Hamiltonian systems

Marc R. Roussel

October 25, 2005

1 Introduction

Today's notes will deviate somewhat from the main line of lectures to introduce an important class of dynamical systems which were first studied in mechanics, namely Hamiltonian systems. There is a large literature on Hamiltonian systems. The intention here is not to comprehensively survey this literature, which would be quite impossible even if we devoted a whole course to the subject, but to discuss some of the key ideas underlying the treatment of this class of systems.

We begin with some definitions:

Definition 1 *Suppose that $H(\mathbf{x},\mathbf{p})$ is a smooth function of its arguments for $\mathbf{x}$ and $\mathbf{p} \in \mathbb{R}^n$. Then the dynamical system*

$$\dot{x}_i = \frac{\partial H}{\partial p_i}, \tag{1a}$$

$$\dot{p}_i = -\frac{\partial H}{\partial x_i} \tag{1b}$$

*($i = 1,2,\ldots,n$) is called a **Hamiltonian system** and H is the **Hamiltonian function** (or just the Hamiltonian) of the system. Equations 1 are called **Hamilton's equations**.*

Definition 2 *The number of **degrees of freedom** of a Hamiltonian system is the number of (x_i, p_i) pairs in Hamilton's equations, i.e. the value of n. The phase space is therefore 2n-dimensional.*

In mechanics, the vector $\mathbf{x}$ represents the generalized coordinates of the components of the system (positions, angles, etc.), while $\mathbf{p}$ is a set of generalized momenta. There is an elaborate theory for constructing generalized momenta and Hamiltonians which we won't go into here. Note that the generalized momenta are not always ordinary linear momenta (mv_i), although that is often the case.

Note that the Hamiltonian function is a contant of the motion:

$$\begin{aligned}\dot{H} &= \sum_{i=1}^{n} \frac{\partial H}{\partial x_i}\dot{x}_i + \sum_{i=1}^{n} \frac{\partial H}{\partial p_i}\dot{p}_i \\ &= \sum_{i=1}^{n} \frac{\partial H}{\partial x_i}\frac{\partial H}{\partial p_i} + \sum_{i=1}^{n} \frac{\partial H}{\partial p_i}\left(-\frac{\partial H}{\partial x_i}\right) = 0.\end{aligned}$$

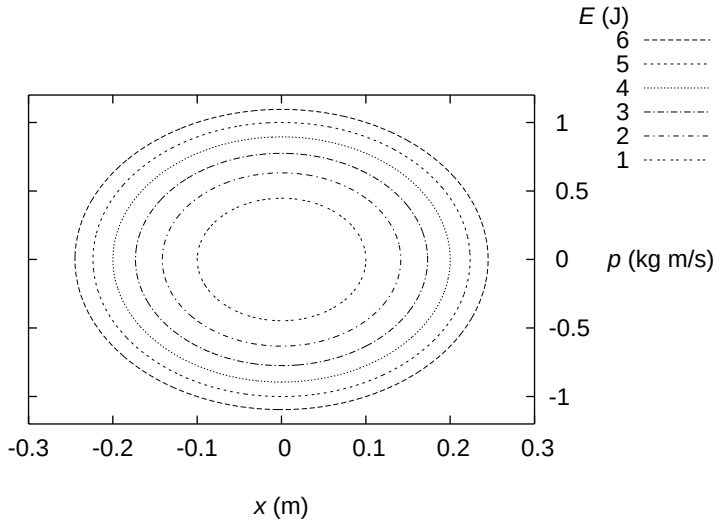

Figure 1: Harmonic oscillator orbits for $k = 200\,\mathrm{N/m}$, $m = 0.1\,\mathrm{kg}$, and various values of the total energy E. Maple code which can be used to draw such a figure is given in Appendix A. (I actually used gnuplot because I think the output looks a little nicer, but the principle is the same.)
Note that these closed orbits are *not* limit cycles since different initial conditions (which in general lead to different values of E) will produce different orbits.

In fact, the Hamiltonian is often just the total energy in mechanical systems, although this isn't always the case.

Let us for the moment specialize the discussion to planar systems, i.e. systems for which $n = 1$. The fact that H is constant is means that the motion is constrained to the curve $H(x, p) = h$, where h is the value of the Hamiltonian function implied by the initial conditions. This curve is of course just the orbit of the system in phase space. If we change the total energy, we get a different orbit.

