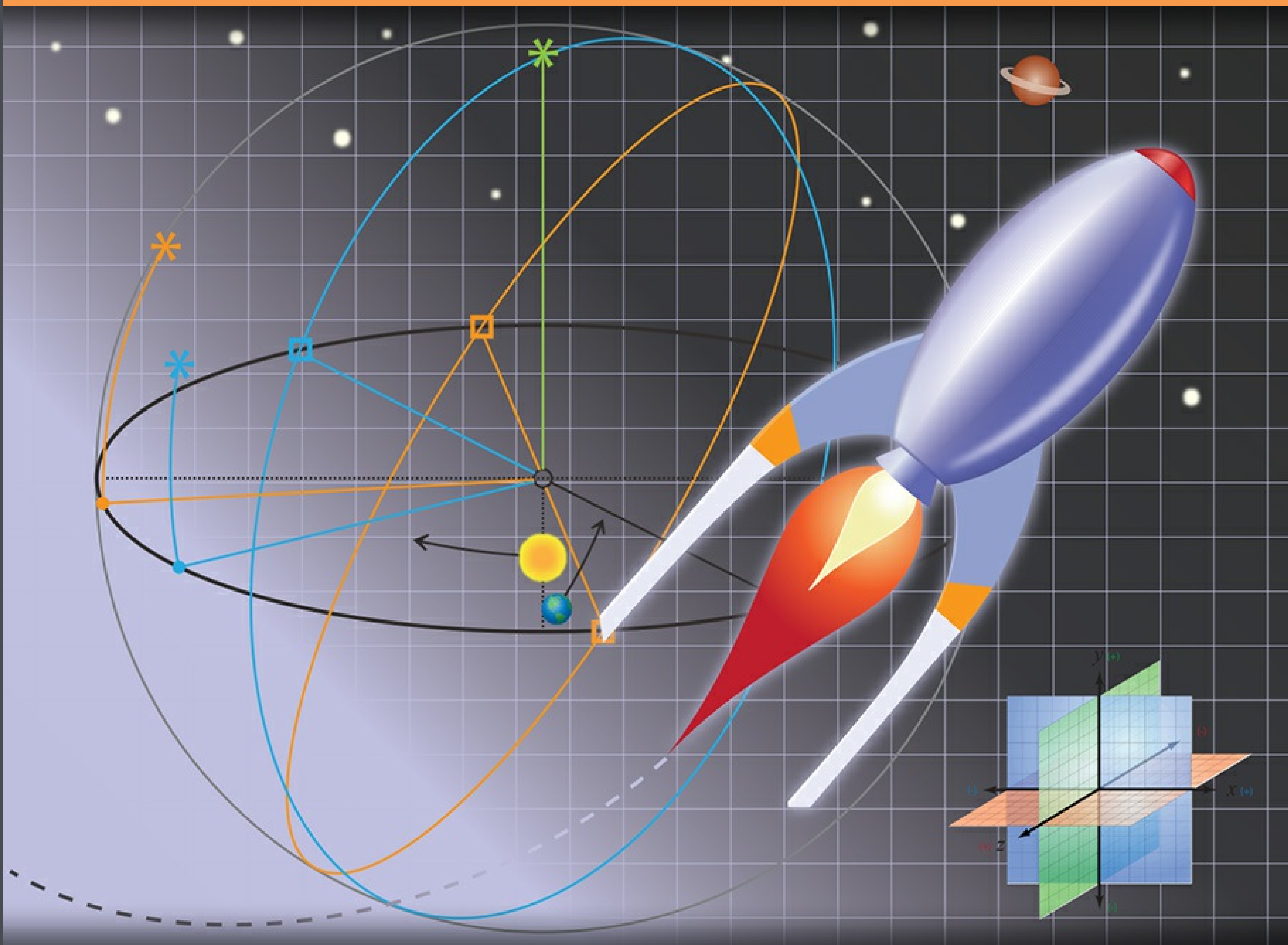
Modern Introductory Mechanics

Walter Wilcox



WALTER WILCOX

MODERN INTRODUCTORY MECHANICS

Modern Introductory Mechanics
2nd edition

ISBN 978-87-403-0855-6

CONTENTS

1	Mathematical Review	5
2	Newtonian Mechanics	45
3	Linear Oscillations	89
4	Nonlinear Oscillations	139
5	Gravitation	178
6	Calculus of Variations	190
7	Lagrangian and Hamiltonian Mechanics	235
	Endnotes	283

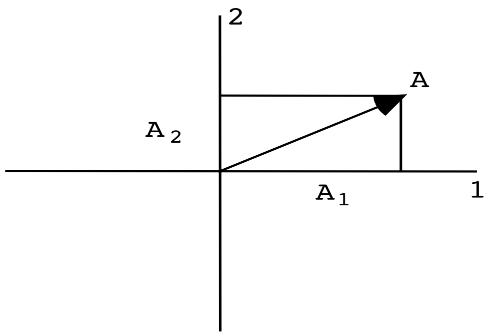
1 MATHEMATICAL REVIEW

TRIGONOMETRY

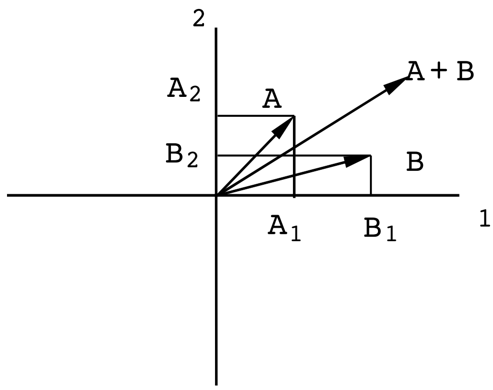
Mathematics is the language of physics, so we must all have a certain fluency. The first order of business is to remind ourselves of some basic relations from trigonometry.

$$\sin\theta = \frac{\text{opposite}}{\text{hypotenuse}},$$
$$\cos\theta = \frac{\text{adjacent}}{\text{hypotenuse}},$$
$$\tan\theta = \frac{\sin\theta}{\cos\theta} = \frac{\text{opposite}}{\text{adjacent}}.$$

For right now just think of a vector as something with both a magnitude and a direction:

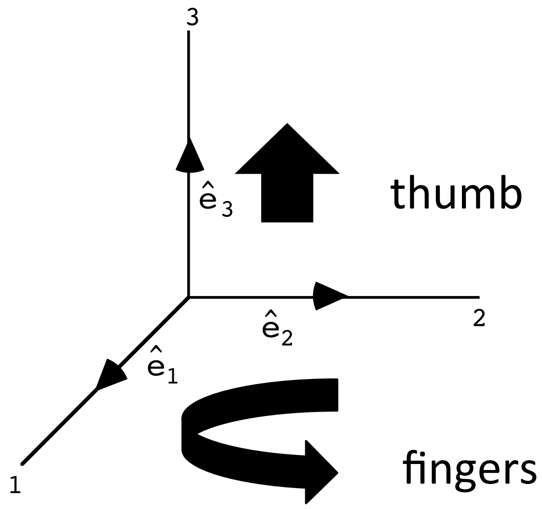


Vector notation: $\vec{A} = (A_1, A_2)$. Vector addition:



$$(A+B)_1 = A_1 + B_1 \ ; \ (A+B)_2 = A_2 + B_2 \ .$$

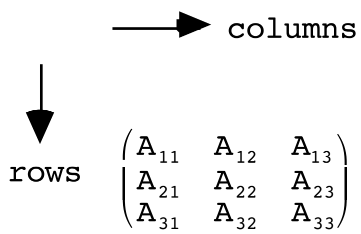
Unit Vectors:



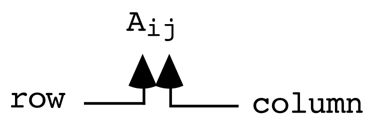
These are handy guys which point along the 1,2 or 3 directions with unit magnitude. We **always** choose a right-handed coordinate system in this class. The **right-hand rule** identifies such a system: curl the fingers of your right hand from $\hat{e}_1$ to $\hat{e}_2$ in the above figure; your thumb will point in the $\hat{e}_3$ direction.

Matrices

A matrix is a collection of entries which can be arranged in row/column form:



A single generalized matrix element is denoted:



Addition of matrices (number of rows and columns the same for both A and B matrices):

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} + \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} A_{11} + B_{11} & A_{12} + B_{12} \\ A_{21} + B_{21} & A_{22} + B_{22} \end{pmatrix}. \quad (1.1)$$

In more abstract language ("index notation") this is just

$$(\mathrm{A}+\mathrm{B})_{ij} = \mathrm{A}_{ij} + \mathrm{B}_{ij}, \tag{1.2}$$

where i and j are taking on all possible values independently. In the above equation i and j are said to be "free" indices. The free indices on one side of an equality must always be the same on the other side.

Multiplication of matrices. (Here we only require that the number of **columns** of A equal the number of **rows** of B):

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} A_{11}B_1 + A_{12}B_2 \\ A_{21}B_1 + A_{22}B_2 \end{pmatrix}. \tag{1.3}$$

Another example:

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}. \tag{1.4}$$

Notice that the result has the same number of **rows** as **A** and the same number of **columns** as B. In index language, these two examples can be written much more compactly as:

$$(AB)_i = \sum_{k=1}^{2} A_{ik}\, B_k \tag{1.5}$$

"dummy" index, k

$$(AB)_{ij} = \sum_{k=1}^{2} A_{ik}\, B_{kj} \tag{1.6}$$

"dummy" index, k

Note that dummy indices are ones which are summed over all of their values. Unlike free indices, which must be the same on both sides of an equation, dummy indices can appear on either side. Also notice that dummy indices always **appear twice** on a given side of an equation. These rules trip up many beginning students of mechanics.

For reference, here is a summary of the understood "index jockey" rules for index manipulations:

1. "Dummy" indices are those which are summed. Each such index always appears exactly **twice**. One can interpret this sum as matrix multiplication only if the indices can be placed directly next to each other. *Separate summation symbols must be used for independent summations.*

2. In general, one can not change the order of indices on an object, such as A_{ij} (Occasionally one knows effect of interchanging indices; see later comments on symmetric and antisymmetric matrices.)
3. Free indices are those that are unsummed. In general, each free index appears once on **both** sides of a given equation.

Identity matrix (3 × 3 context):

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The identity matrix is often simply written as the number "1", or is absent altogether in contexts where its presence is unambiguous. (Physicists must learn to read behind the lines for the meaning!)

We will need three additional matrix operations.

1. Inverse: $AA^{-1} = 1$ (1.7)

   The "1" on the right hand side here means the identity matrix. A Theorem from linear algebra establishes that $AA^{-1} = 1$ implies $A^{-1}A = 1$ (Can you prove this?) Finding A^{-1} in general is fairly complicated. For most of the matrices we will encounter, finding A^{-1} will be easy. (I'm thinking of rotation matrices, which will follow shortly.) Notice:

$$\begin{aligned} &(AB)^{-1}(AB) = 1, \\ &(AB)^{-1}A = B^{-1}, \\ \Rightarrow\ &(AB)^{-1} = B^{-1}A^{-1}. \end{aligned} \tag{1.8}$$

2. Transpose: $A^T_{ij} = A_{ji}$ (1.9)

   Examples of the transpose operation:

$$B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \qquad B^T = (B_1\ B_2)$$

"column matrix" "row matrix"

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},\ A^T = \begin{pmatrix} A_{11} & A_{21} \\ A_{12} & A_{22} \end{pmatrix}$$

Also

$$
\begin{aligned}
(AB)^T_{ij} &= (AB)_{ji} \,, \\
&= \sum_k A_{jk} B_{ki} \,, \\
&= \sum_k B^T_{ik} A^T_{kj} \,, \\
&= (B^T A^T)_{ij} \,, \\
\Rightarrow (AB)^T &= B^T A^T.
\end{aligned}
\tag{1.10}
$$

3. Determinant:

$$
\text{2x2 case: } \det \underbrace{\begin{pmatrix} A_{11} A_{12} \\ A_{21} A_{22} \end{pmatrix}}_{\text{also written as } \begin{vmatrix} A_{11} A_{12} \\ A_{21} A_{22} \end{vmatrix}} = A_{11}A_{22} - A_{12}A_{21}.
$$

$$
\text{x3 case: } \det \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} = A_{11}(A_{22}A_{33} - A_{23}A_{32}) - A_{12}(A_{21}A_{33} - A_{23}A_{31}) + A_{13}(A_{21}A_{32} - A_{22}A_{31}). \tag{1.12}
$$

Note that $\det(AB) = (\det A)(\det B)$ and that $\det A^T = \det A$ and $\det(A^{-1}) = 1/\det(A)$. We'll find a more elegant definition of the determinant later. An important point to realize is that the inverse of a matrix, A, exists only if $\det(A)$ is not zero.

An important point about linear algebra will also be called upon in later chapters. A system of linear (only $x_{1,2,3}$ appear, never $(x_{1,2,3})^2$ or higher powers) homogeneous (the right hand side of the following equations are zero) equations,

$$
\begin{cases}
A_{11}x_1 + A_{12}x_2 + A_{13}x_3 = 0, \\
A_{21}x_1 + A_{22}x_2 + A_{23}x_3 = 0, \\
A_{31}x_1 + A_{32}x_2 + A_{33}x_3 = 0,
\end{cases}
\tag{1.13}
$$

has a **nontrivial** ($x_{1,2,3}$ are not all zero!) solution *if and only if*

$$
\det \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} = 0. \tag{1.14}
$$

One often encounters matrices which are said to be symmetric or antisymmetric. A *symmetric* matrix is one for which

$$A^T = A \text{ or } A_{ji} = A_{ij}$$

An *antisymmetric* matrix has

$$A^T = -A \text{ or } A_{ji} = -A_{ij}$$

Matrices are not generally symmetric or antisymmetric, but such combinations can always be constructed. For example,

$$C = A + A^T,$$

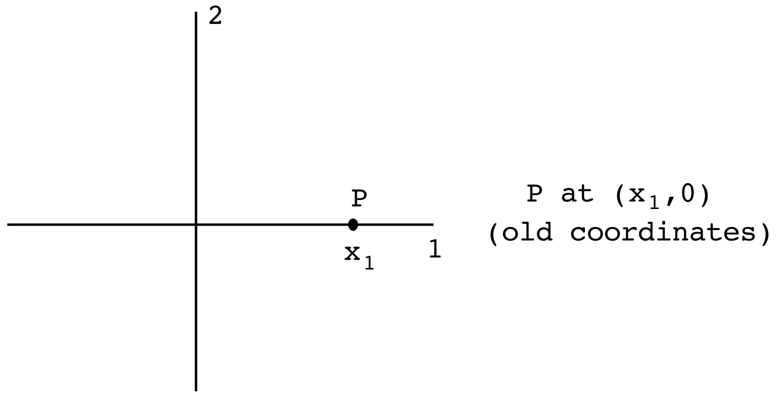
is a matrix combination which is symmetric ($C_{ij} = C_{ji}$) even though A is not itself. Likewise

$$D = A - A^T,$$

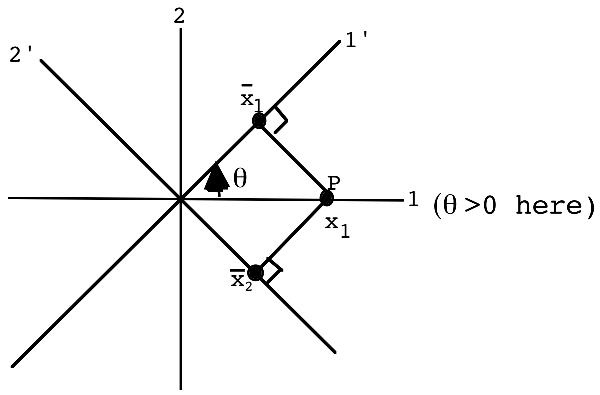
is antisymmetric ($D_{ij} = -D_{ji}$)

Orthogonal Transformations

Let us study a little about the above mentioned rotation matrices. Special case:



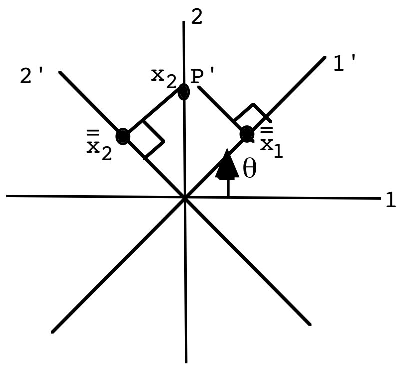
Let's do a "passive rotation," where we rotate the axes and not the point P. (Define $\theta > 0$ for counterclockwise rotations.)



$$P \text{ at } (\bar{x}_1, \bar{x}_2) = (x_1 \cos\theta, -x_1 \sin\theta)$$

$$= (x_1 \cos\theta, x_1 \cos(\theta + \frac{\pi}{2})) \text{ (new coordinates)}$$

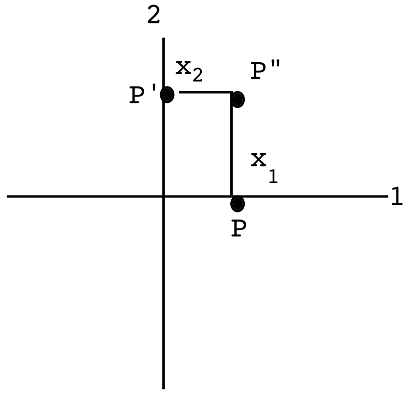
Another special case:



P' at $(0, x_2)$ (old coordinates)

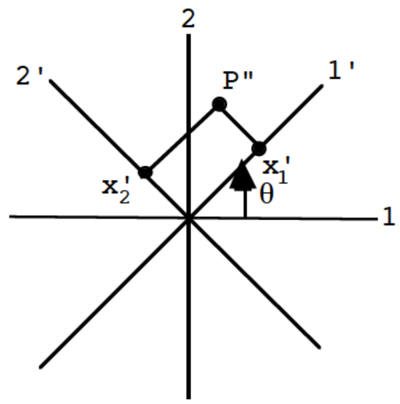
$$P' \text{ at } (\bar{\bar{x}}_1, \bar{\bar{x}}_2) = (x_2 \sin\theta, x_2 \cos\theta) = (x_2 \cos(\frac{\pi}{2} - \theta), x_2 \cos\theta) \text{ (new)}$$

Let's now put these two special cases together into the general case of rotation in a plane:



$$P'' \text{ at } (x_1, x_2) = \underbrace{(x_1, 0)}_{P} + \underbrace{(0, x_2)}_{P'} \text{ (old coordinates)}$$

Now characterize the point P" with respect to rotated axes:



$$P'' \text{ at } (x_1', x_2') = \underbrace{(\bar{x}_1, \bar{x}_2)}_{P} + \underbrace{(\bar{\bar{x}}_1, \bar{\bar{x}}_2)}_{P'}$$

$$= (x_1 \cos\theta, x_1 \cos(\theta + \frac{\pi}{2})) + (x_2 \cos(\frac{\pi}{2} - \theta), x_2 \cos\theta) \text{ (new coordinates)}$$

Therefore, comparing the old and new descriptions, we have

$$\left.\begin{aligned} x_1' &= x_1 \cos\theta + x_2 \cos(\frac{\pi}{2} - \theta), \\ x_2' &= x_1 \cos(\frac{\pi}{2} + \theta) + x_2 \cos\theta. \end{aligned}\right\} \quad (1.15)$$

To simplify the notation, introduce the concept of **direction cosines**:

$\cos(x_i',x_j)$ = cosine of the angle between the x_i' and x_j axes.

↑ start measurement of angle

The angle may be measured in a clockwise or counterclockwise sense but must be consistent. We will adopt counterclockwise as in the above examples. For the above rotation,

$$\begin{cases} \cos(x_1',x_1) = \cos\theta, \\ \cos(x_1',x_2) = \cos(2\pi - \frac{\pi}{2} + \theta) = \cos(\frac{\pi}{2} - \theta), \\ \cos(x_2',x_1) = \cos(\frac{\pi}{2} + \theta), \\ \cos(x_2',x_2) = \cos\theta. \end{cases}$$

So, our general rotation may be written

$$\begin{cases} x_1' = x_1\cos(x_1',x_1) + x_2\cos(x_1',x_2), \\ x_2' = x_1\cos(x_2',x_1) + x_2\cos(x_2',x_2). \end{cases} \tag{1.16}$$

In three dimensions, we have

$$\begin{cases} x_1' = x_1 \cos(x_1', x_1) + x_2 \cos(x_1', x_2) + x_3 \cos(x_1', x_3), \\ x_2' = x_1 \cos(x_2', x_1) + x_2 \cos(x_2', x_2) + x_3 \cos(x_2', x_3), \\ x_3' = x_1 \cos(x_3', x_1) + x_2 \cos(x_3', x_2) + x_3 \cos(x_3', x_3). \end{cases} \tag{1.17}$$

I wrote this out explicitly so that you would notice a pattern in the indices which indicates the above can be written in matrix language. Let us define

$$\lambda_{ij} \equiv \cos(x_i', x_j). \tag{1.18}$$

If we arbitrarily choose to represent the position (x_1, x_2, x_3) as a column matrix, the above relationship can be written,

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}. \tag{1.19}$$

or more abstractly

$$x' = \lambda x, \tag{1.20}$$

where it is understood x' and x are column matrices and λ is a 3×3 matrix. The easiest way to see this is just to do the matrix multiplication on the right-hand side to see that it reproduces the rotation.

Written out in explicit index notation, the above relation may also be written as

$$x_i' = \sum_{j=1}^{3} \lambda_{ij} x_j \tag{1.21}$$

where it is understood that the free index "i" takes on values 1,2,3.

Question: What if we knew the x_i' instead of the x_j? In other words, what is the inverse relationship between these quantities? Using matrix notation we have

$$\begin{aligned} & x' = \lambda x \\ \Rightarrow\ & \lambda^{-1} x' = \lambda^{-1} \lambda x = x, \end{aligned}$$

so[1]

$$x = \lambda^{-1} x'$$

or

$$x_i = \sum_{j=1}^{3} \lambda_{ij}^{-1} x_j' \,. \tag{1.22}$$

As stated above, there is a simple way of getting λ^{-1} (given λ) for rotations. Consider:

$$\sum_i x_i'^{\,2} = x_1'^{\,2} + x_2'^{\,2} + x_3'^{\,2}.$$

By definition

$$\begin{aligned}\sum_i x_i'^{\,2} &= \sum_i \left(\sum_j \lambda_{ij} x_j\right)\left(\sum_k \lambda_{ik} x_k\right) \\ &= \sum_{i,j,k} \lambda_{ij}\lambda_{ik} x_j x_k \,.\end{aligned} \tag{1.23}$$

Require: $\sum_i x_i'^{\,2} = \sum_i x_i^{\,2}$ (length preserved)

$$\Rightarrow \sum_{i,j,k} \lambda_{ij}\lambda_{ik} x_j x_k = \sum_i x_i^{\,2}. \tag{1.24}$$

The coefficient of each $x_j x_k$ term (j,k independent) must be the same on both sides:

$$\Rightarrow \begin{cases} \sum_i \lambda_{ij}\lambda_{ik} = 0, j \neq k, \\ \sum_i \lambda_{ij}\lambda_{ik} = 1, j = k. \end{cases} \tag{1.25}$$

More simply

$$\sum_i \lambda_{ij}\lambda_{ik} = \delta_{jk} \,, \tag{1.26}$$

↑ "Kronecker delta"

where

$$\delta_{ij} \equiv \begin{cases} 0, & i \neq j \,, \\ 1, & i = j \,. \end{cases} \tag{1.27}$$

Notice that the above is not a matrix statement (why?), but that the right-hand side, δ_{jk},, is just index notation for the identity matrix. However, it may be cast into matrix language with the help of the transpose:

$$\lambda^T_{ji} = \lambda_{ij} \;,$$

$$\Rightarrow \quad \sum_i \lambda^T_{ji}\lambda_{ik} = \delta_{jk} \text{ or } \lambda^T\lambda = 1. \qquad (1.28)$$

This last statement establishes that

$$\lambda^T = \lambda^{-1}, \qquad (1.29)$$

for rotation matrices. In fact, the above equation is sometimes taken to **define** such transformations. Actually, the types of coordinate transformations allowed by this equation are a bit more general than simply rotations, as we will see shortly. The name **orthogonal transformations** are usually used for real matrices which satisfy $\lambda^T = \lambda^{-1}$, and therefore preserve the lengths of vectors. Using the above, the relationship between the x'_i and the x_i may now be cast as

$$x' = \lambda x \quad \Leftrightarrow \quad x = \lambda^T x',$$

in matrix language, or

$$x'_i = \sum_j \lambda_{ij}\, x_j \quad \Leftrightarrow \quad x_i = \sum_j \lambda_{ji} x'_j \,,$$

in component language. It is important to note that in place of

$$\sum_i \lambda_{ij}\lambda_{ik} = \delta_{jk} \,, \tag{1.30}$$

we may also write

$$\sum_i \lambda_{ji}\lambda_{ki} = \delta_{jk} \,. \tag{1.31}$$

Deriving one from the other will be a homework problem.

How many independent λ_{ij} elements are there? Notice that the equation $\sum_i \lambda_{ij}\lambda_{ik} = \delta_{jk}$ is symmetric under $j \times k$, so that this system of equations actually represents 6, not 9 equations. (The number of independent elements of a real symmetric 3×3 matrix are 6.) This means the number of independent λ_{ij} is 9 – 6 = 3. This makes sense from the physical point of view of rotations in three dimensions, which require three independent angles, in general.

Let me make three additional points about the λ_{ij}:

1. $$x'' = (\lambda_2\lambda_1)x = \lambda_3 x, \tag{1.32}$$

   λ_3 is an orthogonal transformation if λ_1 and λ_2 also are. (We can view this as a transformation λ_1 followed by a second transformation λ_2.) That is, the product of orthogonal matrices is also an orthogonal matrix. The proof of this statement is left as a problem.

2. $$\begin{cases} x' = \lambda_1 x, \\ x'' = \lambda_2 x' = \lambda_2\lambda_1 x. \end{cases} \tag{1.33}$$

   Do the rotations in the opposite order:

   $$\begin{cases} \bar{x} = \lambda_2 x, \\ \bar{\bar{x}} = \lambda_1 \bar{x} = \lambda_1\lambda_2 x \end{cases} \tag{1.34}$$

   In general $\lambda_1\lambda_2 \neq \lambda_2\lambda_1$, ("noncommutative") and the order in which rotations are performed is important.

3. As stated above, we started out describing rotation, but the λ_{ij} can also represent operations which are **not** rotations. For example,

$$\lambda_{inv} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

This satisfies $\lambda_{inv}^{T}\lambda_{inv} = 1$, but describes a spatial **inversion**. This is not an operation which can be physically carried out on objects in general. We will consider such orthogonal transformations only sparingly.

All orthogonal transformations satisfy

$$\det\lambda = \begin{cases} +1, \text{rotations} \\ -1, \text{inversions} \end{cases} \tag{1.35}$$

λ_{ij} which have det $\lambda = -1$ are called **proper orthogonal transformations**, and those that have det $\lambda = 1$ are called **improper orthogonal transformations**. In general, improper orthogonal transformations can be thought of as an inversion plus an arbitrary rotation since

$$\det(\lambda_{inversion}\ \lambda_{rotation}) = \det(\lambda_{inversion})(\det\lambda_{rotation}),$$

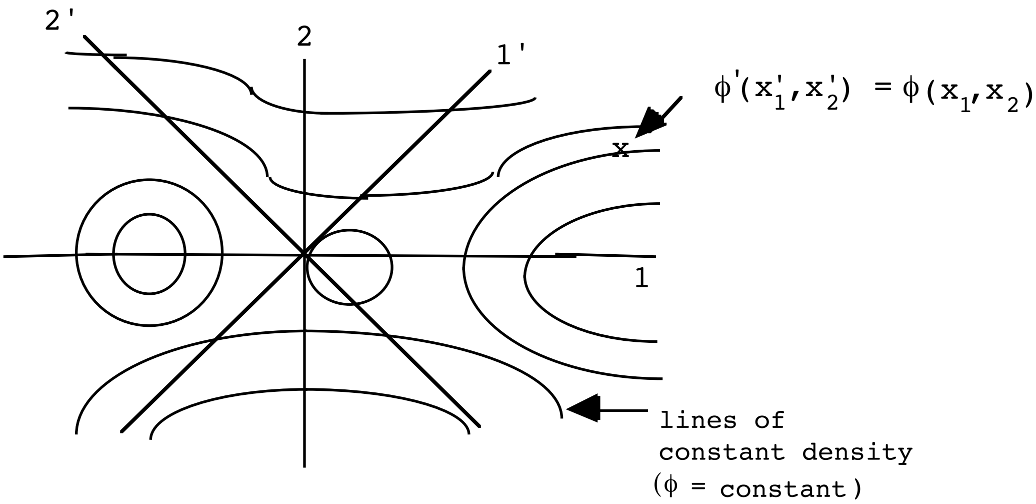
which means the combination is still improper. Proving that $\det(\lambda) = \pm 1$ for orthogonal transformations is left as a homework problem at the end of the chapter.

Scalar and Vector Fields

You probably have an intuitive feeling as to the meaning of a "scalar". It is something which is unchanged under a change in coordinates. Under rotations, a scalar field behaves as

$$\text{Scalar:} \qquad \phi'(x') = \phi(x), \tag{1.36}$$

where $x'_i = \sum_j \lambda_{ij}x_j$. It is understood that the x and x' represent the same point described in two coordinate systems. See the below:



There are also **pseudoscalars**, which behave like scalars under rotations, but acquire an extra minus sign under inversions.

$$\text{Pseudoscalar: } \phi'(x') = (\det\lambda)\,\phi(x). \tag{1.37}$$

Although we've informally defined a vector as a quantity with both magnitude and direction, we need a quantitative definition. Here it is:

$$\text{Vector:} \qquad A'_i(x') = \sum_j \lambda_{ij} A_j(x). \tag{1.38}$$

The λ_{ij} are elements of an orthogonal transformation. Obviously, the transformation law for the coordinates x'_i, x_i is the template for this definition.

There are also **pseudovectors**, which transform as

$$\text{Pseudovector:} \quad A'_i(x') = (\det\lambda) \sum_j \lambda_{ij} A_j(x). \tag{1.39}$$

These quantities do not change direction under inversions, as vectors do. We will see an example of a pseudovector shortly.

Vector Algebra and Scalar Differentiation

Given two vectors, $\vec{A}$ and $\vec{B}$, we may form either a scalar or a vector. Let's study these two possibilities.

$$\underbrace{\vec{A} \cdot \vec{B}}_{\text{denoted}} = \sum_i \underbrace{A_i B_i}_{\text{defined}}. \tag{1.40}$$

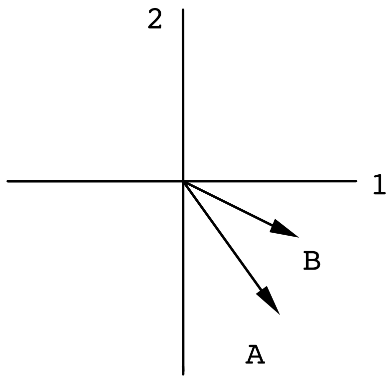
It is also called the "dot product." Here is the proof it is a scalar:

$$\begin{cases} A'_i = \sum_j \lambda_{ij} A_j, \\ B'_i = \sum_k \lambda_{ik} B_k. \end{cases}$$

$$\Rightarrow \vec{A}' \cdot \vec{B}' = \sum_i A'_i B'_i = \sum_i \left(\sum_j \lambda_{ij} A_j\right)\left(\sum_k \lambda_{ik} B_k\right)$$

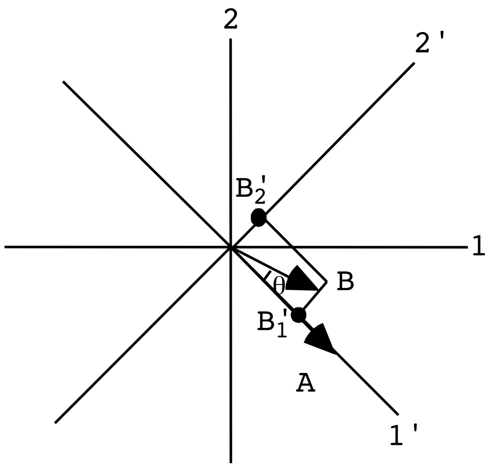
$$= \sum_{j,k} \underbrace{\left(\sum_i \lambda_{ij}\lambda_{ik}\right)}_{\delta_{jk}} A_j B_k = \sum_j A_j B_j = \vec{A} \cdot \vec{B}.$$

Notice I changed the order of the sums in some intermediate expressions, which is always allowed for a finite number of terms. Another, more geometrical, definition of the scalar product follows from the diagrams below:



$$A \cdot B = A_1 B_1 + A_2 B_2$$

Same situation with a new coordinate system:



Therefore

$$\vec{A}' \cdot \vec{B}' = A_1' B_1', \quad |\vec{A}| = \sqrt{\sum_i A_i^2}, |\vec{B}| = \sqrt{\sum_i B_i^2}.$$

$A_1' = |\vec{A}|$ here, so $\vec{A}' \cdot \vec{B}' = |\vec{A}||\vec{B}| \underbrace{\left(\frac{B_1'}{|\vec{B}|}\right)}_{=\ \cos\theta}$. Then given $\vec{A} \cdot \vec{B} =$

$\vec{A}' \cdot \vec{B}'$,

$$\Rightarrow \quad \vec{A} \cdot \vec{B} = |\vec{A}|\,|\vec{B}|\cos\theta. \tag{1.41}$$

The vector product (or cross product) of $\vec{A}$ and $\vec{B}$ is defined as:

$$(\vec{A} \times \vec{B})_i = \sum_{j,k} \varepsilon_{ijk} A_j B_k. \tag{1.42}$$

denoted defined

Define the "permutation symbols", ε_{ijk}:

$$\begin{cases} \varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = 1, \\ \varepsilon_{321} = \varepsilon_{213} = \varepsilon_{132} = -1, \\ \text{all other } \varepsilon_{ijk}\text{'s} = 0. \end{cases} \tag{1.43}$$

Important property: $\varepsilon_{ijk} = \varepsilon_{jki} = \varepsilon_{kij}$, $\varepsilon_{ijk} = -\varepsilon_{ikj} = -\varepsilon_{jik}$. This is known as the cyclic property of the permutation symbols. Using the ε_{ijk}, a more formal definition of the determinant of a 3×3 matrix, **A**, can be written

$$\det A \equiv \sum_{n,\ell,m} \varepsilon_{n\ell m} A_{1n} A_{2\ell} A_{3m}. \tag{1.44}$$

We will use this definition shortly. In order to see how $\vec{A} \times \vec{B}$ transforms, we will need the following identity for orthogonal transformations,

$$\sum_{j,k} \varepsilon_{ijk} \lambda_{j\ell} \lambda_{km} = (\det\lambda) \sum_{n} \varepsilon_{n\ell m} \lambda_{in}, \tag{1.45}$$

This identity will be proven later (end of the chapter.) Let's begin:

$$(\bar{A}' \times \bar{B}')_i = \sum_{j,k} \varepsilon_{ijk} A'_j B'_k = \sum_{j,k,\ell,m} \varepsilon_{ijk} \lambda_{j\ell} A_\ell \lambda_{km} B_m$$

$$= \sum_{j,k,\ell,m} \varepsilon_{ijk} \lambda_{j\ell} \lambda_{km} A_\ell B_m. \tag{1.46}$$

Now use (1.45) on the right-hand side of (1.46):

$$\Rightarrow (\bar{A}' \times \bar{B}')_i = (\det\lambda) \sum_{n,\ell,m} \varepsilon_{n\underset{(j)}{\ell}\underset{(k)}{m}} \lambda_{in} A_{\underset{(j)}{\ell}} B_{\underset{(k)}{m}} = (\det\lambda) \sum_{n,j,k} \lambda_{in} (\varepsilon_{njk} A_j B_k)$$

notice I am
indicating a change
in the dummy indices!

$$= (\det\lambda) \sum_{n} \lambda_{in} (\bar{A} \times \bar{B})_n.$$

Therefore we have that

$$(\bar{A}' \times \bar{B}')_i = (\det\lambda) \sum_{n} \lambda_{in} (\bar{A} \times \bar{B})_n. \tag{1.47}$$

The extra factor of (det λ) indicates that $\bar{A} \times \bar{B}$ actually transforms as a **pseudovector** (assuming $\bar{A}$ and $\bar{B}$ are themselves vectors.)

As an aside, there are many useful identities we can form with the ε_{ijk} by multiplying and summing on indices. For example, let's evaluate: $\sum_{k} \varepsilon_{ijk} \varepsilon_{\ell mk}$. Following the index jockey rules, the only objects we can build out of δ_{ij} and the ε_{ijk} which have 4 free indices are:

$$\delta_{ij}\delta_{\ell m},\quad \delta_{im}\delta_{\ell j},\quad \delta_{i\ell}\delta_{jm}.$$

Therefore, we must have

$$\sum_{k} \varepsilon_{ijk} \varepsilon_{\ell mk} = C_1 \delta_{ij}\delta_{\ell m} + C_2 \delta_{im}\delta_{\ell j} + C_3 \delta_{i\ell}\delta_{jm},$$

where the $C_{1,2,3}$ are unknown constants. One immediately can see that $C_1 = 0$. (Why?) Now multiply both sides of the above by δ_{ij} and sum over i and j:

$$\Rightarrow 0 = C_2\delta_{\ell m} + C_3\delta_{\ell m} \Rightarrow C_2 = -C_3.$$

Now multiply by δ_{jm} and sum over j and m::

$$\Rightarrow \sum_{j,k} \varepsilon_{ijk}\varepsilon_{\ell jk} = -2C_2\delta_{i\ell}.$$

Then, taking a special case, say $i = \ell = 1$, we then see that

$$\sum_{j,k} \varepsilon_{1jk}\varepsilon_{1jk} = 2,$$

so therefore $C_2 = -1$ and

$$\sum_{j,k} \varepsilon_{ijk}\varepsilon_{\ell jk} = 2\delta_{i\ell}, \tag{1.48}$$

$$\Rightarrow \sum_{k} \varepsilon_{ijk}\varepsilon_{\ell mk} = \delta_{i\ell}\delta_{jm} - \delta_{im}\delta_{\ell j}. \tag{1.49}$$

From this last expression, we easily see that

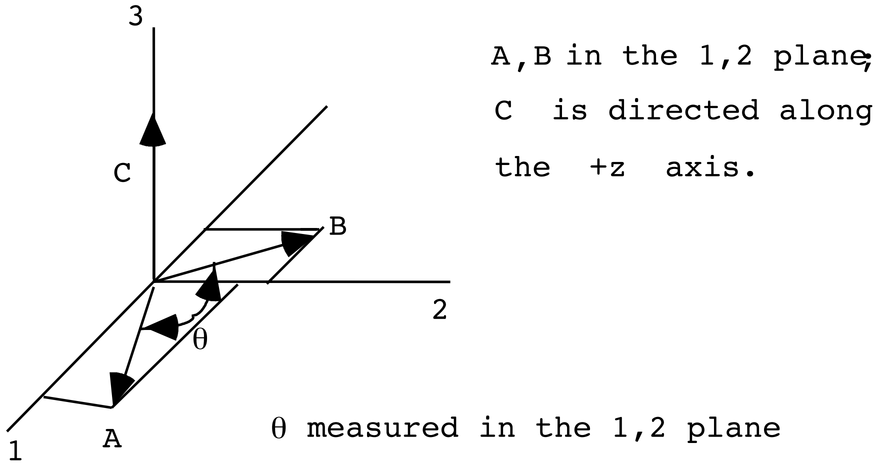
$$\sum_{i,j,k} \varepsilon_{ijk}\varepsilon_{ijk} = \sum_{i,j}(\delta_{ii}\delta_{jj} - \delta_{ij}\delta_{ij}) = 3 \times 3 - 3 = 6. \tag{1.50}$$

There is an alternate definition of the vector product as well. It says:

$$\vec{A} \times \vec{B} = |\vec{A}||\vec{B}|\,|\sin\theta|\,\hat{n}, \tag{1.51}$$

where
$|\vec{A}|$ = magnitude of $\vec{A}$
$|\vec{B}|$ = " " $\vec{B}$
θ = angle between $\vec{A}$ and $\vec{B}$
$\hat{n}$ = unit vector using the right hand rule

It doesn't matter in which direction θ is measured. This angle is measured in the plane defined by $\vec{A}$ and $\vec{B}$ if they aren't colinear. This definition is illustrated below.



Let's now reconnect to the useful concept of unit vectors in this context. The scalar products are summarized as

$$\hat{e}_i \cdot \hat{e}_j = \delta_{ij}. \tag{1.52}$$

Now, using the second definition of the cross product, we find

$$\hat{e}_1 \times \hat{e}_2 = |\sin 90^\circ|\,\hat{e}_3 = \hat{e}_3.$$

Similarly,

$$\hat{e}_3 \times \hat{e}_1 = \hat{e}_2 \ , \ \hat{e}_2 \times \hat{e}_3 = \hat{e}_1.$$

Notice that order matters here since

$$\hat{e}_2 \times \hat{e}_1 = -\hat{e}_3,$$

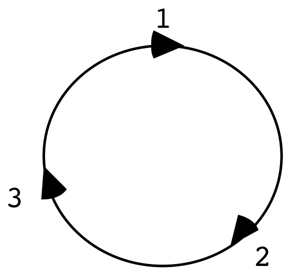
and similarly for the other non zero products. Also

$$\hat{e}_1 \times \hat{e}_1 = \hat{e}_2 \times \hat{e}_2 = \hat{e}_3 \times \hat{e}_3 = 0.$$

We can summarize all properties of unit vectors under the cross product as

$$\hat{e}_i \times \hat{e}_j = \sum_k \varepsilon_{ijk}\,\hat{e}_k, \tag{1.53}$$

and with a cyclic order wheel for permutation symbol indices:



$\Rightarrow \varepsilon_{123} = 1,$

$\Rightarrow \varepsilon_{213} = -1,$ etc.

I can now show the first and second definitions of the cross product are equivalent by resolving the components of $\vec{A}$ and $\vec{B}$ into unit vectors. From the second definition,

$$\vec{A} \times \vec{B} = (A_1\hat{e}_1 + A_2\hat{e}_2 + A_3\hat{e}_3) \times (B_1\hat{e}_1 + B_2\hat{e}_2 + B_3\hat{e}_3)$$

$$= (A_2B_3 - A_3B_2)\,\hat{e}_1 + (A_3B_1 - A_1B_3)\,\hat{e}_2 + (A_1B_2 - A_2B_1)\,\hat{e}_3.$$

On the other hand, from the first definition,

$$(\vec{A} \times \vec{B})_1 = \sum_{j,k} \varepsilon_{1jk} A_j B_k = e_{123}A_2B_3 + e_{132}A_3B_2 = (A_2B_3 - A_3B_2),$$

$$(\vec{A} \times \vec{B})_2 = \varepsilon_{231}A_3B_1 + \varepsilon_{213}A_1B_3, = (A_3B_1 - A_1B_3),$$

$$(\vec{A} \times \vec{B})_3 = \varepsilon_{312}A_1B_2 + \varepsilon_{321}A_2B_1, = (A_1B_2 - A_2B_1).$$

Comparing, we see the two definitions agree. By forming the determinant and comparing with the above definitions for the cross product, one may also show that

$$\vec{A} \times \vec{B} = \begin{vmatrix} \hat{e}_1\hat{e}_2\hat{e}_3 \\ A_1A_2A_3 \\ B_1B_2B_3 \end{vmatrix}, \tag{1.54}$$

where the right hand side is understood as a symbolic determinant.

From either definition of the cross product we have that

$$\vec{A} \times \vec{B} = -\vec{B} \times \vec{A}, \tag{1.55}$$

as well as

$$\vec{A} \times \left(\vec{B} \times \vec{C}\right) = \vec{B}\left(\vec{A} \cdot \vec{C}\right) - \vec{C}\left(\vec{A} \cdot \vec{B}\right). \tag{1.56}$$

The last statement follows from Eq.(1.49) above. In addition, there is the cyclic rule,

$$\vec{A} \cdot \left(\vec{B} \times \vec{C}\right) = \vec{B} \cdot \left(\vec{C} \times \vec{A}\right) = \vec{C} \cdot \left(\vec{A} \times \vec{B}\right), \tag{1.57}$$

which is a direct result of the cylclic nature of the permutation symbols. Given that $\vec{A}$, $\vec{B}$, and $\vec{C}$ are vectors, the quantity $\vec{A} \cdot (\vec{B} \times \vec{C})$ is an example of a pseudoscalar. (This is part of a problem.)

So, we know how to add and multiply vectors together. Question: does it make sense to define division? Consider

$$\vec{A} \times \vec{B} = \vec{C}.$$

Given that $\vec{A}$ and $\vec{C}$ and are known quantities, this equation clearly does not uniquely define $\vec{B}$ since any component of $\vec{B}$ along $\vec{A}$ will not contribute to the cross product. So the answer is in general, no. However, if we know both $\vec{A} \cdot \vec{B}$ and $\vec{A} \times \vec{B}$, we **can** solve for $\vec{B}$, say. (We will have a homework problem along these lines.)

Differentiation of vectors with respect to scalars leads to new vectors.

