DYNAMICS OF GALAXIES

2. Timescales and stellar orbits

Piet van der Kruit
Kapteyn Astronomical Institute
University of Groningen
the Netherlands

Winter 2008/9

Contents

Timescales

Range of timescales

There are a few timescales that are important.

- Crossing time, which is simply the radius divided by the velocity R/V.
  For a galaxy we take some charactristic radius and typical velocity.
  Note that for a uniform sphere with mass M and radius R we have for the typical velocity the circular speed and then

$$V = \sqrt{\frac{GM}{R}} \quad \rho = \frac{3M}{4\pi R^3} \quad t_{\rm cross} = \sqrt{\frac{3}{4\pi G\rho}}$$

- For a galaxy the crossing time is of the order of 10^8 years.
- Hubble time, which is an estimate of the age of the Universe and therefore of galaxies. It is of the order of 10^{10} years.
- The fact that the crossing time is much less than the Hubble time the suggests that we may take the system in dynamical equilibrium.
- Two-body relaxation. This is important for two reasons:
  - Collisions between stars are extremely rare, so collisional pressure is unimportant (contrary to a gas), and
  - Two-body encounters are able to virialize a galaxy so that the kinetic energy of the stars acts as a pressure to stabilize the system, balancing the potential energy.

Two-body relaxation time

Two-body encounters provide processes for a galaxy to come into equilibrium and "virialize", which means that the stellar velocity distribution randomizes.

We will now estimate this relaxation time.

Suppose that we have a cluster of radius R and mass M, made up of N stars with mass m, moving with a mean velocity V.

If two stars pass at a distance r, the acceleration is about Gm/r^2.

Say, that it lasts for the period when the stars are less than the distance r from the closest approach and therefore for a time $2r/V$.

The total change in V^2 is then (acceleration times time)

$$\Delta V^2 \sim \left(\frac{2Gm}{rV}\right)^2$$

The largest possible value of r is obviously R.

For the smallest, we may take $r = r_{\rm min}$, where ΔV^2 is equal to V^2 itself, since then the approximation breaks down. It is not critical, since we will need the logarithm of the ratio $R/r_{\rm min}$.

So we have

$$r_{\rm min} = \frac{2Gm}{V^2}$$

The density of stars is $3N/4\pi R^3$ and the surface density $N/\pi R^2$.

The number of stars with impact parameter r is then the surface density times $2\pi r dr$.

After crossing the cluster once the star has encountered all others. We can calculate the total change in V^2 by integrating over all r

$$(\Delta V^2)_{\rm tot} = \int_{r_{\rm min}}^{R} \left(\frac{2Gm}{rV}\right)^2 \frac{2Nr}{R^2} dr = \left(\frac{2Gm}{RV}\right)^2 2N \ln \Lambda$$

where $\Lambda = R/r_{\rm min}$.

The relaxation time is equal to the number of crossing times it takes for $(\Delta V^2)_{\rm tot}$ to be become equal to V^2.

Since a crossing time is of order R/V and since the virial theorem tells us that $V^2 \sim GNm/R$, we find

$$t_{\rm relax} \sim \frac{RN}{8V\ln\Lambda} \sim \left(\frac{R^3N}{Gm}\right)^{1/2} \frac{1}{8\ln\Lambda}$$

With the expression above for $r_{\rm min}$ we find

$$\Lambda = \frac{R}{r_{\rm min}} = \frac{RV^2}{2Gm} \sim \frac{GNm}{2GRm} \sim \frac{N}{2} \sim N$$

The final expression for the two-body relaxation time then is

$$t_{\mathrm{relax}} \sim \left(\frac{R^3}{GM} \right)^{1/2} \frac{N}{8 \ln N}$$

This ranges from about 10^9 years for globular clusters to 10^{12} years for clusters of galaxies.

Within galaxies encounters are unimportant and they can be treated as collisionless systems.

Violent relaxation

If galaxies are relaxed systems another mechanism must be at work. This is violent relaxation[1].

This occurs when the potential changes on timescales comparable to the dynamical timescale.