Example 1.1 A harmonic oscillator is a mass-spring system with potential energy $\frac{1}{2}kx^2$, where x is the displacement of the spring from equilibrium. For simple systems like this one in which the potential energy simply depends on the position, the Hamiltonian is just the total energy:

$$H(x, p) = \frac{1}{2}kx^2 + \frac{p^2}{2m}, \tag{2}$$

where p is the momentum. Because H is a constant, the orbits are just the family of ellipses

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

The value of E is fixed by the initial conditions. Different values of E correspond to ellipses of different sizes. Figure 1 shows an example.

If we are interested in the equations of motion, we can recover them from Hamilton's equations:

$$\dot{x} = \frac{\partial H}{\partial p} = p/m. \tag{3a}$$

$$\dot{p} = -\frac{\partial H}{\partial x} = -kx. \tag{3b}$$

Since $v = p/m$, the first equation just says that $\dot{x}$ is the velocity of the particle. Recall that $F = \dot{p}$. (We usually write this as $F = ma$ in elementary physics courses.) Thus, the second equation gives us the force law corresponding to the potential energy $\frac{1}{2}kx^2$, which you may recognize from your earlier studies of physics.

As you well know from mechanics, mechanical systems are subject to other conservation laws: momentum, angular momentum, and so on. If we can find enough conservation laws for a system then, as in the example above, the business of determining the behavior of the system becomes rather simple. On the other hand, if we have more degrees of freedom than constants, the behavior can be complex. These issues are discussed briefly in the following sections.

2 Integrable systems

We start with some specialized machinery and definitions pertaining to Hamiltonian systems.

Definition 3 *Let $H(\mathbf{x},\mathbf{p})$ and $L(\mathbf{x},\mathbf{p})$ be differentiable functions of their arguments for $\mathbf{x}$ and $\mathbf{p} \in \mathbb{R}^n$. The **Poisson bracket** of H with L, $\{H,L\}$, is defined by*

$$\{H,L\} = \sum_{i=1}^{n} \left(\frac{\partial H}{\partial p_i}\frac{\partial L}{\partial x_i} - \frac{\partial H}{\partial x_i}\frac{\partial L}{\partial p_i} \right).$$

Definition 4 *A quantity L is called a **first integral** of a Hamiltonian system if it is a constant of the motion, i.e. if $\dot{L} = 0$ under the flow implied by Hamilton's equations.*

Note that the Hamiltonian itself is a first integral according to this definition.

Theorem 1 *The quantity L is a first integral of a Hamiltonian system with Hamiltonian H if $\{H,L\} = 0$.*

The proof of this proposition is similar to the proof that H is a constant, and is left as an exercise to the reader.

Definition 5 *A Hamiltonian system is said to be **completely integrable** if it has n first integrals (including the Hamiltonian itself), where n is the number of degrees of freedom.*

In mechanical systems, the first integrals are often familiar quantities.

Example 2.1 Consider the two-body orbital problem with $m_1 \gg m_2$. In this case, we can treat the body of mass m_1 as being essentially motionless.[1] Moreover, with just two bodies, the coordinate system can always be chosen in such a way that m_1 is at the origin, and that $z = 0$, i.e. the motion is confined to the (x,y) plane. The Hamiltonian is

$$H = \frac{p_x^2 + p_y^2}{2m_2} - \frac{Gm_1m_2}{\sqrt{x^2+y^2}}.$$

Based on our knowledge of mechanics, we expect the angular momentum

$$L = \mathbf{r} \times \mathbf{p} = xp_y - yp_x$$

to be a constant of the motion. If this is so, the Poisson bracket of H with L should be zero:

$$\begin{aligned} \{H,L\} &= \frac{\partial H}{\partial p_x}\frac{\partial L}{\partial x} - \frac{\partial H}{\partial x}\frac{\partial L}{\partial p_x} + \frac{\partial H}{\partial p_y}\frac{\partial L}{\partial y} - \frac{\partial H}{\partial y}\frac{\partial L}{\partial p_y} \\ &= \frac{p_x}{m_2}p_y - \frac{Gm_1m_2x}{(x^2+y^2)^{3/2}}(-y) + \frac{p_y}{m_2}(-p_x) - \frac{Gm_1m_2y}{(x^2+y^2)^{3/2}}x \\ &= 0. \end{aligned}$$

The angular momentum is therefore, as expected, a constant of the motion. This means that the system is confined to the intersection of the hyperplanes $H = E$ and $L = \ell$, where E and ℓ are constants determined by the initial conditions. This intersection is an invariant manifold of the differential equations since all possible trajectories for a given E and ℓ lie within it.

