The difference $f(z) - \sum_{n=0}^N n! z^n$ for $z = -0.1$ and various values of N . Note that the error decreases at first then diverges for larger powers.

•Another example of this phenomenon is the Stirling expansion of the Gamma function. The Gamma function has an essential singularity at infinity, so cannot have a convergent power series expansion valid for large z .

Yet, the Stirling formula gives very accurate numerical values for $|z|$ greater than about 10 . (The Stirling expansion is not quite a power series; but after some leading factors are removed it can be made into one.) Actually there is an irony here. The inverse of the Gamma function is an entire function, hence has a convergent power series expansion for *any* z . So it would seem reasonable that we calculate the value of $\frac{1}{\Gamma(z)}$ by the power series and just find the reciprocal to find the value of $\Gamma(z)$. However, the Stirling series, which is divergent, gives more accurate answers than the convergent series once $|z|$ is moderately large $|z| > 10$.

•So we should not dismiss divergent series as useless. What is going on here is that the usual notion of convergence is too restrictive: it requires the series to give a good approximation for all values of z within a circle. But often we are only interested in a good approximation within a pie-shaped region, for example containing the real axis. Then the appropriate approach is to seek an *asymptotic series*.

•Thus the weakest notion is that of a formal power series: it may not represent any function. Next weakest is that of the asymptotic series, which will determine the function upto one all of whose derivatves at the origin are zero. The strongest is a convergent power series which uniquely determines the function of which it is the Taylor series expansion.

Chapter 14

The Henon-Hiles Potential

•There are a handful of integrabel systems that are usually used to learn classical mechanics. Yet the vast majority of systems are not integrable. We still need to work out specific examples to learn about the behavior of such systems. Examples arise from application to various areas of physics. In this chapter we will study an example arising from astronomy, the *Henon-Hiles system*: (M. Henon and C. Heiles, Astronomical Journal **69**, 73, (1964).)

$$H = \frac{1}{2}[\dot{x}^2 + \dot{y}^2] + \frac{1}{2}[x^2 + y^2] + x^2 y - \frac{1}{3}y^3$$

•This is supposed to describe gravitational potential seen by a star in the center of a globular cluster. Its importance to us is that it is not integrable and is well studied; we will not delve into the underlying astronomy.

•The origin is a stable equilibrium point; the frequencies are equal, so that the system is resonant. Still, because it has only two degrees of freedom, it is formally integrable by the Birkhoff-Gustavson method. Gustavson compared the true behavior (as determined by numerical calculations) to that predicted by the formal power series method. The results are quite instructive and are believed to be representative of all non-integrable systems.

•The phase space is four dimensional. Since the system is conservative, the dynamics takes place on a three dimensional submanifold of constant energy. If $E \leq \frac{1}{6}$ these surfaces of constant energy are bounded. (In general the potential is unbounded below). If we had one more conserved quantity (integral of motion) we would be able to solve the system.

•The Birkhoff-Gustavson procedure provides another such integral as a formal power series. Suppose the hamiltonian has been reduced to normal form

$\Gamma(Q,P)$. Suppose also that there are r linear relationsa among the freqoencies $A\omega = 0$. Then any function of the form $\sum_i \mu_i[Q_i^2 + P_i^2]$ is a conserved quantity if $A\mu_i = 0$. This is not the hamiltonian Γ : there are many additional terms in the hamiltonian itself. Hence if there are r relations among the frequencies (the matrix A is of rank r) there will be $n-r$ independent solutions to the equation $A\mu = 0$, and an equal number formal integrals independent of the hamiltonian.

•Thus any conservative system with two degrees of freedom is formally integrable: there is always the hamiltonian, and even if there is resonance at least one new integral.

•For the Henon-Hiles potential the new integral is just $\frac{1}{2}[Q_1^2+P_1^2+Q_2^2+P_1^2]$ in terms of the new co-ordinates. We must then retrace our steps inverting the canonical transformations to express this in terms of the old variables $(x,y,\dot{x},\dot{y})$.

•It is hard to visualize dynamics in a four-dimensional space; even after imposing conservation of energy we will have three dimensions. We can look a cross-section of the phase space where one of the co-ordonates (say x) is set equal to zero. The solution is a curve that will intersect the two dimensional space at a countable number points.

•This can be viewed as a discrete dynamical system on its own right, determined by a map $(y,\dot{y}) \mapsto \phi(y,\dot{y})$. Given an initial $(y,\dot{y})$ we can find an $\dot{x}$ for which $x = 0$. This initial condition determines a curve and the next time it intersects the plane $x = 0$ determines the new point $\phi(y,\dot{y})$. This map is a canonical transformation.The iterations of this canonical transformation describe a set of points on the plane $(y_{n+1},\dot{y}_{n+1}) = \phi(y_n,\dot{y}_n)$ upon iteration.

•If there is a conserved quantity I (other than the hamiltonian) the points $(y_n,\dot{y}_n)$ will lie along the curves along which I is constant. This suggests a numerical experiment to test the efficacy of the Birkhoff procedure. We calculate the formal integral of motion to some order (eight in the work of Gustavson.) Plot the cuvres along which this polynomial is a constant. Now numerically calculate the orbit $(y_n,\dot{y}_n)$ of a point on such a curve. If the orbit lie on the curve, the system behaves much like an integrable system: we will have an *invariant curve*. If on the other hand the orbit wanders all over some bounded region in the plane then it is 'chaotic'. it does not lie on an invariant curve and there can be no conserved quantity.

•The results of this numerical experiment performed by Gustavson are quite striking. For small values of energy (small departure from integrability) the

orbits lie more or less on the level surfaces. But as the energy grows some small islands (bounded regions) appear, within which the orbits are spread out. However there are still many orbits that still lie on the level curves. As energy grows to the the maximum value of $\frac{1}{6}$ the entire allowed phase space is filled by a single orbit.

•A finite fraction of the initial conditions remain on curves invariant under time evolution for small perturbations. As the perturbations grows this fraction decreases to zero. Eventually there are no invariant curves at all left.

•This suggests that the Birkhoff series is not convergent or even asymptotic to any conserved quantity. But all is not lost: certain curves may remain under time evolution, although they can no longer be thought of as level surfaces of any conserved quantity. This notion of an invariant curve is weaker than that of a conserved quantity and is stable under small perturbations.

•The KAM theorem demonstrates in a systematicaly how to construct such invariant surfaces.

Chapter 15

Area Preserving Maps

•Since the problem of understanding non-integrable systems is quite hard, we should look first at the simplest non-trivial examples. Any conservative hamiltonian system with one degree of freedom is integrable. So the simplest such systems are of two degrees of freedom, like the Henon-Hiles system.

•If time is a discrete rather than a continuous variable, then even a system with one degree of freedom may be chaotic.