If $E(\vec{v},t) = \frac{1}{2}v^2 + \Phi(\vec{x},t)$ then

$$\begin{aligned}
\frac{dE}{dt} &= \frac{dE}{d\vec{v}}\frac{d\vec{v}}{dt} + \frac{d\Phi}{dt} = \vec{v}\frac{d\vec{v}}{dt} + \frac{d\Phi}{dt} \\
&= -\frac{\partial\vec{r}}{\partial t}\frac{\partial\Phi}{\partial\vec{r}} + \frac{\partial\Phi}{\partial t} + \frac{\partial\Phi}{\partial\vec{r}}\frac{d\vec{r}}{dt} \\
&= \frac{\partial\Phi}{\partial t}
\end{aligned}$$

[1] D. Lynden-Bell, MNRAS 136,101 (1967)

Thus a star can change its energy in a collisonless system by a time-dependent potential, such as during the collapse of a galaxy.

The timescale associated with violent relaxation is, according to Lynden-Bell

$$t_{\mathrm{vr}} \sim \langle \frac{\dot{\Phi}^2}{\Phi^2} \rangle$$

So the timescale of violent relaxation is of the order of that of the change of the potential.

A very important aspect is that the change in a star's energy is independent of its mass, contrary to other relaxaton mechanisms, such as two-body encounters, which give rise to mass segregation.

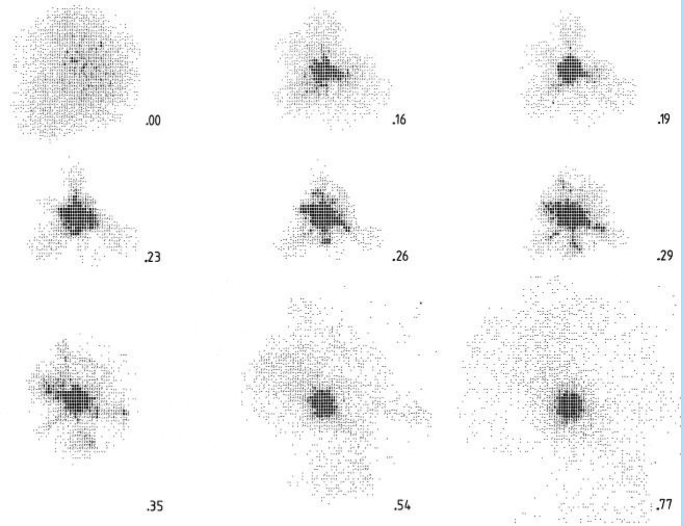
Also some of the information on the initial condition will get lost.

Van Albada[2] was the first to numerically simulate violent relaxation.

He found some remarkable things:

- If the collapse factor was large, irregular initial conditions gave rise to an $R^{1/4}$-law[3] surface density distribution, as observed in elliptical galaxies over a range of up to 12 magnitudes.
- The binding energy of particles before and after collapse correlate, showing that some information on the initial state is not wiped out.

[2] T.S. van Albada, MNRAS 201, 939 (1982)
[3] $\log I(r) = \log I_\circ - 3.33(r/r_e)^{1/4}$.


.00
.16
.19
.23
.26
.29
.35
.54
.77

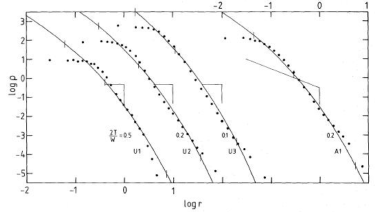
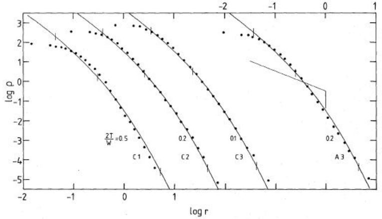
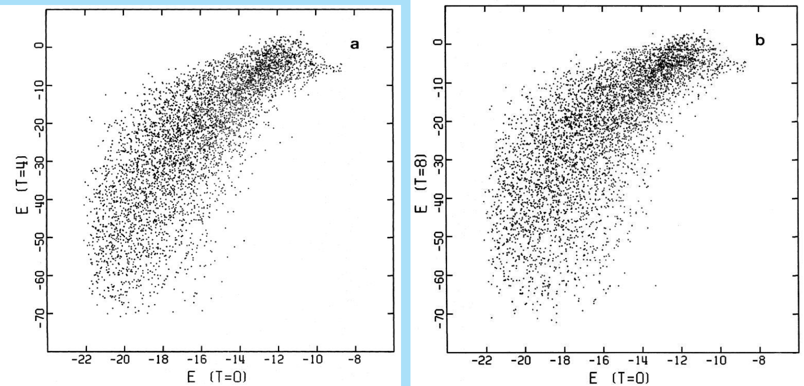