What does this manifold look like? We have two equations connecting our four variables, so these equations implicitly define a two-dimensional surface in the four-dimensional phase space. In this case, this surface is easy to plot using Maple. See Appendix B for sample code. We start by solving for y (say) from the angular momentum:

$$y = (xp_y - \ell)/p_x.$$

We then substitute this expression into the energy equation. Depending on the total energy, we get different kinds of results:

1. The surface is a torus if $E < 0$ (Fig. 2). The trajectories on the torus can either be periodic (which they are for this simple problem) or **quasiperiodic**, i.e. they wind around and around the torus without repeating themselves. In the latter case, a trajectory eventually covers the whole torus and is said to be **ergodic**. In the former case, slightly different initial conditions give different orbits.
2. The surface is essentially a cylinder if $E > 0$ (Fig. 3) or if $E = 0$ (Fig. 4). This corresponds to a case where the lighter particle has enough energy to escape the gravitational attraction of its heavier neighbor. The trajectories wind around this cylinder, either toward $x = \infty$ or toward $x = -\infty$.

[1] A more careful treatment would transform to centre-of-mass and relative coordinates. We would obtain essentially the same Hamiltonian, with a small adjustment in the constants appearing in the equation.

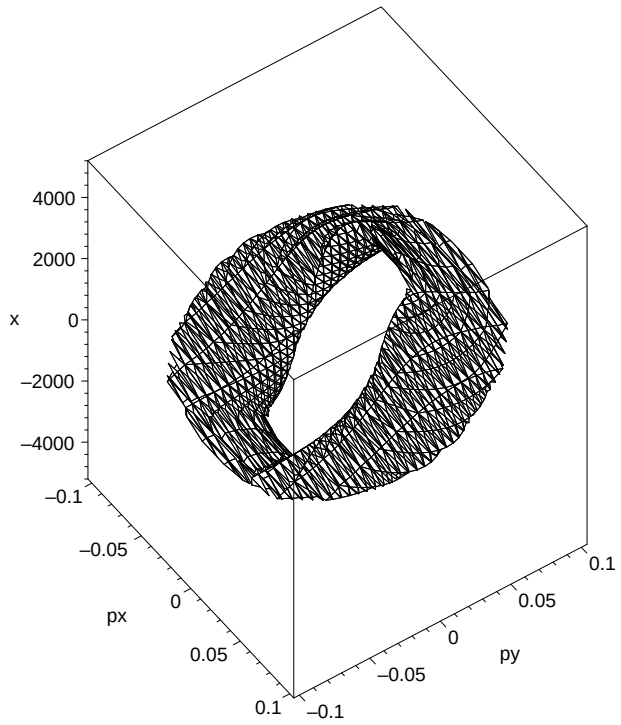

Figure 2: Invariant torus arising from the single-planet orbital problem. This case corresponds to $m_1 = 100\,000\,\mathrm{kg}$, $m_2 = 1000\,\mathrm{kg}$, $E = -1.67 \times 10^{-6}\,\mathrm{J}$, and $\ell = 100\,\mathrm{kg\,m^2/s}$.

The foregoing example brings out important features of integrable Hamiltonian problems: Invariant tori or cylinders are generically seen. The study of these invariant manifolds and how they change when we change the Hamiltonian is an important part of the theory of Hamiltonian dynamical systems. The study of trajectories on these manifolds is of course also interesting and important. In the case of integrable Hamiltonians, the behavior can't be much more complicated than a quasiperiodic trajectory.

3 Numerical integration

We won't go into great detail, but there is an important issue which we must discuss in relation to the numerical integration of Hamiltonian systems: Most numerical methods *do not* preserve the constancy of H. Many numerical methods do a good job of holding H reasonably constant for simple problems, but not for more complex problems, especially if long trajectories are wanted. Numerical integration methods which bound the variation of H are said to be **symplectic**. We will discuss this problem and the construction of symplectic integrators in this section. However, we will not undertake a full implementation, the details being rather messy.

