Classical mechanics

November 15, 2012

Lecture notes from:

- *Mathematical Methods of Classical Mechanics* by V.I. Arnold, Springer, 1989.
- *A brief introduction to Classical, Statistical and Quantum Mechanics* by O. Bühler, Courant Lecture Notes, American Mathematical Society, 2006.

1 "Hurtling through space": Newton's laws of motion

Classical mechanics deals with systems that can be described by a set of *mass points*, that is, by a set of points moving in space, each with an associated mass. Examples are planetary systems, molecular systems, systems of rods and springs, etc. Largely we work in a three-dimensional space plus time, but if the system consists of n points, the problem is formulated in $\mathbb{R}^{3n}$, etc.

A *motion* in $\mathbb{R}^N$ is a differentiable mapping

$$x(t) : I \to \mathbb{R}^N$$

where I is an interval on the real axis. The first derivative of the motion $\dot{x}(t) = dx/dt$ is the *velocity vector*. The second derivative $\ddot{x}(t)$ is the *acceleration vector*.

A *mechanical system* of n points moving in three-dimensional euclidean space is defined as follows: Consider a motion in $\mathbb{R}^3$: $x(t) : \mathbb{R} \to \mathbb{R}^3$. The graph (t, x) of the motion is a curve in $\mathbb{R} \times \mathbb{R}^3$. A motion of n points gives n curves described by n mappings $x_i : \mathbb{R} \to \mathbb{R}^3$, $i = 1, \ldots, n$. The direct product of n copies of $\mathbb{R}^3$ is called the *configuration space* of the system of n points

$$x(t) : \mathbb{R} \to \mathbb{R}^N, \quad N = 3 \cdot n,$$

a motion of n points in the the coordinate system on $\mathbb{R} \times \mathbb{R}^3$.

Newton's second law of motion is of central importance in mechanics:

$$\text{mass} \times \text{acceleration} = \text{force}$$

For example, Newton's apple can be modeled as a mass point. The height of the apple core above the ground is denoted $z(t)$, and its mass is m. The force exerted on the apple is it weight $-mg$, where g is the acceleration of gravity. Newton's second law is

$$m\ddot{z} = -mg.$$

This is a second order ordinary differential equation for $z(t)$ with solution

$$z(t) = z(0) - \dot{z}(0)t - \frac{g}{2}t^2$$

To specify the solution uniquely, we need 2 pieces of data, for instance the height and velocity of the apple at some time $t = 0$. In fact knowing these two quantities at any time t (between the moment the apple falls and the moment it hits the ground!) is enough to completely specify the whole motion.

We define the *state of the system* to be the collection of variables necessary to specify the system in this way. For models based on Newton's Second Law, the state consists of the position and velocity:

$$\text{state} = (\text{position}, \text{velocity}) = (x, \dot{x}).$$

If the motion is well-posed then the state at some time determines the state at all other times

$$\text{state}(0) \to \text{state}(t).$$

Newton's principle of determinacy: The initial state of a mechanical system (the totality of the positions and velocities of its points at some moment of time) uniquely determines all of its motion.

By Newton's determinacy condition, the motion is completely determined (constrained) by the position $x(t)$ and the velocity $\dot{x}(t)$ at some time t. In particular, the acceleration $\ddot{x}(t)$ at time t must be determined by these quantities. In other words,

$$\ddot{x} = F(x, \dot{x}, t).$$

Note that by the existence and uniqueness theorem of initial value problems, the motion of the system now indeed satisfies determinacy if we know $x(t_0)$ and $\dot{x}(t_0)$ at some time t_0.

Multiplying left and right with $\dot{x}$ gives

$$\begin{aligned}
\ddot{x}\dot{x} &= -\frac{\partial U}{\partial x}\dot{x} \\
\frac{d}{dt}\frac{\dot{x}^2}{2} &= -\frac{d}{dt}U(x(t)) \\
0 &= \frac{d}{dt}\left[\frac{\dot{x}^2}{2} + U(x)\right],
\end{aligned}$$

which expresses the conservation of total energy. The quantity $\dot{x}^2/2$ is referred to as the *kinetic energy* for this problem.