Figure 4. Density distribution in final equilibrium model compared with that for the $r^{1/4}$ law (solid line; Young 1976), for models U and A1. Scaling: equilibrium models and $r^{1/4}$ law model have same half-mass radius r_h and same total mass. Short vertical dashes along $\rho(r)$ for $r^{1/4}$ law model indicate radii containing 10, 50 and 99 per cent of the total mass. Density and radius of starting model are indicated by short straight lines.




a
E (T=4)
E (T=0)
b
E (T=8)
E (T=0)
-22 -20 -18 -16 -14 -12 -10 -8
0 -10 -20 -30 -40 -50 -60 -70

Dynamical friction

As a star moves through a background of other stars, the small deflections will give a small overdensity behind the star and consequently induce a drag.

Suppose that a body of mass m moves in a circular orbit with radius R through a background of bodies with mass M at a speed $V_{\rm c}$ and assume that the background is an isothermal sphere[4] with $V_{\rm c}$ the circular speed (and $V_{\rm c}/2$ the velocity dispersion).

[4] An isothermal sphere is a distribution where everywhere the velocity dispersion is constant and isotropic and that is in equilibirium with its own gravity; see later.

Then the loss of angular momentum is about

$$\frac{dJ}{dR} \sim -0.4 \frac{Gm^2}{R} \ln \Lambda$$

where

$$\Lambda = \frac{R_{\rm c} V_{\rm c}^2}{G(m+M)}$$

$R_{\rm c}$ is the core radius of the isothermal sphere (the typical lengthscale of the background density distribution).

The timescale of dynamical friction for the body to spiral into the center is then

$$t_{\rm df} \sim \frac{R^2 V_{\rm c}}{Gm \ln \Lambda}$$

This timescale is large and only relevant for globular clusters in the inner halo or for galaxies in the central parts of clusters.

Stellar orbits

Spherical potentials

The equation of motion in a spherical potential is in vector notation

$$\ddot{\mathbf{R}} = -\frac{d\Phi}{d\mathrm{R}}\hat{\mathbf{e}}_{\mathrm{R}}$$

The angular momentum is

$$\mathbf{R} \times \dot{\mathbf{R}} = \mathbf{L}$$

This is constant and the orbit therefore is in a plane.

We then use polar coordinates in this plane these two equations become

$$\ddot{R} - R\dot{\theta}^2 = -\frac{d\Phi}{dR}$$

$$R^2\dot{\theta} = L$$

Integrating this we get

$$\tfrac{1}{2}\dot{R}^2 + \frac{1}{2}\frac{L^2}{R^2} + \Phi(R) = E$$

The energy E is constant.

If $E < 0$ then the star is bound between radii $R_{\rm max}$ and $R_{\rm min}$, which are the roots of

$$\frac{1}{2}\frac{L^2}{R^2} + \Phi(R) = E$$

The radial period is the interval between the times the star is at $R_{\rm min}$ and $R_{\rm max}$ and back.

$$\begin{aligned} T_{\rm R} &= 2\int_{R_{\rm min}}^{R_{\rm max}} dt = 2\int_{R_{\rm min}}^{R_{\rm max}} \frac{dR}{\dot{R}} = \\ & 2\int_{R_{\rm min}}^{R_{\rm max}} \frac{dR}{\{2[E-\Phi(R)] - L^2/R^2\}^{1/2}} \end{aligned}$$

In the azimuthal direction the angle θ changes in the time $T_{\rm R}$ by

$$\Delta\theta = \int_0^{T_{\rm R}} \frac{d\theta}{dR} dR = 2\int_0^{T_{\rm R}} \left(\frac{L}{R^2}\right) \frac{dR}{\dot{R}}$$

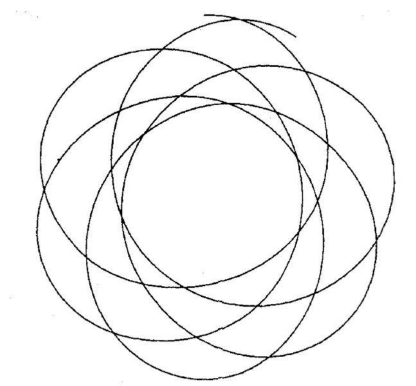
This can be evaluated further in terms of $T_{\rm R}$, which depends upon the particular potential.