As an introduction to this topic, let us discuss the simplest numerical integration method, namely the forward Euler method. The idea behind this algorithm is very simple: The differential

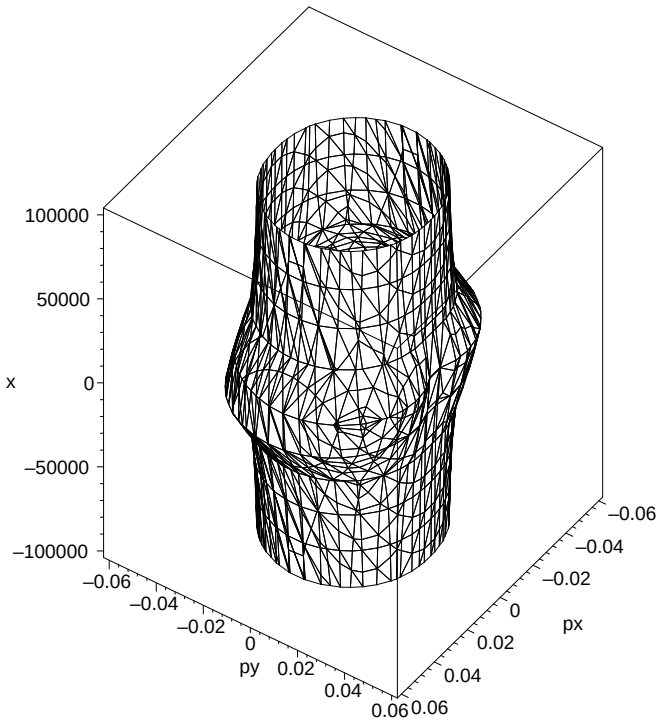

Figure 3: Invariant cylinder arising from the single-planet orbital problem. All parameters are as in Fig. 2, except $E = 5.3 \times 10^{-7}\,\mathrm{J}$ and $\ell = 120\,\mathrm{kg\,m^2/s}$.

equation

$$\frac{d\mathbf{z}}{dt} = \mathbf{f}(\mathbf{z}, t), \tag{4}$$

where $\mathbf{z}$ is the vector of phase-space coordinates, can be interpreted as follows, according to the basic definition of the derivative:

$$\lim_{\Delta t \to 0} \frac{\mathbf{z}(t+\Delta t) - \mathbf{z}(t)}{\Delta t} = \mathbf{f}(\mathbf{z}(t), t).$$

Euler's method simply consists in taking small values of Δt and doing away with the limit. Since Δt is fixed, we will be calculating $\mathbf{z}$ at following sequence of times: $\Delta t, 2\Delta t, 3\Delta t, \ldots, j\Delta t, \ldots$. We therefore label the values of $\mathbf{z}$ by j, the number of multiples of the step size Δt which have passed. Euler's method can thus be summarized by the formula

$$\mathbf{z}_{j+1} = \mathbf{z}_j + \Delta t \mathbf{f}(\mathbf{z}_j, j\Delta t). \tag{5}$$

Equation 5 defines a dynamical system with discrete time (indexed by j) in which the state at the next time increment is calculated from a simple formula. This is called a **map**. In the limit $\Delta t \to 0$, the Euler map converges on the solutions of the differential equation 4. Hopefully, for small values of Δt, Euler's method will give results which are at least reasonably representative of those of the original differential equation.