Such systems arise in several natural ways:

(i) The Poincare section, (or restriction) two a dimensional subspace of a hamiltonian system with more than one degree of freedom is a symplectic (are preserving) map of the plane to itself.

(ii) A time dependent hamiltonian, periodic in time with period T , $H(q, p, t)$ determines a symplectic map of the plane, the time evolution through one period.

(iii) Consider a static magnetic field and a surface orthogonal to the field. The integral curves of the magnetic field will return to the surface, and provide a map of the surface to itself. This maps preserves magnetic flux. If the magnetic field is viewed as a symplectic form on the surface, the map is symplectic.

•Thus we are led to study a map $\phi : R^2 \to R^2$ which preserves area: $\det \partial\phi = 1$. An orbit of such a map is an infinite sequence of points satisfying $(q_{n+1}, p_{n+1}) = \phi(q_n, p_n)$.

• *Example* The map

$$q_{n+1} = q_n + p_{n+1}, \quad p_{n+1} = p_n + \lambda \sin q_n$$

is called the *standard map*. It is easy to check that it is area preserving. We

can regard (q,p) as periodic with period 2π ; then this is a map of a torus to itself.

•This can be thought of as the time evolution of a rotor that is 'kicked' at periodic intervals; the hamiltonian is $H(q,p,t) = \frac{1}{2}p^2 + k\sin q \sum_n \delta(t-n)$. In between two kicks the system evolves like a free particle. During a kick the momentum is changed by $k\sin q$.

•If $\lambda = 0$, this is an integrable system: just a rotor. The orbits lie on circles with constant p . If p is rational $p = \frac{N}{D}$ (with N, D coprime) the orbit is periodic with period D ; then N is the number of rotations around the q -axis in one such period. Even when p is not rational it can be approximated by rational numbers: there will be a sequence of periodic orbits that converge to any orbit.

•A *conserved quantity* for ϕ is a function $I(p,q)$ with non-zero derivative everywhere $dI \neq 0$ and satisfying $I(q,p) = I(\phi(q,p))$. If there is such a conserved quantity, the discrete dynamical system is integrable. If θ is the canonical conjugate of I , the map reduces to $(I,\theta) \mapsto (I, \theta + \alpha(I)$: a rotation of θ by an amount that can depend on I . This is called the *normal form* of the map.

•For example, if we have a conservative hamiltonian system of one degree of freedom, and ϕ is the time evolution through a finite interval, the hamiltonian itself would be such a conserved quantity.Then the co-ordinate system (I,θ) would be the action-angle variables.

•Suppose the discrete system is the restriction to integer values of time of a continuous dynamical system: there are symplectic maps $\phi_t : R^2 \to R^2$ such that $\phi_{t+s} = \phi_t(\phi - s)$. Then the infinitesimal time evolution arises from a hamiltonian: $\frac{dq_t}{dt} = \frac{\partial H}{\partial p}, \frac{dp_t}{dt} = -\frac{\partial H}{\partial q}$. This hamiltonian is then a conserved quantity for ϕ and the system is integrable.

•In general it is not possible to reduce a system to normal form: there is no conserved function. But there may be a conserved quantity in the sense of a power series, which may not converge. *G. D. Birkhoff, Acta. Math.* **43**, *1 (1920)* is the definitive study of such maps. This work has been revived in recent times by Kadanoff for example as a model of deterministic chaos.

•The main idea is that as a formal power series, it is possible to continue orbit (q_n, p_n) to all real values of time n . Then by differentiation we will get a 'hamiltonian' as a formal power series.

•As in the case of continuous systems, we start by expanding the map ϕ in a power series around a fixed point. We can choose this to be at the origin.

In the linear approximation we have

$$\begin{pmatrix} q_{n+1} \\ p_{n+1} \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} q_n \\ p_n \end{pmatrix}, \quad ad - bc = 1.$$

In this approximation the normal form of the map is just the same as the normal form for the above matrix. The simplest case is when the eigenvalues are distinct. Then they will be either some real numbers ρ, ρ^{-1} with $\rho > 1$; or a pair of complex conjugate numbers of modulus one: $e^{i\theta}, e^{-i\theta}$. The first case is that of an unstable fixed point while the second is that of a stable fixed point.

•If we choose normal co-ordinates of the linear map, we will have a pair of power series

$$\begin{aligned} q_1 &= \rho q + \sum_{m+n=2}^{\infty} \phi_{m,n} q^m p^n \\ p_1 &= \rho^{-k} q + \sum_{m+n=2}^{\infty} \psi_{m,n} q^m p^n \end{aligned}$$

for the unstable case and

$$\begin{aligned} q_1 &= q \cos\ \theta - p \sin\ \theta + \sum_{m+n=2}^{\infty} \phi_{m,n} q^m p^n \\ p_1 &= q \sin\ \theta + p \cos\ \theta + \sum_{m+n=2}^{\infty} \phi_{m,n} q^m p^n \end{aligned}$$

for the stable case. These are to be viewed as formal power series.

•Its iteration will give

$$\begin{aligned} q_k &= \rho^k q + \sum_{m+n=2}^{\infty} \phi^k_{m,n} q^m p^n \\ p_k &= \rho^{-k} q + \sum_{m+n=2}^{\infty} \phi^k_{m,n} q^m p^n \end{aligned}$$

and

$$\begin{aligned} q_k &= q \cos\ k\theta - p \sin\ k\theta + \sum_{m+n=2}^{\infty} \phi^k_{m,n} q^m p^n \\ p_k &= q \sin\ k\theta + p \cos\ k\theta + \sum_{m+n=2}^{\infty} \psi^k_{m,n} q^m p^n \end{aligned}$$

respectively. The coefficients $\phi^k_{m,n}, \psi^k_{m,n}$ are determined by recursion relations.

•For example, consider the second order terms $\phi^k_{20}, \phi^k_{11}, \phi^k_{02}$ etc. They satisfy

$$\phi^{k+1}_{20} = \rho^k \phi_{20} + \rho^2 \phi^k_{20}, \quad \phi^{k+1}_{11} = \rho^k \phi_{11} + \phi^k_{11} \quad \phi^{k+1}_{02} = \rho^k \phi_{02} + \rho^{-2} \phi^k_{02}.$$

In general we will have

$$\phi^{k+1}_{mn} = \rho^k \phi_{mn} + \rho^{m-n} \phi^k_{mn} + P_{mn}$$

where P_{mn} is a polynomial in ρ, ρ^{-1} and $\phi^k_{\mu\nu}, \psi^k_{\mu\nu}$ with $\mu + \nu < m + n$.

•Thus, for each order we will be solving linear difference equations of the form

$$y_{k+1} - \sigma y_k = c \lambda^k k^\mu.$$

The solution is a polynomial in k times λ^k . For example, $\phi^k_{20} = \frac{\rho^k - \rho^{2k}}{\rho - \rho^2} \phi_{20}$.