The orbit is closed if

$$\Delta\theta = 2\pi\frac{m}{n}$$

with m and n integers.

This is not generally true and the orbit then has the form of a rosette and can the star visit every point within $(R_{\rm min}, R_{\rm max})$.



Even in the simple case of a spherical potential, the equation of motion of the orbit must be integrated numerically.

The Rosette orbit can be closed by observing it from a rotating frame (see below under resonances), when it is rotating at an angular velocity of

$$\Omega_{\rm p} = \frac{(\Delta\theta - 2\pi)}{T_{\rm R}}$$

We will treat two special cases which can be solved analytically.

The harmonic oscillator

This concerns the potential of a uniform sphere

$$\Phi = \tfrac{1}{2}\Omega^2 R^2.$$

Then we take cartesian coordinates $x = r\cos\theta$, $y = r\sin\theta$ and then

$$\frac{d^2x}{dt^2} = -\Omega^2 x \quad ; \quad \frac{d^2y}{dt^2} = -\Omega^2 y$$

Then

$$x = X\cos(\Omega t + a_{\mathrm{x},\circ}) \quad ; \quad y = Y\cos(\Omega t + a_{\mathrm{y},\circ})$$

The orbits are closed ellipses centered on the origin and $\Delta\theta$ is equal to π in T_{R}.

The Keplerian potential

The potential now is that of a point source in the center and this is the well-known two-body problem[5]:

$$\Phi = -\frac{GM}{R}$$

The orbits are closed ellipses with one focus at the origin:

$$R = \frac{a(1-e^2)}{\{1+\cos(\theta-\theta_\circ)\}}$$

[5] There is a complete derivation of the two-body problem available at (http://www.astro.rug.nl/~vdkruit/jea3/homepage/two-body.pdf).

Here semi-major axis a and excentricity e are related to E and L by

$$a = \frac{L^2}{GM(1-e^2)} \quad ; \quad E = -\frac{GM}{2a}$$

$$R_{\rm max}, R_{\rm min} = a(1 \pm e)$$

$$T_{\rm R} = T_\theta = 2\pi \sqrt{\frac{a^3}{GM}} = T_{\rm R}(E)$$

Now $\Delta\theta = 2\pi$ in $T_{\rm R}$.

Galaxies have mass distributions somewhere between these two extremes, so we may expect that $\Delta\theta$ is in the range π to 2π in $T_{\rm R}$.

Axisymmetric potentials

We now have a potential $\Phi = \Phi(R, z)$, that may be applicable to disk galaxies. The equations of motion are

$$\ddot{R} - R\dot{\theta}^2 = -\frac{\partial \Phi}{\partial R}$$

$$\frac{d}{dt}(R^2\dot{\theta}) = 0$$

$$\ddot{z} = \frac{d^2 z}{dt^2} = -\frac{\partial \Phi}{\partial z}$$

Integration of middle one of these equations gives

$$L_{\mathrm{z}} = R^2\dot{\theta}$$

The motion in the meridional plane then can be described by an effective potential

$$\ddot{R} = -\frac{\partial \Phi_{\rm eff}}{\partial R}$$

$$\ddot{z} = -\frac{\partial \Phi_{\rm eff}}{\partial z}$$

where

$$\Phi_{\rm eff} = \Phi(R,z) + \frac{L_{\rm z}^2}{2R^2}$$

The energy of the orbit is

$$E = \tfrac{1}{2}\dot{R}^2 + \tfrac{1}{2}\dot{z}^2 + \Phi_{\rm eff}(R,z)$$

The orbit is trapped inside the appropriate contour $E = \Phi_{\rm eff}$, which is called the zero-velocity curve.

Only orbits with low $L_{\rm z}$ can approach the z-axis.

The minimum in $\Phi_{\rm eff}$ occurs for $\nabla\Phi_{\rm eff} = 0$, or at $z = 0$ and where

$$\frac{\partial\Phi}{\partial R} = \frac{L_{\rm z}^2}{R^3}$$

This corresponds to the circular orbit with $L = L_{\rm z}$.

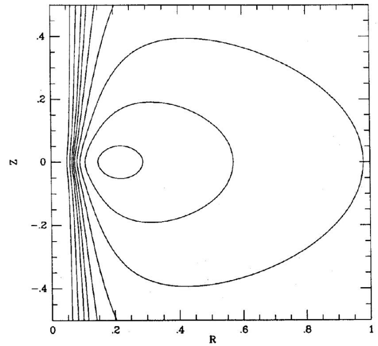
It is the highest angular momentum orbit that is possible for a given E, or in other words, it has all its kinetic energy in θ-motion.