Given:

$$A'_i = \sum_j \lambda_{ij} A_j, \quad s \text{ (a scalar).}$$

Then

$$\frac{dA'_i}{ds'} = \sum_j \lambda_{ij} \frac{dA_j}{ds'} = \sum_j \lambda_{ij} \frac{dA_j}{ds},$$

$$\Rightarrow \frac{dA'_i}{ds'} = \sum_j \lambda_{ij} \left(\frac{dA_j}{ds} \right). \qquad (1.58)$$

A fundamental postulate of mechanics is that the time parameter transforms as a scalar quantity. Given this, these derivative transformations imply that velocity and acceleration are vectors.

Alternate Coordinate Systems

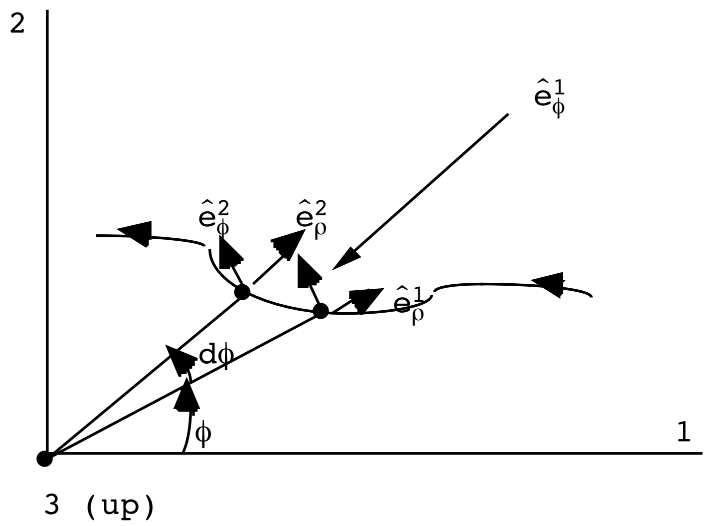
The radius vector, $\vec{r}$, is easy to characterize in the three most common coordinate systems. We have ($\dot{\vec{r}} \equiv \frac{d\vec{r}}{dt}$; t is considered a scalar):

Rectangular: $\vec{r} = x_1\hat{e}_1 + x_2\hat{e}_2 + x_3\hat{e}_3$ (all $\dot{\hat{e}}_i = 0$)

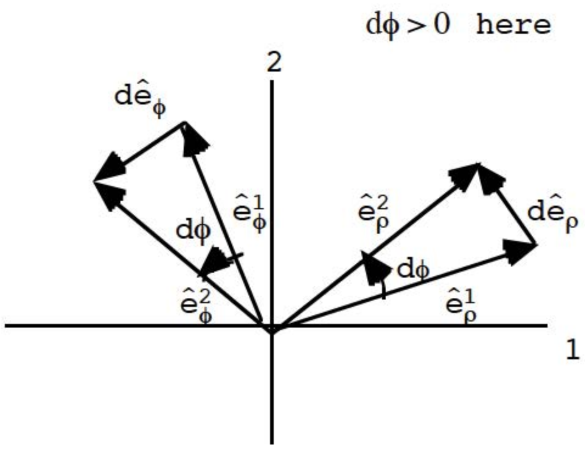
Cylindrical: $\vec{r} = \rho\hat{e}_\rho + z\hat{e}_z$ (only $\dot{\hat{e}}_z = 0$)
Additional unit vector: $\hat{e}_\phi$

Spherical: $\vec{r} = r\hat{e}_r$
Additional unit vectors: $\hat{e}_\phi$, $\hat{e}_\theta$
$\vec{v} = \dot{\vec{r}}$, $\vec{a} = \dot{\vec{v}} = \ddot{\vec{r}}$.

As an exercize, let's work out $\vec{v}$ and $\vec{a}$ in cylindrical coordinates. Follow a particle's trajectory at two closely spaced moments in time:



Displace the unit vectors to the origin:



Can now see that ($d\hat{e}_\rho = \hat{e}^2_\rho - \hat{e}^1_\rho$)

$$|d\hat{e}_\rho| = |d\phi|,$$

$$d\hat{e}_\rho = \text{direction} \times \text{length},$$
$$= \hat{e}_\phi \, d\phi.$$

Likewise $(d\hat{e}_\phi = \hat{e}^2_\phi - \hat{e}^1_\phi)$

$$|d\hat{e}_\phi| = |d\phi|,$$
$$d\hat{e}_\phi = -\hat{e}_\rho \, d\phi.$$

Therefore

$$\frac{d\hat{e}_\rho}{dt} = \frac{d\phi}{dt}\hat{e}_\phi \text{ or } \dot{\hat{e}}_\rho = \dot{\phi}\hat{e}_\phi, \tag{1.59}$$

and

$$\frac{d\hat{e}_\phi}{d\phi} = -\frac{d\phi}{dt}\hat{e}_\rho \text{ or } \dot{\hat{e}}_\phi = -\dot{\phi}\hat{e}_\rho. \tag{1.60}$$

Thus in cylindrical coordinates,

$$\vec{v} = \dot{\rho}\hat{e}_\rho + \rho\dot{\hat{e}}_\rho + \dot{z}\hat{e}_z,$$
$$\Rightarrow \vec{v} = \dot{\rho}\hat{e}_\rho + \rho\dot{\phi}\hat{e}_\phi + \dot{z}\hat{e}_z. \tag{1.61}$$

and

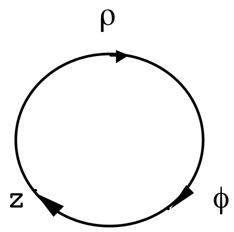
$$\vec{a} = \ddot{\rho}\hat{e}_\rho + \dot{\rho}\dot{\hat{e}}_\rho + \dot{\rho}\dot{\phi}\hat{e}_\phi + \rho\ddot{\phi}\hat{e}_\phi + \rho\dot{\phi}\dot{\hat{e}}_\phi + \ddot{z}\hat{e}_z,$$
$$\Rightarrow \vec{a} = (\ddot{\rho} - \rho\dot{\phi}^2)\hat{e}_\rho + (2\dot{\rho}\dot{\phi} + \rho\ddot{\phi})\hat{e}_\phi + \ddot{z}\hat{e}_z. \tag{1.62}$$

Since the unit vectors in rectangular coordinates do not change, it is very useful to know the decomposition of the unit vectors in other coordinate systems in terms of rectangular coordinates.

Cylindrical:
$$\begin{cases} \hat{e}_\rho = (\cos\phi, \sin\phi, 0), \\ \hat{e}_\phi = (-\sin\phi, \cos\phi, 0), \\ \hat{e}_z = (0, 0, 1). \end{cases} \tag{1.63}$$

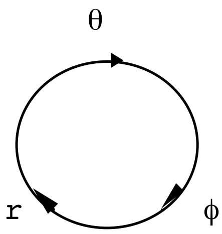
Spherical:
$$\begin{cases} \hat{e}_r = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta), \\ \hat{e}_\theta = (\cos\theta\cos\phi, \cos\theta\sin\phi, -\sin\theta), \\ \hat{e}_\phi = (-\sin\phi, \cos\phi, 0). \end{cases} \tag{1.64}$$

Also



$$\Rightarrow \hat{e}_\rho \times \hat{e}_\phi = \hat{e}_z,$$

cyclically in cylindrical coordinates and

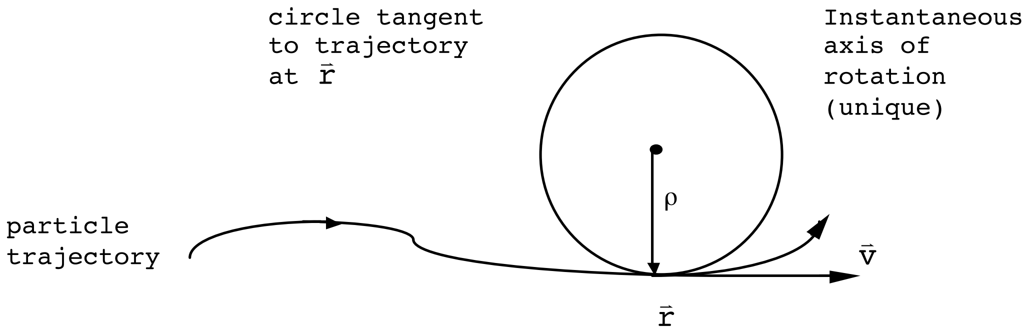


$$\Rightarrow \hat{e}_\theta \times \hat{e}_\phi = \hat{e}_r,$$

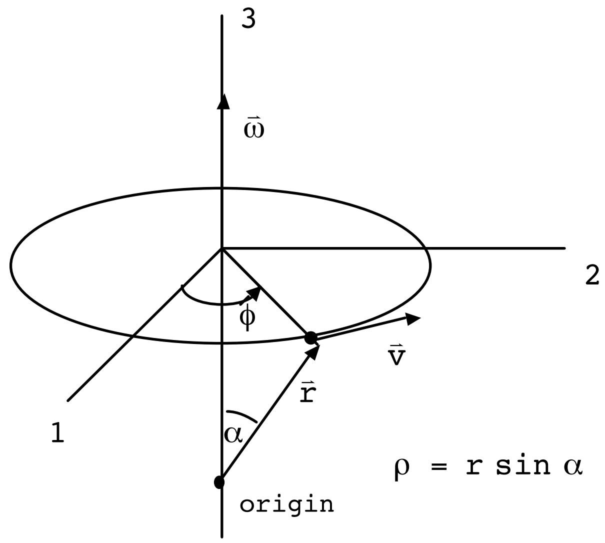
cyclically in spherical coordinates.

Angular Velocity

Another important concept for particle motion is angular velocity. Here we will rely mainly on intuition to understand the concept. First, identify an instantaneous circular path:



Flip this picture on it's side. Choose a cylindrical coordinate system with the origin along the axis of rotation. Use this origin to define the radius vector, $\vec{r}$.



Let $\vec{\omega}$ be directed co-linear to the axis of rotation, which is the third axis above. Angular velocity and velocity are related by the equation

$$\vec{v} = \vec{\omega} \times \vec{r}. \tag{1.65}$$

Notice, using a cylindrical coordinate system,

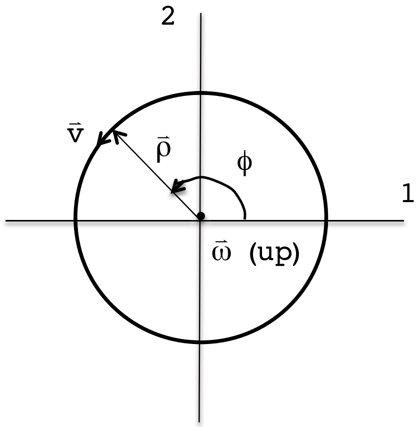
$$\vec{\omega} = \omega_z \hat{e}_z, \quad \vec{v} = \rho\dot{\phi}\hat{e}_\phi, \quad \vec{r} = \rho\hat{e}_\rho + z\hat{e}_z,$$

$$\Rightarrow \vec{v} = \vec{\omega} \times \vec{r} \Rightarrow \rho\dot{\phi}\hat{e}_\phi = \omega_z\hat{e}_z \times \left(\rho\hat{e}_\rho + z\hat{e}_z\right) = \omega_z\rho\hat{e}_\phi.$$

This means

$$\omega_z = \dot{\phi}\,, \tag{1.66}$$

so that ω_z has the meaning of instantaneous time change in the angular position. Notice the direction of $\vec{\omega}$ is given by a right hand rule. (Curl the fingers of you right hand in the direction of and your thumb points in the direction of $\vec{\omega}$). A common case is uniform circular motion:



$\vec{\omega}$ is a constant vector pointed along the 3-axis.

$$\vec{v} = \rho\dot{\phi}\hat{e}_\phi,$$

$$\vec{a} = -\rho\dot{\phi}^2\hat{e}_\rho,$$

$$\vec{\omega} = \dot{\phi}\hat{e}_z.$$

Differential Operators and Leibnitz Rule

A frequently occurring mathematical operation is the **gradient**:

$$\vec{\nabla} \equiv \sum_i \hat{e}_i \frac{\partial}{\partial x_i}\,. \tag{1.67}$$

The x, y, z components of this are just the usual partial derivative operators. When operating on a scalar, it gives a vector:

$$\vec{\nabla}'\phi' = \sum_i \hat{e}_i' \frac{\partial \phi'}{\partial x_i'} \text{ or } \left(\vec{\nabla}'\phi'\right)_i = \frac{\partial \phi'}{\partial x_i'}\,. \tag{1.68}$$

Proof: The chain rule says,

$$\frac{\partial \phi'}{\partial x_i'} = \sum_j \frac{\partial \phi'}{\partial x_j}\frac{\partial x_j}{\partial x_i'} = \sum_j \frac{\partial \phi}{\partial x_j}\frac{\partial x_j}{\partial x_i'}\,.$$

Now remember

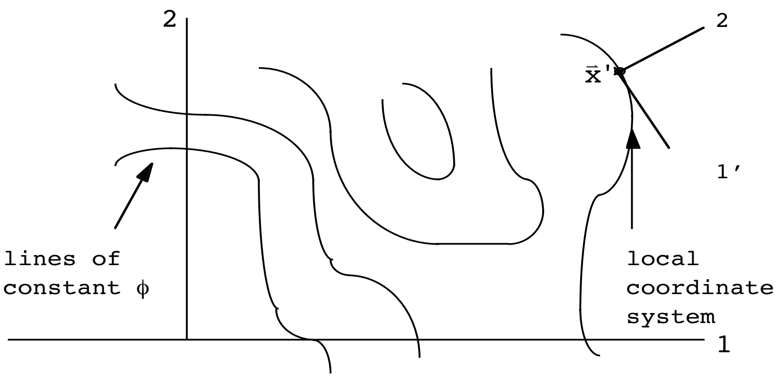
$$x_j = \sum_k \lambda_{kj}\, x_k'\,,$$

$$\Rightarrow \frac{\partial x_j}{\partial x_i'} = \sum_k \lambda_{kj} \underbrace{\frac{\partial x_k'}{\partial x_i'}}_{\delta_{ki}} = \lambda_{ij}\,. \tag{1.69}$$

Therefore

$$\frac{\partial \phi'}{\partial x_i'} = \sum_j \lambda_{ij} \frac{\partial \phi}{\partial x_j} \quad \text{or} \quad \left(\vec{\nabla}'\phi'\right)_i = \sum_j \lambda_{ij} \left(\vec{\nabla}\phi\right)_j , \tag{1.70}$$

which states that $\vec{\nabla}\phi$ transforms as a vector. A physical interpretation of the gradient operator acting on a scalar is given by the following sketch. Think of lines of constant density in 2 dimensions again. Construct the 1' axis tangent to the constant ϕ line, at $\vec{x}\,'$:

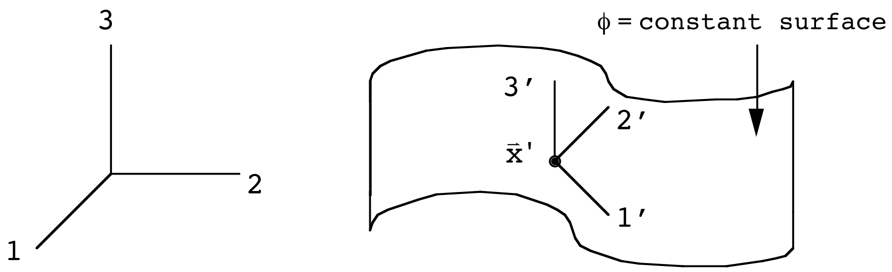


Clearly,

$$\frac{\partial \phi}{\partial x_1'} = 0 \Rightarrow \qquad \vec{\nabla}\phi = \frac{\partial \phi}{\partial x_2'} \hat{e}_2' \text{ only.}$$

We conclude that the gradient at each point is a vector pointing perpendicular to lines of constant ϕ at that point and points in the direction that ϕ is increasing.

In three dimensions we can orient a new coordinate system so that the 1'2' plane is now locally tangent to the ϕ = constant *surface* at $\vec{x}'$:



Again

$$\vec{\nabla}'\phi = \frac{\partial \phi}{\partial x_3'} \hat{e}_3' \text{ only.}$$

So the gradient is perpendicular to lines (2 dimensions) or surfaces (3 dimensions) of constant ϕ, and points in the direction ϕ is increasing. When dotted into a vector field, the divergence gives a scalar. (You should be able to prove this statement.) Thus, $\vec{\nabla} \cdot \vec{A}$ transforms as a scalar if is a vector.

A further mathematical operation, the **curl** of a vector,

$$\text{Curl: } \underbrace{\left(\vec{\nabla} \times \vec{A}\right)}_{\text{denoted}} \equiv \underbrace{\sum_{j,k} \varepsilon_{ijk} \nabla_j A_k}_{\text{defined}} , \tag{1.71}$$

can also be defined using the gradient ∇_i, but will not be used extensively in this course. If $\vec{A}$ is a vector, $\vec{\nabla} \times \vec{A}$ can be shown to transform as a pseudovector.

Another important operator is the **Laplacian**:

$$\vec{\nabla} \cdot \vec{\nabla} = \sum_i \frac{\partial^2}{\partial x_i^2} \equiv \nabla^2 . \tag{1.72}$$

One can show it produces a scalar or vector when acting on a scalar or vector, respectively. Here is the proof it transforms as a scalar when acting on a scalar:

$$\nabla'^2\phi' = \sum_i \frac{\partial^2\phi'}{\partial x_i'^2} = \sum_i \frac{\partial}{\partial x_i'}\frac{\partial}{\partial x_i'}\phi' .$$

Since $\phi' = \phi$, we have

$$\nabla'^2\phi' = \sum_{i,j,k}\left(\frac{\partial x_j}{\partial x_i'}\right)\frac{\partial}{\partial x_j}\left(\frac{\partial x_k}{\partial x_i'}\right)\frac{\partial}{\partial x_k}\phi .$$

As shown above

$$\frac{\partial x_j}{\partial x_i'} = \lambda_{ij},$$

$$\Rightarrow \qquad \nabla'^2\phi' = \sum_{j,k}\left\{\sum_i \lambda_{ij}\,\lambda_{ik}\right\}\frac{\partial}{\partial x_j}\frac{\partial}{\partial x_k}\phi ,$$

$$= \sum_i \frac{\partial^2\phi}{\partial x_i^2} = \nabla^2\phi .$$

The proof that ∇^2 operating on a vector yields another vector is similar.

Finally, there is a useful rule for taking the derivative with respect to a variable which is contained in the upper and lower limits of an integral. This result, called the **Leibnitz rule for differentiation**, states that

$$\frac{d}{dt}\int_{a(t)}^{b(t)} dx\, f(x,t) = f(b(t),t)\dot{b}(t) - f(a(t),t)\dot{a}(t) + \int_{a(t)}^{b(t)} dx\,\frac{\partial f(x,t)}{\partial t}. \quad (1.73)$$

When the differentiating variable appears in the upper and lower limits of an integral as well as in the integrand, one obtains three terms. The first term is the value of the integrand evaluated at the upper limit, $f(b(t),t)$, times the derivative of the upper limit, $\dot{b}(t)$. The second term is similar but represents minus the integrand at the lower limit, $\dot{a}(t)$. times the derivative of the lower limit. The third term represents the contribution from the t-dependence in the integral itself. Eq.(1.73) is the simplest case of the dependence on a single variable, t; we will use a more general version of the Leibnitz rule in a three dimensional context in the next Chapter.

Complex Variables

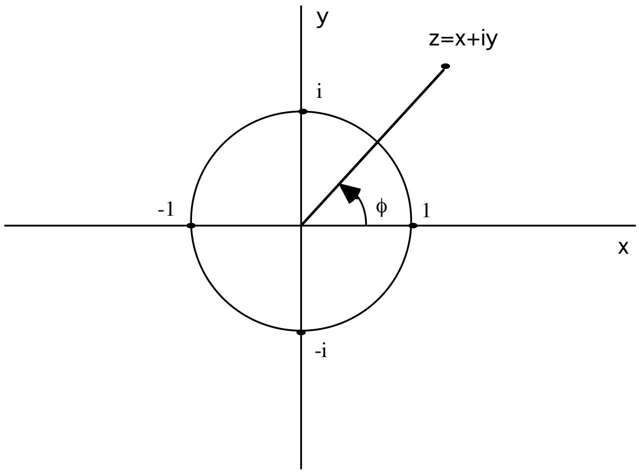
We will have occasion to use complex numbers and variables in our study of mechanics. The imaginary number i is given by

$$i=\sqrt{-1} ,$$

and when raised to various powers gives,

$$i^2 = -1,\ i^3 = -i,\ i^4 = 1.$$

These numbers can be arranged on a unit circle in a plane:



This is called the complex number plane. As the figure suggests, any complex number can be written as a combination of a real and an imaginary number,

$$z = x + iy, \tag{1.74}$$

where x and y are both real numbers. The complex conjugate of z is given by $x - iy$ and is denoted as z^*. These two numbers specify a location in the above plane on the real (x) and imaginary (y) axes. The real and imaginary parts of z are separated off with

$$\mathrm{Re}(z) \equiv x, \ \mathrm{Im}(z) \equiv y,$$

and the distance from the origin to the point z is given as

$$|z| = \sqrt{z z^*} = \sqrt{x^2 + y^2}. \tag{1.75}$$

From plane trigonometry, any complex number on the above unit circle may be written as

$$z = \cos\phi + i \sin\phi, \tag{1.76}$$

where the angle ϕ is measured from the real axis, x. It was Euler who showed this may be written as

$$z = e^{i\phi}, \tag{1.77}$$

for . Given this result we can now represent any number in the complex plane as

$$z = |z| e^{i\phi}. \tag{1.78}$$

We often encounter ratios of complex numbers,

$$z = \frac{z_1}{z_2}.$$

We may always factorize such a ratio into real and imaginary parts by multiplying top and bottom by the complex conjugate of z_2:

$$z = \frac{z_1 z_2^*}{|z_2|^2},$$

which may also be written as

$$z = \left|\frac{z_1}{z_2}\right| e^{i(\phi_1 - \phi_2)}.$$

Complex numbers are extremely useful in solving linear differential equations, which is where we will see them next in this course.

CHAPTER 1 PROBLEMS

1. Let

$$A = \begin{pmatrix} 0 & 1 & 3 \\ -1 & 2 & 0 \\ 2 & 4 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 6 & 2 & -2 \\ 1 & 0 & 4 \\ 4 & 1 & -1 \end{pmatrix}.$$

   a) Find the matrix product AB. (A symbolic manipulator will be handy)
   b) Find det (AB), and show that det $(AB) = (\det A)(\det B)$
   c) Form the matrix product $B^T A^T$ and show that $B^T A^T = (AB)^T$, as it should.

2. Try proving that

$$A^{-1}A=1,$$

   implies that

$$AA^{-1}=1.$$

3. Prove that the inverse of the transpose is the transpose of the inverse of a square matrix (assuming it's inverse exists!).

4. Given $\sum_i \lambda_{ij}\lambda_{ik} = \delta_{jk}$, show that

$$\sum_i \lambda_{ji}\lambda_{ki} = \delta_{jk},$$

   follows immmediately. [Hint: Use matrix notation and remember that $\lambda^{-1} = \lambda^T$ for these matrices.]

5. Starting with (λ is a matrix; T denotes transpose)

$$\lambda^{-1} = \lambda^T,$$

   prove that

$$\det \lambda = \pm 1.$$

[Hint: $\lambda = \det \lambda^T$.]

6. Given that the 3 × 3 matrices λ_1 and λ_2 represent orthogonal transformations on vectors $\mathbf{x}$ and $\mathbf{x}'$

$$\lambda_1^T = \lambda_1^{-1}, \quad \lambda_2^T = \lambda_2^{-1},$$
$$(\mathbf{x}' = \lambda_1 \mathbf{x}, \quad \mathbf{x}'' = \lambda_2 \mathbf{x}')$$

show that the matrix product, $\lambda_2\lambda_1$,, is also an orthogonal transformation.

7. Using the definition given in the text,

$$\det A = \sum_{i,j,k} \varepsilon_{ijk} A_{1i} A_{2j} A_{3k},$$

show that the determinant of a 3 × 3 matrix may also be written in the forms:

a) $\det A = \sum_{i,j,k} \varepsilon_{ijk} A_{i1} A_{j2} A_{k3},$

b) $\det A = \frac{1}{3!} \sum_{i,j,k,\ell,m,n} \varepsilon_{ijk}\varepsilon_{\ell mn} A_{\ell i} A_{mj} A_{nk}$.

8. Argue that the determinant of an antisymmetric 3×3 matrix,

$$A^T = -A,$$

is zero. Is the determinant of an antisymmetric 4x4 matrix necessarily zero?

9. We need to complete the proof of the relation we found necessary when proving that $\vec{A} \times \vec{B}$ actually transformed as a (pseudo) vector. This proof was based upon the supposed identity,

$$\sum_{j,k} \varepsilon_{ijk} \lambda_{j\ell}\lambda_{km} = C \sum_{n} \varepsilon_{n\ell m}\lambda_{in} .$$

By multiplying both sides by λ_{ip} (and summing on i) and choosing a special case, show that this identity may be reduced to Eq.(1.44) of the text (almost; see prob. 7), implying

$$\det \lambda = C.$$

10. (a) Show Eq.(1.56) follows from Eq.(1.49).
(b) Show the cyclic property Eq.(1.57).

11. Show that

$$\vec{A} \cdot \left[\vec{B} \times \left(\vec{A} \times \vec{B}\right)\right] = \left(\vec{A} \times \vec{B}\right)^2 .$$

12. Given that $\vec{A}$, $\vec{B}$ and $\vec{C}$ are vectors, argue that $\vec{A} \cdot \left(\vec{B} \times \vec{C}\right)$ transforms as a pseudoscalar.

13. Does the angular velocity of a particle, $\vec{\omega}$, transform as a vector or pseudovector? Explain.

14. Evaluate the sum:

$$\sum_{k,j,m} \varepsilon_{ijk}\varepsilon_{\ell mk}\varepsilon_{njm} .$$

[Hint: Use results already proven.]

15. (Adapted from Marion and Thorton.) Let's investigate a point raised in the text. Consider

$$\vec{A} \times \vec{B} = \vec{C},$$

where $\vec{A}$ and $\vec{C}$ are considered known quantities. Clearly, this equation does not uniquely define $\vec{B}$ since any component of $\vec{B}$ along $\vec{A}$ will not contribute to the cross product. However, consider

$$\vec{A} \cdot \vec{B} = S,$$

where S is a known scalar. Show that we can solve for $\vec{B}$ and that ($\vec{A}$ considered nonzero)

$$\vec{B} = \frac{1}{\vec{A}^2}\left(S\vec{A} + \vec{C} \times \vec{A}\right).$$

16. Express the spherical unit vector, $\hat{e}_\theta$, in terms of the cylindrical unit vectors $\hat{e}_\rho$, $\hat{e}_\phi$ and $\hat{e}_z$. The x, y, z components of these vectors are

$$\hat{e}_\theta = (\cos\theta\cos\phi, \cos\theta\sin\phi, -\sin\theta) \;,$$

and

$$\hat{e}_\rho = (\cos\phi, \sin\phi, 0),$$

$$\hat{e}_\phi = (-\sin\phi, \cos\phi, 0),$$

$$\hat{e}_z = (0,0,1).$$

17. Show that the velocity vector in spherical coordinates is given by:
$\vec{v} = \dot{r}\hat{e}_r + r\dot{\theta}\hat{e}_\theta + r\sin\theta\,\dot{\phi}\,\hat{e}_\phi.$
[Hint: Use the decomposition of $\hat{e}_r$ found in the text.]

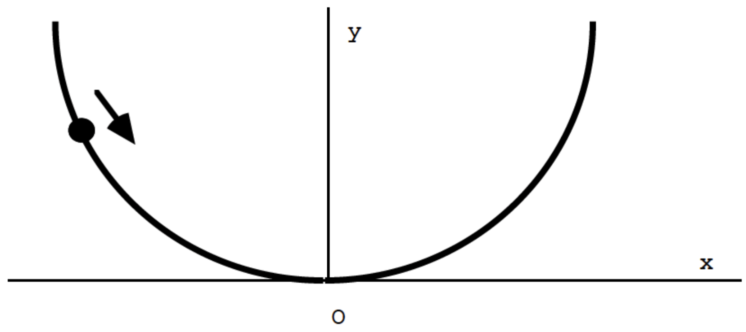
18. In spherical coordinates, show that

a) $\dot{\hat{e}}_\theta = -\dot{\theta}\hat{e}_\theta + \dot{\phi}\cos\theta\,\hat{e}_r,$

b) $\dot{\hat{e}}_\phi = -\ddot{\phi}\cos\theta\,\hat{e}_\theta - \dot{\phi}\sin\theta\,\hat{e}_r.$

[Hint: See decomposition of $\hat{e}_\theta$, $\hat{e}_\phi$ in the text.]

19. A particle follows the trajectory $y = Kx^2$ (K is a constant) with constant **speed**, S, in the xy plane as shown.



a) Show that $\ddot{x} \cdot \dot{x} = 0$ anywhere on the trajectory.
b) Find the acceleration, $\ddot{x}$, when the particle is at the origin, 0.
c) Find the instantaneous angular velocity, $\vec{\omega}$, at 0.

20. Given the vector function $\vec{v} = 20xy\,\hat{i} + 25yz\,\hat{j} + 15xy\,\hat{k}$, find:
(a) $\vec{\nabla} \cdot \vec{v}$ (b) $\vec{\nabla} \times \vec{v}$

21. Given that

$$\nabla_i' = \sum_j \lambda_{ij}\nabla_j ,$$

and

$$A_j' = \sum_k \lambda_{jk}A_k ,$$

prove that:

a) $\vec{\nabla}' \cdot \vec{A}'$ transforms as a scalar,
b) $\vec{\nabla}' \times \vec{A}'$ transforms as a pseudovector.

22. Show that $(\nabla^2 \equiv \sum_j \nabla_j^2)$

$$\nabla^2(f\ g) = g\nabla^2 f + f\nabla^2 g + 2\vec{\nabla} f \cdot \vec{\nabla} g.$$

23. Show that

$$z = \cos\phi + i\ \sin\phi,$$

may be written as

$$z = e^{i\phi}.$$

[Hint: Consider the power series expansion of the exponential.]

24. Show that the angle ϕ in $z = e^{i\phi}$ .is given by $0 \le \phi < 2\pi$

$$\phi = \tan^{-1}\left(\frac{\text{Im}(z)}{\text{Re}(z)}\right).$$

2 NEWTONIAN MECHANICS

REVIEW OF NEWTON'S LAWS

Newton's 3 laws:

1) A body remains at rest or in uniform motion unless acted upon by a force.
2) A body acted upon by a force moves in such a manner that the time rate of change of momentum equals force.
3) If two bodies exert forces on each other, their forces are equal in magnitude and opposite in direction.

We will see that these dynamical laws lead to certain conservation theorems. In fact, it is now known that the conservation laws have a greater range of validity that Newton's laws, so we can regard these as a stairstep toward a more fundamental outlook. More sophisticated formulations of mechanics do not even use the concept of "force". So Newton's laws move us toward a more fundamental outlook in both physical content as well as formalism.

Essentially, Newton's laws are a framework applicable to macroscopic motions due to any kind of physical interaction as long as the characteristic velocities involved are small compared to the speed of light and characteristic length scales are large compare to atomic dimensions. By a "physical interaction" I mean the underlying force of nature responsible for producing the motion of objects. For example, electrodynamics often manifests itself in our macroscopic world in a sort of disguised form in the form of elastic and inelastic (frictional) forces. By elastic, in this context, I mean that no energy is lost in the form of heat during the interaction.

Let me paraphrase Newton's laws, as applied to point objects, to help you remember them better and to be a bit more quantitative.

I. Let $\vec{x}(t)$ be the position of a particle at time t and let $\vec{F}(t)$ be the force on it. Then Newton's first law says

$$\vec{x}(t) = \vec{v}t + \vec{x}_0 \quad \text{if} \quad \vec{F}(t) = 0, \tag{2.1}$$

Of course $\vec{x}(t) = \vec{v}t + \vec{x}_0$ is only going to be true if particle position is measured in a non-accelerating or so-called inertial reference frame. So, we need an initial inertial reference frame to define another one. The inertial frame concept is clearly an idealization, only true to a certain extent in applications here on Earth. (We already are idealizing to point particles as well!) However, this will not prevent us from an intuitive and fruitful application of Newton's laws in real-world situations.

II. Let the momentum of a point particle be defined as

$$\vec{p} \equiv m\vec{v}, \tag{2.2}$$

where m is the particle's mass and $\vec{v}$ is it's velocity. Imagine summing all the individual forces to get the total force on the particle,

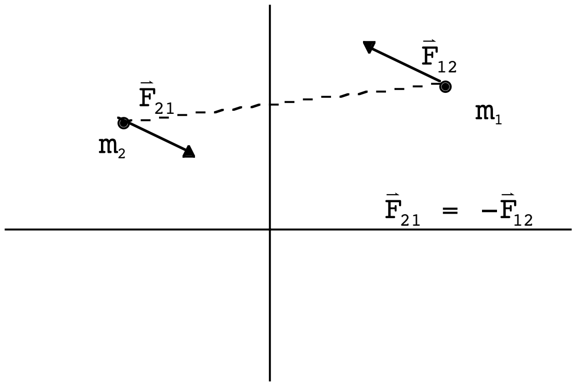
$$\vec{F} = \sum_i \vec{f}_i .$$

Then Newton's second law says

$$\vec{F} = \frac{d\vec{p}}{dt} = \frac{d}{dt}(m\vec{v}) . \tag{2.3}$$

This defines force (undefined in Newton's first law) as long as mass is also defined.

III. I'll use the following picture to help explain Newton's third law:



$\vec{F}_{21}$ and $\vec{F}_{12}$ act on different objects! Note that the forces between particles do not necessarily point along the line connecting them, although it usually does.

I invoked the concept of mass above. How is this defined? It can be regarded simply as the proportionality constant between force and acceleration in some system of units:

$$\vec{F} = m\vec{a}. \tag{2.4}$$

However, this viewpoint is a bit illogical, because we are now using this one equation to define both force **and** mass. If we are going to simply regard mass as a proportionality factor, then we should have a means of defining both force and acceleration. We can always use a gravitational field to independently measure such quantities in the form of the equation,

$$\vec{W} = m_g\vec{g},$$

where $\vec{W}$ is the weight (measured on a spring scale for example) and $\vec{g}$ is the acceleration due to gravity. "m_g" is then the mass as measured using gravity. It is now possible to imagine accelerating the particle using the same spring. If we displace the spring the same amount, we are guaranteed to exert the same force. If we allow the particle to be accelerated, we will now have that

$$\vec{W} = m_a\vec{a},$$

where $\vec{W}$ is the exact same force as before. Although the same force is being used in both situations, it is possible that the accelerations $\vec{a}$ (inertial) and $\vec{g}$ (gravitational) will not be the same and therefore the proportionality constants, mg and ma will not be equal. Let's call ma the inertial mass and mg the gravitational mass. The question

$$m_g \stackrel{?}{=} m_a$$

is an experimental one and has been tested extremely accurately; one such test will be described in Ch.5. As far as is known, however, the equality $m_a = m_g$ is exact. We will use "m" to denote the mass in all future situations.

Let us investigate the solution of Newton's force law for a point particle in two fairly common circumstances. The general form of the force law in one space dimension is just

$$m\ddot{x} = F(x, \dot{x}, t). \tag{2.5}$$

$$\left(\dot{x} \equiv \frac{dx}{dt}, \quad \ddot{x} \equiv \frac{d^2x}{dt^2}\right)$$

What I am suggesting in the notation $F(x, \dot{x}, t)$ is that the force on the particle can be a function of it's position, $x(t)$, it's instantaneous velocity, $\dot{x}(t)$, as well as having explicit time dependence. Without knowing the specific form of $F(x, \dot{x}, t)$ it is not possible to solve Eq.(2.5) in general. However, let's say $F(x, \dot{x}, t) \rightarrow F(t)$ only. Then the first integral of Eq.(2.5) gives

$$\dot{x}(t) = \dot{x}_0 + \frac{1}{m}\int_{t_0}^{t} dt'\, F(t'), \tag{2.6}$$

from which the velocity at time t can be found. In the above, $\dot{x}_0$ is just the constant of integration, which in this case just represents the initial velocity. (Just set $t=t_0$ above to see this.) If our goal is to find $x(t)$, we need only integrate again:

$$x(t) = x_0 + \dot{x}_0\left(t - t_0\right) + \frac{1}{m}\int_{t_0}^{t} dt'\int_{t_0}^{t'} dt''\, F(t''). \tag{2.7}$$

Again, a constant of integration, x_0, has emerged. Clearly, since Eq.(2.5) represents a second-order differential equation, we will always have two such constants to supply. In this case, we are simply requiring

$$\dot{x}\left(t_0\right) = \dot{x}_0, \tag{2.8}$$

$$x\left(t_0\right) = x_0. \tag{2.9}$$

Eqs. (2.8) and (2.9) are often called **initial conditions**. Usually, one chooses $t_0 = 0$.

A common application of Eq.(2.5) is when $F(t) = ma$, where "a" is a constant in time. (This is the case of gravitational attraction of a point mass near the Earth's surface, for example.) Then, we obtain

$$\int_{t_0}^{t} dt' \int_{t_0}^{t'} dt'' = \int_{t_0}^{t} dt' (t' - t_0) = \frac{1}{2}(t - t_0)^2$$

which gives

$$x(t) = x_0 + \dot{x}_0 (t - t_0) + \frac{a}{2}(t - t_0)^2, \tag{2.10}$$

which should be familiar to you as the equation for instantaneous position under constant acceleration.

Another common case is where $F(x, \dot{x}, t) \rightarrow F(x)$ only. Then, we can solve for the motion again, as follows. Integrate both sides of Eq. (2.5) over x. Since

$$\int_{x_0}^{x(t)} dx\, \ddot{x} = \int_{t_0}^{t} dt'\, \dot{x}(t')\, \ddot{x}(t') \tag{2.11}$$

$$= \frac{1}{2}\left(\dot{x}^2(t) - \dot{x}_0^2\right), \tag{2.12}$$

we have

$$\frac{1}{2} m \left(\dot{x}^2(t) - \dot{x}_0^2\right) = \int_{x_0}^{x(t)} dx' F(x'). \tag{2.13}$$

One may confirm that Eq. (2.13) is the first integral of (2.5) by taking the derivative of (2.13) with respect to t.[2] In this case $\dot{x}_0$ is just a constant. Also, Eq. (2.13) is actually just a statement about energy conservation. One recognizes the difference in kinetic energies on the left hand side and the right hand side is just the work done. This will be discussed in more depth later.

It is possible to integrate (2.13) again, as follows. First, solve for $\dot{x}(t)$,

$$\dot{x}(t) = \pm\sqrt{\frac{2}{m}\int_{x_0}^{x(t)} dx' F(x') + \dot{x}_0^2}. \tag{2.15}$$

Then using the general identity (which also applies in more than one dimension if the integration is along the particle's path),

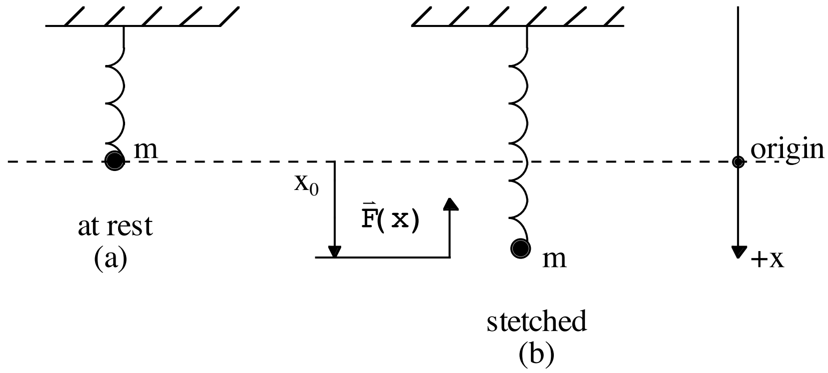
$$\int_{x_0}^{x(t)} \frac{dx}{\dot{x}} = \int_{t_0}^{t} dt,$$

yields

$$\pm\int_{x_0}^{x(t)} \frac{dx'}{\sqrt{\frac{2}{m}\int_{x_0}^{x'} dx'' F(x'') + \dot{x}_0^2}} = (t - t_0). \tag{2.16}$$

The appropriate sign in Eqs.(2.15) and (2.16) must be *chosen* for a given problem. Another constant of integration has been supplied on the right hand side of Eq.(2.16), which in this case is just the initial time, t_0. Notice that while Eq.(2.16) is a formal solution to the problem, it gives $t(x)$ rather than the more usual $x(t)$. However, if the relation of x(t) and t is one-to-one (we can imagine breaking the overall motion of the particle up into submotions for which this is true), then one can in principle invert the relation $t(x)$ to get $x(t)$.

Let us again take an example to illustrate the use of Eq. (2.16). Consider a stretched spring attached to a vertical support, as shown. (Imagine the spring is nearly massless.)



Let us pull the mass m down a distance x_0 and then release it from rest. Let us assume that ("k" is a constant called the "spring constant")

$$F(x) = -kx \ , \ k>0,$$

for small extensions of the spring, a force law for springs known as Hooke's law. (We'll examine this force law in the next Chapter.) The goal will be to use Eq.(2.16) to find $x(t)$ for an appropriate portion of the motion.

Eq.(2.16) reads in our case (choosing $\dot{x}_0 = 0, t_0 = 0$)

$$-\int_{x_0}^{x(t)} \frac{dx'}{\sqrt{-\frac{2k}{m}\int_{x_0}^{x'} x''dx''}} = t, \tag{2.17}$$

where we have chosen the negative sign in Eq.(2.16) for our coordinate system. We now do the integral:

$$-\sqrt{\frac{m}{k}}\int_{x_0}^{x(t)} \frac{dx'}{\sqrt{x_0^2 - x'^2}} = -\sqrt{\frac{m}{k}}\left(\mathrm{Sin}^{-1}\left(\frac{x(t)}{x_0}\right) - \mathrm{Sin}^{-1}(1)\right),$$

$$= -\sqrt{\frac{m}{k}}\left(\mathrm{Sin}^{-1}\left(\frac{x(t)}{x_0}\right) - \frac{\pi}{2}\right). \tag{2.18}$$

Note that the inverse sine is not defined until it's range is specified. I have chosen the principal branch above, symbolized by writing "Sin^{-1}". (Any other branch would do.) Putting (2.17) and (2.18) together now allows us to solve for $x(t)$ as

$$x(t) = x_0 \sin\left(-\sqrt{\frac{k}{m}}\, t + \frac{\pi}{2}\right). \tag{2.19}$$

This type of motion is termed "simple harmonic". We have now solved for the motion of the point mass along the portion of the motion where the connection between t and x is one to one. Obviously, since the motion is simple harmonic we can get $x(t)$ at any time t by just writing the usual "sin" function above, as we have already done. Thus

$$x(t) = x_0 \cos\left(\sqrt{\frac{k}{m}}\, t\right). \tag{2.20}$$

The one to one submotions of the system are connected by "turning points", where the motion of the particle reverses, occuring at $x = \pm x_0$ in this example. We will discuss turning points more extensively in the material to follow. The period of motion of this system is given by the amount of time necessary to complete one spatial cycle of motion and is given by the amount of time necessary for the arument of the cosine in (2.20) to increase by 2π, namely $T_{spring} = 2\pi\sqrt{\frac{m}{k}}$. Note that this period is independent of the amplitude of motion!

Simple Examples Using Newton's Laws

We will look at four simple examples as further applications of the ideas in Newton's laws. These should be considered as purely review.