•Thus the coeffiecients ϕ^k_{mn}, ψ^k_{mn} can be continued to real values of of the 'time' parameter k : each coefficient is an entire function of k . In this way we get a formal power series for the continuous time dynamical system (q_k, p_k) for all real k . As a formal power series, its Lagr will be one: it is area preserving.

•We can now differentiate term by term to get the infinitesimal form of the time evolution $\dot{q} = \left[\frac{dq_k}{dk}\right]_{k=0}$:

$$\dot{q} = q \log \rho + \cdots, \quad \dot{p} = -p \log \rho + \cdots.$$

Since the evolution preserves area, there must be a formal power series ('hamiltonian') such that

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial q},$$

so that

$$H = pq \log \rho + \cdots.$$

This is the conserved quantity.

•There is a parallel set of formule for the case of a stable fixed point.

•Once a formal conserved quantity has been constructed, it is possible to find a canonical transformation to new co-ordinates in which the dynamics is of the normal form (in the unstable case)

$$Q_1 = \rho Q e^{\frac{1}{2}c(PQ)^l}, \quad P_1 = \rho P e^{-\frac{1}{2}c(PQ)^l}$$

for some constant c and natural number l . In the stable case the normal form is

$$\begin{aligned} Q_1 &= Q\cos[\theta + c[P^2+Q^2]^l - P\sin[\theta + c[P^2+Q^2]^l \\ P_1 &= Q\sin[\theta + c[P^2+Q^2]^l + P\cos[\theta + c[P^2+Q^2]^l. \end{aligned}$$

•If the fixed point is marginally stable (the eigenvalues are both equal to one) the normal form is

$$Q_1 = Q_1 + \alpha(P_1), \quad P_1 = P_1.$$

The standard map we discussed earlier is of this type.

•If there is a conserved quantity, the surfaces on which it is constant are invariant under the dynamics. The Birkhoff procedure allows us to construct such curves by power series. We can regard a *formal curve* as given by a pair of series $(f(t), g(t)$; if these were to converge, we would have an analytic curve on the plane parametrized by t . We say that this formal curve is invariant if there is a third formal series $h(t)$ such that $\phi(f(t), g(t)) = (f(h(t)), g(h(t))$. The Birkhoff procedure gives a construction for such series $f(t), g(t), h(t)$.

•For non-integrable systems, the series defining a formal invariant curve will not always converge; but even in this case some of the curves will have a convergent power series if the system is only a small perturbation from an integrable one. This is the content of the KAM theorem in this context. We need to find a way to improve on perturbation theory and get a convergent expansion, at least for some curves.

Chapter 16

Kolmogorov-Arnold-Moser Theorem

See J. Moser *Stable and Random MOtion in Dynamical Systems* Princeton U. Press (1973).

•Suppose we have an integrable system with hamiltonian H_0 . So there are canonical co-ordinates (θ_i, I_i) , and H_0 is independent of θ_i . We assume that the surfaces of constant I are compact in which case they must be tori; θ_i can be chosen to be periodic co-ordinates with period 2π on this torus.

•Since I_i are conserved quantities the motion will lie on these invariant torii. The angle variables depend linearly on time, with frequencies

$$\omega_i = \frac{\partial H_0}{\partial I_i}$$

that depend on the torus. Only for the harmonic oscillator, the frequencies are the same for all torii.

•We ask what happens to these torii under pertrurbations of the hamiltonian. If the perturbed hamiltonian is still integrable, each invariant torus will be deformed into a new invariant torus. But it is known that this is very rare: the subset of such integrable hamiltonians is a countable union of nowhere dense sets in the space of all smooth hamiltonians.

•And yet many of the invariant torii are survive, upto a deformation. We will now make these ideas more precise.

• Kolmogorov-Arnold-Moser Theorem *Let the hamiltonian be a real analytic function of canonical co-ordinates* (θ, I) *and a real paramater,* $H(\theta, I, g)$ *such that* $H(\theta, I, 0) = H_0(I)$ *. Moreover,* $H(\theta, I, g)$ *is periodic with period* 2π *in the angle variables* θ_i *. Consider a torus* $I_i = c_i$ *at which the Hessian*

$$\det \frac{\partial^2 H_0}{\partial I_k \partial I_l}$$

is non-zero. Also, suppose the frequencies $\omega_i = \frac{\partial H_0}{\partial I_i}$ *satisfy*

$$|\sum_i k_i \omega_i| > K \left[\sum_i |k_i|\right]^{-\tau}$$

for some K, τ *and all integers* k_i *. Then for small enough* $|g|$ *, there is exists a torus*

$$\theta_i = \phi_i + u_i(\phi, g), \quad I_i = c_i + v_i(\phi, g)$$

on which the time evolution is given by the same frequencies as for the unperturbed system:

$$\dot{\phi}_i = \omega_i.$$

(Here $u(\phi, g)$ *and* $v(\phi, g)$ *are periodic in* ϕ *and real analytic in* g *; they vanish for* $g = 0$ *.) Moreover, the torus forms a Lagrangian submanifold.*
• Another version of KAM Theorem *Under the same hypothesis as above, there are analytic functions* $w(\theta, g)$ *and* $E(g)$ *, (where* $w(\theta, g)$ *is periodic in* θ *) satisfying the Hamiltoin-Jacobi equation*

$$H(\theta, c + \frac{\partial w}{\partial \theta}, g) = E(g).$$

•The idea is to solve the equations by an iterative scheme analogous to Newton's method. First we will need a norm to measure the distance between two periodic functions. The H-J equation will be a nonlinear equation $f(w) = 0$ in this space. The initial guess for the solution is just $w = 0$. The norm is chosen so that the linear operator $f'(0)^{-1}$ is bounded.

•Let V be the vector space of trigonometric series with no constant term, and with only a finite number of non-zero terms:

$$V = \{\psi(\theta) = \sum_{m\neq 0} \psi_m e^{im\cdot\theta}\}.$$

Let us define a norm on this space

$$||\psi||^2 = \sum_{m\neq 0} |\psi_m|^2 |m|^{2\tau}.$$

We complete V by this norm to get a complex Hilbert space $\mathcal{H}$.

Now, a linear operator $A : \mathcal{H} \to \mathcal{H}$ of the form $A\psi(\theta) = \sum_{m\neq 0} a_m \psi_m e^{im\cdot\theta}$ has norm

$$|A| = \sup_{m\neq 0} |a_m||m|^{-\tau}.$$

•We regard $w(\theta)$ as an element of this Hilbert space[1]: we set the average $\int w(\theta)d\theta = 0$ since only the derivative of w is relevant anyway. We can regard the Fourier coefficients w_m in $w(\theta) = \sum_{m\neq 0} w_m e^{im\cdot\theta}$ as parameters determining w . Similarly the hamiltonian can be expanded

$$H(\theta, I) = \sum_{m\neq 0} H_m(I) e^{im\cdot\theta}.$$

If we put the expansion of w into this the H-J equation becomes

$$H(\theta, c + \frac{\partial w}{\partial \theta}) = E + \sum_{m\neq 0} f_m(w) e^{im\cdot\theta}$$

where

$$f_m(w) = \int e^{-im\cdot\theta} H(\theta, c + i\sum_n n w_n e^{in\cdot\theta}) d\theta.$$

The H-J equation is just $f(w) = 0$.