Earlier in the course we have seen that correct models should be independent of the choice of physical units in which they are expressed (cm, sec, kg, etc.) A related idea is the following: *a correct model should be independent of the coordinates in which it is expressed.* Furthermore, for special systems, there may be even more such considerations. The universe has no absolute coordinate system. The earth rotates about its axis and orbits the sun, which in turn orbits its galaxy, in an expanding universe. Any physical laws that hold universally, should be invariant with respect to all this motion.

Galileo's principle of relativity: there exist inertial coordinate systems such that: (1) all laws of nature at all moments of time are the same in all inertial coordinate systems, and (2) all coordinate systems in uniform, rectilinear motion with respect to an inertial one are themselves inertial.

For example, if you are moving in a straight line at constant speed inside a train (an idealized, smooth, soundproof train, without windows), you cannot detect the motion of the train by conducting experiments inside the car. A *Galilean group* is a group of transformations g of a Galilean space (i.e. $\mathbb{R} \times \mathbb{R}^3$, equipped with a distance function $|\cdot|$ for points in $\mathbb{R}^3$) which preserve its structure. The Galilean transformations are given by

1. Uniform motion with velocity v

$$g(t, x) = (t, x + vt), \quad t \in \mathbb{R}, x, v \in \mathbb{R}^3$$

2. Translation of the origin

$$g(t, x) = (t + s, x + y),$$

3. Rotation of the coordinate axes

$$g(t, x) = (t, Gx),$$

where $G : \mathbb{R}^3 \to \mathbb{R}^3$ is an orthogonal matrix.

Every Galilean transformation can be written as a composition of the above three.

If the motion is Galilean invariant, then the following must hold:

1. Time translation: if $x = \phi(t)$ is a solution of Newton's equations, then so is $x = \phi(t + s)$ for any $s \in \mathbb{R}$. Hence, in an intertial coordinate system, Newton's equations must have the form

$$\ddot{x} = F(x, \dot{x}),$$

i.e. autonomous.

2. Translations in three-dimensional space: if $x_i = \phi_i(t)$, $i = 1, \ldots, n$ are motions of an n-point system, then so are $x_i = \phi_i(t) + r$, $i = 1, \ldots, n$ for any $r \in \mathbb{R}^3$. Consequently, in an inertial coordinate system, Newton's equations may only depend only on relative coordinates $x_j - x_k$. Space is homogeneous. Similarly transformation to a coordinate system moving with constant velocity relative to an inertial one must leave the motion invariant. Hence,

$$\ddot{x}_i = f_i(\{x_j - x_k\}, \{\dot{x}_j - \dot{x}_k\}), \quad i, j, k = 1, \ldots, n.$$

3. Rotations in $\mathbb{R}^3$: If $x_i = \phi_i(t)$, $i = 1, \ldots, n$ is a motion satisfying Newton's equations and $G : \mathbb{R}^3 \to \mathbb{R}^3$ is an orthogonal matrix, then $x_i = G\phi_i(t)$, $i = 1, \ldots, n$ is also motion. This implies there is no preferred direction in Galilean space, and

$$f_i(\{Gx_j - Gx_k\}, \{G\dot{x}_j - G\dot{x}_k\}) = Gf_i(\{x_j - x_k\}, \{\dot{x}_j - \dot{x}_k\}), \quad i, j, k = 1, \ldots, n.$$

Example. Falling from a great height. The previous model has a restricted domain of application. Newton's law of gravitation leads to an improved model

$$\ddot{x} = -g\frac{r_0^2}{r^2}, \quad r = r_0 + x.$$

In this case the potential energy is $U = -gr_0^2/r$.

Example. Motion of a weight under action of a spring

$$\ddot{x} = -\alpha^2 x, \qquad U = \frac{\alpha^2 x^2}{2}.$$

This model is known as the harmonic oscillator.