As an example we take the logarithmic potential

$$\Phi(R, z) = \tfrac{1}{2} V_\circ^2 \ln \left(R^2 + \frac{z^2}{q^2} \right)$$

Here are countours of $\Phi_{\rm eff}$ for the case $q = 0.5$ and $L_z = 0.2$.

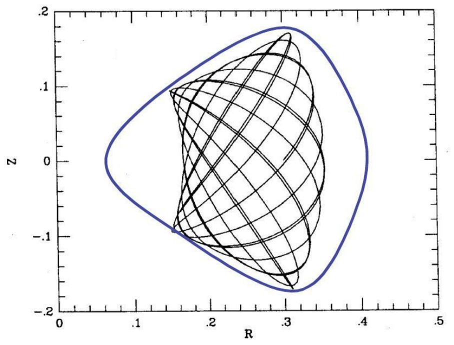
The minimum in $\Phi_{\rm eff}$ occurs where $\nabla\Phi_{\rm eff} = 0$ that is in the plane ($z = 0$).



If E and L_z were the only two isolating integrals the orbits would be able to visit all points within their zero-velocity curves.
In simulations this is often not the case and there must be a third integral.

Here is the case of actual simulated orbits in a slightly flattened logarithmic potential. We show the motion in the meridional plane, rotating along with the angular momentum of the orbit.

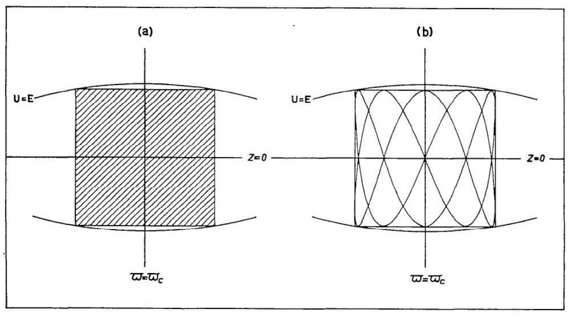
The blue line is the zero-velocity curve corresponding to this orbit.



Third integral and surface of section

Recall that for small deviations from the symmetry plane the energy in the z-direction was a third isolating integral.

Here are two diagrams from an early study by Ollongren[6]. We have either periodic or non-periodic orbits.

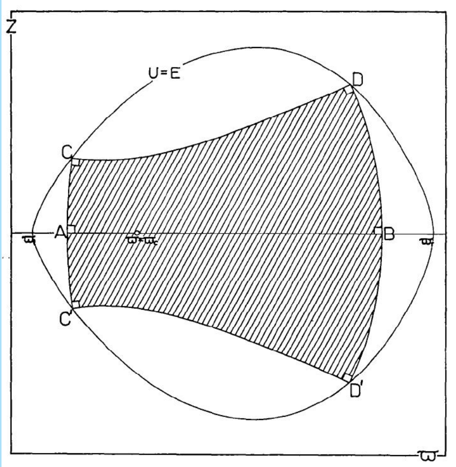


[6] A. Ollongren, B.A.N. 16, 241 (1962)

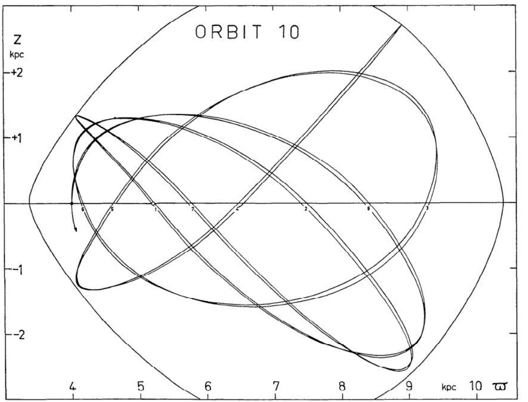
Ollongren did numerical integrations using the potential of a recent model of the mass distribution in the Galaxy by Schmidt[a].

He found that there was a distortion of the box that was covered by the orbit.