•Our first approximation to the solution is just $w = 0$. The derivative of the function f at this point is,

$$f_m(w) \sim m \cdot \omega w_m$$

[1] we suppress the g dependence for convenience

so that $f'(0)^{-1}$ is a linear operator with norm

$$|f'(0)^{-1}| = \sup_{m \neq 0} |m \cdot \omega|^{-1} |m|^{-\tau} \leq K.$$

Now we see the meaning of the constant K in the KAM condition.

•Now we recall that the $f(w, g)$ depends analytically on a parameter g ; when $g = 0$ the root is just $w = 0$. Moreover, $f'(0, 0)^{-1}$ is bounded. Then for small enough g , (the size of g is determined by K) the equation $f(w, g) = 0$ can be solved by Newton's method:

$$w^{k+1} = w^k - f'(w^k)^{-1} f(w^k).$$

See the appendix for more details. The function w produced this way is an element of the Hilbert space $\mathcal{H}$. To show that it is also an analytic function of the angles seems much harder. But that is less important physically.

Appendix: Newton's Method

•Suppose we want to solve an equation $f(x) = 0$ in one real variable. Newton's method is a recursive solution to the equation:

$$x_{k+1} = x_k - [f'(x_k)]^{-1} f(x_k).$$

This converges if the initial guess x_0 is close enough *and* f' does not ever become small.

•The method even applies to the case of several variables, except that x_k are now vectors and f' a matrix. Again, there are problems if the matrix has small eigenvalues; i.e., if its inverse becomes large. Indeed this method even generalizes to infinite dimensional spaces, provided we have a way of measuring the length of a vector; e.g., to Hilbert spaces.

•We can establish convergence of Newton's method by estimating the errors. Let α be the exact solution, $y_k = x_k - \alpha$ .Then

$$0 = f(\alpha) = f(x_k) - f'(x_k)y_k + \frac{1}{2}f''(x_k)y_k^2 + \mathrm{O}\,(y_k^3)$$

so that

$$f'^{-1}(x_k)f(x_k) = y_k - \frac{1}{2}f'^{-1}(x_k)f''(x_k)y_k^2 + \mathrm{O}\,(y_k^3)$$

and using $x_k - x_{k+1} = y_k - y_{k+1} = f'^{-1}(x_k)f(x_k)$

$$y_{k+1} = \frac{1}{2}f'^{-1}(x_k)f''(x_k)y_k^2 + \mathrm{O}\,(y_k^3).$$

(All this makes sense even when f is a map of an infinite dimensional Hilbert space to itself, f' and f'' being tensors of type $(1,1)$ and $(1,2)$ respectively.) Thus we see that

$$|y_{k+1}| < C_k|y_k|^2$$

where $C_k = |f'^{-1}(x_k)f''(x_k)|$. Thus if $f''(x)$ and the inverse $f'^{-1}(x)$ are bounded the procedure will converge. Moreover the convergence is quadratic: $|y_{k+1}| \sim |y_1|^{2^k}$. Trouble can arise if the derivative $f'(x_k)$ does not have a bounded inverse: the problem of small denominators.

•Now suppose $f(x,g)$ depends analytically on a parameter g ; and that when $g = 0$ the root is at the origin: $f(x = 0, g = 0) = 0$. Moreover,

suppose the derivative has bounded inverse at this point: $|f'^{-1}(x=0,g=0)|$ exists. Then for small enough g (determined by K), Newtons method will converge. Thus all torii with small enough K will survive a perturbation. The volume ocuupied by such 'invariant torii' is a finite fraction of the total volume of the energy surface, when g is small enough.

•As the size of the perturbation grows fewer torii will have small enough $f'^{-1}(0)$: more and more of them will dissolve. Eventually, the system should become ergodic: a single orbit can pass arbitrarily close to every point in the energy surface.

Appendix: An Improvement on Newton's Method

•See J. Moser *Stable Random Motions in Dynamical Systems*. We dont need this refinement of Newton's method for our approach to the KAM theorem; it is included for your amusement.

•There is a way to avoid the problem of 'small denominators': use instead the recursion

$$x_{k+1} = x_k - a_k f(x_k), \quad a_{k+1} = a_k + a_k[1 - f'(x_{k+1})a_k].$$

This needs two initial conditons: an x_0 and an x'_0 . We choose x'_0 so that $f'(x'_0)$ is *not* small and set $a_0 = f'(x'_0)^{-1}$. We never invert f' in this method. Essentially we are finding a sequence a_k that converges to f'^{-1} at the root.

•The new method is superior to the Newton method when $f'(x_k)$ is small for some x_k . An example would be to choose $f(x) = x - x^2 + \frac{x^3}{3}, f'(x) = (x-1)^2$. Then the Newton methods will converge only slowly if x_0 is near 1 . For example with $x_0 = 1.15, x'_0 = 2.0$, we will get the sequence

$$1.15, -13.7148, -8.81, -5.5, -3.3, -1. - 9, -.99, -.41, -.1, -0.01, 0.00$$

or

$$2.0, 1.33, -1.78, -0.895, -0.36, -0.085, -0.01, 0$$

by the Newton method and

$$1.15, 0.82, 0.16, -0.01, -0.00$$

by the Kolmogorov method. The first step in the Newton method throws the point far away and then it has to find its way back. In the Kolmogorov method, even when f' vanishes, the corrections remain small at each step.

We can even choose the initial point to be x_0 at which Newtons method will fail but the new method will converge. However Newton's method is simpler and more robust: it is less sensitive to bad choices of initial condition.

- *Proof of Convergence* If we define the 'errors' $y_k = x_k - \alpha$, $z_k = 1 - f'(x_k)a_k$, we get

$$0 = f(\alpha) = f(x_k) - f'(x_k)y_k + \mathrm{O}(y_k^2), \quad z_{k+1} = [1 - a_k f'(x_{k+1})]^2.$$

so that

$$\begin{aligned} y_{k+1} &= y_k - a_k f'(x_k)y_k + \mathrm{O}\,(y_k^2) = (1 - a_k f'(x_k))y_k + \mathrm{O}\,(y_k^2) \\ \Rightarrow |y_{k+1}| &\leq |z_k y_k| + c_1(y_k^2), \end{aligned}$$

Also,

$$\begin{aligned} z_{k+1} &= 1 - a_{k+1}f'(x_{k+1}) = [1 - a_k f'(x_{k+1})]^2. \\ 1 - a_k f'(x_{k+1}) &= z_k - a_k[f'(x_{k+1}) - f'(x_k)] = z_k - a_k f''(y_{k+1} - y_k) = \mathrm{O}\,(|y_k| + |z_k|). \end{aligned}$$

Then the error $\delta_k = \sqrt{[|z_k|^2 + |y_k|^2]}$ satisfies $\delta_{k+1} \leq c_3\delta_k^2$. This leads to quadratic rate of convergence, the same as Newton's.