Let $E^{3n} = E^3 \times E^3 \times \cdots \times E^3$ be the configuration space of a system of n points in Euclidean space $\mathbb{R}^3$. Let $U : E^{3n} \to \mathbb{R}$ be a differentiable function and let $m_1, \ldots, m_n$ be positive numbers. We use the notation $x = (x_1, \ldots, x_n) \in \mathbb{R}^{3n}$ and denote by $\partial U/\partial x$ the gradient of U, i.e.

$$\frac{\partial U}{\partial x} = \left(\frac{\partial U}{\partial x_1}, \ldots, \frac{\partial U}{\partial x_n}\right).$$

The motion of n points of masses $m_1, m_2, \ldots m_n$ in the potential field with potential energy U is given by

$$m_i \ddot{x}_i = -\frac{\partial U}{\partial x_i}, \quad i = 1, \ldots, n$$

All of the models we have seen above have this form (with $m_i = 1$). Computing the scalar product of each side with $\dot{x}_i$ and summing over i shows the conservation of energy

$$\sum_i (\dot{x}_i, m_i \ddot{x}_i) = \sum_i (\dot{x}_i, -\frac{\partial U}{\partial x_i})$$
$$\frac{d}{dt}\sum_i \frac{m_i}{2}\|x_i\|^2 = \frac{d}{dt}U(x)$$
$$0 = \frac{d}{dt}\left[T(\dot{x}) + U(x)\right], \quad T = \sum_i \frac{m_i}{2}\|\dot{x}_i\|^2,$$

where T denotes the kinetic energy.

Example. 3-body problem

$$U = -\frac{m_1 m_2}{\|x_2 - x_1\|} - \frac{m_2 m_3}{\|x_3 - x_2\|} - \frac{m_3 m_1}{\|x_1 - x_3\|}$$

Consider a mechanical system with one degree of freedom (configuration space $\mathbb{R}$)

$$\ddot{x} = f(x), \quad x(t) \in \mathbb{R}, \quad f(x) = -\frac{\partial U}{\partial x},$$

This system can also be written as a first order system

$$\dot{x} = y, \quad \dot{y} = f(x)$$

Consider the phase plane with coordinates (x, y). The right hand sides determine a vector field $V(x, y) = (y, f(x))$ on the plane. Each point corresponds to a state of the system (x and y are Newton's deterministic quantities). A motion $(x(t), y(t))$ traces a curve in the phase plane such that the velocity (of the phase point) at each point is equal to the phase velocity vector V. Conservation of energy says that each phase curve is a level set of the total energy $E(x, y) = T + U = y^2/2 + U(x)$.

2 "The best of all possible worlds": Variational mechanics (19 October 2011)

2.1 Calculus of Variations

The 'calculus of variations' is concerned with the extremals of *functionals*: functions whose domain is the infinite dimensional space, the space of curves.

Example. Let $\gamma = \{(t, x) : x = x(t), t_0 \leq t \leq t_1\}$ be a curve. The *arclength* of γ is an example of a functional:

$$\Phi(\gamma) = \int_{t_0}^{t_1} \sqrt{1 + \dot{x}^2}\, dt,$$

where $\dot{x}$ denotes the slope dx/dt.

Consider now a perturbed curve $\gamma\prime = \{(t,x) : x = x(t) + h(t)\}$, i.e. $\gamma\prime = \gamma + h$, and the increment $\Phi(\gamma + h) - \Phi(\gamma)$. The functional Φ is said to be *differentiable* if $\Phi(\gamma + h) - \Phi(\gamma) = F + R$, where F depends linearly on h:

$$F(h_1 + h_2) = F(h_1) + F(h_2), \quad F(ch) = cF(h),$$

and $R(h,\gamma) = \mathcal{O}(h^2)$:

$$|h| < \varepsilon, \left|\frac{dh}{dt}\right| < \varepsilon \quad \Rightarrow \quad |R| < c\varepsilon^2.$$