[a] M. Schmidt, B.A.N. 13, 15 (1956)

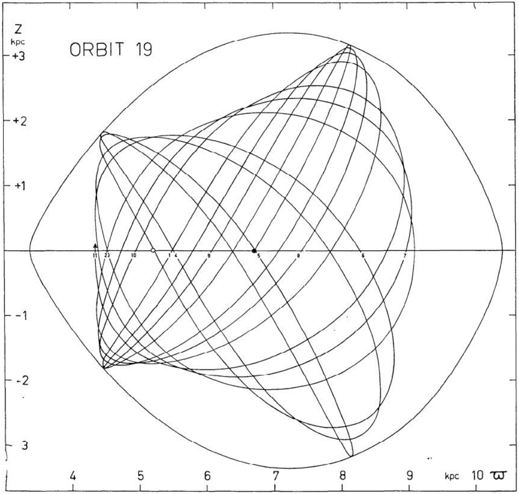


He also found that the most general separable case was in elliptical coordinates, in which a third integral is quadratic in the velocities[7].

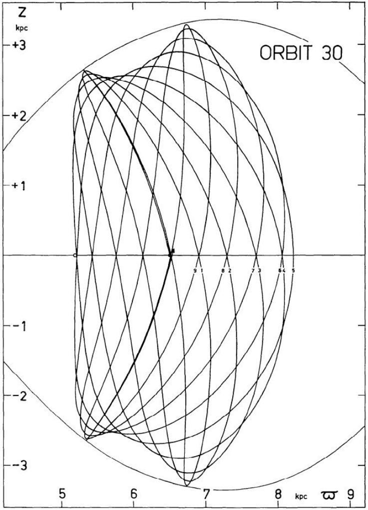


Orbit 10, *low*, *special*, frequency ratio near 14 : 17. In revolution 5 the fundamental point D is approached closely, after which the previous path is nearly retraced back to the starting point ($\varpi = \varpi_A$, $z = 0$).

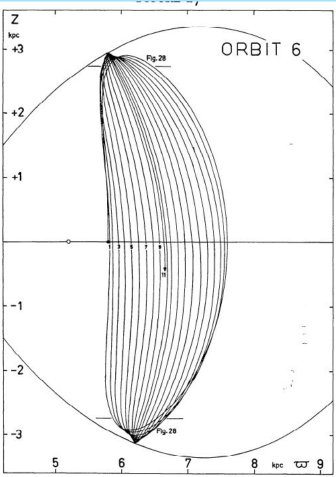
[7] See also H.C. van de Hulst, B.A.N. 16, 235 (1962)



Orbit 19, *intermediate*. In revolution 11 the fundamental point A is approached closely. Motion of orbital plane is shown in Figure 33.



Orbit 30, *intermediate*, close to orbit 3. The trajectory departs only slowly from the trajectory of orbit 3.



Orbit 6, *high*, *special*. Sense of rotation of all revolutions drawn is clockwise, except for revolution 11, which is of the switching type. Region within the boundary for $|z| > 2.75$ enlarged in Figure 28.

Summarizing:

- If E and L_z are the only two isolating integrals, the orbit would visit all points within the zero-velocity curves.
- In practice it was found that there are limiting surfaces that seem to forbid the orbit to fill the whole volume within the zero-velocity curves.
- This behaviour is very common for orbits in axisymmetric potentials, when the combination (E, L_z) is not too far from that of a circular orbit. A third integral is present, although in general its form cannot be explicitly written down.

For each orbit the energy $E(R, z, \dot{R}, \dot{z})$ is an integral, so only three of the four coordinates can be independent, say R, z and $\dot{R}$.

The orbit can visit every point in $(R, z, \dot{R})$-space as far as allowed by E.

Now take a slice through $(R, z, \dot{R})$-space, e.g. at $z = 0$. This is called a *surface of section*.

The orbits' successive crossings of $z = 0$ generate a set of points inside the region $E = \frac{1}{2}\dot{R}^2 + \Phi_{\rm eff}(R, 0)$.

Hénon & Heiles[8] did a famous study of third integrals and surfaces of section. They used a convenient analytical potential in coordinates (x, y):

$$\Psi(x, y) = \frac{1}{2}(x^2 + y^2 + 2x^2 y - \frac{2}{3}y^3)$$

The figure shows consecutive crossings of the surface of section $(y, \dot{y})$.

After an infinite time the full curve will be filled.

This is a signature of a third isolating integral; the orbit is constrained inside the zero-velocity curve.