Appendix: Irrational Numbers

•A *rational number* is a ratio of integers; there is a unique reprsentation of any rational number as a ratio of co-prime integers, $r = \frac{p}{q}$.

•An *algebraic number* is a number that is the solution of an algebraic equation with rational coefficients. The lowest order of such an equation satisfied by x equation is called the order of x . For example $\sqrt{2}$ is an algebraic number of order 2 . The set of algebraic numbers, like the set of rational numbers, is countably infinite.

•A transcendental number is one that does not satisfy an algebraic equation with rational coefficients of any order; i.e, it is not algebraic. Examples of transcendental numbers are are e, π . All except a countable number of real numbers are transcendental.

•Any real number can be approximated arbitrarily closely by a sequence of rational numbers: the subset of rational numbers is dense in the real line. One such sequence can be obtained by the continued fraction method. Let a_0 be the largest integer less than or equal to x . Then $x_1 = (x - a_0)^{-1} > 1$ and let a_1 be the largest integer less than or equal to x_1 and so on: $x_{r+1} = \frac{1}{x_r - a_r}$, a_r the largest integer smaller than x_r . Then we have a sequence

$$\frac{h_n}{k_n} = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \cdots \frac{1}{a_n}}}$$

that converges to x .

•There is a precise notion of rapidity of convergence. The largest number n for which there are an infinite number of co-prime integers satisfying

$$|kx - h| < \frac{c}{k^n}$$

is called the *degree* of x . Rational numbers are of degree 1 , quadratic numbers of degree 2 and so on: the degree is just the order of the algebraic equation the number satisfies. In particular, if n is infinite (the inequality is satisfied for all exponents n) the number is transcendental.

•In mechanics it is interesting to consider the set of real numbers satisfying

$$||k_1 + k_2\omega|| > C[|k_1| + |k_2|]^{-\tau}$$

for all integers k_1, k_2 , but given C, τ . This is a set of finite measure; i.e., a finite fraction of all real numbers satisfy this condition. Clearly this

condition is the opposite of the condition of convergence that we looked at earlier.

Chapter 17

Geodesics on a Riemann Surface

See D. A. Hejhal, *The Selberg Trace Formula for* $PSL_2(R)$, Springer NY 1976.

•Geodesics on a Riemannian manifold (M, g) provide a geometrically inspired mechanical system. Given a curve $x : [0, T] \to M$, its *action* is defined to be

$$S = \int_0^T g(\dot{x}, \dot{x})dt.$$

The extrema of this functional, with fixed endpoints are the geodesics connecting these endpoints.

•The phase space of this system is the co-tangent bundle of the manifold. The velocity is just the tangent vector to the curve and momentum the corresponding one-form. The hamiltonian is just the square of this one-form. The Hamilton-Jacobi equations are

$$g^{ij}\partial_i S \partial_j S = E.$$

•There is a corresponding quantum system, whose Hilbert space is $L^2(M)$ and the hamiltonian is the Laplace operator. The above H-J equation is the classical approximation to the Schrödinger equation

$$-\hbar^2 \Delta \psi = E\psi$$

with $\psi \sim e^{\frac{i}{\hbar}S}$.

•For a compact manifold, the spectrum of the Laplace operator is discrete: this is the *quantum spectrum* of the manifold. There is usually just one closed geodesic in each conjugacy class of the fundamental group. Thus there is a *classical spectrum*; the lengths of the closed geodesics. This is a function from the set of conjugacy classes of the fundamental group to the positive real numbers. There is a deep relation between these two spectra in the semi-classical approximation: the *Gutzwiller trace formula*.

•It is convenient to introduce generating functions that capture the information in the spectrum. Physically the most natural is the *quantum partition function*[1]

$$Z_\Delta(\beta) = \sum_n e^{-\lambda_n \beta} = \frac{1}{\mathrm{Vol}(M)} Tr e^{\Delta\beta}.$$

Its Mellin transform is more convenient for analytic arguments:

$$\zeta_\Delta(s) = \sum_n \frac{1}{\lambda_n^s}.$$

•The simplest example is a circle $S^1 = R/2\pi Z$ .The geodesics are just the images of straightlines. They are all closed. There is one closed geodesic in each homotopy class, labelled by an integer, the *winding number*. The length (action) of this geodesic is just $2\pi n$. eigenfunctions of the Laplace operator are

$$\psi_n(\theta) = \frac{1}{\sqrt{[2\pi]}} e^{in\theta}, \quad -\Delta\psi_n = n^2\psi_n, \quad n \in Z.$$

•In the case of the circle we can see the relation between the classical and quantum spectra through the *Poisson summation formula*. There are two ways to solve the Heat equation on the circle:

$$-\Delta h_\beta(\theta) = \frac{\partial}{\partial\beta} h_\beta, \quad h_0(\theta) = \delta(\theta).$$

•By Fourier analysis we get

$$h_\beta(\theta) = \sum_{n\in Z} e^{-n^2\beta} e^{in\theta}.$$

[1]If an eigenvalue is degenerate, we count it with its multiplicity.

•The solution for the heat equation on the real line is

$$h^R_\beta(x) = \frac{e^{-\frac{x^2}{4\beta}}}{\sqrt{[4\pi\beta]}}.$$

We can get the solution on the circle by summing over all points equivalent to $x : x + 2\pi n$. This makes sense from the physical interpretation of h_β as the probability of diffusion of a particle from the origin to x . We sum over all points in the real line that corresponds to the same point on the circle. It is also easy to verify that this is the solution to the heat equation using the relation between the delta function on the circle to that on the real line.

$$\delta(\theta) = \sum_{n\in Z} \delta_R(\theta + 2\pi n).$$

This is the *Poisson sumamtion formula*.

•Thus we have

$$h_\beta(\theta) = \sum_{n\in Z} \frac{e^{-\frac{[\theta+2\pi n]^2}{4\beta}}}{\sqrt{[4\pi\beta]}} = \sum_{n\in Z} e^{-n^2\beta} e^{in\theta}$$

If we put $\theta = 0$, we get the partition function

$$Z_\Delta(\beta) = \sum_{n\in Z} \frac{e^{-\frac{[2\pi n]^2}{4\beta}}}{\sqrt{[4\pi\beta]}} = \sum_{n\in Z} e^{-n^2\beta}.$$

•The exponent of the first version involves the square of the length of the geodesics. In the limit of small β (the semi-classical limit) the denominator is less important (it is lower order).