Example 1

We illustrate one-dimensional harmonic oscillation in the form of an oscillating spring, but now let us imagine a periodic force is applied to the system, initially at rest. We take

$$m\ddot{x} = F(x, t), \tag{E.1}$$

$$F(x, t) = -kx + F_0 \sin \omega t. \tag{E.2}$$

The equation of motion can be put into the form

$$\ddot{x} + \frac{k}{m}x = \frac{F_0}{m}\sin \omega t. \tag{E.3}$$

Let us solve (E.3) subject to the initial conditions,

$$\dot{x}(0) = 0, \; x(0) = 0. \tag{E.4}$$

Solution:

This problem was chosen because it illustrates some techniques of solving linear differential equations. Let us call $\omega_0^2 = \frac{k}{m}$. Then the complimentary equation is

$$\ddot{x}_c + \omega_0^2 x_c = 0, \tag{E.5}$$

the general solution of which is

$$x_c = A \sin \omega_0 t + B \cos \omega_0 t. \tag{E.6}$$

"ω_0" is now identified as the **angular** frequency of motion, $\omega_0 = 2\pi f_0$, where f0 is frequency. The particular solution can be found by the method of undetermined coefficients. The trial solution is (assume $\omega \neq \omega_0$ initially)

$$x_p = C \sin \omega t + D \cos \omega t. \tag{E.7}$$

Substituting in (E.3) now gives

$$\sin \omega t : \quad C = \frac{F_0}{m(\omega_0^2 - \omega^2)}, \tag{E.8}$$

$$\cos \omega t : \quad D = 0. \tag{E.9}$$

Thus, the general solution is

$$x = x_c + x_p = \frac{F_0}{m(\omega_0^2 - \omega^2)}\sin \omega t + A \sin \omega_0 t + B \cos \omega_0 t. \tag{E.10}$$

It only remains to satisfy the initial conditions. We have

$$x(0) = 0 \quad \Rightarrow \quad B = 0, \tag{E.11}$$

$$\dot{x}(0) = 0 \Rightarrow \quad A = -\frac{F_0 \omega}{m\omega_0 \left(\omega_0^2 - \omega^2\right)}, \tag{E.12}$$

$$\Rightarrow \qquad x = \frac{F_0}{m(\omega_0^2 - \omega^2)}\left(\sin \omega t - \frac{\omega}{\omega_0} \sin \omega_0 t\right). \tag{E.13}$$

Let's now examine the solution for $\omega = \omega_0$. We have the same complimentary solution, (E.6), as before, but now the trial particular solution must be taken as

$$x_p = Ct \sin \omega_0 t + Dt \cos \omega_0 t. \tag{E.14}$$

Again, substituting in (E.3) gives (the coefficients of the $t \sin \omega_0 t$ and $t \cos \omega_0 t$ terms vanish)

$$\sin \omega_0 t : \qquad D = -\frac{F_0}{2m\omega_0}, \tag{E.15}$$

$$\cos \omega_0 t : \qquad C = 0. \tag{E.16}$$

The general solution is thus

$$x = A \sin \omega_0 t + B \cos \omega_0 t - \frac{F_0}{2m\omega_0} t \cos \omega_0 t. \tag{E.17}$$

Again, supplying boundary conditions,

$$x(0) = 0 \Rightarrow \quad B = 0, \tag{E.18}$$

$$\dot{x}(0) = 0 \Rightarrow \qquad A = \frac{F_0}{2m\omega_0^2}. \tag{E.19}$$

This gives finally

$$x = \frac{F_0}{2m\omega_0^2} \sin \omega_0 t - \frac{F_0}{2m\omega_0} t \cos \omega_0 t. \tag{E.20}$$

Notice that as $t \to \infty$, the motion in (E.20) becomes unbounded. We will later study in detail a more realistic case with damping where the unbounded motion for $\omega = \omega_0$ is removed and instead results in a phenomenon known as "resonance".

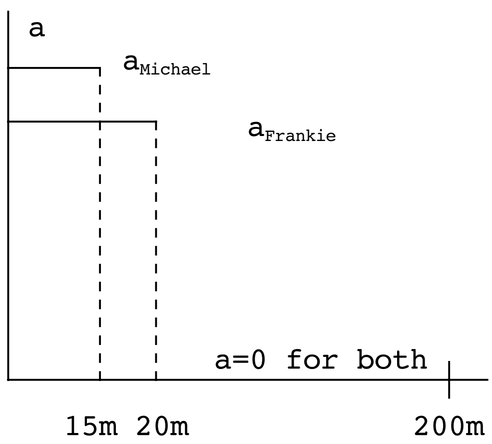
The following is a simple one-dimensional dynamics problem.

Example 2

Two sprinters are running a 200m race. Both runners have the same top speed, 10m/s, at the end of their initial acceleration, but Michael reaches top speed after only 15m while Frankie takes 20m to reach the same top speed. Let's idealize the acceleration of both sprinters as a constant. By how many seconds does Michael beat Frankie in running the race? How far ahead is Michael from Frankie when he crosses the finish line?

Solution:

The acceleration profile of the two runners is as shown.



The velocity and position of the runners during the acceleration period are given by Eq. (2.10) and it's first derivative $\left(x_0 = \dot{x}_0 = t_0 = 0\right)$,

$$v(t) = at, \tag{E.21}$$

$$x(t) = \frac{at^2}{2}. \tag{E.22}$$

Notice that we do not know either the acceleration or the time it takes to reach top speed, but Eq.(E.21) and (E.22) give us two equations in two unknowns. Plugging in $v(t) = 10m/s$ in (E.21) and either $x(t) = 15m$ or $20m$ in (E.22) then gives

$$a_{michael} = \frac{10}{3}\, m/s^2, \tag{E.23}$$

$$a_{Frankie} = \frac{5}{2}\, m/s^2. \tag{E.24}$$

Using these results, solve for the acceleration time from (E.21):

$$\left(t_a\right)_{Michael} = 3s, \tag{E.25}$$

$$\left(t_a\right)_{Frankie} = 4s. \tag{E.26}$$

We now simply solve for the time necessary to run the remaining distance at the constant speed of 10m/s from (different $x(t)$ than (E.22))

$$x(t) = \dot{x}_0 t. \tag{E.27}$$

For Michael $x(t) = 185m$, while $x(t) = 180m$, for Frankie; $\dot{x}_0 = 10\ m/s$. Calling this time t_0, we find

$$(t_0)_{Michael} = 18.5s, \tag{E.28}$$

$$(t_0)_{Frank} = 18.0s. \tag{E.29}$$

Adding t_a and t_0, we then obtain that Michael beats Frankie by only $.5s$. In addition, when Michael crosses the finish line, Frankie will be

$$10\ m/s \times .5s = 5m$$

behind him. (There are much simplier solutions of this problem!)

Of course, in more than one dimension, Newton's law of motion for point objects becomes a vector law,

$$m\ddot{\vec{x}} = \vec{F}\left(\vec{x}, \dot{\vec{x}}, t\right). \tag{2.21}$$

If we assume two dimensional motion and use Cartesian coordinates, this becomes

$$m\ddot{x} = F_x\left(x, y, \dot{x}, \dot{y}, t\right), \tag{2.22}$$

$$m\ddot{y} = F_y\left(x, y, \dot{x}, \dot{y}, t\right). \tag{2.23}$$

An important special case occurs when

$$F_x\left(x, y, \dot{x}, \dot{y}, t\right) \rightarrow F_x\left(x, \dot{x}, t\right) \text{ and } F_y\left(x, y, \dot{x}, \dot{y}, t\right) \rightarrow F_y\left(y, \dot{y}, t\right),$$

because then the two equations can be solved independently, although the initial conditions still tie the two motions together. This is the case for the two dimensional harmonic oscillator (next Chapter) as well as for gravitational force near the Earth's surface. In the latter case we have simply (x is horizontal, y is vertical)

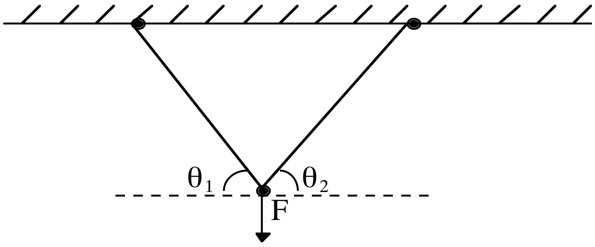
$$x : m\ddot{x} = 0, \tag{2.24}$$

$$y \; : \; m\ddot{y} = -mg, \tag{2.25}$$

where g is the acceleration due to gravity. (We will use the approximate value $9.8 m/s^2$ for g.) The solution of these equations are of course quite easy, but orchestrating the solutions to reach a desired set of end conditions can still be challenging.

We will consider two examples of two dimensional problems, one static and one dynamic.

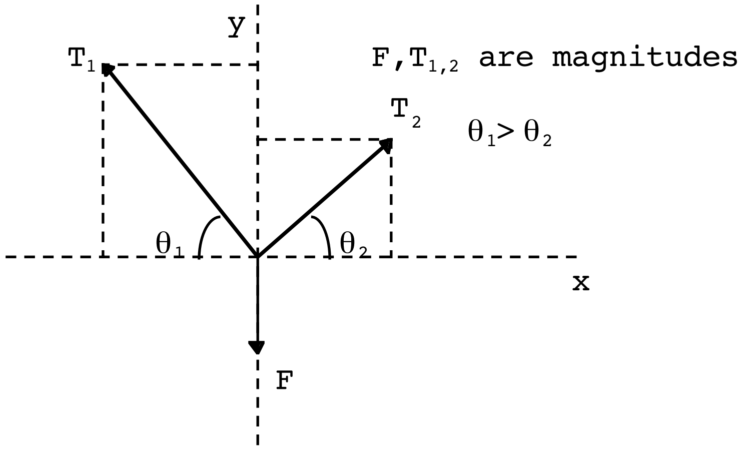
Example 3



A rope is attached to the ceiling and a force $\vec{F}$ is applied to it, as shown. The rope will break when the tension in any portion is greater than a certain magnitude, T'. Find the downward force, F, which breaks the rope. Which side, in general, will break?

Solution:

The force diagram is as below.



We have

$$\sum_i \vec{f}_i = 0,$$

$$\Rightarrow \begin{cases} x : T_1 \cos\theta_1 - T_2 \cos\theta_2 = 0, & \text{E. (30)} \\ y : -F + T_1 \sin\theta_1 + T_2 \sin\theta_2 = 0. & \text{E. (31)} \end{cases}$$

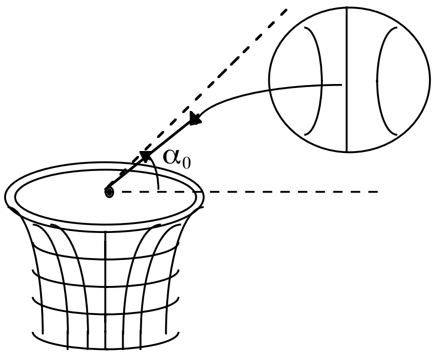
Since in the diagram $\theta_1 > \theta_2$, Eq.(E.30) means that $T_1 > T_2$, so it is the left hand side of the rope which will break first. Given this, use (E.30) to eliminate T_2 from (E.31):

$$F = T_1(\sin\theta_1 + \tan\theta_2 \cos\theta_1). \qquad \text{(E.32)}$$

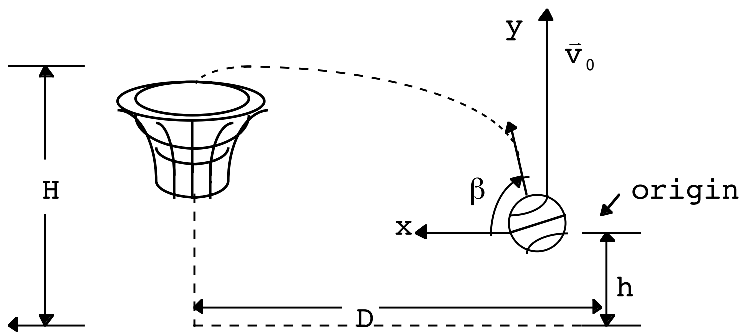
Setting $T_1 = T'$ and using specific values for $\theta_{1,2}$ in (E.32) gives the force which breaks the rope.

Example 4

The rim of a basketball goal at a carnival is designed so it will only allow basketballs to go through the net when the angle of entrance is greater than α_0, as shown. (We are clearly making idealizations in the geometry and the ball's motion.)



The net is at a height, H, above the ground, and the player is located a distance D from the basket. The ball is released a distance h from the ground at an angle β, again as shown.



With what angles, β, can the ball be released and still make it directly through the basket?

Solution:

We will locate the origin of the x,y coordinate system at the point of release of the ball. The instantaneous x and y positions at time t are then $(v_0 > 0)$

$$x(t) = (v_0 \cos\beta)t, \tag{E.33}$$

$$y(t) = (v_0 \sin\beta)t - \frac{1}{2}gt^2 . \tag{E.34}$$

If we require $x = D$ and $y = H-h$, we will have two equations in two unknowns, b and t (the time to reach the hoop),

$$t = \frac{D}{v_0 \cos\beta}, \tag{E.35}$$

$$=> \qquad \frac{H-h}{D} = \tan\beta - \frac{gD}{2v_o^2 \cos^2\beta}. \tag{E.36}$$

The angle at which the ball intersects the hoop is now

$$\tan\alpha \equiv -\frac{v_y}{v_x} = -\tan\beta + \frac{gD}{v_0^2\cos^2\beta}. \tag{E.37}$$

The condition for making a hoop is clearly

$$\tan\alpha > \tan\alpha_0 \Rightarrow \alpha > \alpha_0. \tag{E.38}$$

We can use the trigonometric identity,

$$\cos^2\beta = \frac{1}{1+\tan^2\beta}, \tag{E.39}$$

to put this in the form

$$\tan^2\beta - \frac{v_0^2}{gD}\tan\beta + 1 - \frac{v_0^2}{gD}\tan\alpha_0 > 0 \tag{E.40}$$

(E.36) may now be used to eliminate v_0^2 from (E.40), with the resultw

$$v_0^2 = \frac{gD}{2}\left(\frac{1+\tan^2\beta}{\tan\beta - \frac{(H-h)}{D}}\right), \tag{E.41}$$

$$\Rightarrow \tan^3\beta - 2a\tan^2\beta + \tan\beta - 2a > 0, \tag{E.42}$$

$$a \equiv \frac{(H-h)}{D} + \frac{1}{2}\tan\alpha_0 \tag{E.43}$$

We get a cubic condition in tan β as our general result.

Let's put some numbers in (E.41) and (E.42) to get a feeling for the condition. Let's assume

$$\frac{(H-h)}{D} = \frac{1}{2}, \alpha_0 = 45^\circ. \tag{E.44}$$

$$\Rightarrow \tan^3\beta - 2\tan^2\beta + \tan\beta - 2 > 0 \tag{E.45}$$

There is one real root of the cubic equation $\tan^3\beta_0 - 2\tan^2\beta_0 + \tan\beta_0 - 2 > 0$ It is given by

$$\tan\beta_0 = 2 \Rightarrow \beta_0 \approx 63.4^\circ \tag{E.46}$$

Thus, $\beta > \beta_0$ is required to make a basket in this case. (One must also have the correct v_0 as well, given by (E.41) above.)

Single Particle Conservation Theorems

As I said at the beginning of this Chapter, the conservation laws we will recover from Newton's three laws constitute more general results than Newton's laws themselves. We will now consider conservation laws for a single particle; multi-particle conservation laws will be considered later.

I. The total linear momentum, $\vec{p}$, of a particle is conserved when the total force, $\vec{F}$, on it vanishes.

Paraphrase of I: $\vec{F} = 0 \quad \Rightarrow \quad \vec{p} = \text{constant vector.}$

Proof: The proof follows immediately from Newton's second law,

$$\vec{F} = \frac{d\vec{p}}{dt}, \tag{2.26}$$

$$\vec{F} = 0 \quad \Rightarrow \quad \vec{p}(t) = \vec{p}_0, \tag{2.27}$$

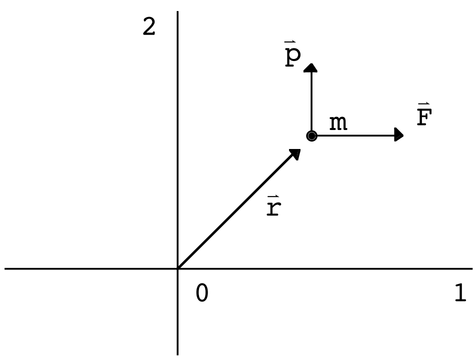
where $\vec{p}_0$ is a constant vector.

Angular momentum, $\vec{L}$, and torque, $\vec{N}$, for a single body are given by

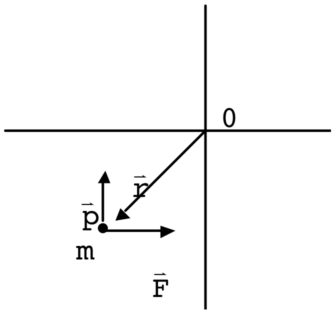
$$\vec{L} = \vec{r} \times \vec{p}, \tag{2.28}$$

$$\vec{N} = \vec{r} \times \vec{F}. \tag{2.29}$$

Clearly, both $\vec{L}$ and $\vec{N}$ are origin-dependent quantities, unlike linear momentum, $\vec{p}$. For example, consider ($\vec{p}$, $\vec{F}$ in a plane, say)



Here, $\vec{r} \times \vec{p}$ is "up" (along 3) and $\vec{r} \times \vec{F}$ is "down." Relocate the origin, 0:



Now $\vec{r} \times \vec{p}$ is "down" while $\vec{r} \times \vec{F}$ is "up." Thus, it is important to maintain a consistent origin when using these concepts.

II. The total angular momentum of a particle, $\vec{L}$, subject to zero total torque, $\vec{N}$, is conserved.

Paraphrase: $\vec{N} = 0 \quad \Rightarrow \quad \vec{L} = \text{constant vector}.$

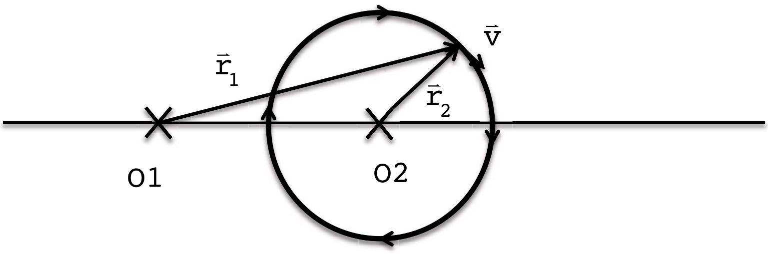
Proof: We have

$$\dot{\vec{L}} = \frac{d}{dt}\left(\vec{r} \times \vec{p}\right) = \dot{\vec{r}} \times \vec{p} + \vec{r} \times \dot{\vec{p}} \tag{2.30}$$

Using Newton's second law gives

$$\dot{\vec{L}} = \vec{r} \times \vec{F} = \vec{N}. \tag{2.31}$$

Notice that angular momentum may be a constant of the motion relative to one origin but not to another for certain force laws. Consider uniform circular motion:



Here, $\vec{L}$ is a constant of the motion measured from O2 (since $\vec{r}$ and $\dot{\vec{p}}$ are co-linear) but not from O1. This will be the case in general if ("central force")

$$\vec{F} = \hat{e}_r f(r), \qquad \left(r = |\vec{r}|\right)$$

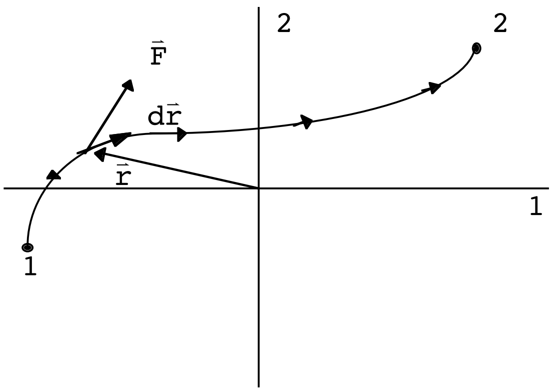
since then

$$\vec{N} = \vec{r} \times \hat{e}_r f(r) = 0.$$

The definition of work done on a single particle in moving it from position 1 to 2 is given as

$$W_{12} = \int_{\vec{r}_1}^{\vec{r}_2} \vec{F} \cdot d\vec{r}\ , \tag{2.32}$$

where $\vec{F}$ is the **total** force acting on the particle. It is understood the integration takes place along some specific path, as below.



(The motion is shown to lie in a plane for convenience.) Since $\vec{F}$ is the total force, by Newton's second law

$$\vec{F} \cdot d\vec{r} = m \frac{d\vec{v}}{dt} \cdot \frac{d\vec{r}}{dt} dt = m \frac{d\vec{v}}{dt} \cdot \vec{v}\ dt = \frac{m}{2} \frac{d}{dt} \left(\vec{v}^2\right) dt,$$

$$\text{Eq.(2.32)} \Rightarrow \quad W_{12} = \frac{1}{2} m\vec{v}^2 \Big|_1^2 = \frac{1}{2} m\left(\vec{v}_2^2 - \vec{v}_1^2\right) = T_2 - T_1, \tag{2.33}$$

where ("kinetic energy")

$$T = \frac{1}{2} m\vec{v}^2. \tag{2.34}$$

Eq.(2.33) generalizes the one-dimensional case, (2.13), above. The left hand side of (2.33) is again the work done.

Let us assume the special case,

$$\vec{F}\left(\vec{x}, \dot{\vec{x}}, t\right) \rightarrow \vec{F}\left(\vec{x}\right).$$

Let us also remember the connection between Newton's force law and energy conservation we saw previously in one dimension: the energy conservation statement was the first (spatial) integral of Newton's law. We have

$$m\ddot{\vec{x}} = \vec{F}\left(\vec{x}\right),$$

$$\Rightarrow m\int_{t_0}^{t} \ddot{\vec{x}} \cdot \dot{\vec{x}}\ dt' - \int_{t_0}^{t} dt' \frac{d\vec{x}}{dt'} \cdot \vec{F}\left(\vec{x}\right),$$

$$\Rightarrow \frac{1}{2} m\left(\dot{\vec{x}}^2 - \dot{\vec{x}}_0^2\right) = \int_{\vec{x}_0}^{\vec{x}} d\vec{x}' \cdot \vec{F}\left(\vec{x}'\right). \tag{2.35}$$

Let us define, ("potential energy")

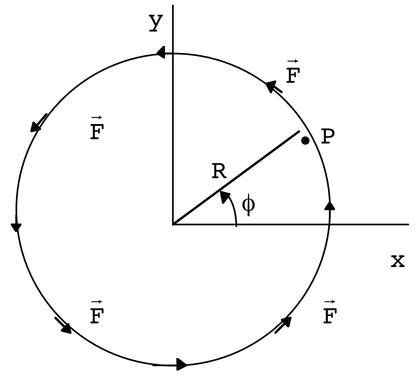
$$U\left(\vec{x}\right) \equiv -\int_{\vec{r}_1}^{\vec{x}} d\vec{x}' \cdot \vec{F}\left(\vec{x}'\right), \tag{2.36}$$

$$\vec{r}_1 : \text{"reference point"}.$$

At this point we will make an assumption about $U\left(\vec{x}\right)$:. We will assume,

$U\left(\vec{x}\right)$: A single-valued function of its argument, $\vec{x}$.

Why do we do this? Consider:



That is, assume a force law of the form,

$$\vec{F} = F\hat{e}_\phi \,.$$

Then

$$\int_{N \text{ turns}} \vec{F} \cdot d\vec{x} = 2\pi RNF.$$

N = number of turns around the circle

Clearly, this force law does not result in a single-valued function of the final position. Thus, a single-valued force law does not necessarily imply a single-valued potential. With this assumption about $U(\vec{x})$, we have

$$\frac{1}{2} m \left(\dot{\vec{x}}^2 - \dot{\vec{x}}_0^2\right) = -U(\vec{x}) + U(\vec{x}_0) \,.$$

Rearranging gives

$$\frac{1}{2} m\dot{\vec{x}}_0^2 + U(\vec{x}_0) = \frac{1}{2} m\dot{\vec{x}}^2 + U(\vec{x})$$

$$\Rightarrow (T + U)\big|_{t_0} = (T + U)\big|_{t} \,. \tag{2.37}$$

This leads to the third conservation law.

III. The total energy a particle,

$$E = T + U(\vec{x}),$$

subject to a single-valued potential, $U(\vec{x})$, is conserved.

Paraphrase: $U(\vec{x})$ single-valued $\Rightarrow \dfrac{dE}{dt} = 0$.

Proof: By the above direct construction we have

$$E(t_0) = E(t).$$

Since t is arbitrary,

$$\frac{dE}{dt} = 0.$$

Note that the potential $U(\vec{x})$ always needs a reference point for it's definition. If we change it, we get

$$U^{new}(\vec{x}) = -\int_{\vec{r}_2}^{\vec{x}} d\vec{x}' \cdot \vec{F}(\vec{x}')$$
$$= \underbrace{-\int_{\vec{r}_1}^{\vec{x}} d\vec{x}' \cdot \vec{F}(\vec{x}')}_{U(\vec{x})} - \underbrace{\int_{\vec{r}_2}^{\vec{r}_1} d\vec{x}' \cdot \vec{F}(\vec{x}')}_{\text{constant in } \vec{x}}. \tag{2.38}$$

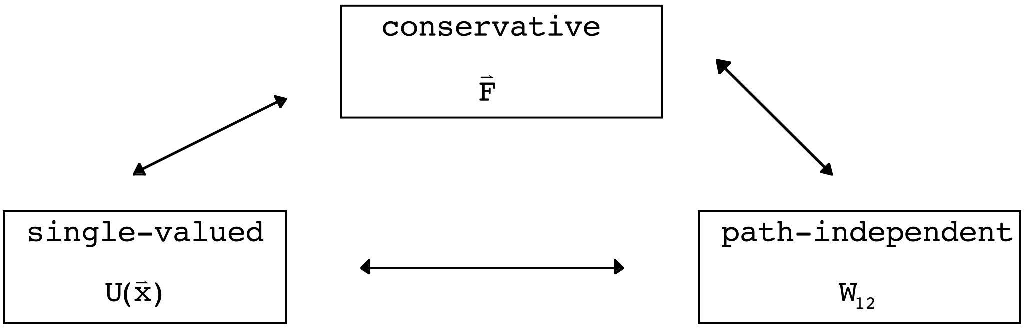
Thus, the old and new potentials just differ by a constant. This means the total energy of the particle, E, is arbitrary up to an overall constant. Clearly, given $U(\vec{x})$, this arbitrary constant does not affect $\vec{F}(\vec{x})$ itself since by the Leibnitz rule,

$$-\vec{\nabla}U(\vec{x}) = \vec{\nabla}\int_{\vec{r}_1}^{\vec{x}} d\vec{x}'\cdot\vec{F}(\vec{x}\,') = \vec{F}(\vec{x}) \,. \tag{2.39}$$

Since E is conserved and using Eq.(2.33), an alternate way of writing the work done in going from from $\vec{x}_1$, to $\vec{x}_2$ is

$$W_{12} = U(\vec{x}_1) - U(\vec{x}_2) \,. \tag{2.40}$$

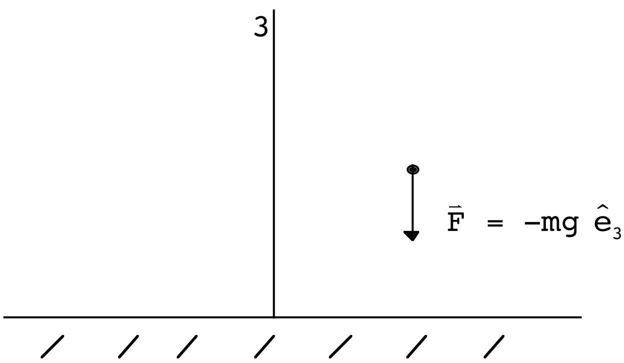
Notice that the work done depends only on the initial and final positions and is therefore **path independent**. This is equivalent to the above assumption of single-valuedness for $U(\vec{x})$. In addition, a force field which can be represented in the form of (2.39) is said to be **conservative**. This is again equivalent to the previous ideas of path independence and single-valuedness. This is summarized as



Each implies (or assumes) the other; any of these concepts can be invoked in writing down conservation law III above.[3]

Potential Energy and Particle Motion

We are familiar with the form of $U(\vec{x})$ near the Earth's surface:



We surmise that $U(\vec{x}) = mgx_3$. Is this correct?

$$\vec{F} = -\vec{\nabla}U = -mg\ \vec{\nabla}x_3 = -mg\hat{e}_3.$$

This is indeed correct. Notice that $U = mgx_3 + \text{const.}$ would do just as well.

With the above connection between $\vec{F}$ and U in a conservative force field, we can recast the solution for the one dimensional case when $F(x, \dot{x}, t) \rightarrow F(x)$ Eq. (2.16), in a different form. We have

$$\int_{x_0}^{x'} dx\ F(x) = -\int_{x_0}^{x'} dx \frac{dU}{dx} = U(x_0) - U(x').$$

Since the total energy, E, is a constant of the motion, we can evaluate it at a single time, say t_0, and it will continue to take on the same value at all later times. Therefore

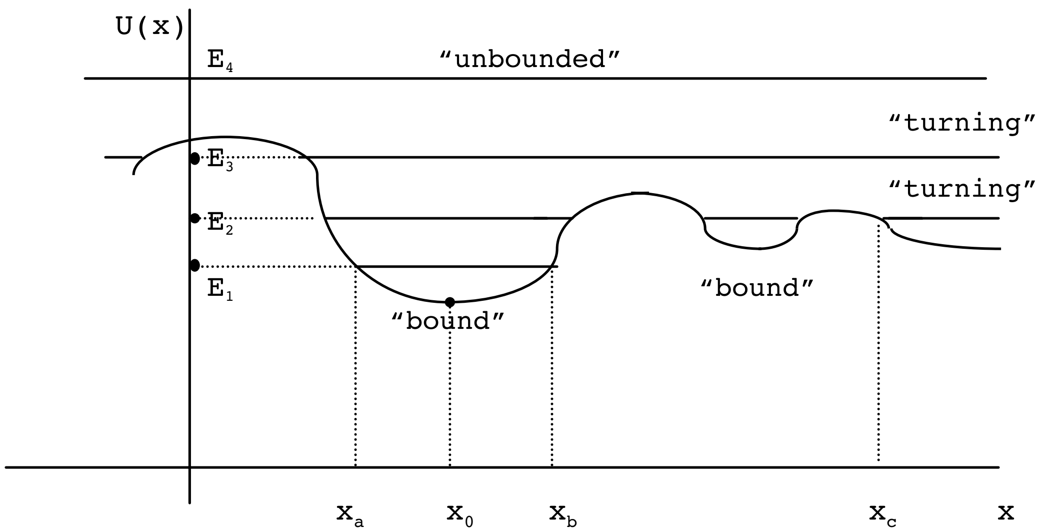
$$E = \frac{1}{2}m\dot{x}_0^2 + U(x_0),$$

$$\Rightarrow \frac{2}{m}\int_{x_0}^{x'} dx\, F(x) + \dot{x}_0^2 = \frac{2}{m}(E - U(x')). \qquad (2.41)$$

and (2.16) becomes

$$\pm\int_{x_0}^{x(t)} \frac{dx'}{\sqrt{\frac{2}{m}(E - U(x'))}} = (t - t_0). \qquad (2.42)$$

Let's find out how to use this equation by relating it to a picture of a general one-dimensional potential, $U(\vec{x})$.



The potential $U(\bar{x})$ is shown as the curvy line. On the same graph I have drawn horizontal lines which represent total energies E_1 through E_4. First of all, let's examine the motion of a particle with total energy E1. Clearly, the particle which has energy $E = E_1$ is bound. The two points labeled x_a and $x_b$are called "turning points", mentioned previously in the harmonic oscillator discussion. They represent, for bound motion, the extreme values in position reached by the particle in it's motion. x_0 is called a "stable equilibrium" point and will be discussed below. Note that a particle with total energy E_2 has both bound motion (in two different locations) and "turning" motion with a single turning point. The turning motion begins at position $x = x_c$; it represents where the velocity of the particle changes sign. In general, turning points are positions where $E = U$, which implying the kinetic energy, T, is zero. The motion of a particle with energy E_3 is separated into two regions with turning points as shown. It is also "turning" in the sense that the motion is not restricted to any finite region in x. The motion of a particle with energy E_4. takes on all values $-\infty < x > \infty$ along the x-axis. I will use the term "unbounded" to describe motion without any turning points.

How would one apply (2.42) to the motion of a particle with energy E_1, say? First of all recall that in deriving (2.42) we assumed that the relation between t and x was one-to-one. This is clearly violated for periodic motion, as we saw previously, but remember we can always piece together various separate motions where the relation is one-to-one to describe the full motion. The rule here is that these periodic motions are separated by turning points. Let us say the initial conditions were $x(0) = x_0$, $\dot{x}(0) > 0$. Thus, clearly one would want to choose the positive sign in (2.42) to describe this part of the motion since $t > 0$ and $x(t) > x_0$:

$$E = E_1, x(t) > x_0 : t = \int_{x_0}^{x(t)} \frac{dx'}{\sqrt{\frac{2}{m}(E - U(x'))}}. \tag{2.43}$$

On the other hand, after the particle reached x_b, it's motion would reverse. For the same starting time as in (2.43) the time, t, associated with a point $x(t)$ on this part of it's motion would be given by:

$$E = E_1, x(t) < x_b : t = t_b - \int_{x_b}^{x(t)} \frac{dx'}{\sqrt{\frac{2}{m}(E - U(x'))}}. \tag{2.44}$$

t_b is the time necessary to reach the first turning point, x_b:

$$t_b = \int_{x_0}^{x_b} \frac{dx'}{\sqrt{\frac{2}{m}(E - U(x'))}}. \tag{2.45}$$

Notice the minus sign in (2.44). It is required by the statement that $t > t_b$ for this part of the motion given that $x(t) < x_b$. Clearly, by patching such equations together in a similar way, one can find the relation between x and t at any time. Similar observations can be made about the motions of particles with energies E_2, E_3 and E_4 corresponding to their various bound or free motions.

If one is interested in the period of bound motion in a one-dimensional potential, one has to add together the time intervals for the $\dot{x} > 0$ motion and the $\dot{x} < 0$ motion. These are given by (again referring to the figure)

$$\Delta t_{ab} \equiv \int_{x_a}^{x_b} \frac{dx'}{\sqrt{\frac{2}{m}(E - U(x'))}}. \tag{2.46}$$

and

$$\Delta t_{ba} = -\int_{x_b}^{x_a} \frac{dx'}{\sqrt{\frac{2}{m}(E - U(x'))}} \tag{2.47}$$

respectively. Obviously, $\Delta t_{ab} = \Delta t_{ba}$,, which is true even if a and b are not turning points, and one has for the total period $\tau = \Delta t_{ab} + \Delta t_{ba}$,

$$\tau = 2\int_{x_a}^{x_b} \frac{dx'}{\sqrt{\frac{2}{m}(E - U(x'))}}, \tag{2.48}$$

where x_a and x_b represent the turning points. The turning points are to be found by setting the kinetic energy equal to zero, which results in

$$E = U(x_{a,b}) \quad , \tag{2.49}$$

where $x_{a,b}$ are the two possible solutions.

One can always use (2.48) to write the period as (take $x_0 = 0$)

$$\tau = 2\int_{x_a}^{0} \frac{dx'}{\sqrt{\frac{2}{m}(E - U(x'))}} + 2\int_{0}^{x_b} \frac{dx'}{\sqrt{\frac{2}{m}(E - U(x'))}}.$$

For a potential even about the origin, $U(x) = U(-x)$,

$$x_a = -x_b, \quad (x_b > 0)$$

and changing $x' \to -x'$ in the first term as the integration variable now shows

$$\tau = 4\int_0^{x_b} \frac{dx'}{\sqrt{\frac{2}{m}(E - U(x'))}}. \qquad (2.50)$$

As an example, consider Hooke's law, where one may take $U(x) = \frac{1}{2}kx^2, k > 0.$. Then the period is given by (2.50) as ($x' = \sqrt{\frac{2E}{k}}y$ is the change of variables)

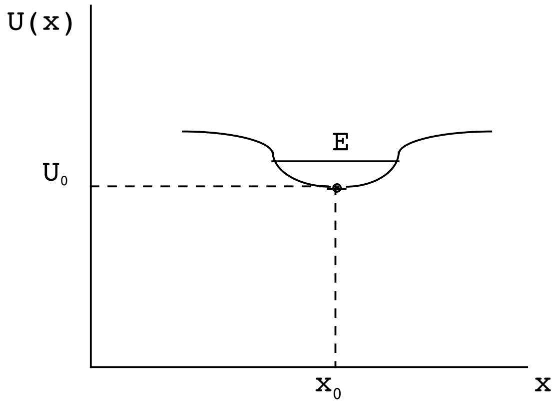
$$T_{spring} = 4\int_0^{\sqrt{\frac{2E}{k}}} \frac{dx'}{\sqrt{\frac{2}{m}\left(E - \frac{1}{2}k\,x'^2\right)}} = 4\sqrt{\frac{m}{k}}\int_0^1 \frac{dy}{\sqrt{1-y^2}}.$$

$$\Rightarrow \qquad T_{spring} = 2\pi\sqrt{\frac{m}{k}}. \qquad (2.51)$$

This is just the result we found earlier in this Chapter for a Hooke's spring.

Equilibrium and Stabilty in One Dimension

In the above discussion, I invoked the idea of a stable equilibrium position. Let us examine the meaning of this concept. We will be asking the question: Is the position of a particle stable under an infinitesimal displacement? The following is an illustration of the concept.



Assuming the potential $U(x)$ may be expanded in a Taylor series about the point x_0 shown above, we have

$$U(x) = U_0 + (x - x_0)\frac{dU}{dx}\Big|_{x_0} + \frac{(x - x_0)^2}{2}\left(\frac{d^2U}{dx^2}\right)\Big|_{x_0} + \ldots \tag{2.52}$$

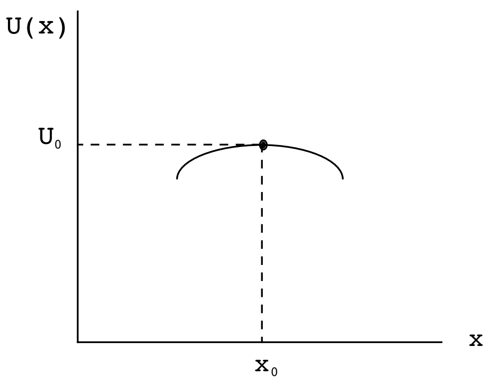
The constant U_0 will not affect any aspect of the particle's motion and can be eliminated by redefining $U(x)$. For the above illustration,

$$\frac{dU}{dx}\Big|_{x_0} = 0 \, , \; \frac{d^2U}{dx^2}\Big|_{x_0} > 0 .$$

Motion with energy E slightly larger than U_0 will then result in bound motion, and we take this as meaning stable equilibrium. If

$$\frac{d^2U}{dx^2}\Big|_{x_0} < 0 \, ,$$

the potential would have looked like:



Clearly, there is no stable equilibrium in this case. A position x_0 where

$$\frac{dU}{dx}\Big|_{x_0} = 0 \,,$$

is called an **extremum** of $U(x)$.

Under the assumption on the expandability of $U(x)$ in a Taylor series, we find that

$$\boxed{\frac{dU}{dx}\Big|_{x_0} = 0 \,,} \tag{C.1}$$

is **necessary** (but not sufficient) for stable equilibrium. If we have in addition the condition,

$$\boxed{\frac{d^2U}{dx^2}\Big|_{x_1} > 0 \,,} \tag{C.2}$$

then (C.1) and (C.2) together constitute **sufficient** conditions for stable equilibrium. (C.2), however, is clearly not necessary itself since we might encounter

$$\frac{d^2U}{dx^2}\Big|_{x_0} = 0 \,.$$

Under these conditions, one must then look at the first nonzero derivative of $U(x)$ at x_0 to decide the matter.

These considerations allow us to solve for the small oscillations near an equilibrium position for a general one dimensional potential. The force at x_0 is of course

$$F = -\frac{d}{dx} U \Big|_{x_0} \,, \tag{2.53}$$

and from (2.52) near equlibrium this gives

$$F \approx -(x - x_0)\frac{d^2U}{dx^2}\Big|_{x_0} \,. \tag{2.54}$$

Choosing the origin at $x = x_0$ now tells us that to lowest order

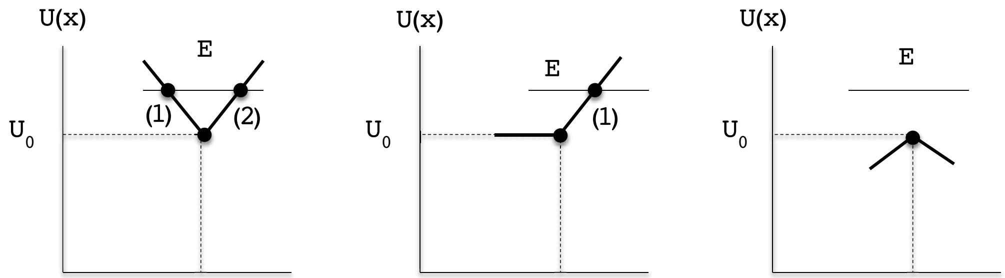
$$m\,\ddot{x} = F \approx -x\,\frac{d^2U}{dx^2}\Big|_{x_0} \,. \tag{2.55}$$

Thus for small oscillations we are back to the Hooke's law equation $m\,\ddot{x} = -k\,x$ where $k \equiv \frac{d^2U}{dx^2}\Big|_{\bar{x}=\bar{x}_0}$ as long as condition (C.2) holds. This means the formula (2.51) holds and that the angular frequency of small oscillations is determined by

$$\omega_0^2 = \frac{k}{m} = \frac{1}{m}\,\frac{d^2U}{dx^2}\Big|_{x_0} \,, \tag{2.56}$$

Although the usual meaning of the x variable is spatial position, it can also represent other things. In an end of Chapter problem it is seen that these ideas are aso useful when x represents an angle.

If a Taylor series expansion can not be made for a point because of discontinuities, one must inquire more generally about the meaning of a stable equilibrium point. The more general meaning is associated with the turning point idea. Let us imagine increasing the energy, E, of the particle infinitesimally higher than U_0 in the three following examples. (Increasing E is the same as giving an initial infinitesimal velocity to the particle.)



2 turning points (a) 1 turning point (b) no turning points (c)

Clearly, for the motion to be stable, there must be two turning points, labeled (1) and (2) as in (a) above, located infinitesimally close to x_0. If there is only a single such turning point, as in (b), the motion will not be bounded in one direction and we have an unstable equilibrium. If there are no turning points, as in (c), this also represents unstable equilibrium. Note that if there are 2 turning points, but one or the other is located a **finite** distance from x_0, resulting in bounded but finite motion, we also consider this as unstable equilibrium.

Given that $\frac{dU}{dx}\Big|_{x_0} = 0$, the more general meaning of stable equilibrium is:

> Under an infinitesimal increase in the total energy $E = U(x_0) + \delta E$, there are two turning points located infinitesmaly close to x_0

A point, x_0, which has $\frac{dU}{dx}\Big|_{x_0} = 0$ but which does not satisfy the above criterion is considered an **unstable equilibrium** point.

There is also the possibility of **neutral equilibrium**, by which I mean a point x_0 where **all** the derivatives of $U(x)$ vanish. Then of course any particle motion is simply uniform in the vicinity of x_0.

Equilibrium and Stability in D Dimensions

What about equilibrium in multidimensional problems? If we again assume that $U(\vec{x})$ can be expanded about the position $\vec{x}_0$ in a Taylor series, we have (understood sums up to dimension D on repeated indices)

$$U(x) = U(\vec{x}_0) + (x - x_0)_i \nabla_i U\big|_{\vec{x}_0} + \frac{1}{2}(x - x_0)_i (x - x_0)_j \nabla_i \nabla_j U\big|_{\vec{x}_0} + \ldots \qquad (2.57)$$

Again, the constant, $U(\vec{x}_0)$, can be ignored in the discussion of particle motion. The analog of the extremum condition (C.1) is now

$$\boxed{\nabla_i U\big|_{\vec{x}_0} \equiv \frac{\partial}{\partial x_i} U\big|_{\vec{x}=\vec{x}_0} = 0.} \qquad (C.3)$$

This is again a necessary, but not sufficient condition for stable equilibrium. Notice that (C.3) represents D conditions in a D-dimensional problem.

What is the analog of (C.2) in this case? We will just generalize the argument of the last section to more than one dimension. Assuming (C.3) holds and using (2.57), the force on the particle at $\vec{x} = \vec{x}_0$ is then given as (I'm using index notation)

$$F_\ell = -\nabla_\ell U \,\big|_{\vec{x}_0} \equiv -\frac{\partial}{\partial x_\ell} U \,\big|_{\vec{x}=\vec{x}_0} ,$$

$$= -\frac{1}{2} \nabla_i \nabla_j U \,\big|_{\vec{x}_0} \left(\underbrace{\frac{\partial (x - x_0)_i}{\partial x_\ell}}_{=\delta_{i\ell}} (x - x_0)_j + (x - x_0)_i \underbrace{\frac{\partial (x - x_0)_j}{\partial x_\ell}}_{=\delta_{j\ell}} \right),$$

$$= -\frac{1}{2} \left(\nabla_\ell \nabla_j U \,\big|_{\vec{x}_0} \right) (x - x_0)_j - \frac{1}{2} \left(\nabla_i \nabla_\ell U \,\big|_{\vec{x}_0} \right) (x - x_0)_i$$

$$= -A_{\ell j} (x - x_0)_j \,, \tag{2.58}$$

where

$$A_{\ell j} \equiv \nabla_\ell \nabla_j U(x) \,\big|_{\vec{x}_0} . \tag{2.59}$$

These set of constants, $A_{\ell j} = A_{j\ell}$, are the analog of the single spring constant, k, in the one-dimensional case. Choosing our origin at $\vec{x} = \vec{x}_0$, the equations of motion are now

$$m\ddot{x}_\ell = F_\ell = -A_{\ell j} x_j \,,$$

$$\Rightarrow m\ddot{x}_\ell + A_{\ell j} x_j = 0. \tag{2.60}$$

This represents a set of **coupled** linear differential equations in the quantities x_l. Let us suppose a solution of the form

$$x_\ell = a_\ell e^{i\omega t} . \tag{2.61}$$

We are now using complex numbers; the a_ℓ are complex (and the $A_{\ell j}$ are real) but it is understood that x_ℓ is represented by the real part of the right-hand side of Eq.(2.61). Then, the equations (2.60) become simply

$$(A_{\ell j} - m\omega^2 \delta_{\ell j}) a_j = 0, \tag{2.62}$$

where the $-\omega^2$ factor in the second term comes from taking the second time derivative in (2.60). Eqs.(2.62) represent D linear, homogeneous equations in the D unknowns, a_j . As was quoted in the first Chapter, the condition for such systems of equations to have nontrivial solutions is that the determinant of the coefficients of the unknowns must vanish. Thus, defining

$$C_{ij} \equiv A_{ij} - m\omega^2\delta_{ij} , \tag{2.63}$$

we must have

$$\det (C) = 0. \tag{2.64}$$

Eq.(2.64) determines (the squares of) the **characteristic angular frequencies**, ω^2, which enter in (2.61) above. In D dimensions, the determinant equation, (2.64), represents D equations, giving D solutions, $\omega_1^2, \omega_2^2, \ldots, \omega_D^2$.