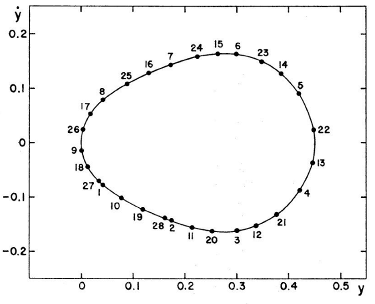
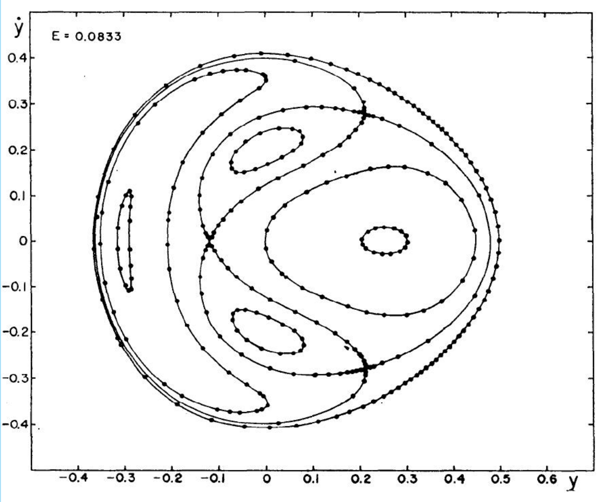


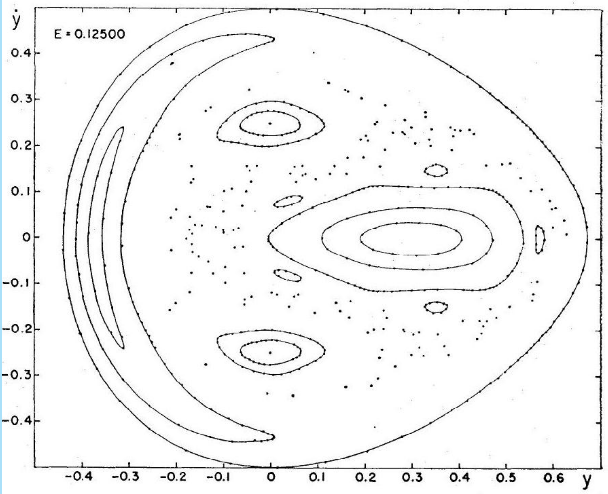
FIG. 3. A typical set of points P_i; $E = 0.08333$.

[8] M. Hénon & C. Heiles, A.J. 69, 73 (1964)

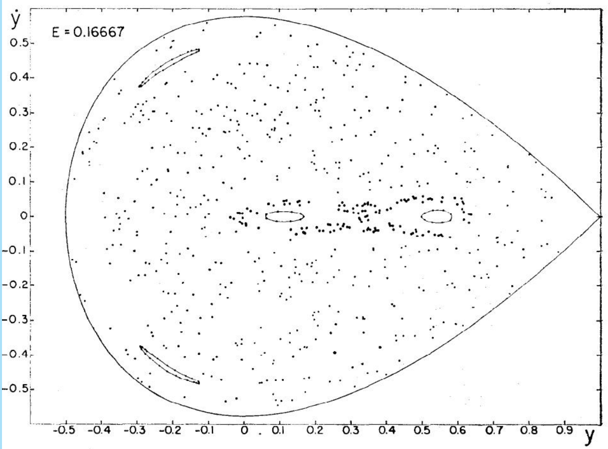
Here are some orbits for $E = 0.08333$. All have a third integral.



Here are orbits for $E = 0.125$. Now some orbits have no third integral.

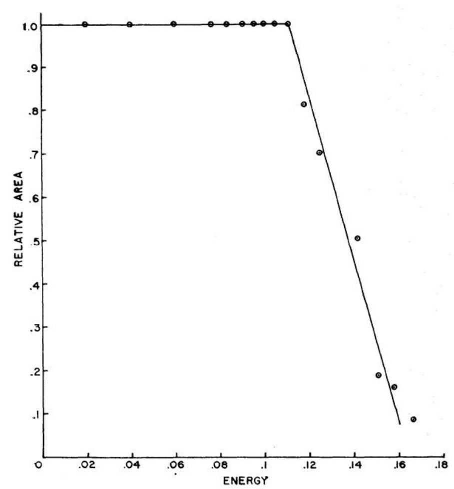


For $E = 0.16667$ almost no orbits have a third integral.



Hénon & Heiles divised a method to derive the fraction of orbits that have a thrid integral for each energy.