•This suggests a general approximate formula

$$\sum_{g\in\widehat{\pi_1(M)}} e^{\frac{-l^2(g)}{4\beta}} \sim \sum_{\lambda\in \mathrm{Spec}\Delta} e^{-\lambda\beta}.$$

The sum on the left is over conjugacy classes of the fundamental group; that on the left is over the spectrum of the Laplace operator. A more precise version of this is the Gutzwiller trace formula.

•To really understand the story we must consider a case where the fundamental group is non-abelian. We look for a case where the heat kernel can be determined exactly in the universal covering space, so that we can a formula valid beyond the semi-classical approximation. Selberg derived such a deep relation, motivated by ideas from number theory. The simplest case of his work is the trace formula for Riemann surfaces.

•A *Riemann surface* is a compact two dimensional manifold. The simplest example is a sphere. If we attach a 'handle' to a sphere we get a torus. Attaching more handles will produce surfaces of greater complexity; the number of handles is called the genus. A compact two dimensional manifold is determined by its genus.

•Riemann surfaces admit metrics of constant curvature. On the sphere this is a positive curvature metric while on the torus it is of zero curvature. The remaining surfaces admit metrics of negative curvature.

•Recall that positive curvature means that two geodesics that start with slightly different initial conditions will tend to converge. For example geodesics on a sphere starting at the same point eventually meet at the anti-podal point.

•If the metric is flat the geodesics neither converge nor diverge. A torus can be thought of a parallelogram with opposite sides identified. The geodesics are straightlines, except for this identification. Although the metric is flat, quite complicated behavior can result: a geodesic can be dense.

•Both of the above examples can be thought of as integrable dynamical systems. The parameter (arc-length) along the curve can be thought of as 'time' and the tangent vector as the 'velocity'.

•The geodesics on a Riemann surface of negative curvature then is an example of a non-integrable dynamical system.Indeed it is the opposite extreme from being integrable, it is an ergodic system.

•Let us go more into details. The basic reference is M. Gutzwiller *Chaos in Classical and Quantum Mechanics*.

•The simplest metric of negative curvature is the Poincare metric on the upper half plane.

$$ds^2 = \frac{dx^2 + dy^2}{y^2}.$$

Its geodesics are circles orthogonal to the real line. •The formula for curvature of a Riemann metric simplifies in the case of a two-dimensional metric $ds^2 = e^{\phi}[dx^2 + dy^2]$ to $R = e^{-\phi}[\phi_{xx} + \phi_{yy}]$. In our case, $R = -1$.

•This metric has an isometry group

$$z \mapsto \frac{az+b}{cz+d}, \quad z = x+iy, \quad a,b,c,d \in R.$$

For example, $x \mapsto x+b, y \mapsto y$ is obviously an isometry. We can take a discrete subgroup of the isometry group and quotient the upper half plane, to produce a compact manifold of constant curvature. This is a non-abelian analogue of the way a torus is constructed as a quotient of the plane by a lattice.

•Any element of $g \in PSL_2(R)$ can be brought to the 'normal form'

$$g : z \mapsto \omega, \frac{\omega-\xi}{\omega-\eta} = N(g)\frac{z-\xi}{z-\eta}.$$

The quantity $N(g)$ is called the multiplier. It is related to the trace : $N(g)^{\frac{1}{2}} + N(g)^{-\frac{1}{2}} = \mathrm{tr} g$. The elements of $PSL_2(R)$ fall into three categories: *elliptic, parabolic* and *hyperbolic* according as whether the trace is of magnitude less than 2, equal to two or greater than 2. The ellptic elements have a fixed point, which can be made to be i by a conjugation. The parabolic elements have a fixed point either at infinity on the real line: upto a conjugation a parabolic element is a translation. A hyperbolic element can be brought to a scaling $z \mapsto \lambda z$ by a conjugation. The distance from z to $h(z)$ is $\log N(h)$, for a hyperbolic element.

•In order that the quotient U/G be a manifold, G must act properly discontinuously on U . This means that every point must have a neighborhood V such that $gV \cap V = \phi$ for all $g \neq 1$. In particular there should be no fixed points or limit points for the group action.

•Hence if a discrete group G acts properly discontinuously it must consist entirely of hyperbolic elements. Such discrete groups can be classified. For each integer $g \geq 2$ there is a subgroup $G_g \subset PSL_2(R)$ generated by hyperbolic elements $A_1, \cdots A_g, B_1, \cdots B_g$, satisfying[2]

$$\prod_i^g [A_i, B_i] = 1.$$

•The manifold U/G is smooth and inherits a metric of constant negative curvature from the Poincare metric on U . To each conjugacy class in the

[2]The group commutator is defined by $[A,B] = ABA^{-1}B^{-1}$.

fundamental group G there is a closed geodesic; all closed geodesics arise this way. A change of γ within a conjugacy class simply yields the same geodesic but with a different starting point. Its length is given by $l(\gamma) = \log N(\gamma)$, in terms of the multiplier. (This can be checked by considering geodesics starting at i).

•An element of G could be the power of a 'primitive' element γ_0 , one that is the not a non-trivial power of ny other element.

•We can now state the Selberg trace formula:

$$\mathrm{tr} e^{[\Delta+\frac{1}{4}]\beta} = \frac{\mu(F)}{4\pi}\int_{-\infty}^{\infty} e^{-\beta k^2} k \tanh k \;\; dk + \sum_{\gamma\in[G]} \frac{l(\gamma_0)}{2\sinh\frac{l(\gamma)}{2}} \frac{e^{-\frac{l^2(\gamma)}{4\beta}}}{\sqrt{[4\pi\beta]}}.$$

Here the sum over the set of all conjugacy classes (except the identity) $[G]$ of the fundamental group.

•The integral in the first term is just the answer in the upper half plane, per unit volume. The remaining sum invloves the lengths of the closed geodesics of the Riemann surface. The normalization factor in the second terms are not important in the small β limit (semi-classical limit) but are necessary to get an exact answer.

•Let us indicate a proof of this formula. Let $K(z', z)$ be the integral kernel of the operator $e^{\beta[\Delta+\frac{1}{4}]}$ on Σ and $k(z', z)$ the corresponding quantity on the covering space U . Then

$$\mathrm{tr} e^{\beta[\Delta+\frac{1}{4}]} = \sum_{\gamma\in G}\int_F k(\gamma z, z) d\mu(z).$$

Any element can be written as $\gamma = \sigma\tilde{\gamma}\sigma^{-1}$, where $\tilde{\gamma}$ labels a conjugacy class: for example it can be chosen to be diagonal. This σ is unique upto right multiplication by elements that commute with γ ; i.e., elements in the centralizer of γ . We now split the sum over G as a sum over the set of conjugacy classes and the sum over $G/Z(\tilde{\gamma})$, where $Z(\tilde{\gamma})$ is the centralizer of $\tilde{\gamma}$.