$F(h)$ is called the *variational derivative* or the *variation.*

For example, let $\gamma = \{(t,x) : x = x(t), t_0 \le t \le t_1\}$ be a curve in the plane, $\dot{x} = dx/dt$, $L = L(a,b,c)$ be differentiable in all three of its arguments. Define

$$\Phi(\gamma) = \int_{t_0}^{t_1} L(x(t), \dot{x}(t), t)\, dt. \tag{1}$$

(The specific choice $L = (1+b^2)^{1/2}$ gives the arc length functional). The functional $\Phi(\gamma)$ is differentiable and its derivative is

$$F(h) = \int_{t_0}^{t_1} \left[\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}}\right] h\, dt + \left.\left(\frac{\partial L}{\partial \dot{x}} h\right)\right|_{t_0}^{t_1}.$$

This follows from

$$\begin{aligned}
\Phi(\gamma + h) - \Phi(\gamma) &= \int_{t_0}^{t_1} L(x(t) + h(t), \dot{x}(t) + \dot{h}(t), t) - L(x(t), \dot{x}(t), t)\, dt \\
&= \int_{t_0}^{t_1} \left[L(x(t), \dot{x}(t), t) + \frac{\partial L}{\partial x} h(t) + \frac{\partial L}{\partial \dot{x}} \dot{h}(t) + \mathcal{O}(h^2)\right] - L(x(t), \dot{x}(t), t)\, dt \\
&= \int_{t_0}^{t_1} \frac{\partial L}{\partial x} h(t) + \frac{\partial L}{\partial \dot{x}} \dot{h}(t) + \mathcal{O}(h^2)\, dt \\
&= \int_{t_0}^{t_1} \frac{\partial L}{\partial x} h(t) - \frac{\partial}{\partial t}\frac{\partial L}{\partial \dot{x}} h(t)\, dt + \left.\frac{\partial L}{\partial \dot{x}} h\right|_{t_0}^{t_1} + \mathcal{O}(h^2) \\
F(h) &= \int_{t_0}^{t_1} \left[\frac{\partial L}{\partial x} - \frac{\partial}{\partial t}\frac{\partial L}{\partial \dot{x}}\right] h(t)\, dt + \left.\frac{\partial L}{\partial \dot{x}} h\right|_{t_0}^{t_1}.
\end{aligned}$$

The extremal of Φ is a curve γ such that $F(h) = 0$ for all h.

The curve $\gamma : x = x(t)$ is an extremal of Φ over the space of curves passing through $x(t_0) = x_0$, $x(t_1) = x_1$, if an only if

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x} = 0 \tag{2}$$

along the curve $x(t)$. (For the class of curves we are considering, the extremum $x(t)$ satisfies the boundary conditions, and $h(t_0) = h(t_1) = 0$. The condition $F(h) = 0$ for all h implies the term in square brackets above is identically zero, since, if $F(h)$ were, say, positive on some interval of positive measure, we could find a nonnegative function h with support on the same interval, such that the integral would be positive.)

The equation (2) above is referred to as the *Euler-Lagrange equation* for the functional Φ defined in (1).

More generally, consider curves in $\mathbb{R}^n$: $\gamma = \{(t, x) : x = x(t) \in \mathbb{R}^n\}$ and a function $L : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$. Then γ is an extremal of

$$\Phi = \int L(x, \dot{x}, t)\, dt, \tag{3}$$

if and only if it satisfies the Euler-Lagrange equation

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}_i}\right) - \frac{\partial L}{\partial x_i} = 0, \quad i = 1, \ldots, n.$$

2.2 Variational Mechanics

The motions of Newton's equations coincide with extremals of the functional

$$\Phi(\gamma) = \int_{t_0}^{t_1} L\, dt, \quad L = T - U, \tag{4}$$

i.e., the difference between the kinetic energy T and potential energy U. For example,

$$T = \frac{1}{2}\sum_i m_i |\dot{q}_i|^2, \quad U = U(q_i), \quad q_i \in \mathbb{R}^3, \quad i = 1, \ldots, n.$$

Then

$$\frac{\partial L}{\partial \dot{q}_i} = m_i \dot{q}_i, \quad \frac{\partial L}{\partial q_i} = -\frac{\partial U}{\partial q_i}.$$

The extremal satisfies

$$\frac{d}{dt} m_i \dot{q}_i = -\frac{\partial U}{\partial q_i}.$$

This is Newton's motion in a potential energy field.