In a problem at the end of the Chapter we will see that the ω^2 are real quantities. Let us examine what would happen if one of these frequencies had a nonzero imaginary part. Notice that since (2.64) is actually a condition on ω^2, not ω, the supposed complex roots occur in pairs,

$$\omega = \pm i\omega_c,$$

where ω_c is real. Thus, in this case the general solution would read

$$x_i = a_i(A\, e^{-\omega_c t} + B\, e^{\omega_c t}),$$

where A and B are constants. Requiring

$$\begin{cases} x(0) = 0 \Rightarrow B = -A, \\ \dot{x}(0) = v_0 \Rightarrow a_i A = -\dfrac{v_0}{2\omega_c}, \end{cases}$$

$$\Rightarrow \text{Real}(x_i) = \frac{v_0}{\omega_c} \sinh(\omega_c t),$$

$$\Rightarrow \lim_{t\to\infty} \text{Real}\, x_i \to \pm\infty .$$

Therefore, the motion is unbounded for $\omega_c \neq 0$.

A sufficient condition for stable equilibrium in higher dimensions is thus,

Given (C.3) holds at $\bar{x}_0$ and the squares of the characteristic angular frequencies, determined by det(C)=0 are positive, the potential possesses stable equilibrium at $\bar{x}_0$.	(C.4)

The characteristic angular frequencies, ω_k (k=1,...,D) are analogous to the single natural frequency, w0, in the Hooke's law case. As we found,

$$\omega_0^2 = \frac{k}{m} = \frac{1}{m}\frac{d^2U}{dx^2}\Big|_{x_0} .$$

Thus the condition in (C.4), "the ω_k^2 are positive", is clearly a generalization of condition (C.2), which required

$$\frac{d^2U}{dx^2}\Big|_{x_0} > 0 ,$$

for stability.

Let us look at one example of the use of (C.4).

Example 5

In two dimensions, a particle near the origin experiences the potential,

$$U(x, y) = \frac{1}{2}k_x x^2 + \frac{1}{2}k_y y^2 + k_{xy}xy,$$

where k_x, k_y, $k_{xy} > 0$. Find the characteristic frequencies. Does the particle posess stable motions?

Solution:

Clearly the origin is an equlibrium position. We compute the nonzero A_{ij}:

$$A_{11} = k_x,\ A_{22} = k_y,\ A_{12} = A_{21} = k_{xy}. \tag{E.49}$$

We have

$$\det\begin{pmatrix} k_x - m\omega^2 & k_{xy} \\ k_{xy} & k_y - m\omega^2 \end{pmatrix} = 0. \tag{E.50}$$

$$\Rightarrow \omega^2 = \frac{1}{2m}(k_x + k_y) \pm \frac{1}{2m}\underbrace{\sqrt{(k_x + k_y)^2 + 4(k_{xy}^2 - k_xk_y)}}_{=\sqrt{(k_x - k_y)^2 + 4k_{xy}^2}} . \tag{E.51}$$

Notice we will have real solutions for ω if and only if

$$k_{xy}^2 < k_x k_y \,, \tag{E.52}$$

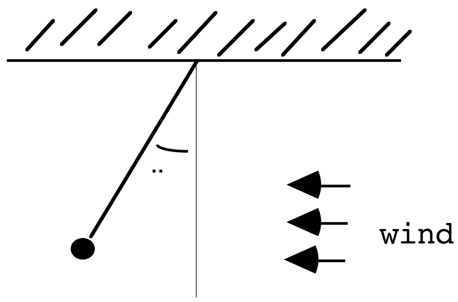
which is the condition for stability.

There are many other considerations related to stability in higher dimensions we do not have the time to cover here. One is often interested in stability under certain assumed types of small perturbations, asssociated with the various squared eigenfrequencies, ω_k^2. One concept seen later is the idea of an "effective potential" for a moving particle in an external field. It is possible to examine such a potential for stability of a given orbit rather than particle position.

CHAPTER 2 PROBLEMS

1. Our sprinter friends Michael and Frankie run another 200m race. In this race both runners accelerate uniformly and then continue to run at different top speeds. If Michael's top speed is 10 m/s achieved after 15m and if Frank's acceleration is only 2/3 of Michael's, how far must Frank accelerate to tie Michael in the race? What is Frank's final top speed?
   [The methods, not the answer, are important, so show me your steps!]

2. A steady wind is blowing on a pendulum hanging from a patio roof. The pendulum bob has mass "m" and the wire is approximately massless. The downward force of gravity is – mg.



The wind exerts a horizontal force, **f**, on the pendulum. Find the equilibrium angle, θ.

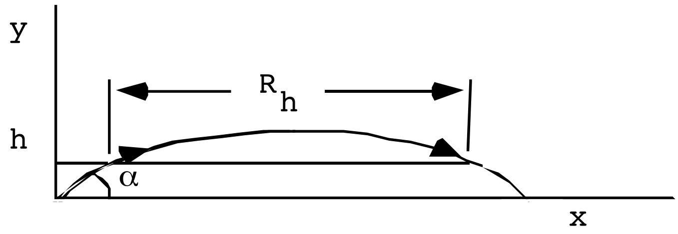
3. A baseball of mass m is thrown vertically during a game of catch. It experiences a downward force $F = -mg - kv$, where $-mg$ is the force due to gravity and $-kv$, represents air friction (v is the baseball's instantaneous velocity, $k > 0$). Given that the baseball starts at the origin of coordinates at $t=0$ with initial upward velocity v_0, solve for the position of the particle at all subsequent times, $x(t)$.

4. A particle is dropped in a medium for which the resistive force is given by $D\dot{x}^2$ $(D>0)$ where $\dot{x}$ is the particle's velocity. The force of gravity is -mg, as usual. Investigate the time taken to fall a distance h from rest. Show that to first order in D, this is given by

$$T \approx \sqrt{\frac{2h}{g}}\left(1 + \frac{Dh}{6m}\right).$$

State the approximation that is being made here.

5. Define R_h = horizontal distance between 2 pts. on the trajectory, a distance h above the horizontal. The particle is projected at an initial velocity of v_0 at an angle α above the horizontal.



a) Find the distance R_h in terms of h, v_0 and α
b) Find the angle of elevation, α, which maximizes the distance R_h for a given velocity v_0, assuming zero air resistance. [Hint: take the derivative of R_h with respect to α.]

6. A particle of mass m is subject to a force (one dimensional motion, $k>0$)

$$F(x,t) = -kx + F_0 \sin(\omega t),$$

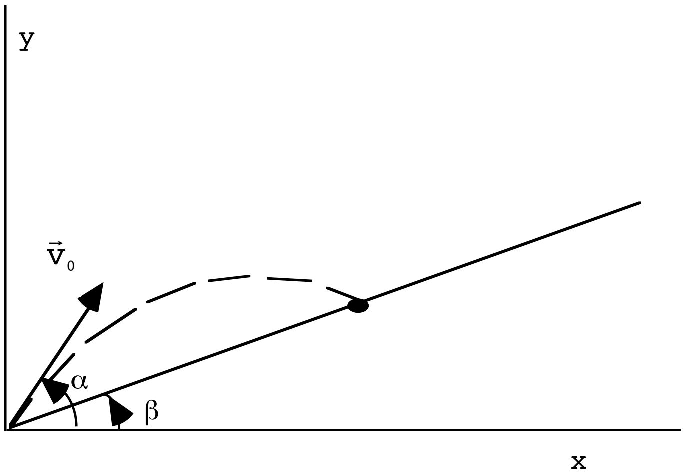
where $\omega^2 \neq \frac{k}{m}$ It also has initial conditions,

$$x(0) = x_0,$$

$$\dot{x}(0) = v.$$

Find $x(t)$ for $t > 0$.

7. Nolan Ryan (in his prime) throws a ball up an incline that makes an angle β with respect to the horizontal, as shown. The angle that the ball is thrown is α. [Neglect air resistance.]

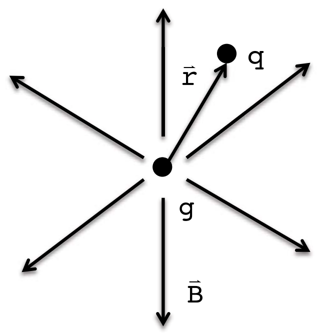


a) Given the initial magnitude of velocity, v_0, find the time it takes to hit an uphill point on the slope.

b) Nolan wants to throw the ball as far as possible up the slope. Given the same initial velocity as in (a), at what angle a should it be thrown? [Ans.: $\alpha = \frac{1}{2}\left(\beta + \frac{\pi}{2}\right)$.]

[Hints: One thing that helps is the idea that maximizing the distance up the hill is the same as maximizing the distance along the x-axis above.]

8. A magnetic monopole is a hypothetical particle with a magnetic rather than an electric charge. Assume a very massive magnetic monopole is sitting at the origin, as shown. A very light electric charge, q, is located at $\vec{r}$ relative to the origin and has a nonzero velocity.



The magnetic field of the monopole is given by

$$\vec{B} = g\frac{\hat{e}_r}{r^2} \quad ,$$

where $\hat{e}_r$ is the unit vector point in the $\vec{r}$ direction and g is the magnetic charge. The force on the charge q is given by (the Lorentz force law)

$$\vec{F} = \frac{q}{c}\dot{\vec{r}} \times \vec{B}.$$

Show that the equations of motion in cylindrical coordinates are

$\left(C \neq \frac{qg}{mc}\right.$ and $\left.r^2 \neq \rho^2 + z^2\right)$:

$$\rho: \quad \ddot{\rho} - \rho\dot{\phi}^2 = \frac{C}{r^3}\rho\dot{\phi}z,$$

$$\phi: \quad 2\dot{\rho}\dot{\phi} + \rho\ddot{\phi} = \frac{C}{r^3}(\dot{z}\rho - \dot{\rho}z),$$

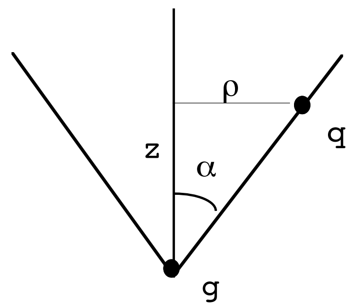
$$z: \quad \ddot{z} = -\frac{C}{r^3}\rho^2\dot{\phi} \quad .$$

[Hint: Recall the formulas for velocity and acceleration in cylindrical coordinates from the last Chapter.]

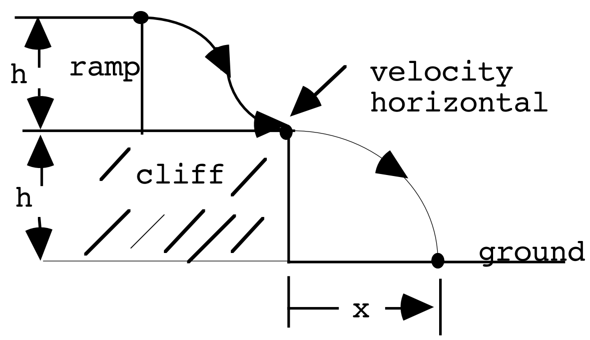
9. Show that the motion of the monopole in problem 8 takes place on a cone, the end of which starts at the stationary magnetic charge g. Hint: A cone is characterized by

$$\rho = z \tan(\alpha),$$

where α is the opening angle:

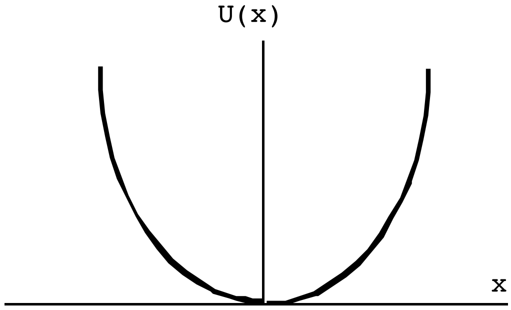


10. A bead is released (at rest) from the top of a ramp of height "h" as shown. It moves without friction until it reaches the edge of the cliff, where it's velocity is purely **horizontal**. The cliff is also of height h.

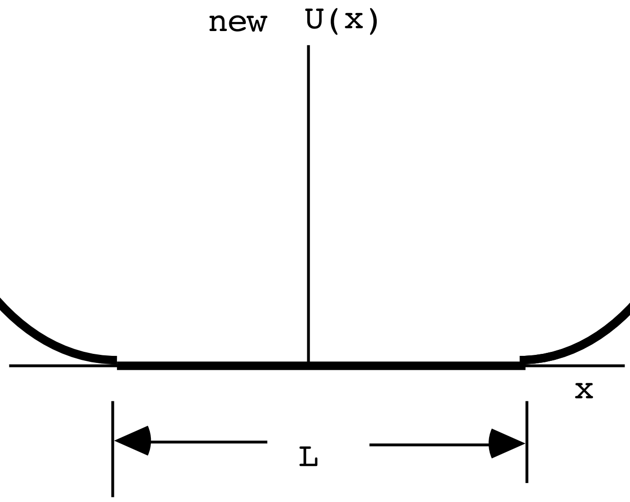


Find the distance, x, that the bead moves horizontally before striking the ground in terms of h.

11. A one dimensional potential, given by $U(x) = \frac{1}{2}kx^2$,



is split in the middle at x=0 into two similar parts by inserting a flat segment of length L, as shown:



Find the period of motion of a particle with mass m and energy E in this new potential.

12. A particle in one dimension experiences a potential, $U(x) = K|x|^n$, where $|x|$ is the absolute value of x, K is a positive constant and $n > 1$.
   a) Find the oscillation period of the particle in the potential for general n. [It is enough to set up the integral which will give this answer.]
   b) If the energy of the particle is increased by 2, by what factor does the period change? [Hint: Get the integral in dimensionless form.]

13. A one-dimensional potential is given by

$$U(x) = \begin{cases} \frac{1}{2}kx^2, & x > 0 \\ -\frac{1}{2}kx^2, & x < 0. \end{cases}$$

Is a point particle placed at the origin stable or unstable? Explain carefully.

14. A one dimensional potential is given by,

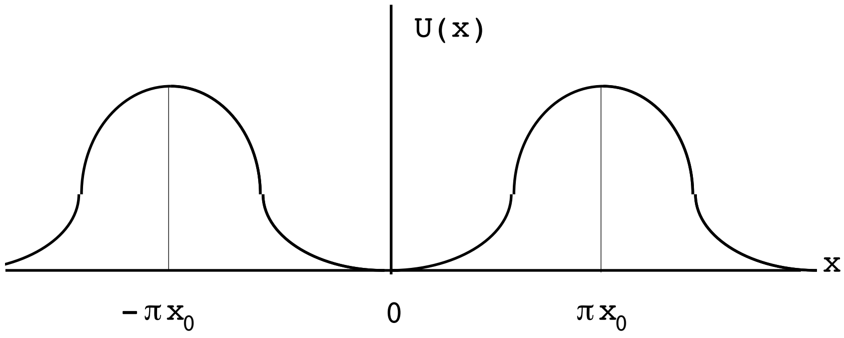
$$U(x) = -\frac{1}{2}kx^2 + \frac{1}{4}dx^4.$$

a) Find the points of stable and unstable equilibrium. Tell me which are stable, which are unstable.
b) Find the turning points for a particle of energy $E>0$.

15. A point particle is restricted to motion along the x-axis. The potential energy of the particle is given by

$$U(x) = U_0\left(1-\cos\left(\frac{x}{x_0}\right)\right).$$

This looks like:



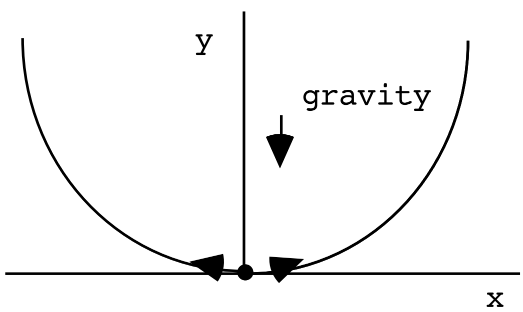
a) Find the points of stable, unstable equilibrium.
b) If the particle is started at $x=0$, what minimum velocity is necessary to reach the region $x > \pi x_0$?
c) Find the frequency of small oscillations about the point $x=0$.

16. The force as a function of position on a particle of mass 1 kg is given by

$$F = \begin{cases} -2, & x > 1 \\ -2x, & -1 < x < 1 \\ 2, & x < -1 \end{cases}.$$

a) Find the potential, $U(x)$.
b) Find the period of motion the particle in this potential if the particle starts at $x_0 = 2$ with zero initial velocity.
c) Find the period of small oscillations of this particle.

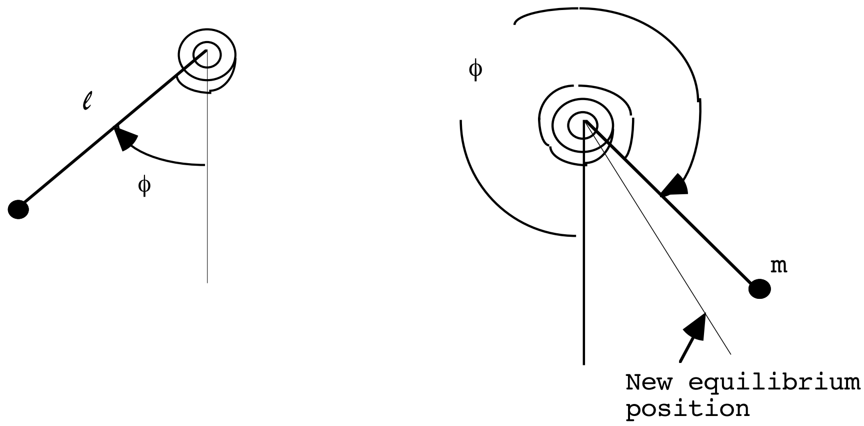
17. A particle of mass m follows the trajectory $y = Kx^2$ (K is a constant) in the xy plane under the influence of gravity as shown. The amplitude of oscillation of the particle is small. Find the period of oscillation of the particle.



18. A weak unstretched watch spring is attached to a pendulum of length ℓ with a mass m at the end. Gravity is present and the equilibrium position is at $\phi=0$. The potential energy of the mass is given by $U(\phi) = mg\ell(1-\cos\phi) + \frac{1}{2}k_s\phi^2$, where k_s is the (rotational) spring constant. Assume $k_s << mg\ell$.

a) Show that the squared angular frequency of small oscillations about the equilibrium angle, $\phi=0$, is

$$\omega^2 = \frac{g}{\ell} + \frac{k_s}{m\ell^2} .$$



b) The spring is twisted approximately 2π radians and let go. Show that the new stable equlibrium angle, ϕ_0, is given by

$$\phi_0 \approx 2\pi\left(1 - \frac{k_s}{mg\ell}\right) .$$

c) Show that the frequency of small oscillations about the new equilibrium position is approximately

$$\omega^2_{new} = \frac{g}{\ell} + \frac{k_s}{m\ell^2} - \frac{2\pi^2 k_s}{gm^2\ell^3} .$$

19. The (squared) characteristic frequencies of a system, ω^2, are determined by the system of equations,

$$A_{\ell j}\, a_j = m\omega^2 a_\ell ,$$

(understood sum on j) where the a_ℓ are complex quantities but the $A_{\ell j}$ are real and symmetric in ℓ and j One may prove that the square of the characteristic frequencies, ω^2, are real quantities by completing the following steps:

1) Multiply by a_ℓ^* on both sides of the above and sum on ℓ.
2) Take the complex conjugate of the result of (1).
3) Now complete the argument that $\left(\omega^2\right)^* = \omega^2$.

20. A particle is subject to the two dimensional potential,

$$U(x,y) = \frac{1}{2}k(x^2 + y^2) + Hyx^2,$$

where k and H are constants $(k>0)$.

a) Find the equilibrium positions of the system (there is more than one).
b) Examine the equilibrium positions in (a) for stability by calculating the characteristic angular frequencies of the system. Which positions are stable?

21. A particle is subject to the three dimensional potential,

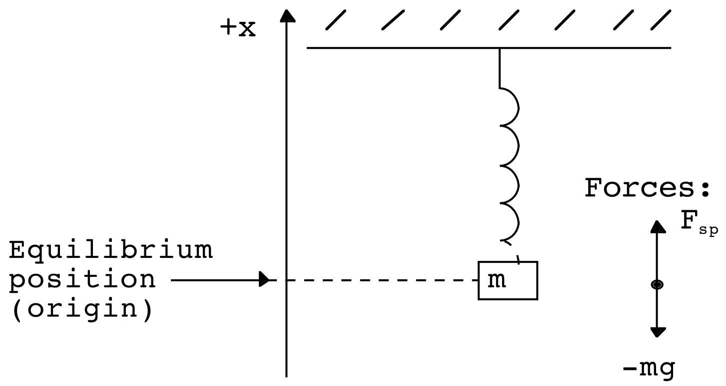
$$U(x,y,z) = \frac{1}{2}k(x^2 + y^2 + z^2) + Hxy + Gyz \ ,$$

where k, H and G are constants ($k>0$). Find the characteristic angular frequencies of the system. Show that $k > \sqrt{G^2 + H^2}$ is necessary for stability.

3 LINEAR OSCILLATIONS

GENERAL RESTORING FORCES IN ONE AND TWO DIMENSIONS

We already looked at a spring example in the last Chapter. Let us imagine setting up such a system in a gravitational field, as below. (Assume the motion is approximately one-dimensional.)



Displace the mass from equilibrium or give it an initial velocity. Newton's law is

$$m\vec{a} = \vec{F}_{tot} \,,$$

or

$$m\ddot{x} = F_{sp}(x) - mg, \tag{3.1}$$

in this case. Do a Taylor series expansion on $F_{sp}(x)$:

$$F_{sp}(x) = F_{sp}(0) + x\left(\frac{\partial F_{sp}}{\partial x}\right)_0 + \frac{x^2}{2}\left(\frac{\partial^2 F_{sp}}{\partial x^2}\right)_0 + \ldots. \tag{3.2}$$

The above coordinate system has been defined so that the total force vanishes at the equilibrium position,

$$F_{sp}(0) - mg = 0. \tag{3.3}$$

Thus

$$m\ddot{x} = x\left(\frac{\partial F_{sp}}{\partial x}\right)_0 + \frac{x^2}{2}\left(\frac{\partial^2 F_{sp}}{\partial x^2}\right)_0 + \ldots. \tag{3.4}$$

Defining ("spring constant" again)

$$k \equiv -\left(\frac{\partial F_{sp}}{\partial x}\right)_0 , \tag{3.5}$$

and neglecting higher-order terms in x in Eq.(3.4), we have the approximate, small oscillation, equation of motion (assuming $k > 0$)

$$\ddot{x} + \frac{k}{m} x = 0 , \tag{3.6}$$

which we have seen and solved previously. The most general solution involves two undetermined constants and can be assumed in the form ($\omega_0 \equiv \sqrt{\frac{k}{m}}$)

$$x(t) = A\cos\omega_0 t + B\sin\omega_0 t, \tag{3.7}$$

or

$$x(t) = C\sin(\omega_0 t - \delta). \tag{3.8}$$

It's easy to find the relationship between the coefficients in (3.7) and (3.8). We have

$$C\sin(\omega_0 t - \delta) = C\cos\delta\sin(\omega_0 t) - C\sin\delta\cos(\omega_0 t),$$

$$\Rightarrow \quad \begin{cases} A = -C\sin\delta \\ B = C\cos\delta \end{cases} \Rightarrow A^2 + B^2 = C^2 . \tag{3.9}$$

(Note that we may always choose $C>0$.) Of course, $D\cos(\omega_0 t - \delta')$ is also a possible general solution.

As usual, the first integral of the force equation with respect to x gives work done and yields the energy conservation equation:

$$\int dt\, \dot{x}\left[\ddot{x} + \omega_0^2 x = 0\right],$$

$$\Rightarrow \frac{1}{2}\dot{x}^2 + \frac{1}{2}\omega_0^2 x^2 = \text{const.} \tag{3.10}$$

Multiply by m and identify the terms:

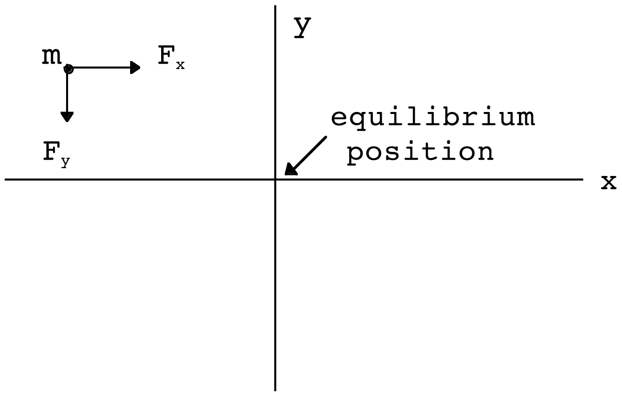
$$\frac{1}{2}m\dot{x}^2 + \frac{1}{2}kx^2 = \text{const.},$$

$$\text{or } T + U = E. \tag{3.11}$$

Usual connections (τ is the period, ν_0 is frequency):

$$\left.\begin{aligned} \omega_0\tau = 2\pi \ , \ \tau = \frac{2\pi}{\omega_0} = 2\pi\sqrt{\frac{m}{k}} \ , \\ \nu_0 = \frac{\omega_0}{2\pi} \ , \ \nu_0 = \frac{1}{\tau} \ . \end{aligned}\right\} \tag{3.12}$$

Let's consider forces in two dimensions.



Let's assume

$$\left.\begin{aligned} F_x(x, y) = F_x(x) \ , \\ F_y(x, y) = F_y(y) \ , \end{aligned}\right\} \tag{3.13}$$

so the motions in the x and y directions are independent. Again, for small amplitudes around equilibrium

$$F_x(x) \simeq \underbrace{F_x(0)}_{0} + x\left(\frac{\partial F_x}{\partial x}\right)\Big|_{x=0} \ ,$$

$$F_y(y) \simeq \underbrace{F_y(0)}_{0} + y\left(\frac{\partial F_y}{\partial y}\right)\Big|_{y=0} \ .$$

Define the independent spring constants

$$k_x = -\left(\frac{\partial F_x}{\partial x}\right)\Big|_{x=0} \ , \ k_y = -\left(\frac{\partial F_y}{\partial y}\right)\Big|_{y=0} \ ,$$

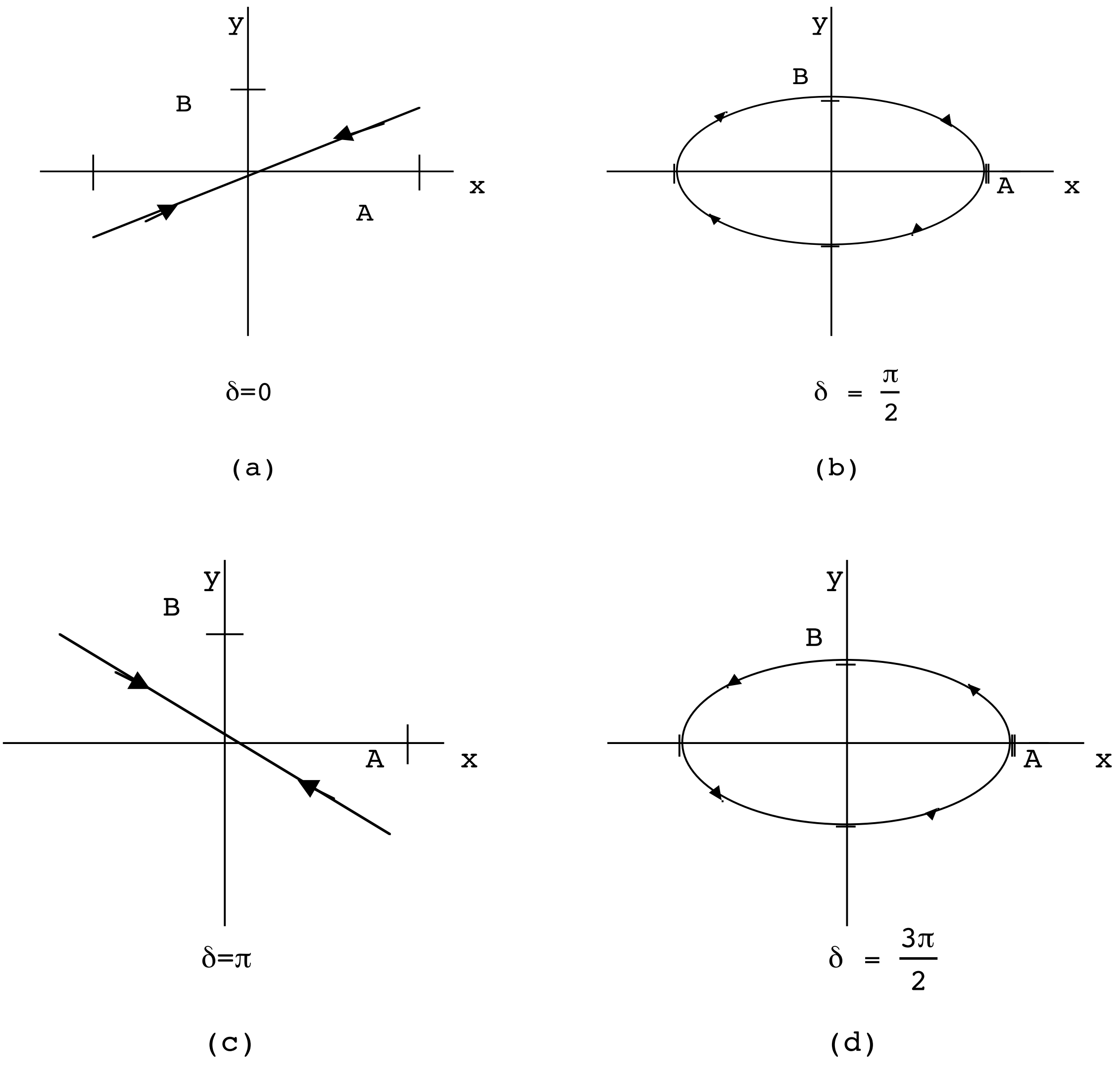
and assume they are positive. Then we have the equations of motion

$$\left.\begin{aligned} \ddot{x} + \omega_x^2\, x = 0 \ , \\ \ddot{y} + \omega_y^2\, y = 0 \ . \end{aligned}\right\} \tag{3.14}$$

where $\omega_x^2 = \dfrac{k_x}{m}$, $\omega_y^2 = \dfrac{k_y}{m}$. The general solutions are

$$\left.\begin{aligned} x(t) &= A\cos(\omega_x t - \alpha), \\ y(t) &= B\cos(\omega_y t - \beta). \end{aligned}\right\} \tag{3.15}$$

Special case: $\omega_x = \omega_y = \omega$, $A,\ B > 0$. Then we get the following behaviors as a function of $\delta = \alpha - \beta$, the phase difference.



Notice the (b) and (d) trajectories are the same except the sense of rotation is opposite. In the equal frequency case, one may show that

$$x^2 + y^2\frac{A^2}{B^2} - 2xy\frac{A}{B}\cos\delta = A^2\sin^2\delta\,. \tag{3.16}$$

Now by introducing new rotated coordinates x' and y',

$$\left.\begin{aligned} x &= x'\cos\gamma - y'\sin\gamma, \\ y &= x'\sin\gamma + y'\cos\gamma, \end{aligned}\right\} \tag{3.17}$$

and substituting in (3.16), one may now show with the choice

$$\tan 2\gamma = \frac{2AB\cos\delta}{A^2 - B^2}, \tag{3.18}$$

that (3.16) can now be written as

$$Cx'^2 + Dy'^2 = 1, \tag{3.19}$$

where

$$C = \frac{B^2\cos^2\gamma - A^2\sin^2\gamma}{A^2B^2\sin^2\delta\cos 2\gamma}, \tag{3.20}$$

$$D = \frac{A^2\cos^2\gamma - B^2\sin^2\gamma}{A^2B^2\sin^2\delta\cos 2\gamma}. \tag{3.21}$$

(3.19) is the equation of an ellipse, if $\sin\delta \neq 0$. In the case $\delta = \pm\frac{\pi}{2}$, we get for example

$$\gamma = 0 \Rightarrow \qquad \frac{x'^2}{A^2} + \frac{y'^2}{B^2} = 1.$$

Let's say now that $\omega_y = \omega_x + \Delta\omega$. Then

$$\left.\begin{aligned} x(t) &= A\cos(\omega_x t - \alpha), \\ y(t) &= B\cos(\omega_x t + (\Delta\omega t - \beta)). \end{aligned}\right\} \tag{3.22}$$

Effectively, $\delta(t) = \alpha - \beta + \Delta\omega t$, a time-dependent phase factor now. If $\Delta\omega << \omega_x$ one can think of an evolution in time from diagrams (a) to (d) above as δ changes from 0 to 2π. This evolution occurs more quickly for larger $\Delta\omega$.

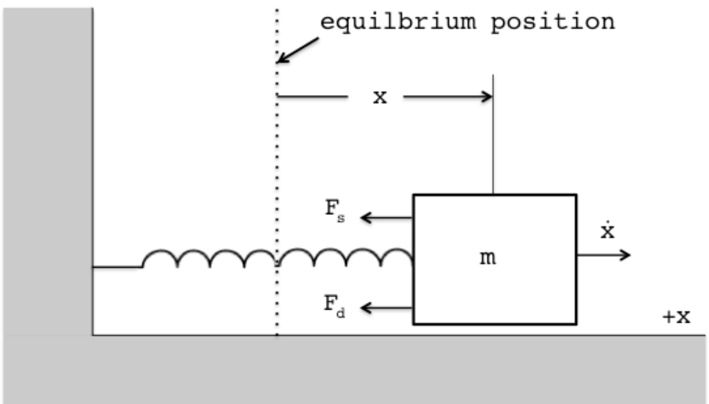
Of course, many more types of motions exist when $\omega_x \neq \omega_y$. Consider two cases:

1. $\dfrac{\omega_x}{\omega_y}$ = rational number. We are guaranteed in this case to generate a closed trajectory. Take the case,
$$\frac{\omega_x}{\omega_y} = \frac{199}{200} \Rightarrow \frac{\tau_y}{\tau_x} = \frac{199}{200},$$
which means it will take 200 repetitions of the τ_y period to equal 199 exact repetitions of τ_x, after which the motion repeats. These are called "Lissajous curves." Some of these are quite beautiful.

2. $\dfrac{\omega_x}{\omega_y}$ = irrational number. In this case, the motion will never repeat and every point in the range $-|A| < x < |A|$, $-|B| < y < |B|$ will be reached by the motion in principle.

Damped oscillations

Next, imagine making the one-dimensional situation more realistic by adding a damping force.



There are now two horizontal forces acting on the mass, the spring, F_{sp}, and the frictional damping force, F_d. We notice that the damping force must reverse sign when the velocity reverses. We will set

$$F_{tot}(x, \dot{x}) = F_{sp}(x) + F_d(\dot{x}), \tag{3.23}$$

$$\left.\begin{array}{l} F_{sp} = -kx \ , \quad k > 0, \\ F_d = -b\dot{x} \ , \quad b > 0. \end{array}\right\} \tag{3.24}$$

Notice that a force proportional to $\dot{x}^2$ would not reverse sign as we desire. Our equation of motion is now of the form

$$m\ddot{x} = -kx - b\dot{x},$$

$$\Rightarrow \quad \ddot{x} + 2\beta\dot{x} + \omega_0^2 x = 0, \tag{3.25}$$

$$2\beta \equiv \frac{b}{m}. \tag{3.26}$$

Because of the damping of the motion, we no longer expect the energy (mass kinetic plus spring potential) to be conserved. However, we can still calculate, as usual, the work done on the system in the usual two ways (W_{12} denotes the work done in changing the system from state 1 to state 2):

$$W_{12} = \int_1^2 F_{tot}(x, \dot{x})\, dx = m\int_{x_1}^{x_2} dx\, \ddot{x} = m\int_{t_1}^{t_2} dt\, \dot{x}\ddot{x}$$

$$\Rightarrow \quad W_{12} = \frac{1}{2} m\left(\dot{x}_2^2 - \dot{x}_1^2\right) \tag{3.27}$$

or

$$W_{12} = -\int_1^2 (kx + b\dot{x})dx = -\frac{1}{2}k\left(x_2^2 - x_1^2\right) - b\int_{t_1}^{t_2} dt\, \dot{x}^2. \tag{3.28}$$

If we use $t=0$ as the starting time of the motion, (3.27) must equal (3.28), and we find

$$E(t_2) + b\int_0^{t_2} dt\, \dot{x}^2 = E(t_1) + b\int_0^{t_1} dt\, \dot{x}^2, \tag{3.29}$$

where $E(t) \equiv \frac{1}{2}m\dot{x}^2 + \frac{1}{2}kx^2$ gives the instantaneous system energy. (This procedure is equivalent to integrating the equation of motion, (3.25), as we did before to get (3.10).) Eq.(3.29) informs us that

$$E_d(t') \equiv b\int_0^{t'} dt\, \dot{x}^2 , \tag{3.30}$$

is the energy lost during the motion because of frictional damping, and the total energy, including heat, is conserved. Thus,

$$E(t) + E_d(t) = \text{const.}, \tag{3.31}$$

and the instantaneous change in the system energy (kinetic plus potential) is

$$\frac{dE(t)}{dt} = -\frac{dE_d}{dt} = -b\dot{x}^2 . \tag{3.32}$$

The right hand side is negative for $b>0$, as it should be.

How does one solve the differential equation (3.25)? We will substitute

$$x = e^{rt} ,$$

$$\Rightarrow \qquad \dot{x} = re^{rt} , \quad \ddot{x} = r^2 e^{rt} ,$$

giving the "characteristic equation",

$$r^2 + 2\beta r + \omega_0^2 = 0. \tag{3.33}$$

The above is a quadratic equation, the roots of which are given by

$$r = -\beta \pm \sqrt{\beta^2 - \omega_0^2} . \tag{3.34}$$

The factor in the square root can be positive, zero, or negative, leading to the following three cases.

(1) $\omega_0 > \beta$ ("underdamped") The roots of (3.34) are complex. Define

$$\omega_1 \equiv \sqrt{\omega_0^2 - \beta^2} > 0 , \tag{3.35}$$

$$\Rightarrow \qquad r = -\beta \pm i\omega_1 . \tag{3.36}$$

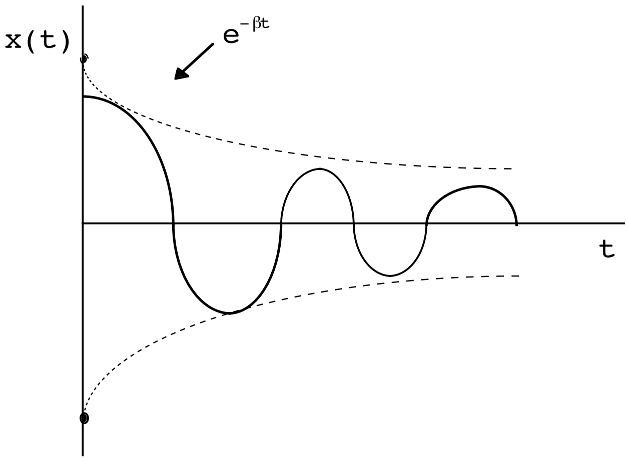
The two solutions are:

$$e^{-\beta t} e^{i\omega_1 t} , \quad e^{-\beta t} e^{-i\omega_1 t}$$

or

$$e^{-\beta t}\cos(\omega_1 t) \quad , \quad e^{-\beta t}\sin(\omega_1 t)\,,$$

When $x(0) > 0$ and $\dot{x}(0) = 0$, The motion looks like:



The system has no precise frequency since the motion is not strictly periodic. However, there is a type of period involved since aspects of the motion reoccur, although the amplitude decreases continuously. In particular the points where $\dot{x} = 0$ seem to reoccur periodically. Where does this happen? Let's build in some B.C.'s to examine the motion in detail. Use:

$$x(0) = x_0, \dot{x}(0) = v_0.$$

Assume

$$x(t) = e^{-\beta t}\left[A_1 e^{i\omega_1 t} + A_2 e^{-i\omega_1 t}\right].$$

$$x(0) = x_0 \Rightarrow x_0 = A_1 + A_2, \tag{3.37}$$

$$x(0) = v_0 \Rightarrow v_0 = e^{-\beta t}\Big\{-\beta\left[A_1 e^{i\omega t} + A_2 e^{-i\omega t}\right] + i\omega_1[A_1 e^{i\omega t} - A_2 e^{-i\omega t}]\Big\}\Big|_{t=0}$$

$$v_0 = -\beta\left[A_1 + A_2\right] + i\omega_1\left[A_1 - A_2\right]. \tag{3.38}$$

Solving (3.37) and (3.38) for A1 and A2 gives

$$A_1 = \frac{v_0 + \beta x_0 + i\omega_1 x_0}{2i\omega_1}, \tag{3.39}$$

$$A_2 = \frac{-v_0 - \beta x_0 + i\omega_1 x_0}{2i\omega_1}. \tag{3.40}$$

Introducing $e^{\pm i\omega_1 t} = \cos\omega_1 t \pm i\sin\omega_1 t$, the solution with the above B.C.'s reads

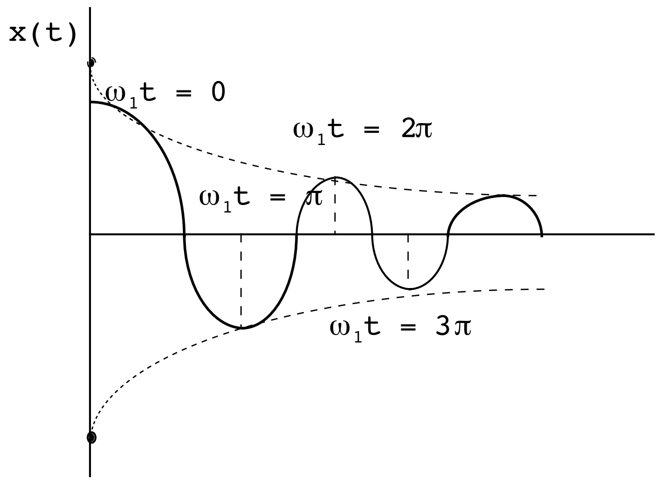
$$x(t) = e^{-\beta t}\Big[\frac{(v_0 + \beta x_0)}{\omega_1}\sin(\omega_1 t) + x_0\cos(\omega_1 t)\Big]. \tag{3.41}$$

Now investigate when $\dot{x} = 0$:

$$0 = e^{-\beta t}\Big[\frac{-\beta(v_0 + \beta x_0)}{\omega_1}\sin\omega_1 t - \beta x_0\cos\omega_1 t$$
$$+(v_0 + \beta x_0)\cos\omega_1 t - x_0\omega_1\sin\omega_1 t\Big],$$

$$\Rightarrow \frac{\beta v_0 + \beta^2 x_0 + x_0\omega_1^2}{\omega_1}\sin\omega_1 t = v_0\cos\omega_1 t. \tag{3.42}$$

When $v_0 = 0$, then $\omega_1 t = n\pi$, $n = 0,1,2,...,$ just like the undamped $(\beta = 0)$ case. The picture is:



Motion is "quasi-periodic." When $v_0 \neq 0$ the motion is still quasi-periodic $(\Delta(\omega_1 t) = \pi)$, but we now get $\omega_1 t = n\pi + \text{const.}$, $n = 0,1,2,...$

(2) $\beta = \omega_0$ ("critically damped") Now one has the repeated root,

$$r = -\beta, \tag{3.43}$$

and so the solutions are

$$e^{-\beta t} \quad , \quad te^{-\beta t} .$$

Solutions are non-oscillatory, although $x(t)$ is permitted one change of sign under certain circumstances. It's fairly easy to see what the condition is. Assume $x_0 > 0$ (β, t already are also). The only way to make $x(t) \leq 0$ is if we choose v_0 negative in the general solution,

$$x(t) = \left[x_0 + \left(v_0 + \beta x_0\right)t\right]e^{-\beta t} . \tag{3.44}$$

If we choose $v_0 \geq -\beta x_0$, we see from the above that $x(t)$ is always positive or zero. However, for $v_0 < -\beta x_0$ we must have $x(t) < 0$ for some values of t since

$$\lim_{t\to\infty} x(t) \to (v_0 + \beta x_0)te^{-\beta t} . \tag{3.45}$$

Thus, by continuity the minimum magnitude of downward v_0 which results in a change in sign of $x(t)$ is $|v_0|_{min} = \beta x_0$, toward the origin.