•Thus we get[3]

$$\mathrm{tr} e^{\beta[\Delta+\frac{1}{4}]} \;\; = \;\; \int_F k(z,z) d\mu(z) + \sum_{\tilde{\gamma}\in\tilde{G}} \sum_{\sigma\in G/Z(\tilde{\gamma})} \int_F k(\sigma\tilde{\gamma}\sigma^{-1} z, z) d\mu(z)$$

[3] $\tilde{G}$ denotes the conjugacy classes of G excluding the identity. Also, $\Phi(0) = k(z,z)$ is a independent of the point z .

$$\begin{aligned} &= \mu(F)\Phi(0) + \sum_{\tilde{\gamma}\in\tilde{G}} \sum_{\sigma\in G/Z(\tilde{\gamma})} \int_{\sigma(F)} k(\tilde{\gamma}z, z) d\mu(z) \\ &= \mu(F)\Phi(0) + \sum_{\tilde{\gamma}\in\tilde{G}} \int_{F(Z(\tilde{\gamma}))} k(\tilde{\gamma}z, z) d\mu(z). \end{aligned}$$

The first equality follows by change of variables $z \mapsto \gamma z$ and the invariance of k under G. In the second, $F(Z(\tilde{\gamma})) = \bigcup_{\sigma\in G/Z(\tilde{\gamma})} \sigma(F)$ is the disjoint union of copies of F under the action of the various σ 's. A moment's thought will show that this is just the fundamental region of the upper half plane under the action of the group $Z(\tilde{\gamma})$.

•It is easy to determine this centralizer. It is the set of all matrices that commute with the diagonal matrices $\tilde{\gamma}$. Let γ_0 be the primitive of γ : i.e., $\gamma = \gamma_0^m$ for some integer and γ_0 itself is not a power of any other element. Then $Z(\tilde{\gamma})$ is the infinite cyclic group generated by γ_0 : γ_0^n are the only elements that commute with γ_0^m. (We are using the fact that these are hyperbolic). It acts on U is a simple way: $z \mapsto N(\gamma_0)z$. A fundamental region of $Z(\tilde{\gamma})$ is the strip $\{(x, y)| < 1 <\leq N(\gamma_0)$.

•Thus we have a more explicit formula

$$\mathrm{tr} e^{\beta[\Delta+\frac{1}{4}]} = \sum_{\tilde{\gamma}\in\tilde{G}} \int_{-\infty}^{\infty} dx \int_1^{N(\gamma_0)} k(N(\gamma_0)z, z) \frac{dy}{y^2}.$$

•Next we have to determine this integral $E(N, N_0) = \int_{-\infty}^{\infty} dx \int_1^{N_0} k(Nz, z) \frac{dy}{y^2}$ over the heat kernel of the upper half plane. Being a function invariant under action of the isometry group, $k(z', z)$ can only depend on the points through the combination $\frac{|z-z'|^2}{y'y}$. Hence there is a function of one variable Φ such that

$$k(z', z) = \Phi(z', z).$$

Putting this in and evaluating the integrals we get

$$E(N, N_0) = \frac{\log N_0}{N^{\frac{1}{2}} - N^{-\frac{1}{2}}} g(\log N)$$

where

$$g(u) = \int_x^{\infty} \frac{\Phi(t)dt}{\sqrt{[t-x]}}, \quad x = e^u - e^{-u} - 2.$$

•The fractional integral equation above for Φ can be solved to give

$$\Phi(t) = -\frac{1}{\pi}\int_t^\infty \frac{dQ(x)}{\sqrt{[x-t]}}, \quad Q(e^u - e^{-u} - 2) = g(u).$$

In particular

$$\Phi(0) = -\frac{1}{\pi}\int_0^\infty \frac{dg(u)}{e^{\frac{u}{2}} - e^{-\frac{u}{2}}} = \frac{1}{2\pi}\int_0^\infty h(k)\; k \tanh \pi k \; dk$$

where

$$h(k) = \int_{-\infty}^\infty g(u) e^{iku} du.$$

•Thus we have the formula

$$\mathrm{tr} e^{[\Delta + \frac{1}{4}]\beta} = \frac{\mu(F)}{4\pi}\int_{-\infty}^\infty h(k)\; k \tanh k \;\; dk + \sum_{\gamma \in [G]} \frac{l(\gamma_0)}{2\sinh \frac{l(\gamma)}{2}} g(l(\gamma_0)).$$

All that remains is to determine $g(u)$ and $h(u)$.

•For this we note that

$$\Delta y^s = s(s-1)y^s, \quad \Delta = y^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right).$$

Thus

$$\int k(z', z) y^s d\mu(z) = e^{\beta[s(s-1)+\frac{1}{4}]} y'^s.$$

•Now put $y' = 1, s = \frac{1}{2} + ik$ in the above equation and express $k(i, z)$ in terms of Φ . Evaluating the integrals we will get (using previous relation between Φ and g)

$$e^{-\beta k^2} = \int_{-\infty}^\infty g(u) e^{iku} du = h(u).$$

This completes the proof of the Selberg trace formula.

PHY 411 Advanced Classical Mechanics (Chaos)
Problem set 1 Spring 2000
Due Jan 24 2000

Please put your solutions in my mail box by the end of the day on Monday.

1. A pair of bodies of masses m_1 and m_2 move in each other's gravitational field.

(i) Find all (i.e., a complete set) the constants of motion of the system.

(ii) Derive the three laws of Kepler on planetary motion.

2. In the two body problem show that a collision occurs if and only if the angular momentum in the center of mass reference frame vanishes. Find the behavior of the position as a function of time in the immediate neighborhood of the time of collision.

3. Consider the differential equation

$$\frac{dx}{dt} = Ax$$

where $x : R \to R^2$ is a vector valued function and A is a constant 2×2 matrix. What is the condition on A in order that all the solutions $x(t)$ remain bounded for all real values of t ?

PHY 411 Advanced Classical Mechanics (Chaos)
Problem set 2 Spring 2000
Due Jan 31 2000

Please put your solutions in my mail box by the end of the day on Monday.

4. Derive the equations of motion of a spherical pendulum; i.e, a particle of mass m suspended at the end of a rigid rod of length R moving in the Earth's gravitational field.

5. Derive the exact solution of the pendulum above when it is constrained to move in a vertical plane. Show that the solution is doubly periodic, with one real period and one imaginary period.

6. Derive the conservation law due to rotation invariance around the vertical axis in the spherical pendulum.

7.* A pendulum with natural frequency ω is suspended from a support which is itself oscillating in the vertical plane with frequency ν. Show that, in the limit $\nu >> \omega$, the pendulum can be stable with its mass above the point of suspension.