In other words, if $q = (q_1, \ldots, q_n)$ denote coordinates of the configuration space of a system of n mass points, then the evolution of q with time is subject to the Euler-Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} = 0.$$

This fact is known as the *principle of least action.* The function $L(q, \dot{q}, t) = T - U$ is referred to as the *Lagrangian.* The associated functional (4) is called the *action integral.*

Examples.

1. A free mass in $\mathbb{R}^3$:
$$L = T = \frac{1}{2} m |\dot{q}|^2.$$
The Euler-Lagrange equation is
$$m\ddot{q} = 0.$$
Introducing the generalized momentum $p = \partial L / \partial \dot{q} = m\dot{q}$, The Euler-Lagrange equation says
$$\frac{dp}{dt} = 0.$$
The momentum is constant and the motion is in a straight line.

2. Coordinate invariance. Consider planar motion in polar coordinates:
$$q^1 = r\cos\theta, \quad q^2 = r\sin\theta,$$
and
$$\dot{q}^1 = \dot{r}\cos\theta - r\sin\theta\dot{\theta}, \quad \dot{q}^2 = \dot{r}\sin\theta + r\cos\theta\dot{\theta}.$$
The kinetic and (radial) potential energies are
$$T(r, \theta) = \frac{m}{2}|\dot{q}|^2 = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2), \quad U(r, \theta) = U(r).$$
The Euler-Lagrange equations are
$$m\ddot{r} - mr\dot{\theta}^2 = -\frac{\partial U}{\partial r}, \qquad \frac{d}{dt} r^2 \dot{\theta} = 0.$$
The second equation reduces to $\dot{\theta} = c/r^2$, from which the first can be written in terms of r alone.

2.3 Noether's theorem

Consider the mapping $\phi : \mathbb{R}^n \to \mathbb{R}^n$. The Lagrangian $L(q, \dot{q})$ *admits the mapping* ϕ if $L(\phi(q), \frac{d}{dt}\phi(q)) = L(q, \dot{q})$ for all q.

If L admits a one parameter group of diffeomorphisms $\phi^s : \mathbb{R}^n \to \mathbb{R}^n$, $s \in \mathbb{R}$, then the Euler-Lagrange equations have a first integral I:
$$I(q, \dot{q}) = \frac{\partial L}{\partial \dot{q}} \frac{\partial \phi^s(q)}{\partial s}\bigg|_{s=0}.$$

To see this, let $q(t)$ be a solution of the Euler-Lagrange equations. Since L admits ϕ^s, the translation $\phi^s(q(t))$ also satisfies the Euler-Lagrange equations, for all s. Consider the family of solutions $Q(s,t) = \phi^s(q(t))$. By hypothesis L is invariant with respect to s:

$$\frac{\partial L}{\partial s} = 0 = \frac{\partial L(Q, \frac{\partial Q}{\partial t})}{\partial s} = \frac{\partial L}{\partial q}\frac{\partial Q}{\partial s} + \frac{\partial L}{\partial \dot{q}}\frac{\partial^2 Q}{\partial t \partial s}. \tag{5}$$

But Q also satisfies the Euler-Lagrange equation for any s. Hence,

$$\frac{d}{dt}\left[\frac{\partial L}{\partial \dot{q}}\left(Q, \frac{\partial Q}{\partial t}\right)\right] = \frac{\partial L}{\partial q}\left(Q, \frac{\partial Q}{\partial t}\right). \tag{6}$$

Substituting (6) into (5) for $\partial L/\partial q$:

$$\begin{aligned} 0 &= \frac{d}{dt}\left[\frac{\partial L}{\partial \dot{q}}\left(Q, \frac{\partial Q}{\partial t}\right)\right]\frac{\partial Q}{\partial s} + \frac{\partial L}{\partial \dot{q}}\frac{\partial^2 Q}{\partial t \partial s} \\ &= \frac{d}{dt}\left[\frac{\partial L}{\partial \dot{q}}\frac{\partial Q}{\partial s}\right] = \frac{dI}{dt}. \end{aligned}$$

This important result links symmetries in the Lagrangian to conserved quantities of the Euler-Lagrange equations, and is referred to as Noether's theorem.