(3) $\omega_0 < \beta$. ("overdamped"). Define

$$\omega_2 \equiv \sqrt{\beta^2 - \omega_0^2} > 0,$$

$$\Rightarrow r = -\beta \pm \omega_2 , \tag{3.46}$$

so there are two real roots. Solutions are:

$$e^{-\beta t}e^{\omega_2 t} \quad , \quad e^{-\beta t}e^{-\omega_2 t} .$$

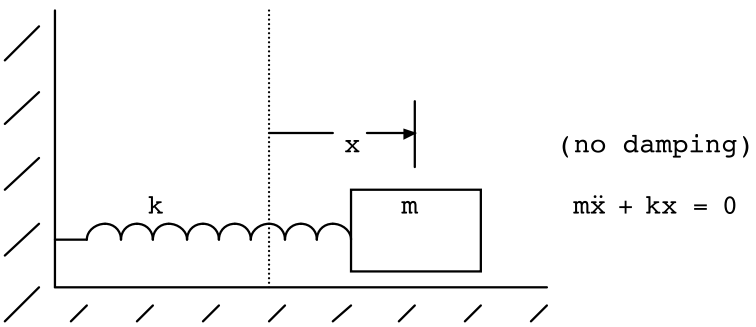
Also non-oscillatory; ω_2 is not the frequency of any motion (as ω_1 was in case (1)). $x(t)$ is again permitted one change of sign under certain conditions. I will refrain from writing down the general solution in terms of x_0 and v_0 because this will be a problem. The general form is

$$x(t) = e^{-\beta t}\left[A_1 e^{\omega_2 t} + A_2 e^{-\omega_2 t}\right]. \tag{3.47}$$

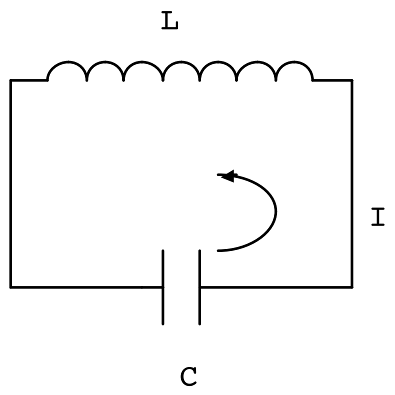
Circuit/Oscillator Analogy

We will now point out a useful analogy between mechanical oscillator and electrical circuit combinations. This will be deduced from the differential equations that describe these different situations.

The analog of



is the LC circuit (define positive direction of current to be counterclockwise)



$$V_L : L\frac{dI}{dt},$$

$$V_c : \frac{q(t)}{C} = \frac{1}{C}\int I\,dt.$$

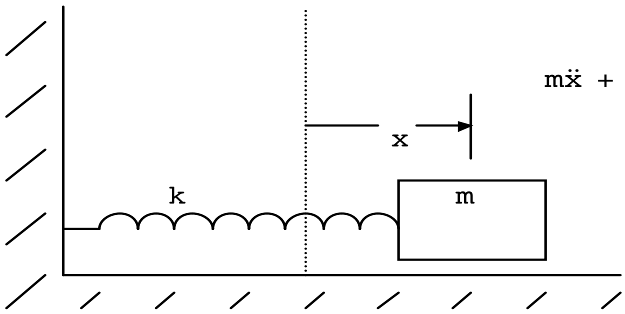
Kirchoff's law is:

$$V_L + V_c = 0\,,$$

$$\Rightarrow L\dot{I} + \frac{1}{C}\int I\,dt = 0 \qquad (I = \dot{q})$$

$$\Rightarrow L\ddot{q} + \frac{1}{C}q = 0. \tag{3.48}$$

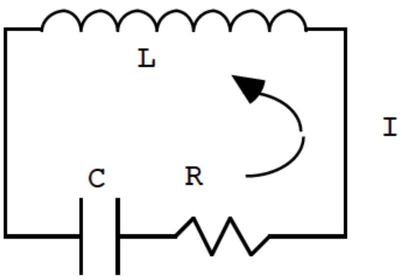
Another example:



$$m\ddot{x} + b\dot{x} + kx = 0.$$

(damping)

Compare with:



$$V_R : RI\,.$$

$$\Rightarrow L\dot{I} + RI + \frac{1}{C}\int I\,dt = 0,$$

$$\Rightarrow L\ddot{q} + R\dot{q} + \frac{1}{C} q = 0. \tag{3.49}$$

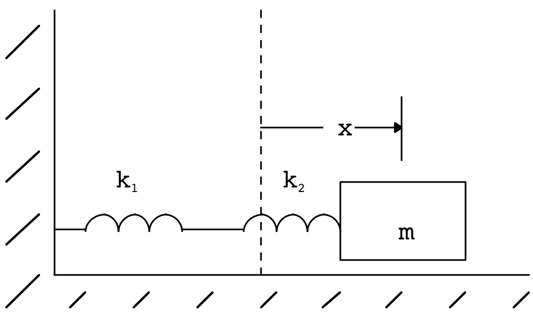
Here are the analogous quantities:

Mechanical	Electrical
force, F	voltage, V
x (mass or spring displacements)	q (inductance or capacitive charge)
$\dot{x} = v$	$\dot{q} = I$
m	L
k	$\frac{1}{C}$
b	R

By "displacements" or "charges" above I mean values measured from equilibrium. The above Table now gives:

$$\left.\begin{aligned} \omega_0 &= \sqrt{\frac{k}{m}} \leftrightarrow \frac{1}{\sqrt{LC}}, \\ \beta &= \frac{b}{2m} \leftrightarrow \frac{R}{2L}. \end{aligned}\right\} \tag{3.50}$$

Let's find the circuit analog of some mechanical situations. Consider:



$$\begin{cases} x_1 : \text{extension of spring 1} \quad , \quad F = -k_1 x_1 . \\ x_2 : \text{extension of spring 2} \quad , \quad F = -k_2 x_2 . \end{cases}$$

The same force, F, is exerted on both springs, which implies

$$k_1 x_1 = k_2 x_2 . \tag{3.51}$$

Since

$$x = x_1 + x_2\,, \tag{3.52}$$

we have

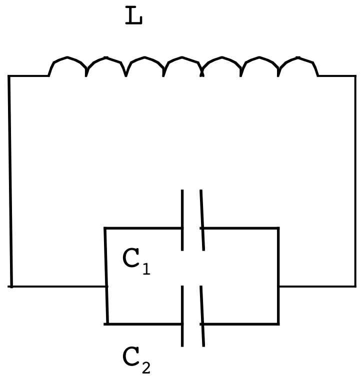
$$x = -F\left(\frac{1}{k_1} + \frac{1}{k_2}\right),$$

$$\Rightarrow \quad m\ddot{x} + \underbrace{\frac{k_1 k_2}{k_1 + k_2}}_{\equiv\, k_s} x = 0\,. \tag{3.53}$$

"k_s" is the equivalent spring constant of k_1 and k_2, attached in series. Using the above analogy, this equation becomes

$$L\ddot{q} + \frac{1}{(C_1 + C_2)} q = 0\,, \tag{3.54}$$

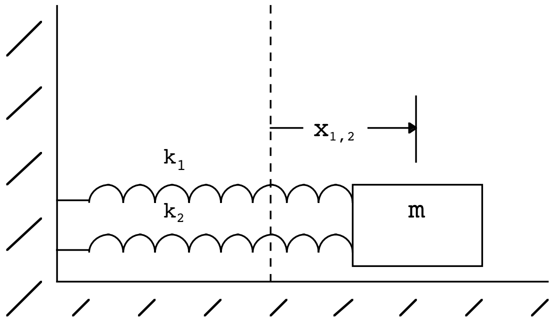
which represents a circuit with capacitors in parallel:



$$V_c : \frac{1}{C_s} q$$

$$C_p = C_1 + C_2 \; .$$

What about springs connected parallel to one another?



Here we have

$$F_1 = -k_1 x_1 \;,\; F_2 = -k_2 x_2 \;,\; x_1 = x_2 = x,$$

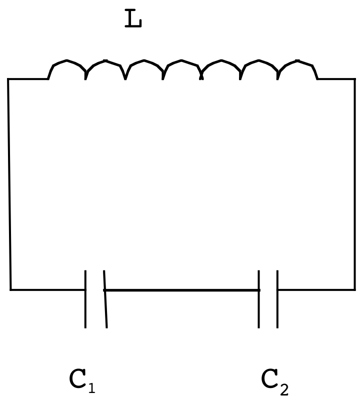
$$\Rightarrow F = F_1 + F_2 = -(k_1 + k_2)x \;,$$

$$\Rightarrow m\ddot{x} + \underbrace{(k_1 + k_2)}_{\equiv k_p} x = 0. \tag{3.56}$$

Again, substituting from the mechanical/electrical analogy, the above is equivalent to

$$L\ddot{q} + \left(\frac{1}{C_1} + \frac{1}{C_2}\right) q = 0 \;, \tag{3.57}$$

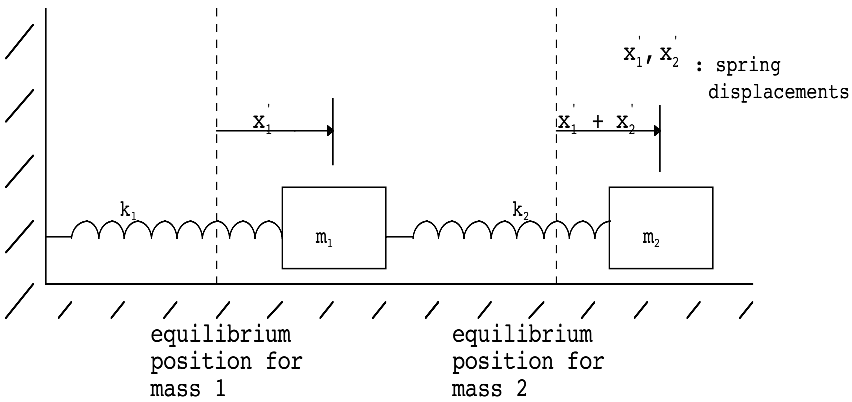
which represents a circuit with capacitors in series:



$$V_c : \frac{1}{C_p} q ,$$

$$\frac{1}{C_s} = \frac{1}{C_1} + \frac{1}{C_2} .$$

Let's see if we can find the electrical analog of one more case, this time with two masses:



I will derive the equations of motion by deriving the potential energy of the system and using Newton's second law. Since there are two springs in the system, the potential energy has the form

$$U = \frac{1}{2} k_1 x_1'^2 + \frac{1}{2} k_2 x_2'^2 . \tag{3.59}$$

However, notice that while x_1' and x_2' are the spring displacements, the corresponding mass displacements from equilibrium are x_1' and $x_1' + x_2'$. Let us define

$$\left.\begin{aligned} x_1 &= x_1' , \\ x_2 &= x_1' + x_2' , \end{aligned}\right\} \tag{3.60}$$

leading to

$$U = \frac{1}{2} k_1 x_1^2 + \frac{1}{2} k_2 (x_2 - x_1)^2 . \tag{3.61}$$

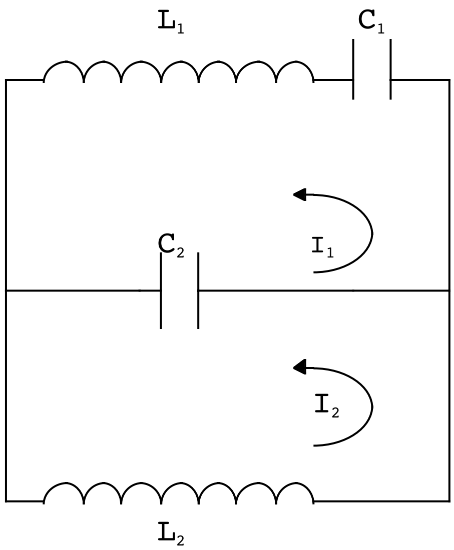
The reason for the redefinition is that Newton's laws refer to **particle**, not spring, displacements. Therefore, we have

$$\left.\begin{aligned} m_1\ddot{x}_1 &= -\frac{\partial U}{\partial x_1} = -k_1x_1 + k_2(x_2 - x_1)\,, \\ m_2\ddot{x}_2 &= -\frac{\partial U}{\partial x_2} = -k_2(x_2 - x_1)\,. \end{aligned}\right\} \quad (3.62)$$

These are our coupled differential equations. We will not attempt to solve them but will simply use our mechanical/electrical analogy to turn Eqs.(3.62) into

$$\left.\begin{aligned} L_1\ddot{q}_1 + \frac{1}{C_1}q_1 - \frac{1}{C_2}(q_2 - q_1) &= 0\,, \\ L_2\ddot{q}_2 + \frac{1}{C_2}(q_2 - q_1) &= 0\,. \end{aligned}\right\} \quad (3.63)$$

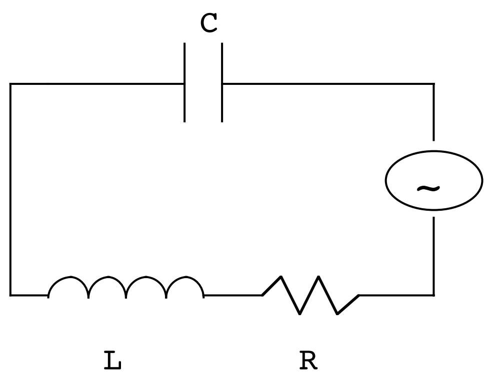
Using Kirchoff's law, this is easily seen to be equivalent to two circuits across a capacitor:



If we rewrite Eqs.(3.62) in terms of the spring diplacement variables instead of mass diplacement variables, we would have found the same circuit but with individual currents running across the capacitors rather than the inductors.

Driven Harmonic Oscillations

I now wish to discuss driven oscillations. In the electrical case, we will consider a driven LRC circuit,



$$\Rightarrow \qquad L\ddot{q} + R\dot{q} + \frac{1}{C}q = V(t). \tag{3.64}$$

We already know the mechanical analog is a driven, damped oscillator:

$$m\ddot{x} + b\dot{x} + kx = F(t),$$

or

$$\ddot{x} + 2\beta\dot{x} + \omega_0^2 x = A(t), \tag{3.65}$$

where $A(t) = \frac{F(t)}{m}$. Notice that "$F(t)$" is the external force on the particle, not the total force.

We will talk about two methods to solve Eq.(3.65):

A. Fourier Series
B. Green functions

Method "A" will work for periodic functions, $F(t)$, while method "B" works in the periodic or non-periodic case (but is more difficult to carry out).

Let's begin by talking about a special case for $F(t)$, which, however, will lead us smoothly into a discussion of Fourier series solutions. The equation we will consider is ($A_0 = \frac{F_0}{m}$ where $F(t)=F_0\cos(\omega t)$)

$$\ddot{x} + 2\beta\dot{x} + \omega_0^2 x = A_0 \cos(\omega t). \tag{3.66}$$

Assume a solution of the form,

$$x_p(t) = D(\omega_0, \omega, \beta)\cos(\omega t - \delta(\omega_0, \omega, \beta)), \tag{3.67}$$

and substitute into Eq.(3.66). After expanding the sine and cosine in the quantities ωt and δ, and equating the coefficients of sin ωt and cos ωt, we have the equations,

cosine: $$D(\omega_0^2 - \omega^2)\cos\delta + 2D\beta\omega\sin\delta = A_0\,, \tag{3.68}$$

sine: $$D(\omega_0^2 - \omega^2)\sin\delta - 2D\beta\omega\cos\delta = 0\,. \tag{3.69}$$

From (3.69),

$$\tan\delta = \frac{2\beta\omega}{\omega_0^2 - \omega^2}\,, \tag{3.70}$$

from which we find

$$\sin\delta = \frac{2\beta\omega}{\sqrt{(2\beta\omega)^2 + (\omega_0^2 - \omega^2)^2}}\,, \tag{3.71}$$

$$\cos\delta = \frac{\omega_0^2 - \omega^2}{\sqrt{(2\beta\omega)^2 + (\omega_0^2 - \omega^2)^2}}\,, \tag{3.72}$$

where I have chosen the signs in (3.71) and (3.72) such that $\delta > 0$. From (3.68), (3.71) and (3.72),

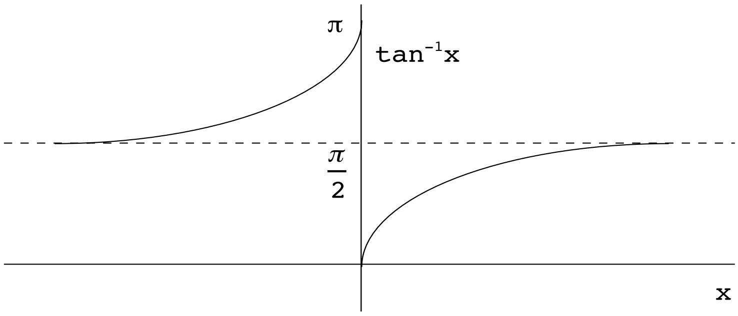
$$D(\omega) = \frac{A_0}{\sqrt{(\omega_0^2 - \omega^2)^2 + (2\beta\omega)^2}}, \tag{3.73}$$

Therefore the particular solution can be written very simply as

$$x_p(t) = D(\omega)\cos(\omega t - \delta), \tag{3.74}$$

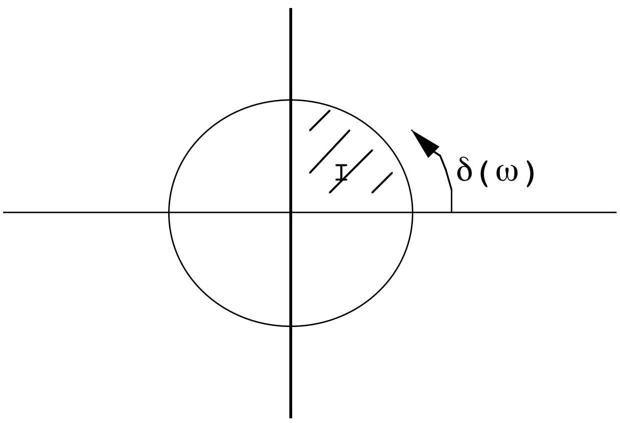
$$\delta = \tan^{-1}\left(\frac{2\beta\omega}{\omega_0^2 - \omega^2}\right). \tag{3.75}$$

Our choice $\delta > 0$ means that we do not use the principle branch of the inverse tangent, but instead use:



Notice that the particular solution, (3.74) is completely periodic; it represents the steady-state solution after all initial transients have damped out. We can now add to (3.74) the complementary solution to meet a given set of initial conditions. (The complementary solution corresponds to one of the three types of solutions for the damped oscillator studied previously.)

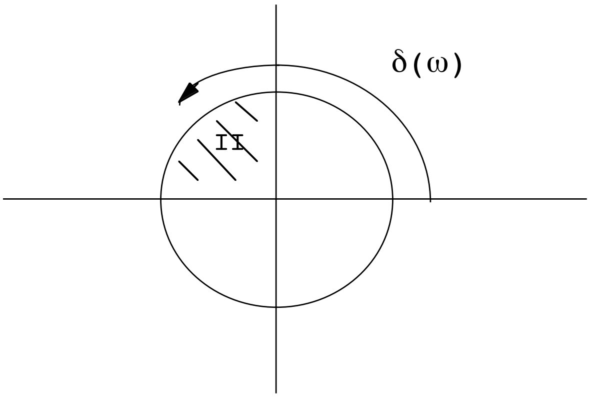
Notice that when $\omega < \omega_0 \Rightarrow \sin\delta, \cos\delta > 0$,



values of δ are in I. (We required $\delta > 0$ and we have $\delta \to 0$ as $\omega \to 0$.)

$\delta(\omega)$ increases as ω increases

When $\omega_0 < \omega \Rightarrow \sin\delta > 0$ but $\cos\delta < 0$,



values of δ are in II. since δ(ω) must be a continuous function of ω.

δ(ω) continues to increase as ω increases

In any case we see that

$$0 \le \delta \le \pi.$$

Comparing the cosines of the driving and resulting motion:

driving: $\cos(\omega t)$,
motion: $\cos(\omega t - \delta)$.

We see that δ is the phase difference between the applied force and the steady-state motion.

A quick and easy way of solving harmonic electrical/mechanical problems is to use complex numbers. Let's redo the solution in this faster manner. Consider,

$$m\ddot{x} + b\dot{x} + kx = F(t). \tag{3.76}$$

Let's assume (real parts understood)

$$\left.\begin{aligned} F(t) &= F_0 e^{i\omega t}, \\ x(t) &= x_0 e^{i\omega t} \end{aligned}\right\} \tag{3.77}$$

where F_0 is real but x_0 is complex. Substituting these above gives

$$\left[-m\omega^2 + ib\omega + k\right]x_0 = F_0,$$

$$\Rightarrow \qquad x_0 = \frac{A_0}{\left[\left(\frac{k}{m} - \omega^2\right) + i\,\frac{b}{m}\,\omega\right]},$$

or ($\omega_0^2 = \frac{k}{m}$, $2\beta = \frac{b}{m}$)

$$x_0 = \frac{A_0}{\sqrt{\left(\omega^2 - \omega_0^2\right)^2 + \left(2\beta\omega\right)^2}}\, e^{-i\delta}, \tag{3.78}$$

$$\delta = \tan^{-1}\left(\frac{2\beta\omega}{\omega_0^2 - \omega^2}\right). \tag{3.79}$$

The full solution is thus (again, take the real part)

$$x(t) = D(\omega)\, e^{i(\omega t - \delta)}, \tag{3.80}$$

the same as (3.74) above.

An important aspect of this solution is contained in the amplitude, $D(\omega)$. We will see a graph of it shortly. This graph has a maximum where

$$\left.\frac{dD(\omega)}{d\omega}\right|_{\omega = \omega_R} = 0. \tag{3.81}$$

This is called "amplitude resonance". It happens when (algebra is a problem)

$$\omega_R = \sqrt{\omega_0^2 - 2\beta^2}. \tag{3.82}$$

Thus, unlike the undamped oscillator in Chapter 2 which was seen to have an infinite amplitude (as $t \to \infty$) at $\omega = \omega_0$, the damped oscillator displays a more physical behavior.

What is the value of $D(\omega)$ at $\omega = \omega_R$?

$$D(\omega_R) = \frac{A_0}{\sqrt{\left(\omega_0^2 - \omega_R^2\right)^2 + 4\omega_R^2\beta^2}},$$

$$= \frac{A_0}{\sqrt{4\beta^2\left(\omega_0^2 - \beta^2\right)}} \underset{\beta \ll \omega_0}{\simeq} \frac{A_0}{2\beta\omega_0}. \tag{3.83}$$

Thus, for fixed driving force, $F(t)$, when $\beta \ll \omega_0$, the amplitude at resonance is inversely proportional to the damping constant, β.

Now let's look at $D(\omega)$ at the shifted frequencies, $\omega = \omega_R \pm \beta$. To the same approximation (leaving out terms β^2 or higher),

$$\omega_R \simeq \omega_0, \tag{3.84}$$

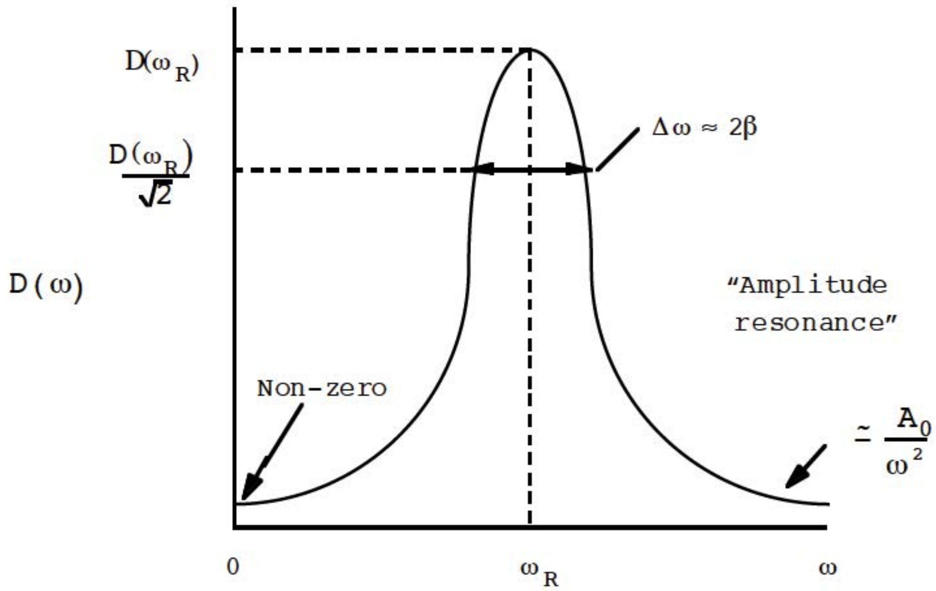
$$\Rightarrow D(\omega_R \pm \beta) \simeq \frac{A_0}{\sqrt{\left(\omega_0^2 - (\omega_0 \pm \beta)^2\right)^2 + 4\omega_0^2\beta^2}},$$

$$\simeq \frac{A_0}{\sqrt{8\beta^2\omega_0^2}} = \frac{A_0}{\sqrt{2}\,(2\beta\omega_0)}. \tag{3.85}$$

Comparing (3.85) and (3.83) shows that

$$D(\omega_R \pm \beta) \simeq \frac{1}{\sqrt{2}} D(\omega_R), \quad \beta << \omega_0. \tag{3.86}$$

The graph we obtain for $\beta << \omega_0$ is then:



Thus when $\beta << \omega_0$, the width of the amplitude response graph is given approximately by 2β. This is called the "full width" of the resonance response curve.

An important figure of merit for such oscillating systems is the so-called Q (quality) factor, defined as

$$Q_A \equiv \frac{\omega_R}{(\Delta\omega)_{Full}}, \tag{3.87}$$

where $(\Delta\omega)_{Full}$ is the full width. When $Q >> 1$ the resonance curve displays a very high, narrow peak. In a problem we will see that Q_A is proportional to the average total energy divided by the energy lost during a cycle of driven oscillation, when $\omega = \omega_R$ and $Q_A >> 1$.

An additional type of resonance is called "velocity resonance". The velocity response in the steady state is

$$\dot{x}_p(t) = -\omega D(\omega)\sin(\omega t - \delta), \tag{3.88}$$

from which we define

$$D_v(\omega) \equiv \omega D(\omega), \tag{3.89}$$

and

$$\left.\frac{dD_v(\omega)}{d\omega}\right|_{\omega=\omega_v} = 0. \tag{3.90}$$

One can show that

$$\omega_v = \omega_0 \quad . \tag{3.91}$$

so that velocity resonance occurs at the undamped angular frequency. In addition, the equality

$$Q_v \equiv \frac{\omega_v}{(\Delta\omega)_{Full}} = \frac{\omega_0}{2\beta}, \tag{3.92}$$

is exact in this situation. (One only has $Q_A \simeq \frac{\omega_0}{2\beta}$, $Q_A >> 1$, for amplitude resonance.)

When discussing driven motion, it is often convenient to introduce time-averaged quantities. We define

$$\begin{aligned} < \cdots > &\equiv \frac{1}{\tau}\int_0^{\tau} dt\,(\cdots), \\ &= \frac{\omega}{2\pi}\int_0^{\frac{2\pi}{\omega}} dt\,(\cdots). \end{aligned} \tag{3.93}$$

Therefore,

$$\left\langle \sin^2(\omega t - \delta)\right\rangle = \frac{\omega}{2\pi}\int_0^{\frac{2\pi}{\omega}} dt\,\sin^2(\omega t - \delta),$$

$$= \frac{1}{2\pi}\int_{-\delta}^{2\pi-\delta} dx\,\sin^2 x = \frac{1}{2\pi}\left(\frac{x}{2} - \frac{1}{4}\sin 2x\right)\Bigg|_{-\delta}^{2\pi-\delta},$$

$$= \frac{1}{2\pi}\left[\frac{2\pi - \delta}{2} - \frac{1}{4}\sin(4\pi - 2\delta) - \left(\frac{-\delta}{2} - \frac{1}{4}\sin(-2\delta)\right)\right],$$

$$= \frac{1}{2}. \quad \text{(independent of } \delta\text{)}$$

Similarly for $\cos^2(\omega t - \delta)$. Therefore, we have for the kinetic and potential energies,

$$\left.\begin{aligned} \langle T\rangle &= \frac{1}{4} m\omega^2 D^2(\omega), \\ \langle U\rangle &= \frac{1}{4} m\omega_0^2 D^2(\omega). \end{aligned}\right\} \tag{3.94}$$

Notice from the above that $\langle T\rangle >> \langle U\rangle$ for $\omega >> \omega_0$, so that most of the energy is kinetic. Anyone who has pumped their feet too quickly on a swing will understand this statement! Also notice that since $\langle T\rangle \propto D_v^2(\omega)$, the kinetic energy resonates at $\omega = \omega_V$ (velocity resonance) while potential energy resonates at $\omega = \omega_R$ (amplitude resonance).

Fourier Series Methods

Our solution, Eqs.(3.74) and (3.75), to the differential equation, (3.66), now affords a general method of solving (3.65) when the driving force, $F(t)$, is periodic. Our differential equation can be written in operator form as

$$\vartheta\, x(t) = A(t), \tag{3.95}$$

where

$$\vartheta \equiv \frac{d^2}{dt^2} + 2\beta\frac{d}{dt} + \omega_0^2. \tag{3.96}$$

A linear differential operator is such that

$$\vartheta(x_1 + x_2) = \vartheta x_1 + \vartheta x_2. \tag{3.97}$$

Eq.(3.97) holds for any derivative operator of any order, which can be proven by induction. Thus, (3.96) is a linear operator. This would not be the case, for example, for a damping term proportional to $\dot{x}^2$, since

$$\left(\frac{d(x_1 + x_2)}{dt}\right)^2 \neq \left(\frac{dx_1}{dt}\right)^2 + \left(\frac{dx_2}{dt}\right)^2.$$

The linearity of (3.96) now implies that the steady state solution to (ϕ is an arbitrary phase)

$$F(t) = \sum_n \alpha_n \cos(\omega_n t - \phi), \tag{3.98}$$

is just

$$x_p(t) = \frac{1}{m}\sum_n \frac{\alpha_n}{\sqrt{(\omega_0^2 - \omega_n^2)^2 + 4\omega_n^2\beta^2}}\cos(\omega_n t - \phi - \delta_n), \tag{3.99}$$

$$\delta_n = \tan^{-1}\left(\frac{2\omega_n\beta}{\omega_0^2 - \omega_n^2}\right). \tag{3.100}$$

The reason this is important is because of the Fourier series theorem. Any periodic function $F(t) = F(t + \tau)$ (subject to some nonrestrictive continuity conditions) can be represented by the infinite series,

$$F(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n \cos(n\omega t) + b_n \sin(n\omega t)\right), \tag{3.101}$$

where

$$\left.\begin{aligned} a_n &= \frac{2}{\tau}\int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} dt' F(t') \cos\left(\frac{2\pi n t'}{\tau}\right), \\ b_n &= \frac{2}{\tau}\int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} dt' F(t') \sin\left(\frac{2\pi n t'}{\tau}\right). \end{aligned}\right\} \tag{3.102}$$

One can actually integrate over any time interval of length τ in doing the integrals (3.102). The particular solution to (3.65) is then

$$x_p(t) = \frac{a_0}{2\omega_0^2 m} + \frac{1}{m}\sum_{n=1}^{\infty} D(n\omega)\left(a_n \cos\left(n\omega t - \delta(n\omega)\right) + b_n \sin\left(n\omega t - \delta(n\omega)\right)\right), \tag{3.103}$$

where

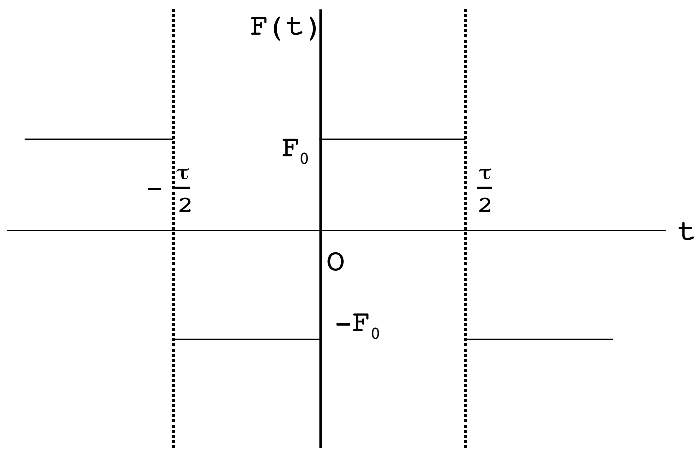
$$D(n\omega) \equiv \frac{1}{\sqrt{\left(\omega_0^2 - n^2\omega^2\right)^2 + 4n^2\omega^2\beta^2}}, \tag{3.104}$$

$$\delta(n\omega) = \tan^{-1}\left(\frac{2n\omega\beta}{\omega_0^2 - n^2\omega^2}\right). \tag{3.105}$$

Let us do a Fourier series decomposition as an example.

Example 1

$$F(t) = \begin{cases} F_0, & 0 < t < \frac{\tau}{2} \\ -F_0, & -\frac{\tau}{2} < t < 0. \end{cases}$$



Solution:

Since this is an odd function of t, we have

$$a_n = 0,$$

for all n. For b_n we have

$$b_n = -\frac{2F_0}{\tau}\int_{-\frac{\tau}{2}}^{0} dt'\sin\left(\frac{2\pi nt'}{\tau}\right) + \frac{2F_0}{\tau}\int_{0}^{\frac{\tau}{2}} dt'\sin\left(\frac{2\pi nt'}{\tau}\right)$$

$$\Rightarrow b_n = \frac{2F_0}{\pi n}(1 - \cos(\pi n)) = \begin{cases} 0, & n \text{ even} \\ \frac{4F_0}{\pi n}, & n \text{ odd}. \end{cases}$$

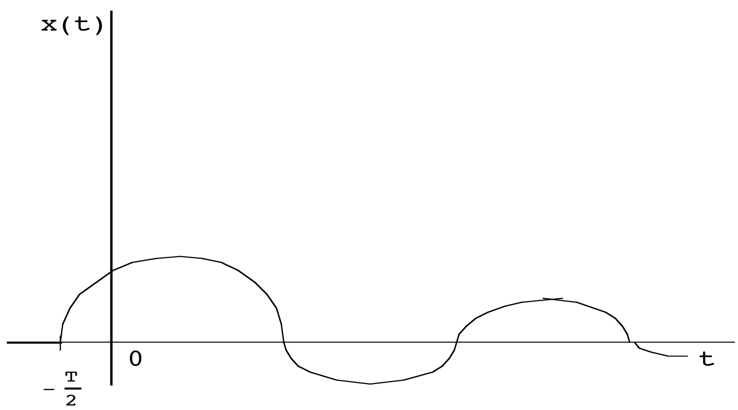
The explicit solution to the damped oscillator equation with this force is

$$x_p(t) = \frac{4F_0}{\pi m}\sum_{n=1,3,5\ldots}^{\infty}\frac{D(n\omega)}{n}\sin(n\omega t - \delta(n\omega)),$$

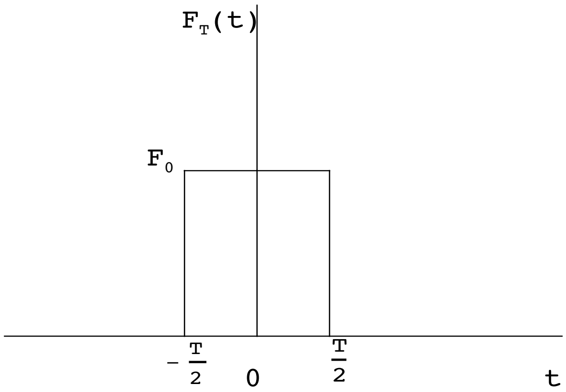
where $\delta(n\omega)$ is given by (3.105) as usual.

Green Function Methods

As mentioned above, there is another, more general method of solving the driven mechanical circuit equation (3.65). We will start the explanation of the method by asking what happens when we "kick" an underdamped oscillator. Qualitatively, we get:



We have "kicked" it at a time beginning at $t = -\frac{T}{2}$. The above is in response to an idealized force,



This is given by

$$F_T(t) = \begin{cases} 0 \quad , \quad t < -\dfrac{T}{2} \\ F_0 \quad , \quad -\dfrac{T}{2} < t < \dfrac{T}{2} \\ 0 \quad , \quad t > \dfrac{T}{2}. \end{cases} \tag{3.106}$$

Let's idealize this situation further to a very short, sharp kick so that the resulting motion is independent of the details of $F_T(t)$. Define

$$\vec{I} = \int dt\, \vec{F}(t), \tag{3.107}$$

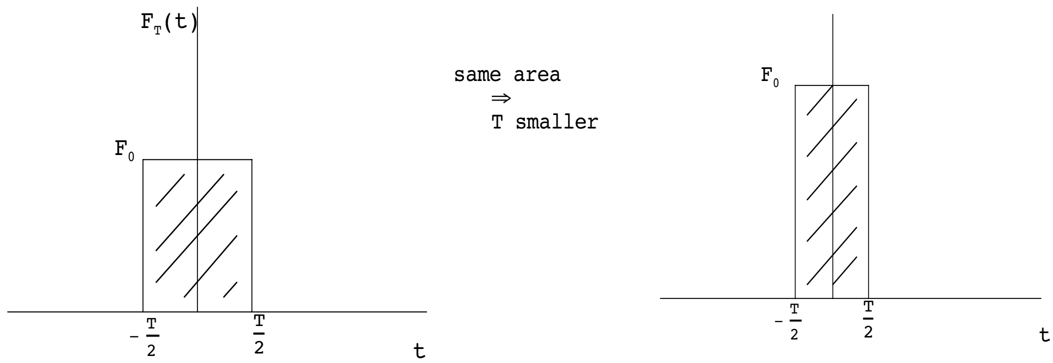
which we will call the **impulse**. It is momentum transfer. Clearly, from our form (3.106) we have

$$I = F_0 T. \tag{3.108}$$

What happens to the oscillator during this time, T? Under the conditions $T \ll \frac{1}{\beta}$ and $T \ll \frac{2\pi}{\omega_0}$, the oscillator is approximately "free". So, we simply have

$$\begin{aligned} m\ddot{x} &\simeq F_0 , \\ \dot{x} &\simeq \frac{1}{m}\int_{-\frac{T}{2}}^{\frac{T}{2}} F_0 dt = \frac{F_0 T}{m} = \frac{I}{m}, \\ \Rightarrow \quad x(T) &\simeq \frac{1}{2}\frac{F_0}{m}T^2 = \frac{1}{2}\frac{I}{m}T. \end{aligned} \tag{3.109}$$

Therefore, in the limit $T \to 0+$ ("0+" means to approach zero from positive numbers), $I = \text{const.}$, our impulse function serves **only** to give the oscillator an initial velocity, given by (3.109). This limit produces an increasingly narrow force profile:



Let's define the "kicking function",

$$\Im(t) = \lim_{\substack{T\to 0+ \\ I=\text{const.}}} F_T(t). \tag{3.110}$$

We have $(a < 0 < b)$

$$\int_a^b \Im(t)dt = \int_a^b \lim_{\substack{T\to 0+ \\ I=\text{const.}}} F_0\, dt = \lim_{\substack{T\to 0+ \\ I=\text{const.}}} \int_{-\frac{T}{2}}^{\frac{T}{2}} F_0\, dt = I, \tag{3.111}$$

where I have assumed the limit and the integral can be interchanged. Also notice for an arbitrary but continuous function $A(t)$,

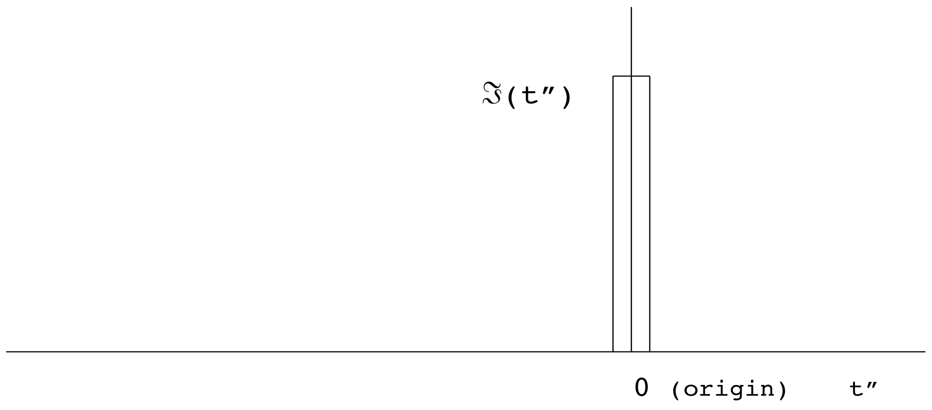
$$\int_a^b \Im(t)\, A(t)dt = \lim_{\substack{T\to 0+ \\ I=\text{const.}}} F_0 \underbrace{\int_{-\frac{T}{2}}^{\frac{T}{2}} A(t)\, dt}_{\simeq A(0)T} = I\, A(0)\,, \tag{3.112}$$

where I have made the same assumption. In addition,

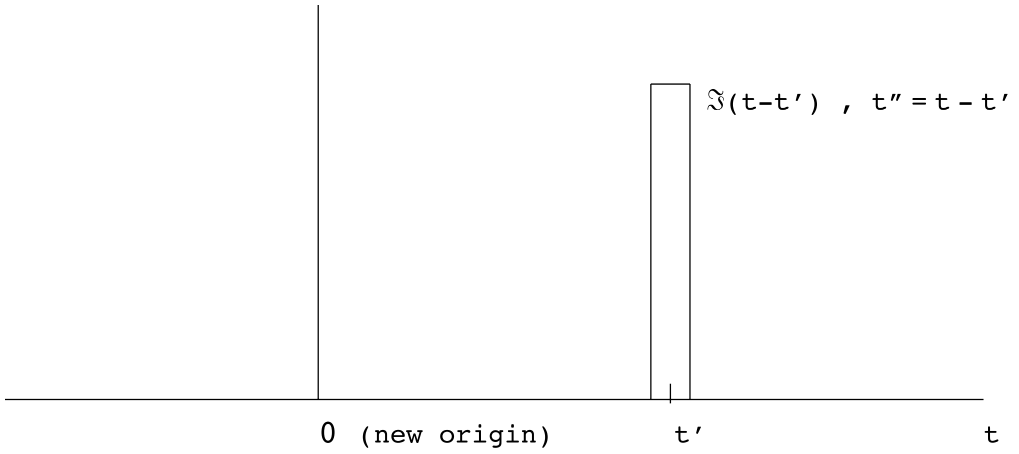
$$\Im(t) = \Im(-t)\,, \tag{3.113}$$

by construction.

Now let us say we kick the damped oscillator at time $t = t'$ instead of at $t = 0$. Here is the original situation:



Here is the same situation described with a shifted time axis:



Therefore, $t - t'$ is the appropriate shifted argument of the kicking function, $\Im$.

Let us solve

$$\ddot{x} + 2\beta\dot{x} + \omega_0^2 x = \frac{1}{m}\Im(t - t'), \tag{3.114}$$

subject to the initial conditions,

a. $x(t') = 0,$

b. $\lim_{t \to t'+} \dot{x}(t) = \frac{I}{m}.$

("$\lim_{t \to t'+}$" means t approaches t' from the positive side.) The second condition is just supplying the initial velocity condition, Eq.(3.109). Whenever $t \neq t'$ the right hand side of (3.114) is zero and so is easy to solve. (Let us assume $\beta < \omega_0$):

$$x(t) = \begin{cases} A' e^{-\beta t} \sin(\omega_1 t + \alpha'), & t > t' \\ 0, & t < t' \end{cases} \tag{3.115}$$

Notice the two undetermined constants, A' and α'. For convenience, let's take

$$\alpha' = -\omega_1 t' + \alpha,$$

$$A' = Ae^{\beta t'},$$

since t' is a constant as far as our differential equation is concerned. Using these redefined parameters in (3.115) becomes

$$x(t) = \begin{cases} Ae^{-\beta(t - t')} \sin(\omega_1(t - t') + \alpha), & t > t' \\ 0, & t < t'. \end{cases} \tag{3.116}$$

Now apply initial condition "a" above when $t = t'$:

$$\Rightarrow \quad A \sin(\alpha) = 0,$$

which means we can choose $\alpha = 0$. The remaining parameter, A, is determined by condition "b". To see how this condition comes about, let's integrate both sides of the differential equation, (3.114), over a short time interval, 2ε ($\varepsilon > \frac{T}{2}$), centered about $t = t'$:

$$\lim_{\varepsilon \to 0+} \int_{t'-\varepsilon}^{t'+\varepsilon} \left[\ddot{x} + 2\beta\dot{x} + \omega_0^2 x = \frac{1}{m}\Im(t - t')\right] dt$$

$$\Rightarrow \quad \lim_{\varepsilon \to 0+} (\underbrace{\dot{x}(t' + \varepsilon)}_{v_0} - \underbrace{\dot{x}(t' - \varepsilon)}_{0}) = \frac{I}{m}.$$

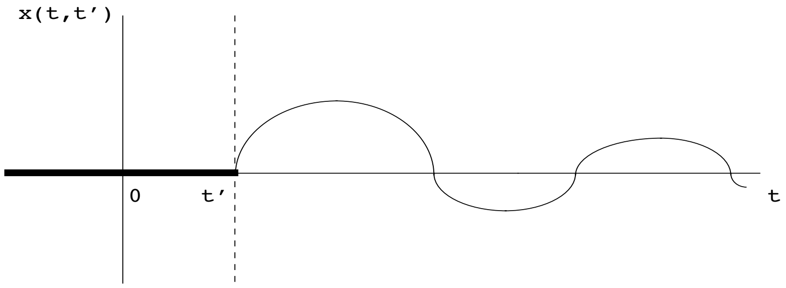
Using the explicit $t > t'$ solution then shows

$$\lim_{\varepsilon \to 0+} \dot{x}(t' + \varepsilon) = A\omega_1 \Rightarrow A = \frac{I}{m\omega_1}.$$

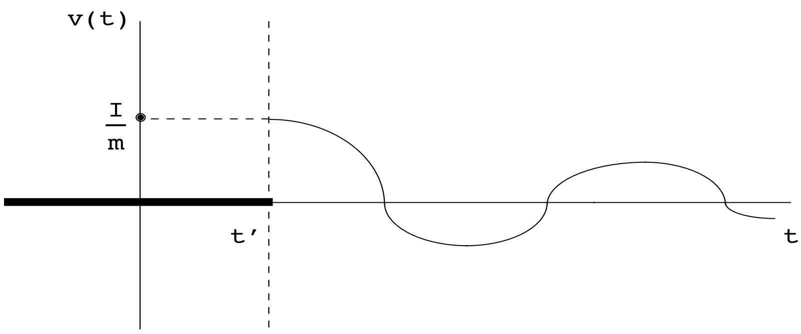
Everything is now determined. Using a new notation for our solution, $x(t) \rightarrow x(t,t')$, we have

$$x(t,t') = \begin{cases} \dfrac{I}{m\omega_1} e^{-\beta(t-t')} \sin[\omega_1(t - t')] & , t \geq t' \\ 0 & , t \leq t'. \end{cases} \tag{3.117}$$

This solution looks like we thought it would (assuming $I > 0$ so that $v_0 > 0$):



Notice that $x(t,t')$ is continuous across $t = t'$ (condition "b") but that it's first derivative is not. The velocity profile is



In solving (3.114) we have done much more than solve a rather isolated, special example. Just like the solution we formed for the harmonic force, we can now build on this solution to solve a wide variety of other problems. The comparison with the Fourier series solution is quite enlightening, and I present analogous quantities in the two situations in the following Table:

	Fourier Series	Green Function
Initial condition of oscillator	periodic motion $(x(t+\tau)=x(t))$	quiescent $(x(t,t')=0,\ t<t')$
"primitive" force	$F_0\cos(\omega t-\phi)$	
"primitive" solution	$D(\omega)\cos(\omega t-\phi-\delta)$	$x(t,t')\rvert_{t'\ \text{fixed}}$
Nature of F(t) decomposition	discrete	continuous
F(t) representation	$\sum_n \alpha_n \cos(n\omega t-\phi)$	$\int_{-\infty}^{\infty} dt'\left(\frac{F(t')}{I}\right)\Im(t-t')$
General solution for x(t)	$\sum_n \alpha_n D(n\omega)\cos(n\omega t-\phi-\delta(n\omega))$	?