The problems with a star are probably harder.

PHY 411 Advanced Classical Mechanics (Chaos)
Problem set 3 Spring 2000
Due Feb 7 2000

Please put your solutions in my mail box by the end of the day on Monday.

8. Find the normal modes of oscillations of three equal masses moving in a plane, connected to each other by springs of equal strength.

9. Find the difference in times in going from $x = -\infty$ to $x = \infty$ between a free particle of energy E and the same particle under the influence of a potential

$$V(x) = \frac{A}{\cosh^2 x}.$$

10*. Three bodies of masses M, m, μ are moving in each others gravitational fields. (Let us think of them as the Earth, the Moon and a satellite.) They may be all be assumed to move in the same plane.

(i) Obtain the Lagrangian of this system.

(ii) Assume that the satellite is of infinitesimal mass: $\mu << m, M$. Solve for the motion of the Earth and the Moon ignoring the effect of the satellite.

(iii) Now obtain the Lagrangian for the motion of the satellite in the gravitational field produced by the Earth and the Moon.

This is the *restricted three body problem* of Lagrange.

PHY 411 Advanced Classical Mechanics (Chaos)

Problem set 4 Spring 2000

Due Feb 21 2000

11(i) Show that

$$\det e^A = e^{\mathrm{tr}\, A}$$

for any matrix.

(ii) For a matrix valued function $g(t)$ show that

$$\frac{d}{dt} g^{-1}(t) = -g^{-1}(t) \frac{dg(t)}{dt} g^{-1}(t).$$

12(i) The anharmonic oscillator has hamiltonian

$$H = p^2 + \lambda x^4 + \omega^2 x^2.$$

Find the curves in the (x, p) plane that correspond to the orbits. Find the area $I(E)$ enclosed by each orbit as a function of the energy.

(ii) Show that the area $I(E)$ is the action variable of this system; find the 'angle' variable canonically conjugate to it.

(iii) Show that the period of the orbit is given by the derivative of the action with respect to energy.

13(i) Find the equations of motion of the three-body problem with gravitational forces. The masses are not necessarily equal, or infinitesimal. Assume that all three bodies move in the same plane, and choose the center of mass as the origin of the co-ordinate system.

(ii) ** Show that there is an exact solution (!) where the three bodies are on the vertices of an equilateral triangle, which is rotating around the center of mass at a constant angular velocity. (Lagrange's solution). Find the relation between the length of the side of this triangle and the frequency of its rotation.

PHY 411 Advanced Classical Mechanics (Chaos)
Problem set 5 Spring 2000
Due Feb 28 2000

14. Solve the Kepler problem by separation of variables of the Hamilton-Jacobi equation. (Find the orbit, not just the solution of the H-J equation.)

15. Consider a particle in a hyperbolic orbit in the Kepler problem. Find the time delay $\Delta T(E, L)$ of the orbit as a function of energy and angular momentum. Also the difference $\Delta W(E, L)$ between the action of the orbit and that of the free particle. What is the relatonship between these two quantities?

16. Consider the restricted three-body problem of celestial mechanics; i.e., one of the bodies is of infinitesimally small mass. Assume that the two large bodies are in a circular orbit around the center of mass. Write the equation of motion of the small body in a co-ordinate system that rotates with the massive bodies.

(i) Find the five static solutions of Lagrange.

(ii) Show that two of them are stable with respect to infinitesimal perturbations.

(iii) Find the position of the two stable Lagrange points (L_4 and L_5) for the Earth-Moon system.

PHY 411 Advanced Classical Mechanics (Chaos)
Problem set 6 Spring 2000
Due Mar 13 2000

17. Define the function $f(z) = \int_0^\infty \frac{e^{-t}}{1-tz} dt$; the integral converges for $\arg(z) > 0$.

(i) Find an asymptotic expansion for $f(z)$ valid in the above wedge.

(ii) Numerically calculate (plot or tabulate) the error in $f(z)$ for small values of $|z|$ such as $z = -0.1$ keeping $k = 0, 1, \cdots 25$ terms in the sum.

(iii) Give an estimate of the optimum number of terms to keep for small values of $|z|$.

18. Let G be the set of formal power series of the form $z + \sum_2^\infty f_n z^n$. Define an operation $f \circ g(z) = f(g(z))$.

(i) Show that $f \circ g$ is also a formal power series.

(ii) Show that every power series f in G has an inverse g in G ; i.e., $f(g(z)) = z$.

(iii) * Thus G is a group. Is there a corresponding Lie algebra of power series?. If so find its commutation relations.

19. (i) Determine the Birkhoff normal form for two coupled oscillators with frequencies in irrational ratio ω :

$$H = \frac{1}{2}(p_1^2 + q_1^2) + \frac{1}{2}\omega(p_2^2 + q_2^2) + \lambda(q_1^2 + q_2^2)^2.$$

At a minimum, find the answer to two orders in λ .

(ii) Determine a formal power series representing a conserved quantity other than the hamiltonian. (Again, at least to second order in λ .)

PHY 411 Advanced Classical Mechanics (Chaos)

Problem set 7 Spring 2000

Due Apr 3 2000

20. (i) Use Newton's method to solve numerically the transcendental equation $\cos\frac{x}{2} = 0$ with an accuracy of 10 decimal places.

(ii) Solve the equation $\sin\frac{x}{2} - g = 0$ for small values of g by Newton's method and compare with the power series of $2\arcsin(g)$. The initial point $x_0 = 0$. What is the region of convergence of each method?.

21. Let $f(w, g)$ be an analytic function of two variables, such that $f(0, 0) = 0$. Show that for sufficiently small g , the Newton method for solving $f(w, g) = 0$ converges with the initial choice $w_0 = 0$. Estimate the radius of convergence in terms of the magnitudes of the derivatives of f at the origin and $|f'(0, 0)^{-1}|$.

22. Consider a sequence of points defined by the recursion relation

$$I_{n+1} = I_n + K\sin\theta_n, \quad \theta_{n+1} = \theta_n + I_{n+1}$$

where K is a constant. Both θ and I are thought of as periodic with period 2π .

(i) Show that the map $(\theta_n, I_n) \mapsto (\theta_{n+1}, I_{n+1})$ prserves the area.

(ii) Find the invariant curves when $K = 0$.

(iii) What is the condition on K in order that the fixed point at $\theta = 0$ is stable?.

23* It can be shown that the KAM circles of the standard map are given by the difference equation

$$q(\theta + \omega) - 2q(\theta) + q(\theta - \omega) = K\sin[q(\theta)].$$

(i) For small values of K find a KAM torus by approximately solving this equation.

Hint Turn it into a differential equation.

(ii) For $\omega = \frac{1+\sqrt{5}}{2}$, find a value of K at which there is an invariant torus, by some suitable approximation method.

Hint Find a variational principle for K .