Example. Consider the *translation symmetry* $Q(s,t) = \phi^s q(t) = q(t) + sv$, for some direction vector $v \in \mathbb{R}^n$. This leads to the conserved quantity

$$I(q, \dot{q}) = \frac{\partial L}{\partial \dot{q}}\frac{\partial Q(s,t)}{\partial s}\bigg|_{s=0} = \frac{\partial L}{\partial \dot{q}} \cdot v$$

For a system of particles with $T(q) = \sum_i \frac{m_i}{2}|q_i|^2$, and v equivalent to translation along x-coordinate axis $e_1 = (1,0,0)$, i.e. $v = (e_1, \ldots, e_1)$, we find

$$I = \sum_i m_i \dot{q}_i \cdot e_1 = e_1 \cdot \sum_i p_i,$$

which is just the x-coordinate of the total momentum. So invariance of the Lagrangian to translation implies conservation of total momentum.

2.4 Constraints

Sometimes it is desirable to place an algebraic constraint on the configurations. It may make sense to work with variables in Cartesian coordinates, for instance, even though it is known that system is defined on some sub-manifold. For example, one possible model for the planar pendulum is a free mass acted upon by gravity, but constrained such that its distance from the origin is fixed. Adding an algebraic constraint to a model described by a variational principle is easy to achieve, by simply adding the constraint relation to the action integral, pre-multiplies by a Lagrange multiplier.

Consider a system with general action integral (3), and suppose we wish to add an algebraic constraint of the form $g(x) = 0$, $g : \mathbb{R}^n \to \mathbb{R}$. Introducing the Lagrange multiplier $\lambda \in \mathbb{R}$, the action integral is modified as follows:

$$\Phi = \int L(x, \dot{x}) + \lambda g(x) \, dt.$$

Note that for motions $x(t)$ satisfying $g(x(t)) = 0$, the value of the action integral is unchanged. Also, considering λ as a new variable, the Euler-Lagrange equation for λ gives simply $g(x) = 0$, so the constraint relation appears explicitly in the equations of motion. In addition, the Euler-Lagrange equation for x is modified to include a term in the direction of the gradient of the constraint, and the combined Euler-Lagrange equations become:

$$-\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) + \frac{\partial L}{\partial x} + \lambda \frac{\partial g}{\partial x} = 0, \qquad g(x) = 0.$$

Example. We construct a model for the planar pendulum. For this example we consider a mass m with horizontal position $x(t)$ and vertical position $z(t)$. The gravitational acceleration is assumed to be uniform, such that the potential energy becomes $U = mgz$. The Lagrangian of the unconstrained mass is

$$L = T - U = \frac{m}{2}(\dot{x}^2 + \dot{z}^2) - mgz$$

However the Euler-Lagrange equations for this model just describe an mass falling to earth. To model a pendulum, we need to constrain the distance between the mass and the axis of the pendulum. We define the function $g(x, y) = x^2 + z^2 - 1$, such that $g = 0$ defines the manifold of possible configurations of the pendulum. The modified variational principle is

$$\phi(x, \lambda) = \int \frac{m}{2}(\dot{x}^2 + \dot{z}^2) - mgz + \lambda(x^2 + z^2 - 1) \, dt.$$

The associated Euler-Lagrange equations are

$$\begin{aligned} m\ddot{x} &= 2\lambda x \\ m\ddot{z} &= -mg + 2\lambda z \\ 0 &= x^2 + z^2 - 1. \end{aligned}$$

Example. As a final example of the concepts of this chapter, we consider a reduction of the two-body gravitational problem, in the context of the earth and the sun. The mass and position of the sun are denoted by M and $X \in \mathbb{R}^3$, respectively, and those of the earth by m and $x \in \mathbb{R}^3$. Newton's gravitational potential energy is given by

$$U(x, X) = -\frac{GmM}{\|x - X\|}.$$

The Lagrangian for the the two-body problem is

$$L = \frac{m}{2}\|\dot{x}\|^2 + \frac{M}{2}\|\dot{X}\|^2 + \frac{GmM}{\|x - X\|}.$$