In either case we are decomposing the given $F(t)$ into either a discrete series of harmonic functions (Fourier) or a continuous density of "kicks" (Green function) and using linearity and the known "primitive" solutions to get the response, $x(t)$. I have deliberately left a blank in the table where the Green function general solution is supposed to be. We can guess the form of the solution based upon the analogy. The table informs us that the following quantities or operations are analogous:

$$\sum_n \Leftrightarrow \int_{-\infty}^{\infty},$$

$$\alpha_n \Leftrightarrow \frac{F(t')}{I}dt',$$

$$D(n\omega)\cos(n\omega t-\phi-\delta(n\omega)) \Leftrightarrow x(t,t').$$

Based upon this, we will hypothesize the form of the general solution using $x(t,t')$:

$$x(t) = \int_{-\infty}^{\infty} dt'\left(\frac{F(t')}{I}\right)x(t,t'). \tag{3.118}$$

That is, the full solution is just given by integrating the force per impulse, $\frac{F(t')}{I}$, over the solutions, $x(t,t')$ For convenience, let us introduce (finally, the "Green function"!)

$$G(t,t') \equiv \frac{x(t,t')}{I}. \tag{3.119}$$

Since $G(t,t')=0$ for $t<t'$. we can write the general solution as

$$x(t) = \int_{-\infty}^{t} dt' \, G(t, t') \, F(t'). \tag{3.120}$$

Let us verify the correctness of (3.120) by substituting it back into (3.65). First, to make stonger contact with standard terminology, let us introduce

$$\delta(t - t') \equiv \frac{1}{I} \Im(t - t'), \tag{3.121}$$

called the **Dirac delta function**. We now have solved

$$\left(\frac{d^2}{dt^2} + 2\beta \frac{d}{dt} + \omega_0^2\right) G(t, t') = \frac{1}{m} \delta\left(t - t'\right),$$

where the delta function has the property ($A(t)$ is some arbitrary smooth function and $a < 0 < b$)

$$\int_a^b \delta(t)\, A(t)\, dt = \begin{cases} A(0), & a < 0 < b \\ 0, & \text{otherwise}. \end{cases}$$

To verify we have a solution, putting $x(t) = \int_{-\infty}^{\infty} dt' \, G(t, t') \, F(t')$ into (3.65) gives

$$\ddot{x} + 2\beta\dot{x} + \omega_0^2 x = \frac{1}{m}F(t),$$

$$\Rightarrow \int_{-\infty}^{\infty} dt' \underbrace{\left[\frac{d^2}{dt^2} + 2\beta\frac{d}{dt} + \omega_0^2\right] G(t,t')}_{=\frac{1}{m}\delta(t-t')} F(t') = \frac{1}{m}F(t),$$

$$\Rightarrow \frac{1}{m}F(t) = \frac{1}{m}F(t).$$

We can now automate the solution to (3.65):

(1). Solve the linear differential equation,

$$\vartheta\, G(t,t') = \frac{1}{m}\delta(t-t'), \tag{3.122}$$

for some second-order linear differential operator, $\vartheta = \frac{d^2}{dt^2} + \ldots,$ subject to the conditions,

a. $G(t,t') = 0,\ t \le t',$

b. $\lim_{t\to t'+} \dot{G}(t,t') = \frac{1}{m}.$

(2). Do the integral

$$x(t) = \int_{-\infty}^{t} dt'\, G(t,t')\, F(t'). \tag{3.123}$$

(3). One can add the complementary solution to reach some desired initial conditions.

Although our "kicked" solution, $x(t,t')$ represents an initially quiescent oscillator, this is not necessarily true for $x(t)$ in (3.123) because we can let $F(t')$ be defined all the way to $t' \to -\infty$. If we do this with a periodic $F(t')$, the resulting $x(t)$ from (3.123) will correspond to the particular solution. In addition, in step **(3)** one can add to (3.123) a complementary solution which will then match any desired initial conditions, quiescent or not. Thus, the Green function method works in essentially all cases; the price to be paid for this generality is a sometimes difficult integral to perform.

Let us do some examples to illustrate this method.

Example 2

A driven undamped oscillator's equation of motion is given by $(\omega \neq \omega_0)$

$$m\ddot{x} + kx = F_0 \sin\omega t.$$

Find $x(t)$ for an initially quiescent oscillator at $t = 0$:

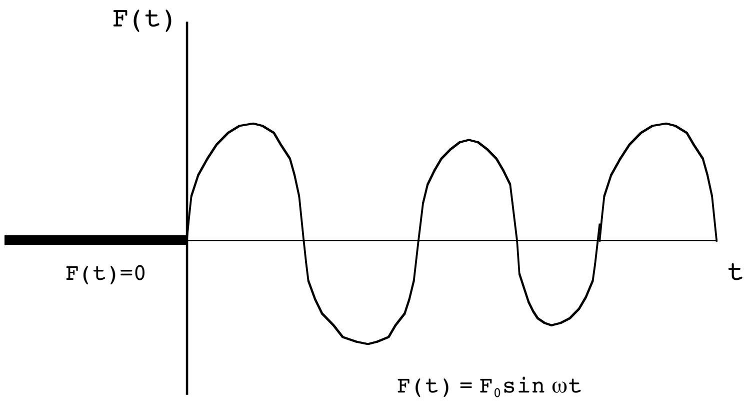
$$x(0) = 0\,, \quad \dot{x}(0) = 0\,.$$

Solution:

This is actually **Example 1** from Ch.2. The appropriate Green funtion is just the $\beta \to 0$ limit of the $G(t,t')$ in (3.117). We will use

$$G(t,t') = \begin{cases} \dfrac{1}{m\omega_0}\sin\left[\omega_0(t - t')\right], & t \geq t' \\ 0, & t \leq t'. \end{cases}$$

Since we are not interested in the solution before $t = 0$, we may take the force profile to be:



Thus

$$x(t) = \int_0^t dt'\, G(t,t')\, F(t')$$

$$= \frac{F_0}{m\omega_0}\int_0^t dt' \sin\omega_0(t - t')\sin\omega t'.$$

We can write this as

$$x(t) = \frac{F_0}{m\omega_0}\left\{\sin\omega_0 t\int_0^t dt' \cos\omega_0 t' \sin\omega t' - \cos\omega_0 t\int_0^t dt' \sin\omega_0 t' \sin\omega t'\right\}$$

Some useful integrals are, $a \neq b$:

$$\int \sin ax \sin bx \, dx = \frac{\sin(a-b)x}{2(a-b)} - \frac{\sin(a+b)x}{2(a+b)},$$

$$\int \sin ax \cos bx \, dx = -\frac{\cos(a+b)x}{2(a+b)} - \frac{\cos(a-b)x}{2(a-b)},$$

$$\int \cos ax \cos bx \, dx = \frac{\sin(a+b)x}{2(a+b)} + \frac{\sin(a-b)x}{2(a-b)}.$$

$a = b$ case:

$$\int \sin^2 ax \, dx = \frac{1}{2}\left(x - \frac{1}{2a}\sin 2ax\right),$$

$$\int \sin ax \cos ax \, dx = \frac{1}{2a}\sin^2 ax,$$

$$\int \cos^2 ax \, dx = \frac{1}{2}\left(x + \frac{1}{2a}\sin 2ax\right).$$

Using these above and doing the algebra gives

$$x(t) = \frac{F_0}{m}\frac{1}{\omega_0^2 - \omega^2}\left\{\sin \omega t - \frac{\omega}{\omega_0}\sin \omega_0 t\right\}.$$

This is the same as we had previously. We are "cracking a walnut with a sledgehammer."

Example 3

Redo the harmonically driven, damped oscillator using the Green function technique. That is, solve the differential equation

$$m\ddot{x} + b\dot{x} + kx = F_0 \cos(\omega t),$$

where the right hand side defines the force for all t, $-\infty < t < \infty$.

Solution:

Since the force has been defined for $t \to -\infty$, we should recover the particular solution, Eqs.(3.74) and (3.75) above. We have

$$x(t) = \int_{-\infty}^{t} dt' G(t, t') F_0 \cos(\omega t')$$

$$= \frac{F_0}{m\omega_1}\int_{-\infty}^{t} dt' e^{-\beta(t-t')} \sin[\omega_1(t - t')] \cos \omega t'.$$

This integral looks nasty. However, notice

$$\cos \omega t' \sin[\omega_1(t - t')] = \frac{1}{2}\{\sin[\omega t' + \omega_1(t - t')] - \sin[\omega t' - \omega_1(t - t')]\},$$

so that

$$x(t) = \frac{F_0}{2m\omega_1}\left\{\int_{-\infty}^{t} dt'\, e^{-\beta(t-t')} \sin[(\omega - \omega_1)t' + \omega_1 t] - \int_{-\infty}^{t} dt'\, e^{-\beta(t-t')} \sin[(\omega + \omega_1)t' - \omega_1 t]\right\}.$$

One can now do the two integrals above with the use of

$$\int e^{ax} \sin x dx = \frac{e^{ax}}{a^2 + 1}(a \sin x - \cos x).$$

After changing variables, the result is

$$x(t) = \frac{F_0}{2m\omega_1}\left[\frac{1}{(\omega - \omega_1)} \frac{1}{\left(\frac{\beta}{\omega - \omega_1}\right)^2 + 1} \left(\frac{\beta}{\omega - \omega_1} \sin \omega t - \cos \omega t\right) - \frac{1}{(\omega + \omega_1)} \frac{1}{\left(\frac{\beta}{\omega + \omega_1}\right)^2 + 1} \left(\frac{\beta}{\omega + \omega_1} \sin \omega t - \cos \omega t\right)\right],$$

$$x(t) = \frac{F_0}{2m}\left\{\frac{4\omega\beta}{\left(\omega^2 - \omega_0^2\right)^2 + 4\beta^2\omega^2} \sin \omega t - \frac{2\left(\omega^2 - \omega_0^2\right)}{\left(\omega^2 - \omega_0^2\right)^2 + 4\beta^2\omega^2} \cos \omega t\right\}.$$

This can be written

$$x(t) = \frac{A_0}{\sqrt{\left(\omega^2 - \omega_0^2\right)^2 + 4\beta^2\omega^2}} \left\{ \left(\underbrace{\frac{2\omega\beta}{\sqrt{\left(\omega^2 - \omega_0^2\right)^2 + 4\beta^2\omega^2}}}_{= \sin\delta} \right) \sin\omega t \right.$$

$$\left. + \left(\underbrace{\frac{\left(\omega_0^2 - \omega^2\right)}{\sqrt{\left(\omega^2 - \omega_0^2\right)^2 + 4\beta^2\omega^2}}}_{= \cos\delta} \right) \cos\omega t \right\}$$

$$\Rightarrow x(t) = \frac{A_0}{\sqrt{\left(\omega^2 - \omega_0^2\right)^2 + 4\beta^2\omega^2}} \cos\left(\omega t - \delta\right).$$

This is precisely our previous solution to the problem; the phase angle δ having the same meaning as before.

CHAPTER 3 PROBLEMS

1. Substituting Eqs.(3.17) into (3.16), show that Eq.(3.19) (an ellipse) results, where C and D are given by Eqs.(3.20) and (3.21), and tan 2γ is given by Eq.(3.18). [Hints: After substituting (3.17) into (3.16), require that the coefficient of the $x'y'$ cross term vanish. After this, use (3.18) to eliminate the cos δ term and do the algebra leading to Eqs.(3.20) and (3.21).]

2. Show that the coefficients C and D in Eqs.(3.20) and (3.21) are both positive. Do this as follows:

   a) Show that requiring $C>0$ is the same as

   $$B^2 + A^2 > \frac{A^2 - B^2}{\cos 2\gamma}.$$

   b) Similarly, show that $D > 0$ implies

   $$B^2 + A^2 > \frac{B^2 - A^2}{\cos 2\gamma}.$$

   c) Show that (a) and (b) together imply the true statement,

   $$\cos^2 \delta < 1,$$

   when the explicit form for cos 2γ is used. [The other three cases, $C>0$ $D<0$, $C<0$ $D>0$, and $C<0$ $D<0$ may be eliminated one by one.]

3. For the out-of-phase two dimensional oscillator, study the special case where $A=B$ and show that in this case we can choose

   $$C = \frac{1 - \cos\delta}{A^2 \sin^2\delta}, \quad D = \frac{1 + \cos\delta}{A^2 \sin^2\delta},$$

   and that $\gamma = \pm\frac{\pi}{4}$. [Note: It is not correct to simply put $A=B$ in Eqs.(3.20) and (3.21)]

4. (a) The general solution of the damped oscillator,

   $$\ddot{x} + 2\beta\dot{x} + \omega_0{}^2 x = 0,$$

   when $\beta > \omega_0$ (overdamped case) is given by

   $$x(t) = e^{-\beta t}\left[A_1 e^{\omega_2 t} + A_2 e^{-\omega_2 t}\right],$$

   where $\omega_2 = \sqrt{\beta^2 - \omega_0^2}$ and $A_{1,2}$ are constants. For initial conditions

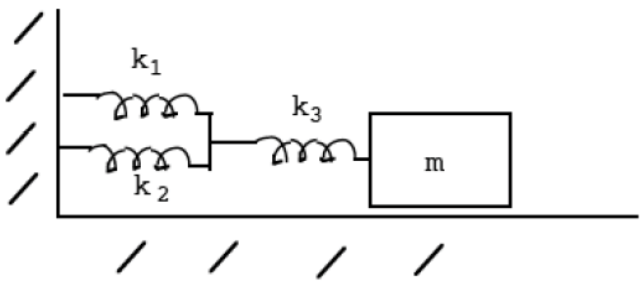
   $$x(0) = x_0, \quad \dot{x}(0) = v_0,$$

find A_1 and A_2.

b) Find conditions on x_0, v_0 which permit a change in sign of $x(t)$.

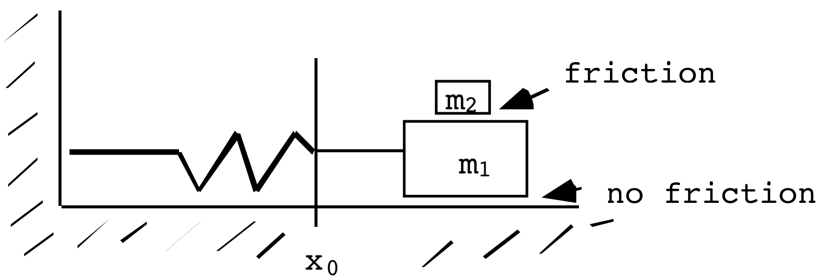
5. In the analogy developed between mechanical oscillators and electrical circuits we used Kirchoff's circuit law, which says the sum of voltages around a closed loop is zero. Which of Newton's laws of motion is analogous to this circuit law? Can you formulate circuit statements analagous to the other two of Newton's laws?

6. A mass is attached to a wall with the three springs shown. The system is without friction.



Using the circuit/mechanical analogy, find the equivalent circuit to this mechanical oscillator.

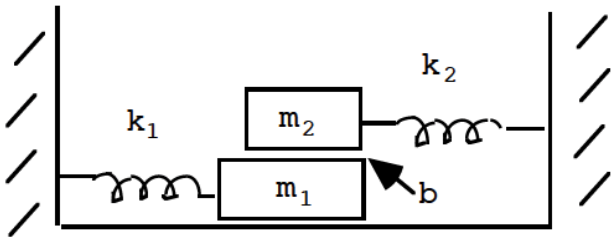
7. Consider the system with two masses:



$x_{1,2}$ locate the positions of masses $m_{1,2}$ relative to m_1's equilibrium position, x_0. Assume there is a frictional force proportional to velocity between m_1 and m_2, but that there is no frictional force between m_1 and the ground and that the oscillations are small.

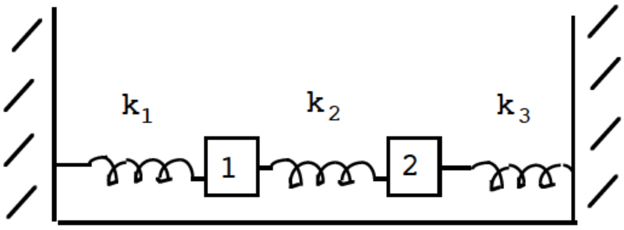
a) Write down Newton's equations of motion for this system.
b) Find the electrical circuit analogous to the mechanical situation in (a).

8. Consider the system of two masses shown. One mass rides on top of the other and the two springs have different spring constants, k_1 and k_2. There is a friction term linear in velocity between m_1 and m_2 but not between m_1 and the ground; the oscillations are small.



a) Write out Newton's equations of motion for this system.
b) Find and draw the equivalent electrical circuit to this mechanical problem.

9. Consider the system of two masses and three springs shown.



a) Write out Newton's equations of motion for this system.
b) Find and draw the equivalent electrical circuit to this mechanical problem.

10. Consider the driven, damped harmonic oscillator:

$$\ddot{x} + 2\beta\dot{x} + \omega_0{}^2 x = A_0 \cos \omega t.$$

a) Show that the instantaneous energy of the oscillator $\left(E(t) \equiv \frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\omega_0{}^2x^2\right)$ satisfies

$$\frac{dE(t)}{dt} = -2\beta m\dot{x}^2 + mA_0\,\dot{x}\,\cos \omega t.$$

[Hint: Take the derivative of $E(t)$ and use the above differential equation.]

b) Give an interpretation of the two terms on the right.

11. a) Show that "amplitude resonance" occurs at angular driving frequency

$$\omega_R = \sqrt{\omega_0^2 - 2\beta^2}.$$

b) Show that **velocity resonance** takes place at the system's natural angular frequency,

$$\omega_V = \omega_0.$$

c) Find the angular frequency of **accleration resonance**. That is, find the driving frequency, ω_A, such that the acceleration amplitude of the oscillator is a maximum.

12. Show that the velocity quality factor, defined by

$$Q_V \equiv \frac{\omega_V}{2\beta}$$

($\omega_v = \omega_0$ from problem 11(b)) can also be written **exactly** as

$$Q_V = \frac{\omega_0}{\Delta\omega},$$

where $\Delta\omega =$ full width of the **velocity** resonance curve at $\frac{1}{\sqrt{2}}$ x maximum velocity.

13. Can a driven, overdamped oscillator undergo amplitude resonance? What about velocity resonance? Explain why or why not.

14. Show that the Q_A factor for light damping for the driven oscillator may be written as

$$Q_A \approx \pi N_e,$$

where N_e is the number of oscillations of the free oscillator which occur in the time it takes for the amplitude of the undamped oscillator to decrease to $1/e$ of it's initial value.

15. (a) Show that

$$\frac{<E>}{\Delta E} = \frac{\omega^2 + \omega_0^2}{8\pi\beta\omega},$$

where $<E>$ is the time averaged energy of a driven oscillator and ΔE is the energy lost during a cycle of driven oscillation.

b) Using (a), prove that for light damping the amplitude quality factor, Q_A, is thus given by

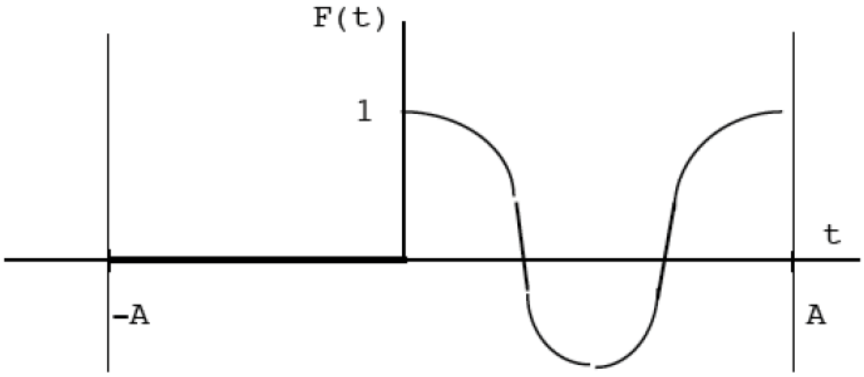
$$Q_A \approx 2\pi \frac{< E >}{\Delta E},$$

16. Referring back to Prob.10 above, explicitly compute the time average of both terms in the energy change of the damped oscillator over 1 cycle, $< \frac{dE(t)}{dt} >$. Does the answer you get surprise you?

17. Obtain the Fourier series representing the periodic function,

$$F(t) = \begin{cases} 0, & -A < t < 0, \\ \cos(\frac{2\pi}{A} t), & 0 < t < A. \end{cases}$$

It looks like:



"A" is an arbitrary constant.

18. A damped, driven harmonic oscillator with natural frequency ω_0 and damping constant β is subject to the force (for all times times, t),

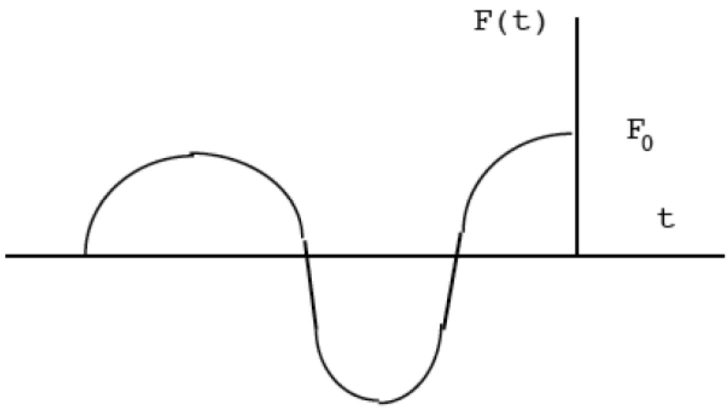
$$F(t) = \alpha_1 \cos(\omega t) + \alpha_2 \cos(2\omega t),$$

where α_1 and α_2 are constants. Find the average kinetic, $\langle T \rangle$, and potential, $\langle U \rangle$, energies of this system. [Some of the integrals given in Example 2 above may be useful.]

19. A underdamped harmonic oscillator (damping constant b and natural angular frequency ω_0) is subject to a force as a function of time given by

$$F(t) = \begin{cases} F_0 \cos \omega t, & t < 0 \\ 0, & t > 0 \end{cases}$$

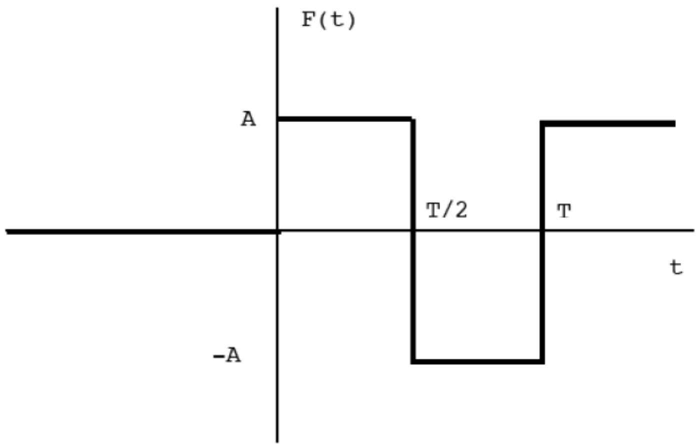
This looks like:



Find the motion of the oscillator, x(t), for t>0. [Hint: There is a simple way of doing this problem considering that you should know x(t), for t<0.]

20. An initially quiescent underdamped harmonic oscillator, with damping constant β amd natural frequency ω_0, is subject to an infinite, periodic square wave of amplitude A and period T, starting at time $t=0$. It's differential equation is:

$$\left(\frac{d^2}{dt^2} + 2\beta\frac{d}{dt} + \omega_0^2\right)x(t) = \frac{1}{m} F(t).$$



Find the response of the system, $x(t)$, for all times t. [Hint: Use Example 1 above, making sure the initial conditions $x(0) = 0$, $\dot{x}(0) = 0$ are met at $t=0+$.]

21. a) Solve (Green's function for a free particle)

$$\frac{d^2}{dt^2} G(t,t') = \frac{1}{m} \delta(t-t'),$$

subject to: 1. $G(t,t') = 0, \; t \le t'$,

2. $$\frac{dG(t,t')}{dt} = \frac{1}{m} \quad \text{for } t = t'+.$$

b) Using

$$x(t) = \int_{-\infty}^{\infty} dt' G(t,t')F(t'),$$

find $x(t)$ for

$$F(t') = \begin{cases} F_0, & t' > 0 \\ 0, & t' < 0 \end{cases}$$

where F_0 is a constant. The result should be very familiar!

22. Solve the harmonic oscillator equation,

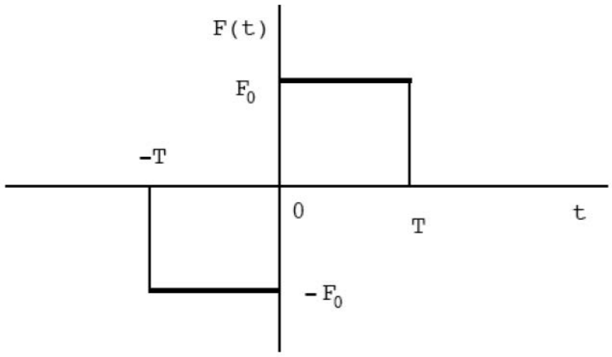
$$m\ddot{x} + kx = \begin{cases} 0, & t < 0 \\ F_0 \sin(\omega t), & t > 0 \end{cases},$$

for an initially quiescent oscillator using the Green's function technique for the case $\omega=\omega_0$ ($\omega_0^2 = \frac{k}{m}$). Use the known Green's function for this problem from the text. [Hint: The answer for $x(t)$ is given somewhere in the text.]

23. An initially quiescent, undamped oscillator with mass m and natural angular frequency ω_0 is subjected to a force,

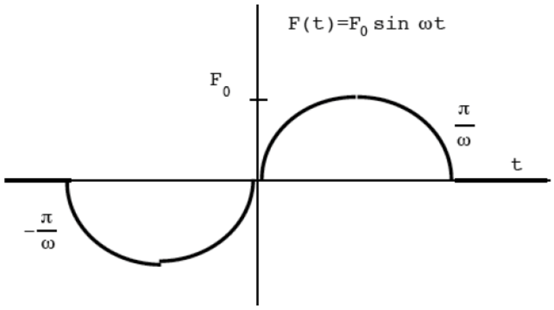
$$F(t) = \begin{cases} 0, & t < -T, \\ -F_0, & -T < t < 0, \\ F_0, & 0 < t < T, \\ 0, & t > T. \end{cases}$$

This looks like:



a) Using the Green's function method, find $x(t)$ for all times, t. [Hint: Consider the regions $-T<t<0$, $0<t<T$ and $t>T$ separately.]
b) Show that $x(t) = 0$ results for $t > T$ when $\omega_0 = \frac{2\pi n}{T}$, $n = 1,2,3....$

24. Consider the following force profile on an undamped (damping constant $\beta = 0$) oscillator of natural angular frequeny, ω_0. The driving angular frequency is $\omega \neq \omega_0$.



The force is zero for $t > \frac{\pi}{\omega}$ and for $t < -\frac{\pi}{\omega}$. Find the response of the system, $x(t)$ for $t > \frac{\pi}{\omega}$. [Ans: For $t > \frac{\pi}{\omega}$, I get $x(t) = -\frac{F_0}{m\omega_0}\frac{2\omega}{\omega^2-\omega_0^2}\sin(\frac{\pi\omega_0}{\omega})\cos(\omega_0 t)$.] **Extra Credit**: Also find $x(t)$ for $-\frac{\pi}{\omega} < t < \frac{\pi}{\omega}$.

25. The Green function for a damped oscillator satisfies

$$\left(\frac{d^2}{dt^2} + 2\beta\frac{d}{dt} + \omega_0^2\right)G(t,t') = \frac{1}{m}\delta(t-t'),$$

where $\delta(t-t')$ is the Dirac delta-function. Solve this equation and construct the appropriate Green function, $G(t,t')$, for a **critically damped** oscillator, $\omega_0 = \beta$ (The solution in the notes and in the book is for the underdamped case, $\omega_0 > \beta$.)

26. One can use the Green function to also build in arbitrary initial conditions for an oscillator. Newton's equation for an undamped oscillator is,

$$\left(\frac{d^2}{dt^2} + \omega_0^2\right)x(t) = \frac{1}{m}f(t),$$

where $f(t)$ is the external force. It's Green function is

$$G(t,t') = \begin{cases} \frac{1}{m\omega_0}\sin\left[\omega_0\left(t-t'\right)\right], & t \geq t' \\ 0, & t \leq t'. \end{cases}$$

For an oscillator in an arbitrary initial state at $t=0$, show that $x(t)$ is given by (x_0 and v_0 are the initial position and velocity, respectively, at $t=0$),

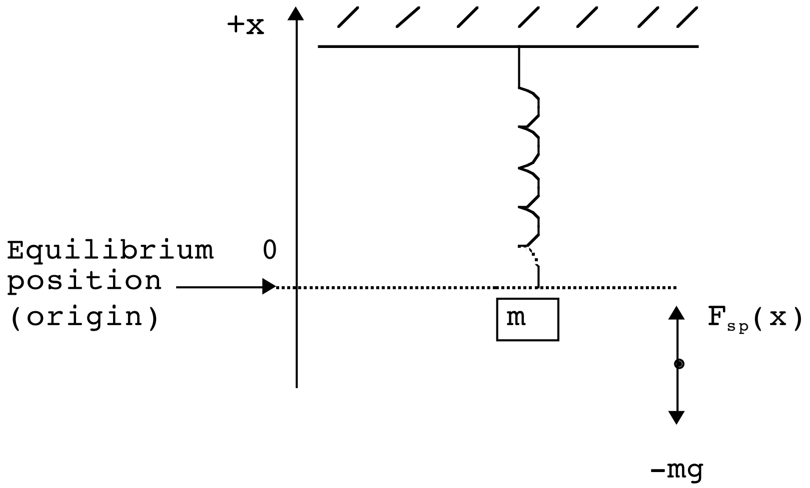
$$x(t) = \int_0^\infty dt' G(t,t')f(t') + m\left(G(t,0)v_0 - x_0\frac{dG(t,t')}{dt'}\Big|_{t'=0}\right)$$

[Hint: There are special techniques to show these sorts of results. For our purposes, it's enough to show that $x(t)$ is a solution of the above differential equation with the correct initial position and velocity, x_0 and v_0. One must aso show that $G(t,t') = G(-t',-t)$.]

4 NONLINEAR OSCILLATIONS

THE ANHARMONIC OSCILLATOR

Let us go back to how we started Chapter 3.



We have

$$m\ddot{x} = F_{tot} = F_{sp}(x) - mg,$$

$$F_{sp}(x) = \underbrace{F_{sp}(0)}_{mg} + x\underbrace{\left(\frac{\partial F_{sp}}{\partial x}\right)_0}_{\equiv -k} + \frac{x^2}{2}\underbrace{\left(\frac{\partial^2 F}{\partial x^2}\right)_0}_{\equiv 2\varepsilon} + \cdots \quad . \tag{4.1}$$

So, a more accurate description of the motion for a general spring is

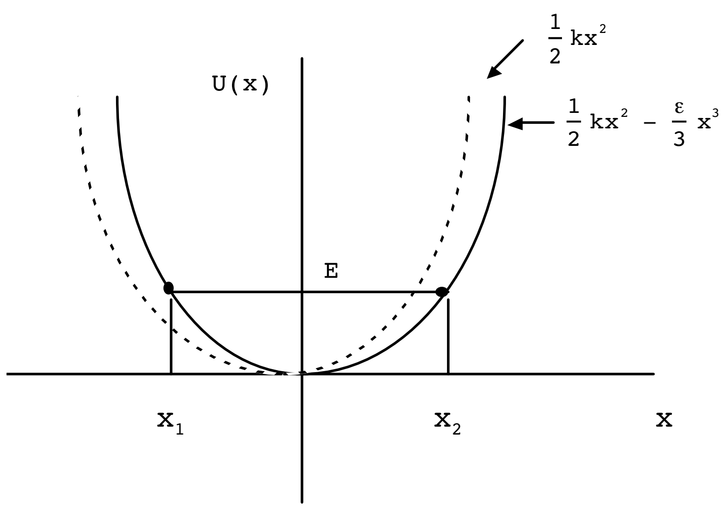
$$m\ddot{x} = -kx + \varepsilon x^2 ; \text{ "anharmonic oscillator"} \tag{4.2}$$

this is a non-linear differential equation. It is still second order (requiring two initial conditions) and still conservative.

Even though (4.2) is non-linear, we can calculate t(x) for portions of the motion not separated by turning points as before. The potential energy function can be taken as

$$U(x) = \frac{1}{2}kx^2 - \frac{1}{3}\varepsilon x^3 , \tag{4.3}$$

and looks like (assuming $\varepsilon > 0$)



Thus, the potential is no longer symmetric about the equilibrium position. In fact, the magnitude of the force is slightly greater when it compressed ($x < 0$) than when it is stretched.

The formal solution to $t(x)$ is given by Eq.(2.42):

$$(t - t_0) = \pm\int_{x_0}^{x(t)} \frac{dx'}{\sqrt{\frac{2}{m}\left(E - U(x')\right)}} , \tag{4.4}$$

with $U(x')$ given by Eq.(4.3). Although we have a formal solution, turning $t(x)$ into $x(t)$, which much more useful, is quite difficult, except in special cases. Even finding closed form expressions for the turning points, x_1, x_2, is awkward since one is examining the cubic equation,

$$U(x) - E = 0, \tag{4.5}$$

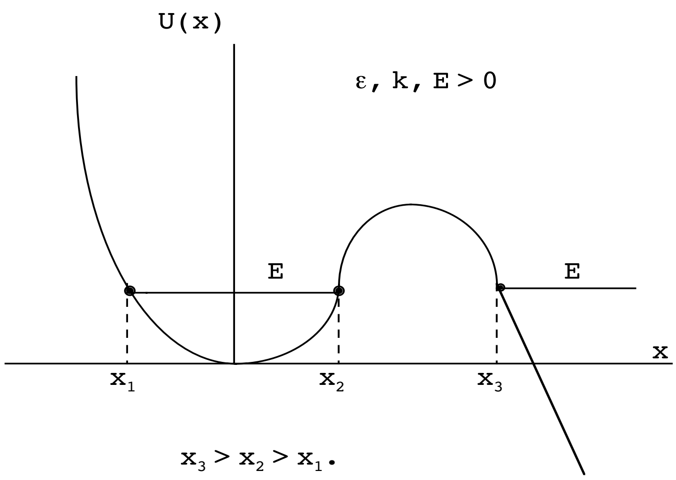
$$\Rightarrow \; x^3 - \frac{3}{2}\left(\frac{k}{\varepsilon}\right)x^2 + 3\left(\frac{E}{\varepsilon}\right) = 0, \tag{4.6}$$

for the real roots x_1, x_2 and x_3

We can relate the anharmonic oscillator equation of motion to a special function called an elliptic integral. If we agree to choose $t_0 = 0$ when the spring is released at rest after compression, we will have.

$$t = \int_{x_1}^{x(t)} \frac{dx'}{\sqrt{\frac{2}{m}\left(E - \frac{1}{2}kx'^2 + \frac{\varepsilon}{3}x'^3\right)}}, \tag{4.7}$$

where x_1 is the negative root, which is a turning point of (4.6), assuming $\varepsilon > 0$. The three roots of (4.6) can be pictured as the positions where the E=const. line cuts the potential energy:



The integral in (4.7) can be related to an elliptic integral of the first kind:

$$F(z,k) \equiv \int_0^z \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}}. \tag{4.8}$$

Using the integral,

$$\int_c^u \frac{dx'}{\sqrt{(a - x')(b - x')(x' - c)}} = \frac{2}{\sqrt{a - c}} F(\gamma, q), \tag{4.9}$$

where $a > b \geq u > c$ and

$$\gamma = \sqrt{\frac{u - c}{b - c}}, \tag{4.10}$$

as well as

$$q = \sqrt{\frac{b - c}{a - c}}. \tag{4.11}$$

In our case, we have

$$a = x_3,\ b = x_2,\ c = x_1,\ u = x(t),$$

which gives

$$t = 2\left(\frac{3m}{2\varepsilon}\right)^{1/2} \frac{F(\gamma, q)}{\sqrt{x_3 - x_1}}. \tag{4.12}$$

If one is interested in the period of motion, it is necessary only to let $x \rightarrow x_2$ in (4.12) and multiply by two:

$$\tau = 4\left(\frac{3m}{2\varepsilon}\right)^{1/2} \frac{F(1, q)}{\sqrt{x_3 - x_1}}. \tag{4.13}$$

Sometimes the expression $F(1,q)$ is called a "complete elliptic integral of the first kind."

Exact expressions for the roots of (4.6) are available; however, to first nontrivial order in ε , one can show that

$$\left.\begin{aligned} x_1 &\simeq -\left(\frac{2E}{k}\right)^{1/2}\left(1 - \frac{1}{3}\left(\frac{2E}{k^3}\right)^{1/2}\varepsilon\right), \\ x_2 &\simeq \left(\frac{2E}{k}\right)^{1/2}\left(1 + \frac{1}{3}\left(\frac{2E}{k^3}\right)^{1/2}\varepsilon\right), \\ x_3 &\simeq \frac{3}{2}\left(\frac{k}{\varepsilon}\right)\left(1 - \frac{8}{9}\left(\frac{E}{k^3}\right)\varepsilon^2\right). \end{aligned}\right\} \tag{4.14}$$

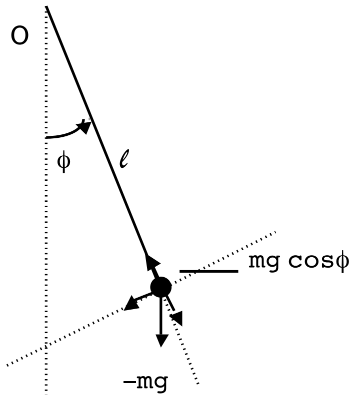
Using (4.13) and (4.14) and an approximate form for $F(1,q)$ (see Eq.(4.27) below and the HW problem), we can find an approximate expression for τ for small amplitude motion, which will be a power series in ε. I find that[4]

$$\tau \simeq \underbrace{2\pi\left(\frac{m}{k}\right)^{1/2}}_{=\tau_0}\left(1+\frac{5}{6}\left(\frac{E}{k^3}\right)\varepsilon^2\right), \tag{4.15}$$

so that the period is slightly in excess of τ_0, for fixed E and K.

The Plane Pendulum

Let us examine another intrinsically nonlinear system, the plane pendulum.



Assume *planar* motion of the mass m and take the origin of coordinates (cylindrical) to be at 0, the point of attachment. (z-axis is out of the page.) Note the attachment is such that the pendulum can swing a full 360° and go "over the top".

For angular variables, the appropriate concept is torque:

$$\dot{\vec{L}} = \vec{N},$$

$$\vec{N} = \vec{r}\times\vec{F} = mg\ell\sin\phi\left(-\hat{e}_z\right),$$

$$\vec{L} = \vec{r}\times\vec{p} = m\vec{r}\times\vec{v}.$$

In cylindrical coordinates, remember

$$\vec{v} = \underbrace{\dot{\rho}\hat{e}_\rho}_{0} + \rho\dot{\phi}\hat{e}_\phi + \underbrace{\dot{z}\hat{e}_z}_{0},$$

$$\Rightarrow \vec{L} = \left(m\,\rho\hat{e}_\rho\right) \times \left(\rho\dot{\phi}\hat{e}_\phi\right) = m\rho^2\dot{\phi}\hat{e}_z\,,$$

$$\Rightarrow \dot{\vec{L}} = m\,\ell^2\ddot{\phi}\,\hat{e}_z \quad (\rho = \ell \text{ here}).$$

The equation of motion is thus

$$m\ell^2\ddot{\phi} = -mg\ell\sin\phi\,. \tag{4.16}$$

For small oscillations $\sin\phi \simeq \phi$ and we have harmonic motion

$$\ddot{\phi} + \frac{g}{\ell}\phi \simeq 0\,, \tag{4.17}$$

with angular frequency $\omega_0^2 = \frac{g}{\ell}$.