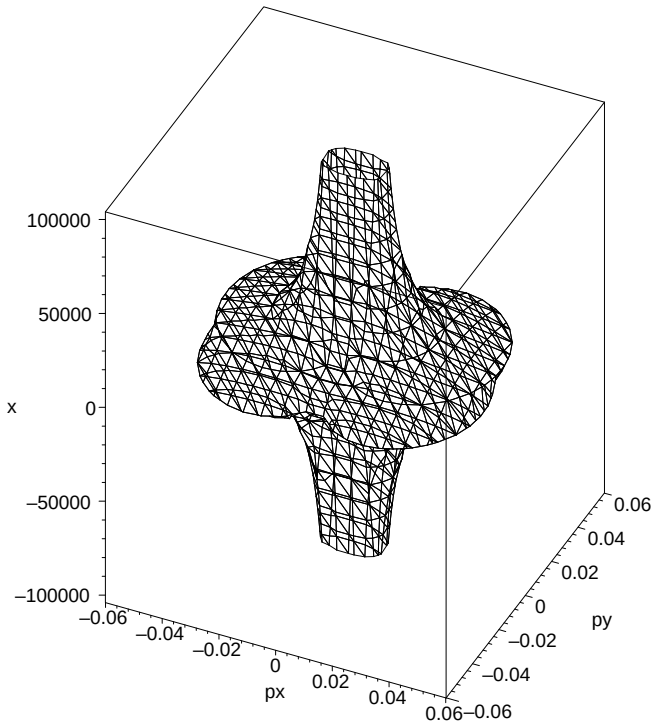

Figure 4: Invariant surface arising from the single-planet orbital problem with $E = 0$ and $\ell = 115\,\mathrm{kg\,m^2/s}$, all other parameters being set as in Fig. 2.

As you can imagine, Euler's method is a very naïve technique, and tends not to be very accurate. Nevertheless, it has its uses, among them pedagogical applications. We will use it here to illustrate the problem you run into with ordinary numerical methods and Hamiltonian systems.

Example 3.1 Suppose that we want to solve the differential equations 3 which arise from the Hamiltonian 2 by Euler's method. Then we have

$$\begin{aligned} x_{j+1} &= x_j + \Delta t\, p_j/m, \\ p_{j+1} &= p_j - \Delta t k x_j. \end{aligned}$$

The value of the Hamiltonian at time index $j+1$ is

$$\begin{aligned} H(x_{j+1}, p_{j+1}) &= \frac{1}{2}kx_{j+1}^2 + \frac{p_{j+1}^2}{2m} \\ &= \frac{k}{2}\left(x_j + \Delta t\, p_j/m\right)^2 + \frac{1}{2m}\left(p_j - \Delta t k x_j\right)^2 \\ &= \frac{1}{2}kx_j^2 + \frac{p_j^2}{2m} + \frac{k}{m}(\Delta t)^2\left(\frac{1}{2}kx_j^2 + \frac{p_j^2}{2m}\right) \\ &= H(x_j, p_j)\left(1 + \frac{k}{m}(\Delta t)^2\right). \end{aligned}$$

The value of the Hamiltonian therefore grows by a factor of $1+\frac{k}{m}(\Delta t)^2$ at every step. The problem is less serious the smaller we make Δt, but it remains that it is often just not acceptable to have H grow (or decrease) with time.

Most of the common numerical methods you may have heard of (Runge-Kutta, Gear, etc.) are *not* symplectic. For some problems, some of these methods give reasonable results. However, if you rely on these methods for Hamiltonian systems, you will sooner or later get burned.

Euler's method is an **explicit** method: Given the current point in phase space, we can calculate the next point just by plugging numbers into a formula. Unfortunately, all general-purpose symplectic integration methods are **implicit**: We can't in general write down a formula for the next point. Rather, we have to solve an equation (usually numerically) to calculate $(\mathbf{x}_{j+1}, \mathbf{p}_{j+1})$. There are however some explicit methods for special forms of the Hamiltonian which we will not discuss here.

The key to maintaining the constancy of the Hamiltonian is to derive a numerical method directly from Hamilton's equations 1. Let's start with equation 1a. We can estimate the time derivative on the left-hand side as we did in Euler's method:

$$\frac{dx_i}{dt} \approx \frac{x_{i,j+1} - x_{i,j}}{\Delta t}.$$

In this equation, $x_{i,j+1}$ is the value of x_i at time step j. The right-hand side of equation 1a is also a derivative. In each time step, we will increment the variables, for instance incrementing p_i from $p_{i,j}$ to $p_{i,j+1}$. Since we only take small time steps, p_i should only change by a small amount. This gives us an opportunity to evaluate $\partial H/\partial p_i$ by using another finite-difference approximation. You can probably imagine what we need to do, but it's a bit of a mess to write down without some special notation. Let $\mathbf{p}_j \leftarrow p_{i,j+1}$ be the vector obtained by replacing $p_{i,j}$ by $p_{i,j+1}$ in the vector of momenta at time step k, $\mathbf{p}_j$. Then,

$$\frac{\partial H}{\partial p_i} \approx \frac{H(\mathbf{x}_j, \mathbf{p}_j \leftarrow p_{i,j+1}) - H(\mathbf{x}_j, \mathbf{p}_j)}{p_{i,j+1} - p_{i,j}}.$$