For $E < 0.11$ all orbits have a third integral, but for $E > 0.17$ almost none do.



If there is no other integral then these points fill the whole region.

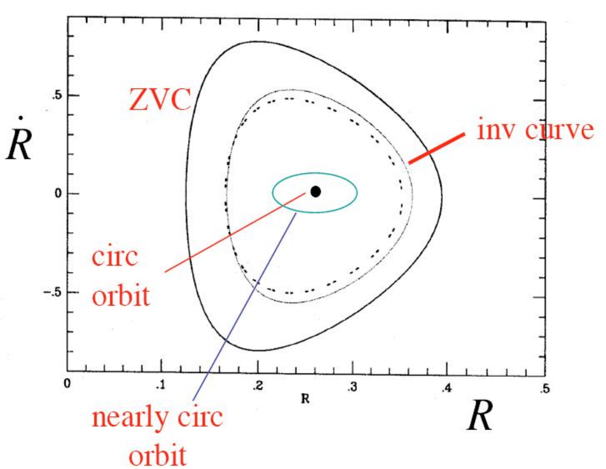
If there is another integral, then its surface $I_{\mathrm{R}}(R, z, \dot{R})$ cuts the plane in a *curve* $I_{\mathrm{R}}(R, 0, \dot{R}) = \text{constant}$.

A *periodic* orbit is a *point* or a set of points on the $(R, \dot{R})$ surface of section.

Such curves and points are called *invariant*, because they are invariant under the mapping of the surface of section onto itself generated by the orbit.

Invariant points often have closed invariant curves around them on the surface of section. These represent *stable* periodic orbits. Ones where invariant curves cross are *unstable* periodic orbits.

This diagram (taken from Ken Freeman) summarizes the points.



Rotating non-axisymmetric potentials

In cases of bars or some elliptical galaxies we may consider a potential that rotates with a rigid angular velocity Ω.

Then the equation of motion is

$$\ddot{\mathbf{r}} = -\nabla\Psi - 2(\mathbf{\Omega} \times \mathbf{r}) - \mathbf{\Omega} \times (\mathbf{\Omega} \times \mathbf{r})$$

The second term on the right is the Coriolis force and the third one the centrifugal force.

Then we can define an effective potential, so that

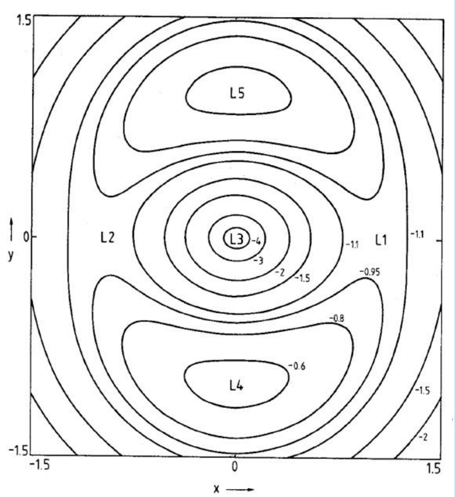
$$\ddot{\mathbf{r}} = -\nabla\Psi_{\mathrm{eff}} - 2(\mathbf{\Omega} \times \mathbf{r})$$

Such a potential has equipotential curves in the $z = 0$ plane that show neutral points.

L_1 and L_2 are saddle points and are unstable.

L_3 is a minimum and is stable.

L_4 and L_5 are maxima that can either be stable or unstable.

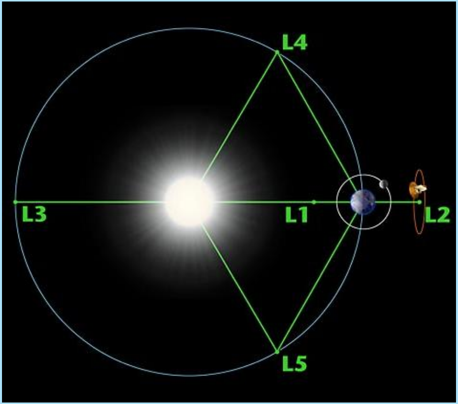


These point should in spite of their notation not be confused with Lagrange points in the restricted three-body problem, although there is some similarity.

There are two bodies (here Sun and Earth) in circular orbits.

The Lagrange points L_1, L_2 and L_3 are saddle points and unstable.

L_4 and L_5 are stable.



Stars describe orbits that reinforce the bar potential.