Going beyond this approximation, we get from (4.16) the energy conservation equation by integration

$$\Rightarrow \frac{1}{2} m\ell^2\dot{\phi}^2 - mg\ell\cos\phi = \text{const.} \tag{4.18}$$

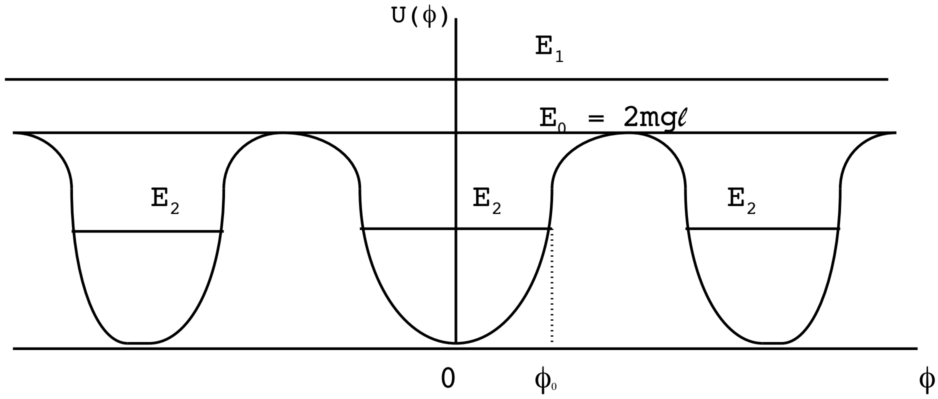
We can always add or subtract a constant in defining the total energy, E. We will choose

$$\text{const.} \rightarrow E - mg\ell,$$

which corresponds to choosing the zero of potential energy at the point $\phi = 0$. We have

$$E = \underbrace{\frac{1}{2} m\ell^2\dot{\phi}^2}_{=T} + \underbrace{mg\ell\left(1 - \cos\phi\right)}_{=U(\phi)}. \tag{4.19}$$

The periodic potential looks like:



The diagram tells us that for energy values such that $E < 2mgl$ (or $\phi_0 < \pi$) the pendulum motion is bounded in ϕ. this would correspond to an energy E_2 in the above plot. We can get the relation between t and ϕ by again relating the motion integral to an elliptic integral. From (4.18) we have

$$\dot{\phi} = \pm\left(\frac{2}{m\ell^2}\right)^{1/2}\left(E - U(\phi)\right)^{1/2}. \tag{4.20}$$

Choosing $t_0 = 0$ at $\phi = 0$ now gives

$$t = \left(\frac{m\ell^2}{2}\right)^{1/2}\int_0^{\phi(t)}\frac{d\phi'}{\left(E - U\left(\phi'\right)\right)^{1/2}}, \tag{4.21}$$

where we can integrate only up to the first turning point, $\phi = \phi_0$. At maximum extension, $\phi = \phi_0$, the energy, E, is given by

$$E = mg\ell\left(1 - \cos\phi_0\right) = 2mg\ell\sin^2\left(\frac{\phi_0}{2}\right). \tag{4.22}$$

Also, we can write

$$U(\phi) = 2mg\ell \sin^2 \frac{\phi}{2}, \tag{4.23}$$

which now gives

$$t = \frac{1}{2}\left(\frac{\ell}{g}\right)^{1/2} \int_0^{\phi(t)} \frac{d\phi}{\left(\sin^2 \frac{\phi_0}{2} - \sin^2 \frac{\phi}{2}\right)^{1/2}}. \tag{4.24}$$

Defining

$$x \equiv \frac{\sin \frac{\phi}{2}}{\sin \frac{\phi_0}{2}} \quad , \quad k \equiv \sin \frac{\phi_0}{2},$$

we have

$$dx = \frac{\cos \frac{\phi}{2}\, d\phi}{2 \sin \frac{\phi_0}{2}} = \frac{\sqrt{1 - k^2x^2}}{2k}\, d\phi,$$

$$\Rightarrow t = \frac{1}{2}\left(\frac{\ell}{g}\right)^{1/2} \int_0^{k^{-1}\sin\frac{\phi}{2}} \frac{2k\,dx}{\left(1 - k^2x^2\right)^{1/2}} \frac{1}{\left(k^2 - k^2x^2\right)^{1/2}}$$

$$= \left(\frac{\ell}{g}\right)^{1/2} F\left(k^{-1} \sin \frac{\phi}{2}, k\right), \tag{4.25}$$

where we again see the elliptic integral of the first kind. The period, τ, is now given by setting $\phi = \phi_0$in (4.25) and multiplying by four,

$$\tau = 4\left(\frac{\ell}{g}\right)^{1/2} F(1, k), \tag{4.26}$$

We may get an approximate expression for t for small oscillations, $k << 1$, by expanding $F(1, k)$:

$$F(1, k) \simeq \int_0^1 \frac{dx}{\sqrt{1 - x^2}} \frac{1}{\sqrt{1 - k^2x^2}},$$

$$= \int_0^1 \frac{dx}{\sqrt{1 - x^2}} \left(1 + \frac{1}{2}k^2x^2 + \frac{3}{8}\left(k^2x^2\right)^2 + \ldots\right). \tag{4.27}$$

Doing the integrals in (4.27), we find

$$\tau = 4\left(\frac{\ell}{g}\right)^{1/2}\left(\frac{\pi}{2} + \frac{\pi}{8}k^2 + \frac{9\pi}{128}k^4 + \ldots\right),$$
$$= \underbrace{2\pi\left(\frac{\ell}{g}\right)^{1/2}}_{=\tau_0}\left(1 + \frac{1}{4}k^2 + \frac{9}{64}k^4 + \ldots\right).$$

Also

$$k = \sin\frac{\phi_0}{2} \simeq \frac{\phi_0}{2} - \frac{\phi_0^3}{48} + \ldots,$$
$$k^2 \simeq \frac{\phi_0^2}{4} - \frac{\phi_0^4}{48} + \ldots,$$
$$k^4 \simeq \frac{\phi_0^4}{16} + \ldots.$$

Collecting terms, and to order ϕ_0^4, we thus find

$$\tau = \tau_0\left(1 + \frac{\phi_0^2}{16} + \frac{11}{3072}\phi_0^4 + \ldots\right). \tag{4.28}$$

Again, the period is slightly increased.

One can understand why the plane pendulum's period is nearly independent of amplitude from the smallness of the ϕ_0^2 , ϕ_0^4 coefficients. For example, choosing $\phi_0 = \frac{\pi}{2}$ we get (from a numerical evaluation of $F\left(1, \frac{1}{\sqrt{2}}\right)$)

$$\tau \simeq \tau_0(1.1803),$$

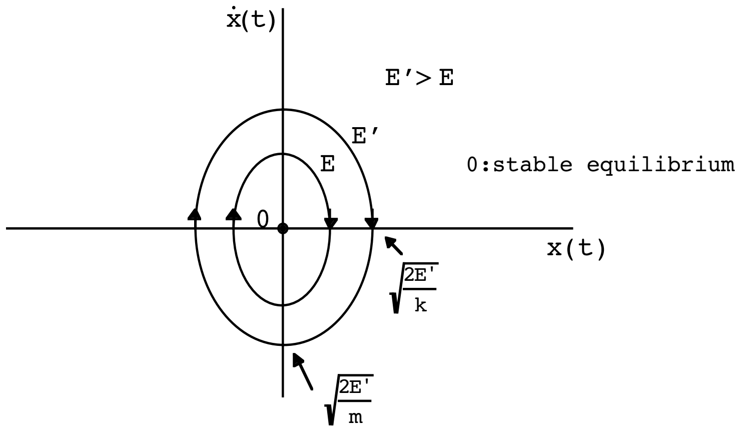
or about an 18% change.

Phase Diagrams and Nonlinear Oscillations

An important qualitative technique for the understanding of both linear and nonlinear systems is the construction of **phase diagrams**. Let us return to the harmonic oscillator. It's energy is

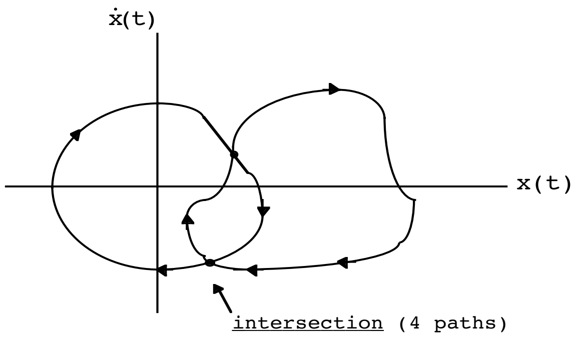
$$E = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}kx^2. \tag{4.29}$$

Plotting $\dot{x}(t)$ and $x(t)$ simultaneously gives a series of ellipses for various energies:



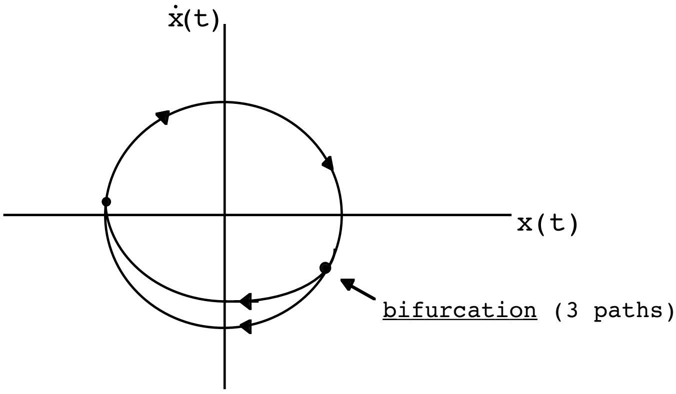
This is called "phase space." The arrows show the evolution of the system in time.

Of course, the trajectories will never intersect for $E' \neq E$ This is true for any linear, second order, system. This can be seen by the following argument. Assume such a system had a phase diagram where the trajectories crossed, as in



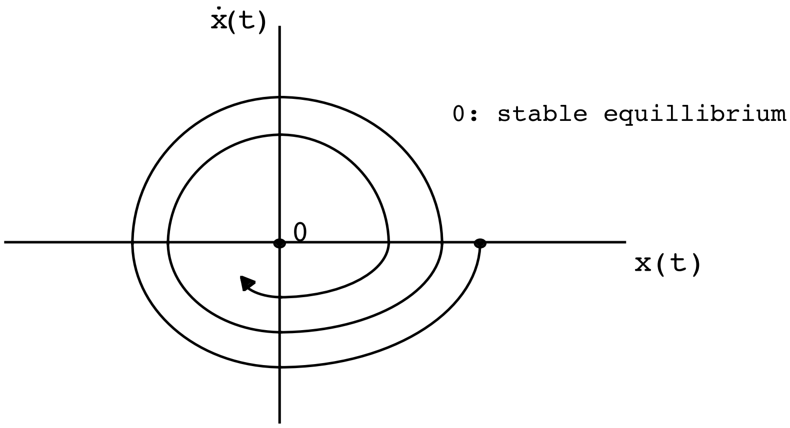
Imagine starting a particle at the above intersection point. If the above could happen, this would mean the particle at the intersection has two possible solutions. However, the solutions of linear, second order, differential equations are unique, given initial conditions on x and $\dot{x}$. We will see shortly that a special type of intersection is actually allowed for nonlinear systems.

Another type of behavior not allowed for linear, second order systems in phase space is a **bifurcation**, as in the following.



Actually, it is clear that no **deterministic** system can display such behavior. Although bifurcations in phase space do not occur, bifurcations in parameter space descriptions can and do occur, as we will also see shortly.

Since the underdamped oscillator looses energy as time progresses, it's phase diagram will look like:

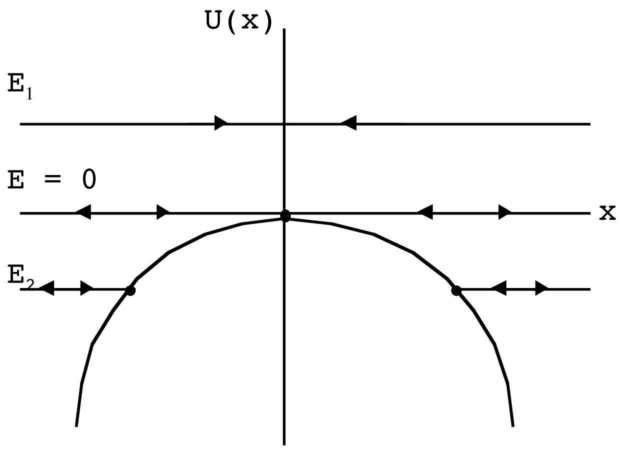


It spirals in toward the point of stable equilibrium taking an infinite amount of time to come to rest.

Now turn around the sign of k. Take

$$U(x) = -\frac{1}{2}kx^2 \quad , \quad k > 0. \tag{4.30}$$

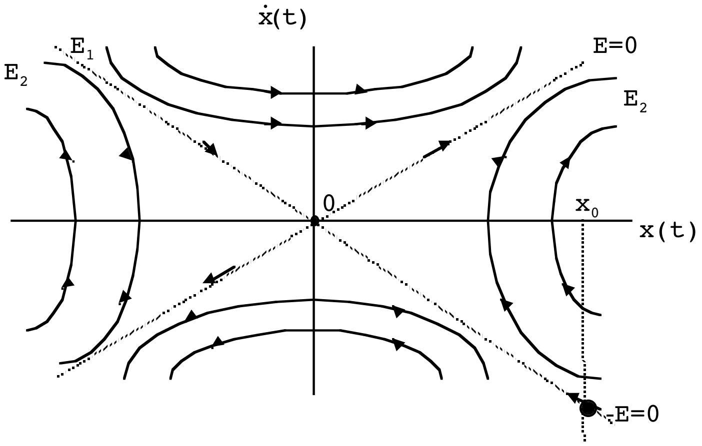
The potential is a downward opening parabola:



The lines represent trajectories of different energies. The energy equation is given by (4.29) with the sign of **k** reversed:

$$E = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2. \tag{4.31}$$

This is the equation of a hyperbola. The phase diagram is



0: unstable equilibrium

Let's try to follow the trajectory of the particle shown with $x_0 > 0$, $\dot{x}(0)<0$ and $E = 0$. We have

$$\dot{x} = \pm\sqrt{\frac{k}{m}}\,x\,, \tag{4.32}$$

which is the equation of a straight line. By integrating, we have

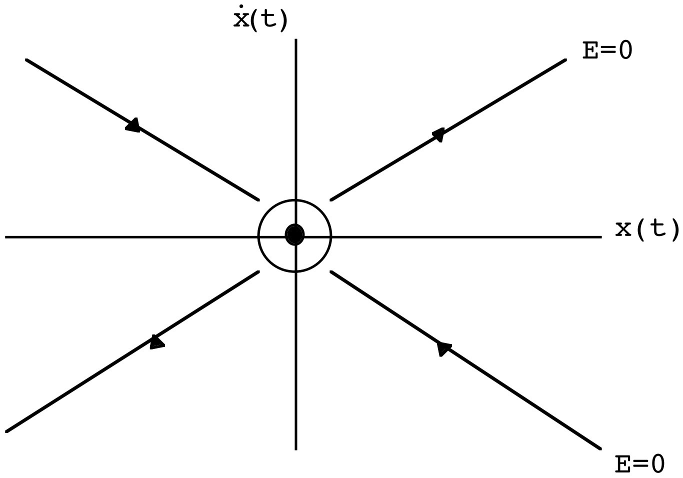
$$\int_{x_0}^{0+} \frac{dx}{x} = -\sqrt{\frac{k}{m}} \int_0^{\tau} dt\,,$$

$$\Rightarrow \ln\Big|_{x_0}^{0+} = -\sqrt{\frac{k}{m}}\;\tau\,. \tag{4.33}$$

This means it takes an infinite amount of time to reach the origin. Notice, however, that $x(t) = 0$ is still a (trivial) solution the equation of motion,

$$\ddot{x} - \frac{k}{m}x = 0\,. \tag{4.34}$$

This leads to the picture for E = 0 trajectories:



I am suggesting that there is an isolated $E = 0$ solution at the origin, which can be reached only after an infinite amount of time in any continuous manner. I will symbolize this by drawing a circle around the limiting point of the motion. Thus, the equation of motion, (4.34), which is a linear second order equation, avoids the possibility of an intersection by making the time necessary to reach the intersection point infinite! Such points always correspond to unstable equilibrium positions.

Nonlinear second order systems **can** have intersections in $\dot{x}$, x phase space. Consider the potential ($\bar{k} > 0$)

$$U(x) = -\frac{1}{2}\bar{k}\,|\,x\,|^{3/2}. \tag{4.35}$$

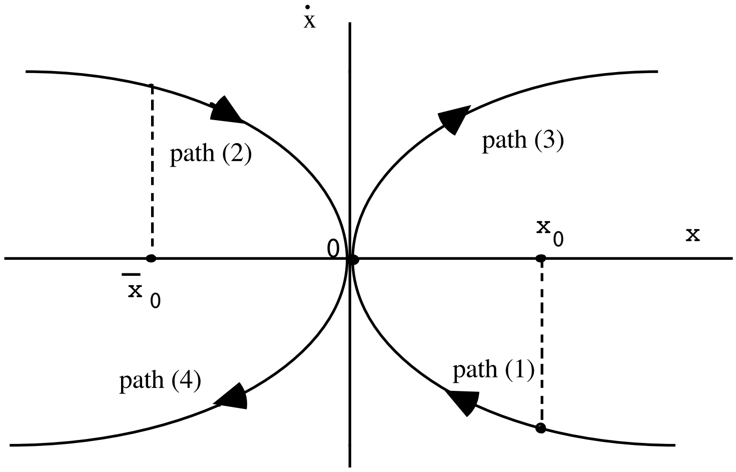
It corresponds to the (highly nonlinear) second order differential equation,

$$m\ddot{x} - \frac{3}{4}\bar{k}\frac{x}{|x|^{1/2}} = 0. \tag{4.36}$$

For $E = 0$ trajectories,

$$\dot{x} = \pm\sqrt{\frac{\bar{k}}{m}}\,|x|^{3/4}. \tag{4.37}$$

The situation looks like:



The time necessary to reach the origin on path (1) or (2) is now finite. Integrating (4.37) in the two cases gives

$$\text{path (1):} \qquad x = \left(x_0^{1/4} - \sqrt{\frac{\bar{k}}{m}}\,\frac{t}{4} \right)^4 , \tag{4.38}$$

$$\text{path (2):} \qquad x = -\left(|\bar{x}_0|^{1/4} - \sqrt{\frac{\bar{k}}{m}}\,\frac{t}{4} \right)^4 . \tag{4.39}$$

Setting $x = 0$ gives the time necessary to reach the origin,

$$\text{path (1):} \qquad \tau_0 = 4\, x_0^{1/4} \left(\frac{m}{\bar{k}} \right)^{1/2} , \tag{4.40}$$

$$\text{path (2):} \qquad \bar{\tau}_0 = 4\, |\bar{x}_0|^{1/4} \left(\frac{m}{\bar{k}} \right)^{1/2} . \tag{4.41}$$

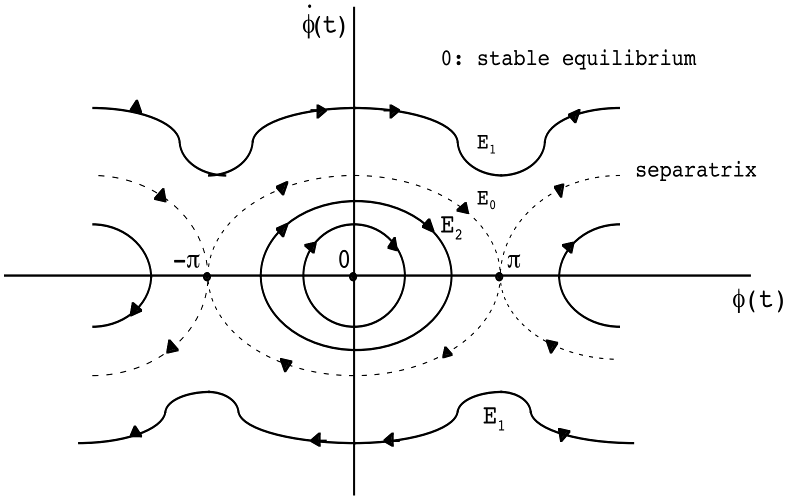
Notice that at the origin on either path, we have

$$\text{path (1):} \qquad \dot{x}\,\big|_{t=\tau_0} = 0 , \tag{4.42}$$

$$\text{path (2):} \qquad \dot{x}\,\big|_{t=\bar{\tau}_0} = 0 . \tag{4.43}$$

Thus the two paths **do** touch at the origin, 0, which is now reached in a finite amount of time. Note that this does **not** say such equations are not deterministic. It simply means, mathematically speaking, that the solutions, given initial x, $\dot{x}$ data are not unique. In fact, one can show that path (3) is the continuation of path (1) and path (4) is the continuation of path (2).

Let's now examine the phase diagram for the plane pendulum. Looking back at the potential, $U(\phi)$, we find the qualitative picture:



Notice the periodicity of the phase diagram. Notice also that when the full range of ϕ is allowed, this phase diagram has both the bound motion seen in the harmonic oscillator phase diagram ($E = E_2$) as well as the unbound motion of the potential with the downward opening parabola ($E = E_1$). The trajectory which separates these two types of motion has $E = E_0$ and is called the **separatrix**.

Let's again examine the trajectories for this special energy to see whether there is an intersection of trajectories at $\phi = \pm n\pi,\ n = 0,1,2,3\ldots$ Setting $E_0 = 2mg$ in (4.19), we have

$$\frac{1}{2}m\ell^2\dot{\phi}^2 + 2mg\ell\sin^2\left(\frac{\phi}{2}\right) = 2mg\ell\,, \tag{4.44}$$

$$\Rightarrow \dot{\phi} = \pm 2\sqrt{\frac{g}{\ell}}\cos\left(\frac{\phi}{2}\right). \tag{4.45}$$

Setting $\phi = \pi + \delta \ (\delta << 1)$, we have

$$\cos\left(\frac{\pi + \delta}{2}\right) \simeq -\frac{\delta}{2}\,,$$

and the equation of motion in the close vicinity to $\phi = \pi$ is given by

$$\dot{\delta} \simeq \mp\sqrt{\frac{g}{\ell}}\,\delta\,. \tag{4.46}$$

This is the same type of equation of motion for the unstable system with $U(x) = -\frac{1}{2}kx^2$, Eq.(4.32). Thus, although the plane pendulum is nonlinear, in the vicinity of the point $\phi = n\pi,$ the system behaves increasingly linearly. Thus, the picture near $\phi = \pi$ is essentially the same as before, and the system with energy E_0 takes an infinite amount of time to reach this point.

The Logistic Difference Equation

The purpose of this Chapter is to introduce you to aspects of nonlinear systems. This is a vast subject and so we will only get a flavor of some of the new phenomenon and concepts which emerge. The best way to introduce this material is through a study of some properties of simple one-dimensional difference equations. These mappings are analogous to dynamical systems in ways we will discuss.

A one-dimensional difference equation mapping is given by

$$x_{n+1} = f(x_n) \tag{4.47}$$

Let us consider the example (the "logistic map")

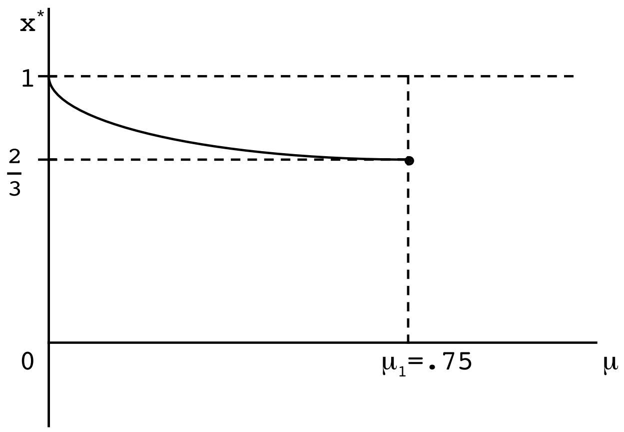
$$x_{n+1} = 1 - \mu x_n^2 \tag{4.48}$$

for m in the open interval (0,2) and xn in the closed interval [-1,1]. A fixed point, x^*, of such a mapping is a value of x_n such that $x_{n+1} = x_n$:

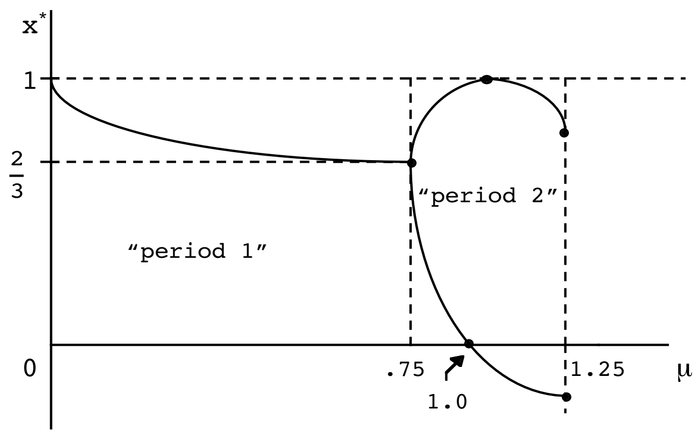
$$x^* = 1 - \mu x^{*2},$$

$$\Rightarrow \quad x^* = -\frac{1}{2\mu} + \frac{1}{2\mu}\sqrt{1 + 4\mu}\,. \tag{4.49}$$

This looks like:



However, a very interesting thing happens when μ is equal to or slightly larger than .75. We have a bifurcation in the fixed point diagram. The map after the first bifurcation looks like:

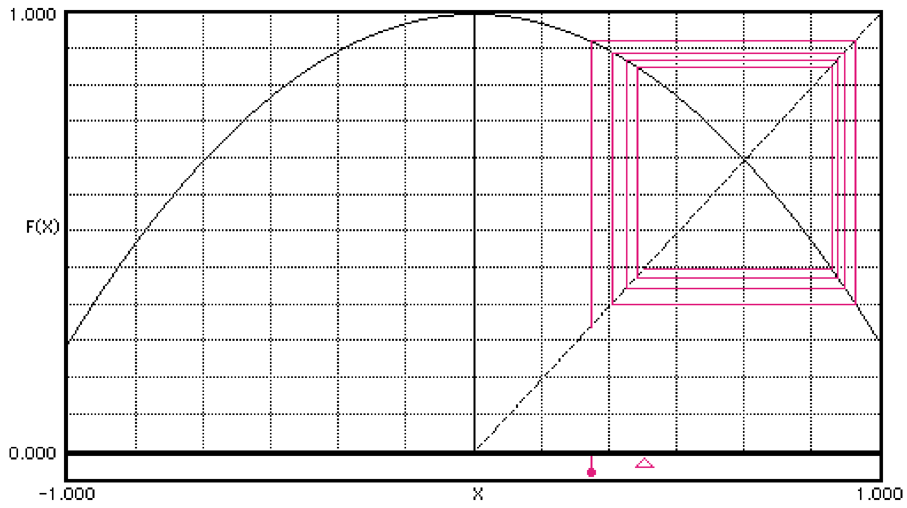


We encounter other bifurcations for larger values of μ as indicated:

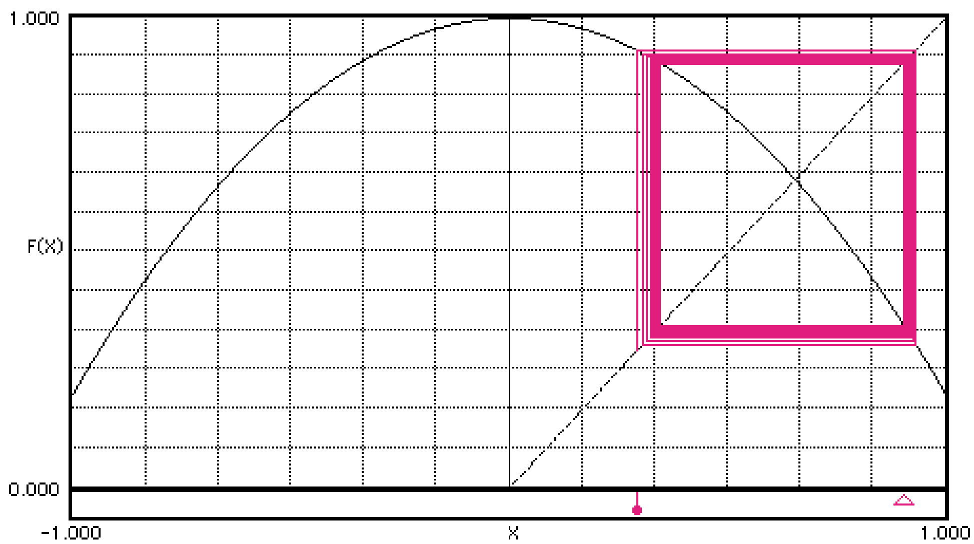
μ_1 = .75	first bifurcation
μ_2 = 1.25	2^{nd} "
μ_3 = 1.3680989...	3^{rd} "
μ_4 = 1.3940461...	4^{th} "

This process of bifurcation continues until at $\mu_\infty \cong 1.401155\ldots$ an infinite number of bifurcations occurs. This result is very surprising because we are seeing an infinite amount of structure from a simple quadratic mapping.

The bifurcation seen at $\mu = .75$ (where $x^* = \frac{2}{3}$) is easily understood from the function we are evaluating, a graph of which is below.



Notice that there is a perpendicular crossing of the function and the 45° line for this value of μ. It is in this situation that a bifurcation occurs. Some iterative lines, which use x_n as an input and show x_{n+1} as an output, are shown in the above figure. Notice how they tend to spiral inward toward x^*, the fixed point. We will see in the context of a problem below that such a crossing becomes unstable if the magnitude of the slope of the line becomes greater than 1 at the fixed point. This is in fact what occurs for μ slightly in excess of .75. Then the crossing point between the function and the 45° line becomes unstable and a stable square of iterative lines develops around this unstable point; see the below which is for $\mu = .80$, for which $x^* \cong 0.345$ and 0.905.



There are certain universal features of quadratic mappings, one of which is the fact that (discovered by M. Feigenbaum)

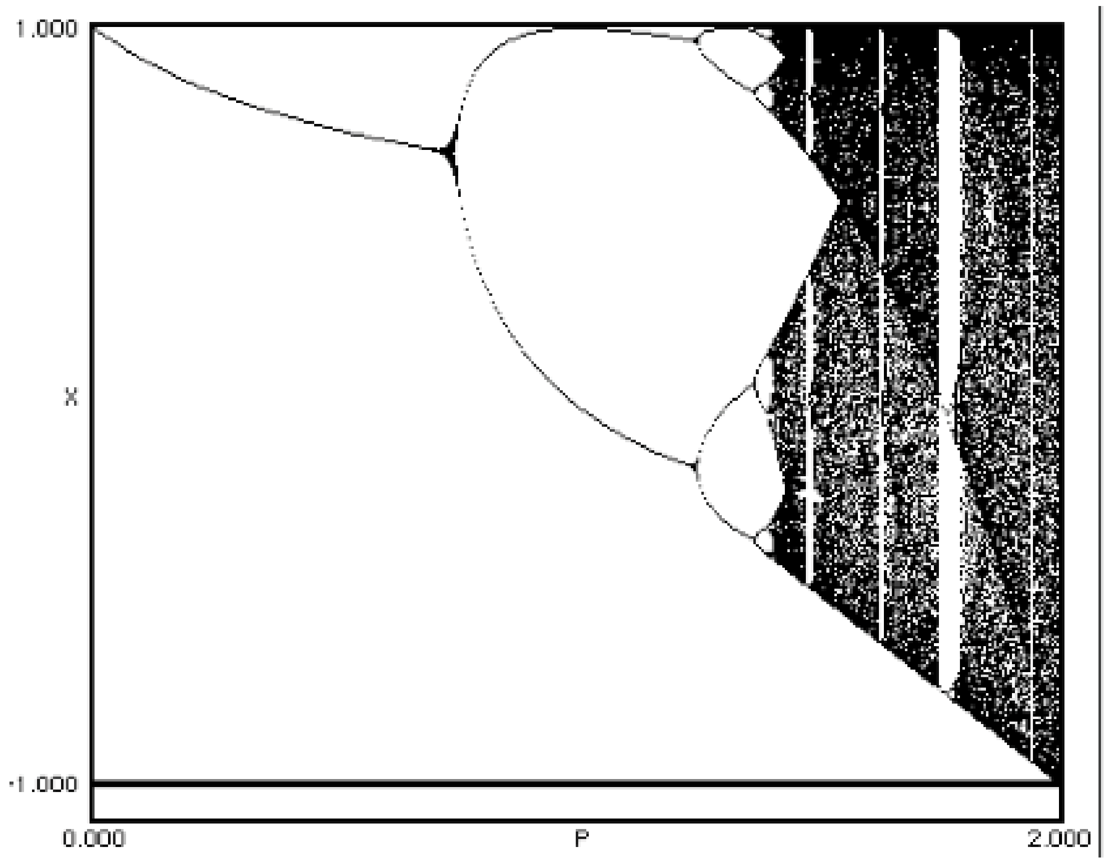
$$\mu_n \simeq \mu_\infty - \frac{A}{\delta^n}, \tag{4.50}$$

for $n>>1$ where $\delta \simeq 4.669201609\ldots$ (called the Feigenbaum constant). One can show that this implies that

$$\lim_{n\to\infty} \frac{\mu_n - \mu_{n-1}}{\mu_{n+1} - \mu_n} = \delta. \tag{4.51}$$

which says that the ratio of bifurcation intervals tends to a fixed limit, although the ratios themselves are shrinking to zero. The reason these things were not discovered previously is because computers were not available to "explore" this relation, and because nobody thought such interesting structure would emerge from such a simple mapping.

I do not mean to give you the impression that all is chaos for $\mu > \mu_\infty$ Within the range $[\mu_\infty ,2]$ are also regions of order. For example, for m slightly in excess of 1.75 is a region of period 3. The structure of the fixed point diagram is both infinitely complex and astoundingly simple. The fixed point map is illustrated below.



Fractals

An elementary example of a fractal is given by the "Cantor set of the middle third", in which the middle third of a series of line segments is removed, as follows:

step	segments	size
0	1	L
1	2	L/3
2	4	L/9
3	8	L/27

This is an example of a fractal set of zero length.

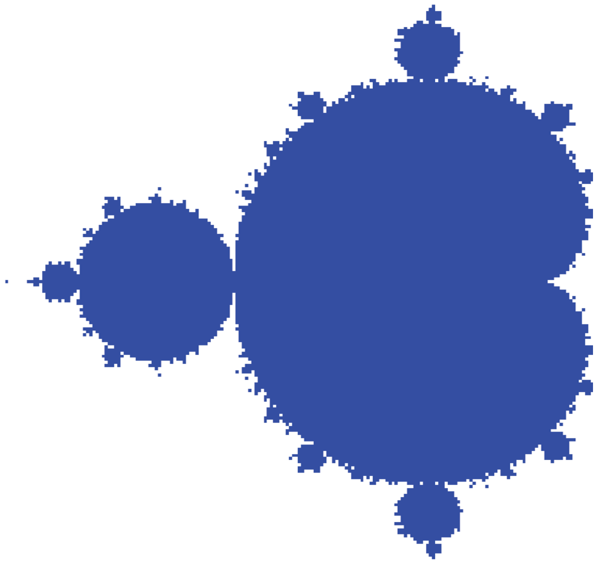
As another example of a fractal, consider the complex generalization of the quadratic mapping considered above, the so-called "Mandelbrot set". It is generated by simply iterating

$$z_{n+1} = z_n^2 + c \ , |c| \leq 2, \tag{4.52}$$

in the complex plane. For each value c, one chooses $z_0 = 0$ and iterates the result. We then implement the algorithm:

iteration	test	result
k	$\left\|z_k^2 + c\right\| > 2$	stop
	$\left\|z_k^2 + c\right\| < 2$	continue

If, after **N** iterations we still have $\left|z_k^2 + c\right| < 2$, one colors the point black; if not, some other color denoting how many iterations have been done. The surface of this infinitely structured object is the Mandelbrot set; it is a connected set. It is an example of a fractal which has infinite length! In the below picture the real axis is horizontal and the imaginary axis is vertical.



Since the Mandelbrot set is generated from a difference equation which is just a generalization to complex numbers of the logistic equation, we might expect that they would be closely related. They are related in the following way. As one cuts through the Mandelbrot diagram along the real axis (going from right to left in the above diagram), one is looking at stable numbers on the real axis. The big bulb above corresponds to the stable single period line of fixed points in the bifurcation diagram. The first bifurcation corresponds to the appearance of the smaller blub to the left, the next bifurcation the the next smaller bulb and so forth. The point of infinite number of bifurcations corresponds to the point in the above axis where an infinite number of bulbs are stacked. Thus going along the real axis from right to left in the Mandelbrot set there is a one to one correspondence with the bifurcation cascade seen in the logistic equation fixed point map going from left to right. The Mandelbrot set has been called the most beautiful and complicated object in mathematics – one gets a glimpse of the infinite on a sheet of paper!

Chaos in Physical Systems

Why talk about fractals and maps in a course on particle dynamics? Because the phase space of driven, dissipative systems can be thought of as an iterative mapping, like our difference equation, and the trajectories are often fractal objects. Let us consider a system of first order equations,

$$\frac{dx_1}{dt} = v_1(\vec{x}), \frac{dx_2}{dt} = v_2(\vec{x}), \frac{dx_3}{dt} = v_3(\vec{x}),$$

or

$$\frac{d\vec{x}}{dt} = \vec{v}(\vec{x}). \tag{4.53}$$

Because these equations are the equations for particle motion, usually the x_1, x_2, etc. turn out to be generalized velocities or coordinates. The first such system which displayed chaotic motion was studied by Edward Lorenz, a meterologist, with an early computer called a "Royal McBee." His system of equations were

$$\frac{dx}{dt} = \delta(y - x),$$

$$\frac{dy}{dt} = rx - y - xz,$$

$$\frac{dz}{dt} = xy - bz.$$

Here δ, r and b are constants and obviously $x_1=x$ $x_2=y$ and $x_3=z$. The trajectories in the three dimensional phase space of this system result in chaotic motion and possess a fractal structure.

The damped driven oscillator may be characterized as in (4.53). One may write

$$\left.\begin{aligned}\frac{dx}{dt'} &= y,\\ \frac{dy}{dt'} &= -cy - \sin x + F\cos z, (c > 0)\\ \frac{dz}{dt'} &= \omega. (\text{a constant})\end{aligned}\right\} \tag{4.54}$$

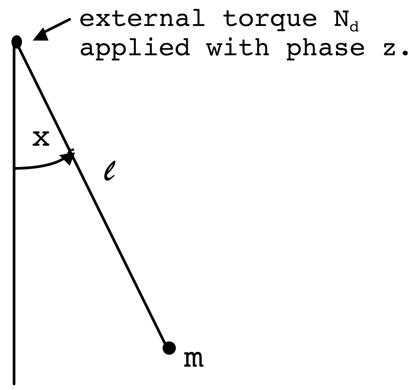
The physical interpretation of each term is

z: **driving phase**

x: **angular position of pendulum**

y: **dimensionless angular velocity**

and the above represents:



The above first order equations are in convenient, dimensionless form, with c, F, t' and ω related to m, ℓ, b (damping constant), N_d (external torque constant) and ω_d (driving frequency) by

$$c = \frac{b}{mg\ell}, \qquad \left(\omega_0^2 = \frac{g}{\ell}\right)$$

$$F = \frac{N_d}{mg\ell}, \ t' = \omega_0 t, \ \omega = \frac{\omega_d}{\omega_0}.$$

The three equations, (4.54), are equivalent to the single second order differential equation,

$$m\ell^2\ddot{x} = -b\dot{x} - mg\ell\sin x + N_d\cos(\omega_d t + \varphi), \tag{4.55}$$

which is just Newton's law for the damped, driven pendulum.

One necessary condition for chaotic behavior is that the equations in the above first order form contain a nonlinear term coupling one variable to another. This is seen in the $-\sin x$ term in the second equation. Surprisingly, a system of three such dynamical equations is already sufficient to display chaotic behavior.

The true phase space of the damped, driven pendulum is three-dimensional because of the three first order equations above. Physically, this is because it is no longer enough to give $x(t)$ and $\dot{x}(t)$ to specify the initial state of the system. It is necessary also to specify the phase of the driving term, which is given by $(\omega_d t+ \phi)$ in (4.55). However, all the trajectories in this three-dimensional space may be projected back on the old $\dot{x}(t)$, $x(t)$ plane. This projected version of phase space can display many types of motion for a given set of parameters C, F and ω and a given set of initial conditions. Analytic methods are of limited usefulness here and computer simulations must be used in order to gain insight into the types of motions that occur.

There are motions where the pendulum moves in a completely periodic manner, although the period of motion may not be the same as the period of the driving force. These are called **closed trajectories**, an example of which is the harmonic oscillator phase space above. There are also motions which display no periodicity at all, although there may be submotions which continue to reoccur. This is chaotic motion. After initial transients die out the damped pendulum evolves to a steady state trajectory called an **attractor**. The simplest example of an attractor is the point of stable equilibrium in the two-dimensional phase space of the damped, free pendulum. An attactor can also be a closed trajectory.

For chaotic motion in the driven, damped pendulum, the projected $x(t)$, $\dot{x}(t)$ phase space diagram is usually shown as a filled-in portion of phase space. Not **all** velocities, $\dot{x}(t)$, are available to such a system because the system is dissipative (has a positive damping constant.) Of course, it can never reach every point of projected phase space in a finite amount of time. Actually, what one sees after a simulation is done is a tight, packed set of intersecting and non-intersecting trajectories. (These lines do **not** intersect in the full three-dimensional phase space.) The infinitely structured geometrical object that the system tends toward in the steady state when chaos is involved is called a **strange attractor**, and is another example of a fractal.

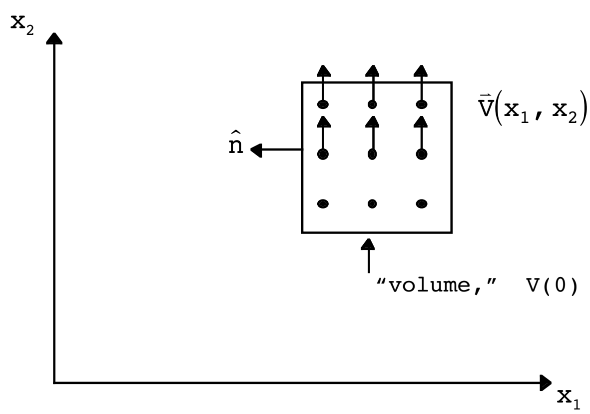
Because of the intricacy of the projected phase space plots and because the full three-dimensional phase space is hard to show, one needs tools to help uncover structures present in the phase space. One such tool is called a **Poincairé section**. Such a plot shows an overlap of a large number of "snapshots" of the $\dot{x}(t)$, $x(t)$ state of the system. If we agree to take these snapshots at a period $\tau_d = \frac{2\pi}{\omega}$, where w is the driving angular frequency, we can see at once whether the system is periodic or chaotic.

The simplest thing one can observe is a finite number of points. This immediately implies that the motion is a closed, periodic trajectory. Of course, the location of the points in the Poincairé section is a function of the phase, ϕ, of the projection, so this does not have any particular meaning. If τ points appear, the resulting motion could have been completed in a period, τ, given by $n\tau_d$, $\frac{n}{2}\tau_d$, $\frac{n}{3}\tau_d$, ..., where $\frac{n}{2}$, $\frac{n}{3}$,... itself is not an integer. If in fact the closed trajectory has $\tau = 2\tau_d$, this is called "period 2 motion". The transition from period one to period two motion as the parameters of the pendulum are changed is called **period doubling** and is the analog of the bifurcations seen in the fixed-point difference equation mapping. A Poincairé section can also be regarded as a mapping, using particle dynamics instead of a quadratic formula, from a state n to $n + 1$ The trajectories that the system settles down to in the steady state for the pendulum are the analog of the fixed points, x^*, in the difference equation. Such mappings, we have learned, can involve chaotic results and fractals in either context.

Since chaotic dynamical systems are still deterministic, true bifurcations in phase space do not exist. However, this bit of wisdom is essentially useless for true chaotic motion since neighboring trajectories, no matter how close to begin with, eventually diverge exponentially. (More on how to characterize this divergence in a moment.) However as I said above, bifurcations do exist in their parameter space descriptions. For example, if one measures $\dot{x}$ (angular velocity) as a function of the driving torque, F, on a Poincairé map, one sees bifucations appear as F is increased. These bifucations will be present either because a change in the initial conditions produces a new steady state or because a true period doubling trajectory has evolved. Eventually the period doublings merge into chaos, just as in the one-dimensional mappings. This period doubling route to chaos is again characterized by δ, the Feigenbaum number and has been seen in physical systems.

Dissipative Phase Space

How can chaos in dynamical systems be identified? Let us first try to understand the evolution of dissipative systems in phase space. Consider a simple case of two-dimensional particle trajectories in phase space as shown below:



Each particle has been assigned a “velocity”, $\vec{v}(x_1, x_2) = \frac{d\vec{x}}{dt}$, which is a function of position. The general formula for the instantaneous flux out of the (static) volume $V(0)$ is

$$f = \oint_S ds \ \rho(x_1, x_2) \ \vec{v} \cdot \hat{n} , \tag{4.56}$$

where the closed “surface”, S, is simply a rectangular perimeter. We have introduced in (4.56) a position dependent density of points, $\rho(x_1, x_2)$. Setting the density of points to one per unit volume, $\rho(x_1, x_2)=1$ and using Gauss’ law in two dimensions we obtain,

$$f = \int_V dV \ \vec{\nabla} \cdot \vec{v} . \tag{4.57}$$

The result (4.57) applies to an arbitrary number of dimensions if we generalize the meanings of the surface "S" and the volume “V” in (4.56) and (4.57). We will allow a dynamic volume, $V(t)$, to expand or contact to contain the same number of particles. The instantaneous change in this volume is just given by

$$\delta V(t) = \delta t \ f(t) , \tag{4.58}$$

$$\Rightarrow \qquad \frac{dV}{dt} = \int_V dV \ \vec{\nabla} \cdot \vec{v} . \tag{4.59}$$

For example, for the system (4.54), and using dimensionless time, t'

$$\vec{v} = \left(x_2, -cx_2 - \sin x_1 + F \cos x_3, \omega\right) ,$$

$$\Rightarrow \quad \vec{\nabla} \cdot \vec{v} = -c ,$$

$$\Rightarrow \frac{dV}{dt'} = -cV(t') ,$$

$$\Rightarrow \quad V(t') = V(0)e^{-ct'} . \tag{4.60}$$

From this we learn that the phase space is exponentially contracting $(c > 0)$ This is a general characteristic of dissipative systems. $\vec{\nabla} \cdot \vec{v} = 0$ characterizes a nondissipative system whose phase space volume can stretch and change, but whose volume is fixed. One can show that systems which possess a Hamiltonian (Ch.7) behave in this manner; this is called **Liouville’s theorem**.

Lyapunov Exponents

Even though the phase space volume is contracting for these systems, chaos emerges as a stretching of the phase space volume so that nearby phase space points eventually exponentially diverge. We can characterize the divergence of trajectories in a phase space plots by the **Lyapunov exponents**. They are related to the previous phase space considerations via the expression (ε_i is initial separation in the i^{th} direction between x_i' and x_i at $t=0$)

$$\lim_{t\to\infty} \left|x_i'(t) - x_i(t)\right| = \varepsilon_i \, e^{\lambda_i t} \tag{4.61}$$

Since the overall volume is given by (imagine a rectangle)

$$\lim_{t\to\infty} V(t) \sim \prod_i \left|x_i'(t) - x_i(t)\right| \tag{4.62}$$

($\prod_i$ is the product symbol over $i=1,2,3$) we find

$$\lim_{t\to\infty} V(t) \sim \prod_i \varepsilon_i e^{\lambda_i t} = \underbrace{(\prod_i \varepsilon_i)}_{V(0)} e^{\sum_i \lambda_i t}, \tag{4.63}$$

and the total volume grows with an exponent which is a sum of the individual Lyapunov exponents times the time. Thus, for the damped oscillator,

$$\sum_{i=1}^{3} \lambda_i = -c, \tag{4.64}$$

and the sum of the exponents is just the negative of the damping constant. Although the sum of exponents is assured negative, only one exponent has to be positive to indicate an exponential stretching of the phase space volume, $V(t)$. Thus, all of the exponents must be examined before a possible chaotic motion can be identified. (The calculation of the *individual* Lyapunov exponents for this system is much more difficult than getting the sum as above.)