Putting the two sides together, we have

$$\frac{x_{i,j+1} - x_{i,j}}{\Delta t} = \frac{H(\mathbf{x}_j, \mathbf{p}_j \leftarrow p_{i,j+1}) - H(\mathbf{x}_j, \mathbf{p}_j)}{p_{i,j+1} - p_{i,j}}. \tag{6a}$$

We can repeat this procedure for equation 1b, with one minor variation:

$$\frac{p_{i,j+1} - p_{i,j}}{\Delta t} = -\frac{H(\mathbf{x}_j \leftarrow x_{i,j+1}, \mathbf{p}_j \leftarrow p_{i,j+1}) - H(\mathbf{x}_j, \mathbf{p}_j \leftarrow p_{i,j+1})}{x_{i,j+1} - x_{i,j}}. \tag{6b}$$

The reason that we used the updated momentum in equation 6b is that this decision makes the scheme symplectic. To see this, rearrange equations 6a and 6b to

$$\frac{1}{\Delta t}\left(x_{i,j+1} - x_{i,j}\right)\left(p_{i,j+1} - p_{i,j}\right) = H(\mathbf{x}_j, \mathbf{p}_j \leftarrow p_{i,j+1}) - H(\mathbf{x}_j, \mathbf{p}_j),$$

and

$$\frac{1}{\Delta t}\left(x_{i,j+1} - x_{i,j}\right)\left(p_{i,j+1} - p_{i,j}\right) = -\left[H(\mathbf{x}_j \leftarrow x_{i,j+1}, \mathbf{p}_j \leftarrow p_{i,j+1}) - H(\mathbf{x}_j, \mathbf{p}_j \leftarrow p_{i,j+1})\right].$$

Now subtract these two equations:

$$0 = H(\mathbf{x}_j \leftarrow x_{i,j+1}, \mathbf{p}_j \leftarrow p_{i,j+1}) - H(\mathbf{x}_j, \mathbf{p}_j),$$

or

$$H(\mathbf{x}_j \leftarrow x_{i,j+1}, \mathbf{p}_j \leftarrow p_{i,j+1}) = H(\mathbf{x}_j, \mathbf{p}_j).$$

In other words, the update rule consisting of equations 6 keeps the value of H constant for each pair of conjugate variables (x_i, p_i) to which it is applied. Thus, the method is symplectic.

The tradeoff is that we have an implicit method: We need values of the Hamiltonian with some of the coordinates evaluated at the unknown next time point. We therefore have $2n$ equations in $2n$ unknowns. Unless the Hamiltonian is of a particularly simple form, we won't be able to reduce this problem to a simple mapping.

Not all symplectic integration methods hold the value of the Hamiltonian exactly constant. Rather, their effect is to limit the variation in the value of the Hamiltonian. The drift in the Hamiltonian encountered with ordinary numerical methods can be disastrous. Holding the Hamiltonian within tight bounds is often enough to give good results, and one is often content to trade off exact constancy of the Hamiltonian against step size in order to accelerate long integrations.

Appendices

A Maple code for drawing the orbits of the harmonic oscillator

```
k := 200;
m := 0.1;
H := (x,p) -> pˆ2/(2*m) + k*xˆ2/2;
with(plots):
contourplot(H(x,p),x=-0.3..0.3,p=-1.2..1.2,contours=[$1..6]);
```

B Maple code for drawing the invariant tori or cylinders of Hamiltonians with two degrees of freedom

```
y := (x*py - ell)/px;   # Calculate y from angular momentum
H := (x,y,px,py) -> (pxˆ2+pyˆ2)/(2*m2) - G*m1*m2/sqrt(xˆ2+yˆ2);
m1 := 1e11;     # A slightly different set of masses than in the notes
m2 := 1000/1e6;
G := 6.67e-11;
x0 := 1000;     # Initial point (x,y,px,py) = (x0,0,0,p0)
p0 := 0.1/sqrt(1e6);
En := H(x0,0,0,p0);     # Set the energy
ell := x0*p0;   # Set the angular momentum
with(plots):
implicitplot3d(H(x,y,px,py)=En,px=-p0..p0,py=-p0..p0,x=-5*x0..5*x0,
axes=BOXED,grid=[10,40,40],orientation=[-35,50],shading=NONE);
```