The associated Euler-Lagrange equations can be computed to give

$$\begin{aligned} m\ddot{x} &= \frac{GmM}{\|x - X\|^3}(x - X) \\ M\ddot{X} &= -\frac{GmM}{\|x - X\|^3}(x - X) \end{aligned}$$

Adding these two equations together gives

$$m\ddot{x} + M\ddot{x} = 0$$

which is the statement of total momentum conservation. In an inertial coordinate system moving with our two bodies, we can take without loss of generality

$$m\dot{x} + M\dot{X} = 0$$

In terms of the (very) small parameter $\varepsilon = m/M$ we can write this as

$$\varepsilon\dot{x} + \dot{X} = 0.$$

To lowest order of approximation this amounts to $\dot{X} = 0$. Which means we can simplify the model to the heliocentric model

$$\ddot{x} = -\frac{GM}{\|x\|^3}x.$$

This system can also be derived from the variational principle with Lagrangian

$$L = \frac{m}{2}\|\dot{x}\|^2 + \frac{mMG}{\|x\|}.$$

To simplify even more we note that this Lagrangian is invariant to rotations of the coordinate system, i.e. if Q is an orthogonal matrix, then $L(Qx, Q\dot{x}) = L(x, \dot{x})$. We can construct a one-parameter group of rotation matrices as follows:

First, we remark that if B is a skew-symmetric matrix $B^T = -B$, then the matrix exponential

$$Q = \exp B = I + B + \frac{1}{2}B^2 + \frac{1}{3!}B^3 + \cdots$$

is an orthogonal matrix. (The matrix exponential is a complete function on the class of square matrices.) What is more, the vector cross product generates an isomorphism on R^3. Specifically, to every vector $b = (b_1, b_2, b_3) \in R^3$, the matrix

$$\hat{b} = \begin{bmatrix} 0 & -b_3 & b_2 \\ b_3 & 0 & -b_1 \\ -b_2 & b_1 & 0 \end{bmatrix}$$

effects the cross product with b, i.e. for any vector $y \in \mathbb{R}^3$:

$$b \times y = \hat{b}y$$

So to any vector $b \in \mathbb{R}^3$ we can associate a one-parameter family of orthogonal transformations

$$\phi^s(x) = \exp(s\hat{b})x$$

to which our Lagrangian is invariant.

Applying Noether's theorem to our variational principle yields the conserved quantity

$$I(x, \dot{x}) = \left(\frac{\partial L}{\partial \dot{x}}\right) \cdot \frac{d\phi^s(x)}{ds}\bigg|_{s=0} = m\dot{x} \cdot \frac{d}{ds}\bigg|_{s=0} \Big[(I + s\hat{b} + \ldots)x\Big] = m \cdot \dot{x}(b \times x)$$

Some manipulations yield

$$I = \dot{x} \cdot (b \times x) = -\dot{x} \cdot (x \times b) = -\dot{x} \cdot \hat{x}b = b \cdot \hat{x}\dot{x} = b \cdot (x \times \dot{x}).$$

Taking the time derivative of I shows that

$$\frac{dI}{dt} = 0 = b \cdot \frac{d}{dt}(x \times \dot{x}).$$

But recall that the vector b is arbitrary. In particular we could choose the canonical basis vectors $e_1 = (1, 0, 0)^T$, $e_2 = (0, 1, 0)^T$ and $e_3 = (0, 0, 1)^T$. This shows that the term in parentheses above must be zero, i.e. we have the conservation law

$$\frac{d}{dt}(x \times \dot{x}) = 0,$$

and $x \times \dot{x} = c = \mathit{const}$. It follows that the vector c is always normal to both the position and velocity vectors of the earth motion. But this means the motion is confined to a plane. Hence the problem could again be reformulated in $\mathbb{R}^2$.