The Lyapunov exponent for one-dimensional mappings can be numerically calculated. For small enough ε one assumes

$$f(x_0 + \varepsilon) - f(x_0) \simeq \varepsilon \frac{df}{dx}\Big|_{x_0}, \tag{4.65}$$

where $x_{n+1} = f(x_n)$ Next, one has

$$f^{(2)}(x_0 + \varepsilon) - f^{(2)}(x_0) \simeq \varepsilon \frac{df^{(2)}}{dx}\Big|_{x_0}. \tag{4.66}$$

where $f^{(2)}(x_0)$ is called the second iterate of $f(x_0)$,

$$f^{(2)}\left(x_0\right) \equiv f\left(f\left(x_0\right)\right). \tag{4.67}$$

Let us call $x_1 = f(x_0)$, $x_2 = f(x_1) = f(f(x_0)$, etc. Taking the derivative of $f^{(2)}(x_0)$ now gives by the chain rule,

$$\Rightarrow \frac{df^{(2)}}{dx}\bigg|_{x_0} = \frac{df^{(2)}}{dx}\bigg|_{f(x_0)}\frac{df}{dx}\bigg|_{x_0} = \frac{df}{dx}\bigg|_{x_1}\frac{df}{dx}\bigg|_{x_0}. \tag{4.68}$$

After n iterations,

$$f^{(n)}\left(x+\varepsilon\right) - f^{(n)}\left(x\right) = \varepsilon\frac{df}{dx}\bigg|_{x_{n-1}}\frac{df}{dx}\bigg|_{x_{n-2}}\cdots\frac{df}{dx}\bigg|_{x_0}. \tag{4.69}$$

The definition of the Lyapunov exponent is essentially the same as before, with the iterate number, n, playing the role of time ($x_0^{'} \equiv x_0 + \varepsilon$),

$$\lim_{n\to\infty}\left|\frac{x_n^{'} - x_n}{\varepsilon}\right| = \lim_{n\to\infty}\left|\frac{f^{(n)}\left(x_0+\varepsilon\right) - f^{(n)}\left(x_0\right)}{\varepsilon}\right| = e^{n\lambda}, \tag{4.70}$$

$$\Rightarrow \lambda = \lim_{n\to\infty}\frac{1}{n}\ln\left[\left|\frac{df}{dx}\bigg|_{x_{n-1}}\cdots\frac{df}{dx}\bigg|_{x_0}\right|\right],$$

$$= \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln \left[\left| \frac{df}{dx} \right|_{x_i} \right]. \tag{4.71}$$

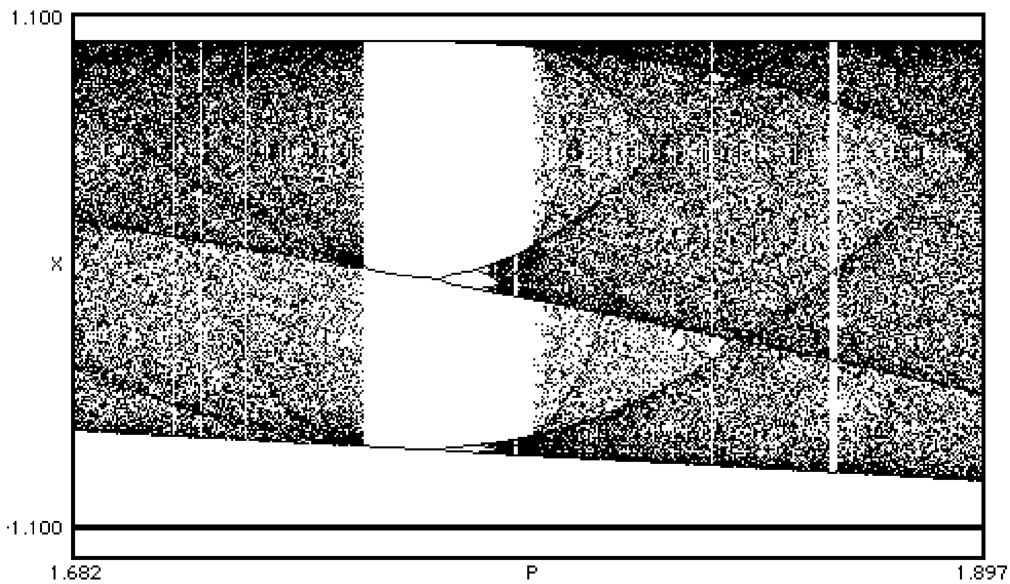
This can be numerically calculated as a function of the parameters of the mapping and chaotic regions can be identified.

If one of the Lyapunov exponents is positive, this is a necessary but not sufficient condition for dynamical chaos to emerge. For example, if we (somewhat unrealistically) changed the damping constant to be negative for the harmonic oscillator, we would have a system which is exponentially sensitive to initial conditions, but which is linear and so is not chaotic. On the other hand, even though the sum of the Lyapunov exponents for the damped pendulum is negative, only one has to be positive for this nonlinear system to exhibit chaotic motion. It is a combination of the nonlinear behavior and the exponential growth in at least one direction which are characteristic of dynamical chaos.

The Intermittent Transition to Chaos

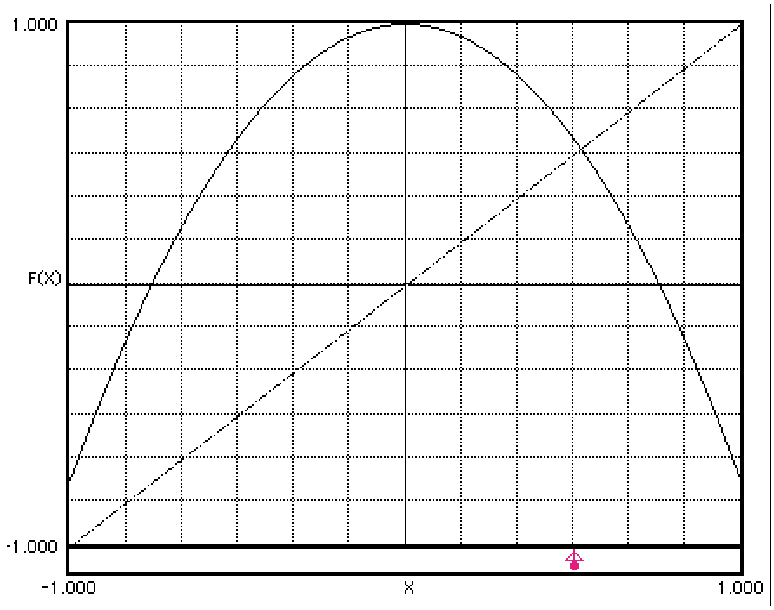
We have learned about one path or way of transitioning from simple, predictable periodic motion to chaotic motion. It was called period doubling and is illustrated in the logistic fixed point map above. However, there are other ways systems can transition from simple periodic motions to chaotic ones. A system may appear to be periodic or well behaved for a long time and then diplay "bursts" of nonperiodic behavior. As one changes the system parameters closer to the chaotic regime, these bursts take place more frequently until fully chaotic behavior results.

This *intermittent transition* to chaos can also be understood from our one dimensional logistic map. In the above fixed point map near the region of the large period 3 window for m slightly above 1.75, the system is in a three period, that is, after a few iterations it settles down is a stable system with three fixed points.

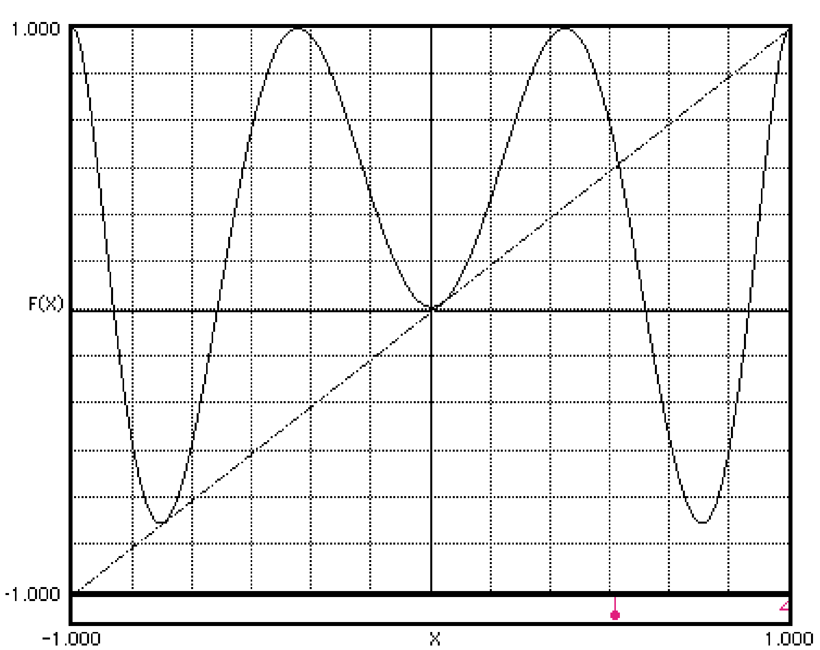


However, notice the wall of chaos to the left of this periodic region. When we consider values just slightly inside this chaotic region, we almost have a three period. That is, the system behaves for a long time as if it is in a three period, but with occasional bursts of chaos. The further we imbed ourselves in this chaotic region, the less we see of the three period and the more chaos we see. How can this sort of behavior be understood?

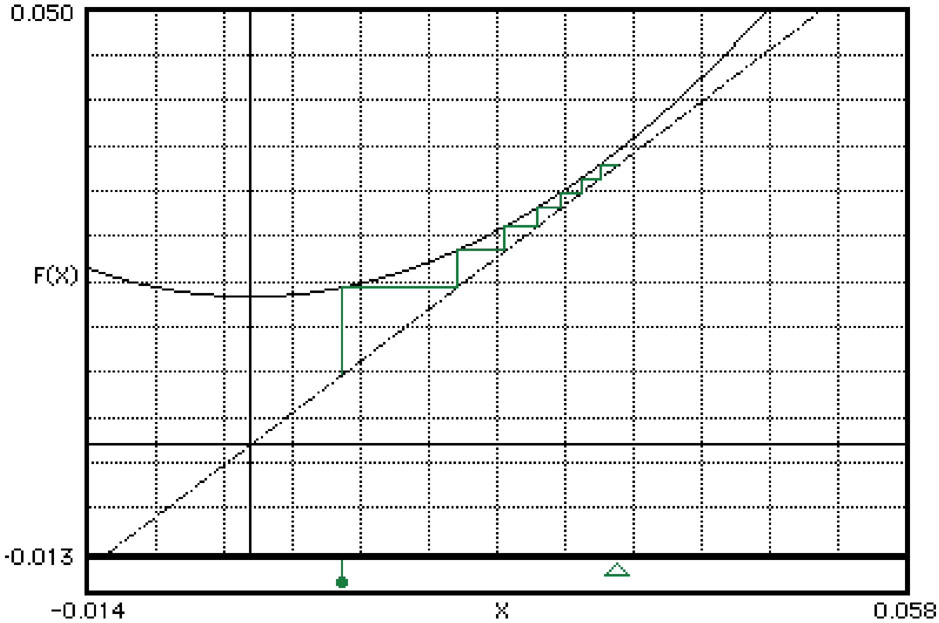
Amazingly, this behavior may be understood very simply. The function $f(x_n) = 1 - \mu x_n^2$ with $\mu=1.75$ looks like the following.



However, the third iterate of this function, $f^{(3)}(x_n) = f(f(f(x_n)))$, looks like the following:



Notice the four values in x where $f^{(3)}(x_n)$ touches the $45°$ line. One of these is unstable (which one?); the other three are fixed points of the mapping, x^*, and correspond to the three lines of bifurcation in the nonchaotic window. Let us zoom in on the middle these fixed point regions near $x=0$ when μ has a value of 1.7495, slightly into the the chaotic region. Notice that a "tangency channel" has developed where the function is close to, but not actually touching the $45°$ line.



Also shown in this figure are some iterates of $f^{(3)}(x_n)$ going into the tangency channel. Notice the the movement of the iterates through the channel becomes quite slow in the bottleneck part of the channel. This corresponds to the situation of having a nearly periodic three cycle system. However, eventually the iterates tunnel through the channel, reaching the chaotic region outside, corresponding to the "bursts" of chaotic motion. Put another way, the closer we are to true tangency (at $\mu=1.75$), the narrower the channel and the more time the system displays regular, predictable motion. The bursts occur more often the further we go into in the chaotic region because the tangency channel is wider.

The number of iterations that the system stays in the channel, displaying quasi-periodic motion, can be understood from the above difference equation. In the vicinity of the tangency, the third iterate, , may be expanded in a Taylor series to read

$$f^{(3)}(x_n) = x_n + a_c(x_n - x^*)^2 + b_c(\mu_c - \mu), \tag{4.72}$$

where $\mu_c = 1.75$ and x^* is one of the three fixed point values. After the transformation $y_n = (x_n - x^*)/b_c$, the recusion relation $x_{n+3} = f^{(3)}(x_n)$ in the vicinity of the channel may be written

$$y_{n+3} - y_n = ay_n^2 + \varepsilon, \tag{4.73}$$

where $a = a_c b_c$ and $\varepsilon = \mu_c - \mu > 0$. When the number of steps in the channel is large, (4.73) may be approximated by a differential equation,

$$\frac{dy}{dn} = ay^2 + \varepsilon. \tag{4.74}$$

This differential equation may be solved by separation of variables for n:

$$n(y_{out}, y_{in}) = \frac{1}{\sqrt{a\varepsilon}}\left[\tan^{-1}\left(y_{out}\sqrt{\frac{a}{\varepsilon}}\right) - \tan^{-1}\left(y_{in}\sqrt{\frac{a}{\varepsilon}}\right)\right], \tag{4.75}$$

where y_{in} and y_{out} are the entrance and exit values of y in the channel. Notice that for y_{out}, y_{in} fixed, $n(y_{out}, y_{in})$ goes to infinity like $\frac{1}{\sqrt{\varepsilon}}$ as $\varepsilon \to 0$.

This describes the transition to chaos that one is seeing in the logistic map near $\mu=1.75$ and many other regions. The transition to chaos in a period doubling manner and in an intermittent manner are very closely related. We have seen that period doubling occurs when there is a perpendicular crossing of the 45° line. Intermittent chaos emerges when instead the crossing is a tangent to the line. In another sense they are also related. Notice in the above graph of the 3 period we have been discussing, that the other side of the periodic window displays a period doubling cascade to chaos. Such windows always displays these two, complimentary, transition routes.

CHAPTER 4 PROBLEMS

1. Verify that the change of variables

   $$x = \sqrt{\frac{x' - c}{b - c}},$$

   converts Eq.(4.8) into (4.9).

2. By making the substitution $x = \sin\theta$ in (4.27), show that the elliptic integral $F(1,k)$ may be approximated by,

   $$F(1,k) = \frac{\pi}{2}\left(1 + \frac{1}{4}k^2 + \frac{9}{64}k^4 + \ldots\right),$$

   for $k<<1$.

3. A free (not driven) underdamped oscillator is released or set in motion under the following initial conditions:

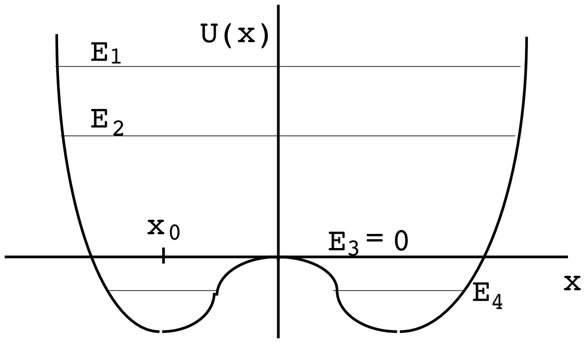
   a) $x(0) = x_0$, $\dot{x}(0) = 0$ $(x_0 > 0)$
   b) $x(0) = 0$, $\dot{x}(0) = \dot{x}_0$ $(\dot{x}_0 > 0)$.

Sketch phase diagrams of the resulting motion in the two cases. Make sure your phase trajectories have arrows.

4. a) The potential function $(A, \lambda > 0)$

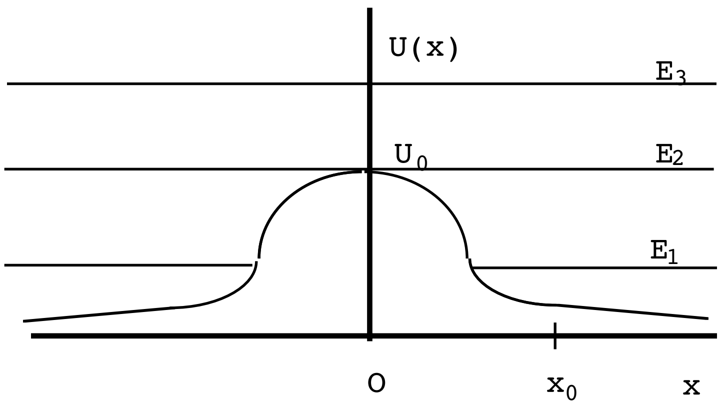
$$U(x) = A(x^4 - \lambda^2 x^2),$$

looks like



Sketch a phase space diagram showing qualitative trajectories for energies $E_{1,2,3,4}$. $(E_3 = 0.)$ (These trajectories must have a direction.) Identify points of stable and unstable equilbrium in your sketch.

5. A one-dimensional potential has the form, $U(x) = U_0 e^{-ax^2}$



a) Sketch a qualitative phase space diagram corresponding to system total energies E_1, E_2 and E_3 $(E_2 = U_0.)$
b) Starting from $x = x_0$ with total energy $E = U_0$, find the time necessary to reach the origin, $x = 0$. Is it finite or infinite?

6. A one-dimensional potential energy, $U(x)$ is given by $U(x) = -K|x|^n$, where $|x|$ is the absolute value of x, K is a positive constant and $n > 1$ For what values of "n" does a particle, starting at $x_0 \neq 0$ take a **finite** time to reach the top of the potential hill when the total energy, $E = \frac{1}{2} m\dot{x}^2$ is zero?

7. The steady-state response of a damped oscillator to a periodic force,

$$m\ddot{x} + b\dot{x} + kx = F_0 \cos(\omega t),$$

is

$$x(t) = D(\omega)\cos(\omega t - \delta),$$

where $\left(2\beta = \frac{b}{m},\ \omega_0^2 = \frac{k}{m},\ A_0 = \frac{F_0}{m}\right)$

$$D(\omega) = \frac{A_0}{\sqrt{(\omega_0^2 - \omega^2)^2 + (2\omega\beta)^2}},\quad \tan\delta = \frac{2\beta\omega}{\omega_0^2 - \omega^2}.$$

a) Show that the trajectories in (projected) $\dot{x}$, x phase space are ellipses.
b) Show that the driven oscillation total energy is constant in time if $\omega = \omega_0$.

c) We know that the (unprojected) phase space of the same oscillator when it is free (i.e., undamped, nondriven) is also an ellipse. Show that these two ellipses coincide only if $\omega^2 = \omega_0^2$ and their total energies are equal.

8. Given the one-dimensional mapping (all $x_n \geq 0$),

$$x_{n+1} = \sqrt{x_n} \ ,$$

a) Find all of the fixed points (there are two) and show that one is stable, the other is not.
b) Calculate the unique Lyapunov exponent for this mapping.

9. We investigated the mapping $(\mu \in (0,2), x \in [-1,1])$

$$x_{n+1} = 1 - \mu x_n^2$$

for fixed points. We found period doubling at $\mu = .75, 1.25, \ldots$. Show that this equation is equivalent to the mapping

$$y_{n+1} = \alpha y_n (1 - y_n).$$

for $\alpha \in (0,4)$ and $y \in [0,1]$. At what value of α does the first period doubling occur?

10. When n becomes large, the period doubling values, μ_n, for the mapping in prob.9 obey

$$\mu_n \approx \mu_\infty - \frac{A}{\delta^n} \quad ,$$

where μ_∞, δ and A are constants. From this show that

$$\lim_{n \to \infty} \frac{\mu_n - \mu_{n-1}}{\mu_{n+1} - \mu_n} = \delta .$$

This says that the ratio of intervals tends to a fixed limit (although the intervals themselves are shrinking to zero).

11. a) A real linear mapping is given by

$$x_{n+1} = \beta x_n + c,$$

where $|\beta| < 1$ and c is a real constant. Show that the single fixed point of the mapping is given by $x^* = \frac{c}{1 - \beta}$. What happens for $|\beta| > 1$?

b) Explain the relevance of the result in (a) for a linear mapping to the idea of bifurcations seen in nonlinear mappings.

12. As shown in problem 11, the real linear finite difference equation,

$$x_{n+1} = \beta x_n + c,$$

with $|\beta| < 1$ and c a constant has a single (stable) fixed point at $x^* = \frac{c}{1 - \beta}$. Let's say we decide to find x^* by brute force iteration. Picking an initial value x_0, we iterate the difference equation until we obtain $|x_N - x^*| < \varepsilon$, where e is a small positive number. Show that the number of iterations needed to achive this accuracy is given by

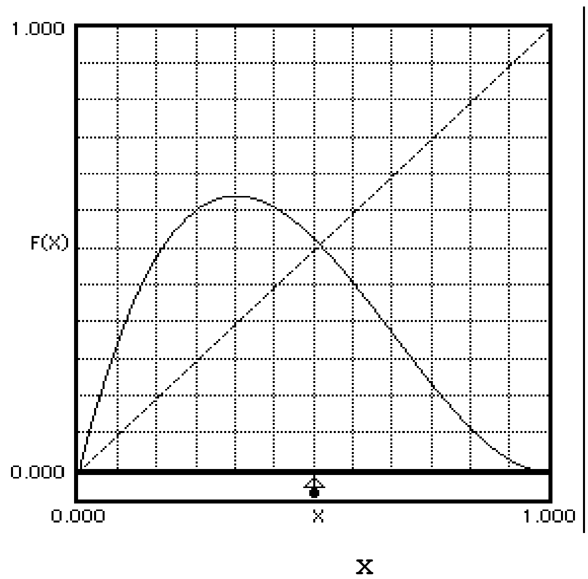
$$N = \left| \frac{\ln \frac{e}{|x_0 - x^*|}}{\ln |b|} \right|,$$

showing convergence is worst for $|\beta| – 1$.

13. Consider the non-linear difference equation given by

$$x_{n+1} = F(x_n) = \alpha x_n (1 - x_n)^2,$$

for $0 < x_n < 1$ and $\alpha > 1$. It looks like:



a) Find the location of the fixed points of the mapping for $0 \leq x_n \leq 1$.
b) At what value of a does the first "period doubling" bifurcation occur?

14. Consider the "tent map" given by ($0<\alpha<1$)

$$x_{n+1} = 2\alpha x_n, \quad 0<x_n<1/2,$$
$$x_{n+1} = 2\alpha(1-x_n), \quad 1/2<x_n<1.$$

a) Find the values of the fixed points of this map. (There are two.)
b) Examine each point in (a) for stabilty. For what values of α are the fixed points stable? [Hint: Think about slope.]

15. a) Show that the transformation,

$$y_n = \frac{2}{\pi} \sin^{-1}(\sqrt{x_n}),$$

in $x_{n+1} = a\, x_n(1-x_n)$ for $a=4$ converts it into a tent map,

$$y_{n+1} = 2y_n, \quad 0<y_n<1/2,$$
$$y_{n+1} = 2(1-y_n), \quad 0<y_n<1/2.$$

b) Therefore calculate the Lyupanov exponent of the above transformation for $a=4$. (Ans: $\ell n(2)$.)

16. Consider the three coupled first-order differential equations that Edward Lorenz considered (same as in this text),

$$\frac{dx_1}{dt} = \delta(x_2 - x_1),$$

$$\frac{dx_2}{dt} = rx_1 - x_2 - x_1x_3,$$

$$\frac{dx_3}{dt} = x_1x_2 - bx_3,$$

where δ, r and b are constants. The rate at which a volume of the configuration space changes is given by

$$V(t) = V(0)e^{\lambda t}.$$

Find λ. Under what conditions is this system dissipative (does the configuration space shrink in time)? Given the values $\delta=10,\ r=-2,\ b=-10,$ is the system dissipative? Can you tell if the system is chaotic? Answer the same two questions for $\delta=10,\ r=-2,\ b=-12.$

17. Show that when the three equations in (4.54) are combined, they give the single equation, (4.55).

18. Evaluate the Lyapunov exponent of the linear mapping,

$$x_{n+1} = \beta x_n + c.$$

Show that it is negative only for $|\beta| < 1$.

19. By numerical calculation, find the values of the three stable fixed points of the mapping $f^{(3)}(x_n) = f(f(f(x_n)))$, $f(x_n) = 1 - \mu x_n^2$, for $\mu=1.76$.

20. Show that when Eq.(4.72) and the transformation $y_n = (x_n - x^*)/b_c$ are used in the recusion relation $x_{n+3} = f^{(3)}(x_n)$, the recursion relation in the vicinity of a tangency channel may be written

$$y_{n+3} - y_n = ay_n^2 + \varepsilon.$$

21. (a) Use the relation Eq.(4.71) and the result (4.73) to show that the Lyapunov exponent for a single pass through the tangency channel is approximately given by

$$\lambda(y_{out}, y_{in}) \cong \frac{1}{n(y_{out}, y_{in})} \int_{y_{in}}^{y_{out}} dy\, \frac{\ln(1 + 2ay)}{ay^2 + \varepsilon}.$$

[Hint: Make the same sort of approximation in going from a sum to an integral that we did in the text in solving (4.74).]

b) Make the approximation that 2ay is small inside the integral above and do the integral. From this predict the behavior of $\lambda(y_{out}, y_{in})$ on ε as $\varepsilon \to 0$ (y_{in}, y_{out}, held constant).

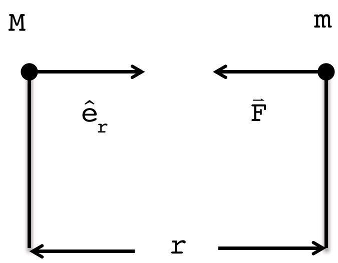
5 GRAVITATION

NEWTON'S LAW OF GRAVITATION

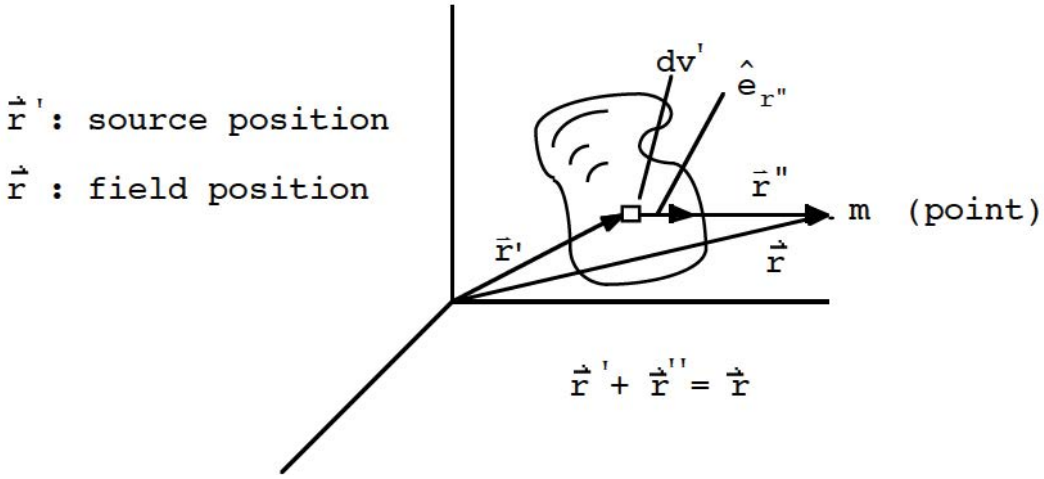
Newton's law of gravitation for point masses is (force on m due to M):

$$\vec{F} = -G\frac{mM}{r^2}\hat{e}_r\ . \tag{5.1}$$

Picture: (M usually more massive than m)



Obviously an idealization. (Masses are not points. Also, forces are not transmitted *instantaneously*.) More general form for a mass distribution:



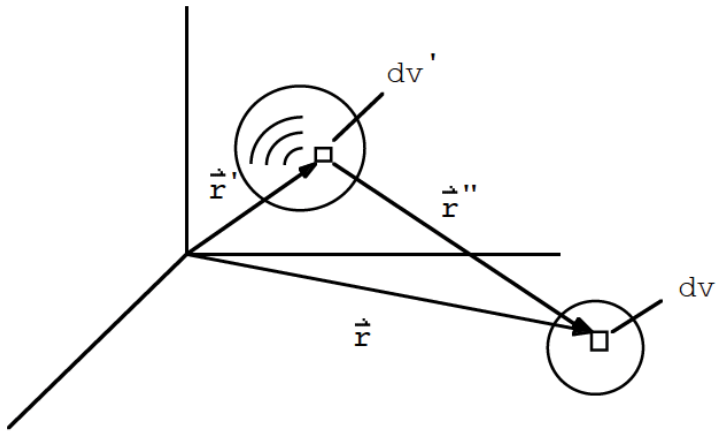
$$d\vec{F} = -Gm \frac{\rho(\vec{r}\,')}{r''^2} \hat{e}_{r''} dv' , \quad (5.2)$$

$$\Rightarrow \quad \vec{F}(\vec{r}) = -Gm \int \frac{\rho(\vec{r}\,')}{r''^2} \hat{e}_{r''} dv' . \quad (5.3)$$

where

$\vec{r}'' = \vec{r} - \vec{r}'$ (can not be taken outside the integral in general)

Even this is an idealization for, say, two planets:



$$dM \rightarrow \rho_1(\vec{r}')dv' , \quad dm \rightarrow \rho_2(\vec{r})dv ,$$

$$\vec{F} = -G \iint dv' dv \frac{\rho_1(\vec{r}\,')\rho_2(\vec{r})}{r''^2} \hat{e}_{r''} . \quad (5.4)$$

Gravitational Potential

Things are beginning to get complicated. We need some simplifications. First simplification: the idea of a gravitational potential. Go back to point masses case. Notice

$$\vec{\nabla}\left(-G\,\frac{Mm}{r}\right) = -GMm\,\vec{\nabla}\left(\frac{1}{r}\right),$$

$$\frac{\partial}{\partial x_1}\left(\frac{1}{\sqrt{x_1^2 + x_2^2 + x_3^2}}\right) = -\frac{1}{2}\,\frac{2x_1}{\left(x_1^2 + x_2^2 + x_3^2\right)^{3/2}} = -\frac{x_1}{r^3}.$$

Likewise for x_2, x_3.

$$\Rightarrow \quad \vec{\nabla}\left(-\frac{GMm}{r}\right) = \frac{GMm}{r^2}\,\hat{e}_r = -\vec{F}. \qquad (5.5)$$

Therefore, with $U(r) \equiv -G\,\dfrac{Mm}{r}$ (makes $U=0$ at $r=\infty$), we have

$$\vec{F} = -\vec{\nabla}U. \qquad (5.6)$$

Define $\Phi \equiv \dfrac{U}{m}$ and call it the "gravitational potential". It represents the work/mass to move a test mass from one point to another. In the case of M being an extended mass, it is easy to see that

$$U(\vec{r}) = -Gm \int \frac{\rho\left(\vec{r}\,'\right)}{r\,''}\,dv\,', \qquad (5.7)$$

$$\Rightarrow \quad \Phi(\vec{r}) = -G \int \frac{\rho\left(\vec{r}\,'\right)}{r\,''}\,dv\,'. \qquad (5.8)$$

Verification:

$$-\vec{\nabla}U = Gm\int \rho\left(\vec{r}\,'\right)\vec{\nabla}\,\frac{1}{\left|\vec{r}\,' - \vec{r}\right|}\,dv\,'.$$

We have

$$\frac{\partial}{\partial x_1}\left(\frac{1}{\sqrt{\left(x_1' - x_1\right)^2 + \left(x_2' - x_2\right)^2 + \left(x_3' - x_3\right)^2}}\right) = \frac{\left(x_1' - x_1\right)}{r\,''^3},$$

$$\Rightarrow -\vec{\nabla}U = Gm\int dv\,'\,\rho\left(\vec{r}\,'\right)\underbrace{\frac{\left(\vec{r}\,' - \vec{r}\right)}{r\,''^3}}_{-\frac{\hat{e}_{r''}}{r\,''^2}}.$$

The point here is that Φ is much simpler to calculate than $\vec{F}$. One usually defines

$$\vec{g} = \frac{\vec{F}}{m} \quad \Rightarrow \quad \vec{g} = -\vec{\nabla}\Phi \,. \tag{5.9}$$

Near the Earth's surface:

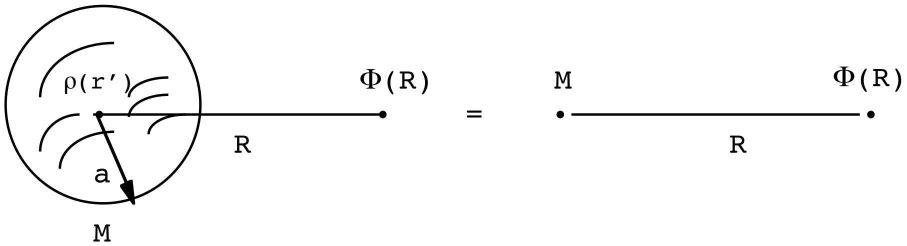
$$|\vec{g}| \simeq 980 \text{ cm/sec}^2 \qquad \text{direction: "downward"}$$

Modifications for Extended Objects

We can recover the usual form of gravitational potential energy near the Earth's surface, with the help of a result we will prove momentarily. The result we need is that for a spherical mass distribution

$$\Phi = -G\int \frac{\rho(r')}{r''}\,dv' = -\frac{GM}{R}, \qquad \left(U = -\frac{GMm}{R}\right) \tag{5.10}$$

where R is the distance to the center of the sphere. Pictorially,



Then we have $(R = a + h \;,\; h \ll a)$

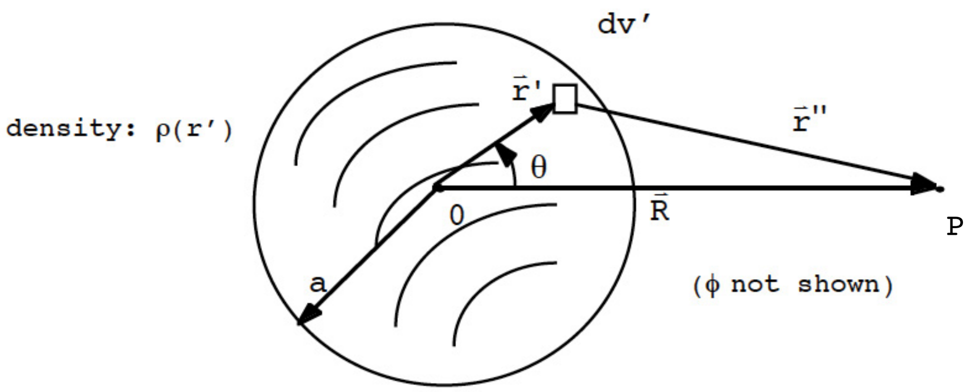
$$\frac{1}{R} = \frac{1}{a+h} = \frac{1}{a\left(1+\frac{h}{a}\right)} \simeq \frac{1}{a}\left(1 - \frac{h}{a}\right)$$

$$\Rightarrow U \simeq -\frac{GMm}{a}\left(1 - \frac{h}{a}\right) = \underbrace{-\frac{GMm}{a}}_{\text{const.}} + mgh$$

$$g \equiv \frac{GM}{a^2} \left(= 980 \text{ cm / sec}^2\right) \qquad \text{5.11)}$$

↑(assuming a perfectly spherical Earth)

Let's now prove the result that the potential outside a spherically symmetric distribution of matter is independent of the size of the distribution and acts as if all the mass were concentrated at the sphere's center. (This is the problem which spurred Newton to the creation of calculus. Note that $\vec{r} \to \vec{R}$ here.)



0 : origin of coordinates

$$\Phi(R) = -G\int dv' \frac{\rho(r')}{r''}, \quad (dv' = r'^2 \sin\theta \, d\theta \, d\varphi \, dr') \tag{5.12}$$

$$\Rightarrow \Phi(R) = -2\pi G \int_0^a dr' \rho(r') r'^2 \int_0^\pi \frac{\sin\theta}{r''} d\theta \, . \tag{5.13}$$

Now

$$\vec{R} = \vec{r}' + \vec{r}'', \tag{5.14}$$

$$\Rightarrow r''^2 = r'^2 + R^2 - 2Rr'\cos\theta, \tag{5.15}$$

$$\Rightarrow \Phi(R) = -2\pi G \int_0^a dr' \frac{\rho(r') r'^2}{2r'R} \int_0^\pi \frac{2r'R\sin\theta d\theta}{\sqrt{r'^2 + R^2 - 2Rr'\cos\theta}} \, . \tag{5.16}$$

In the later form, the θ integral is a perfect differential. We have

$$\Phi(R) = -2\pi G \int_0^a \frac{r'dr'}{2R} 2\sqrt{r'^2 + R^2 - 2Rr'\cos\theta} \Big|_0^\pi \tag{5.17}$$

Notice that the above now involves absolute values

$$\Phi(R) = -\frac{2\pi G}{R} \int_0^a dr' r' \rho(r')(|r' + R| - |R - r'|) \, . \tag{5.18}$$

If the point P is outside the mass distribution, we have $|R - r'| = R - r'$, and the above becomes

$$\Phi(R) = -\frac{4\pi G}{R} \int_0^a dr' r'^2 \rho(r') \, . \tag{5.19}$$

However, the total mass of the distribution is

$$M = 4\pi \int_0^a dr' r'^2 \rho(r'), \tag{5.20}$$

so therefore

$$\Phi(R) = -\frac{GM}{R} \quad , \tag{5.21}$$

and thus

$$\vec{F} = -\frac{GMm}{R^2}\hat{e}_R \ , \tag{5.22}$$

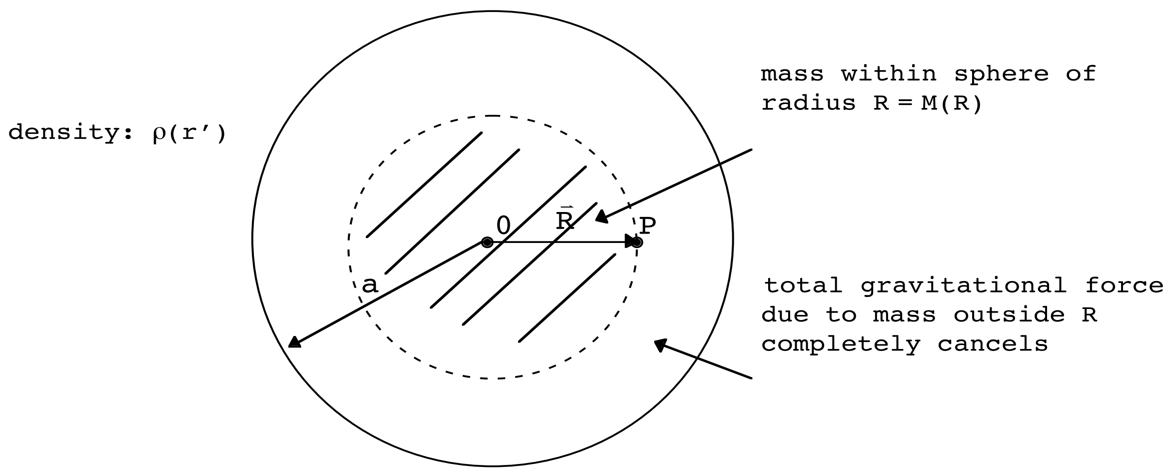
which says that the distributed mass just behaves as if it were concentrated at a point at the sphere's center. Interestingly, the same essential form holds even if the point P is located within the mass distribution, $r' < R$. Then, one may show (HW problem) that the same form holds,

$$\vec{F} = -\frac{GM(R)m}{R^2}\hat{e}_R \ , \tag{5.23}$$

where

$$M(R) \equiv 4\pi\int_0^R dr'\, r'^2\, \rho(r') \ , \tag{5.24}$$

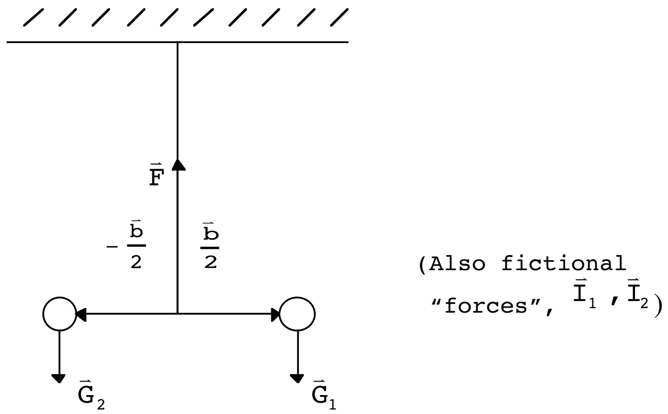
is the total mass within a sphere of radius R. See the following picture of the situation.



Eötvös Experiment on Composition Dependence of Gravitational Forces

I would now like to pick up the thread of the argument concerning the supposed equality between gravitational and inertial mass (Chapter 2). This equality can be tested. Let me describe the Eötvös experiment. (Follows J. Weber, "General Relativity and Gravity Waves"). This experiment was done in 1905 and established the equivalence of these two quantities for specific materials, to about one part in 100 million.

Here is the basic idea of a torsion balance:



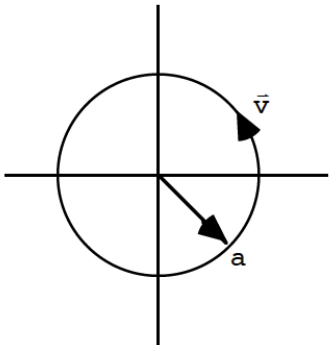
$(\text{Gravitational mass})_{1,2} = M_{1,2}$

$(\text{Inertial mass})_{1,2} = m_{1,2}$

We will define the ratio of gravitational to inertial mass to be

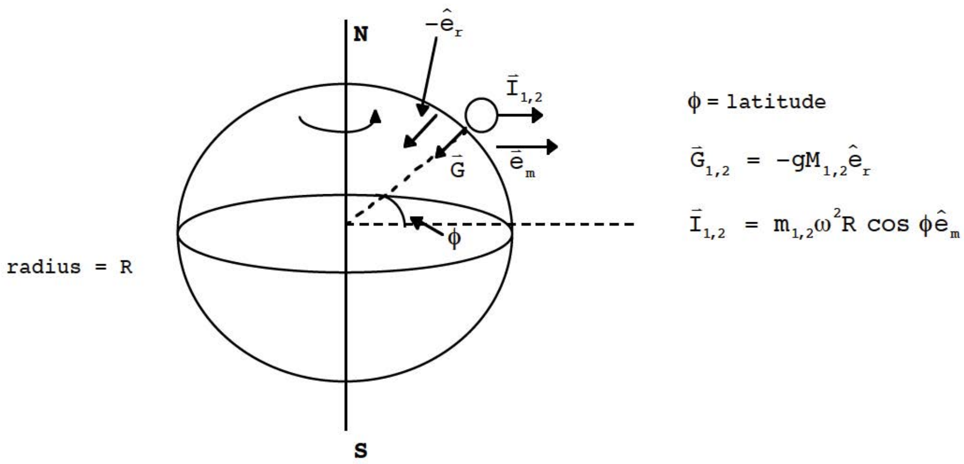
$$\alpha_{1,2} \equiv \frac{M_{1,2}}{m_{1,2}}.$$

Origin of "fictional" forces is the spinning Earth:



$$|\text{force}| = \frac{mv^2}{a} = m\omega^2 a$$

Apply to the Earth:



Net torque on torsion balance:

$$\vec{T} = \frac{\vec{b}}{2} \times \left(\vec{G}_1 + \vec{I}_1\right) - \frac{\vec{b}}{2} \times \left(\vec{G}_2 + \vec{I}_2\right), \tag{5.25}$$

$$\vec{F} = \vec{G}_1 + \vec{G}_2 + \vec{I}_1 + \vec{I}_2, \text{ (along wire)} \tag{5.26}$$

$$T \equiv \frac{\vec{F}\cdot\vec{T}}{|\vec{F}|}, \text{ (crucial thin+g to calculate)}$$

$$\Rightarrow T = \frac{\left(\vec{G}_1 + \vec{G}_2 + \vec{I}_1 + \vec{I}_2\right)\cdot\left[\frac{\vec{b}}{2}\times\left(\vec{G}_1 + \vec{I}_1\right) - \frac{\vec{b}}{2}\times\left(\vec{G}_2 + \vec{I}_2\right)\right]}{|\vec{F}|},$$

$$\Rightarrow T = \frac{\vec{b}\cdot\left[\left(\vec{G}_1 + \vec{I}_1\right)\times\left(\vec{G}_2 + \vec{I}_2\right)\right]}{|\vec{F}|}. \tag{5.27}$$

The second term above just doubles the result from the first. We have the reduction:

$$\left(\vec{G}_1 + \vec{I}_1\right)\times\left(\vec{G}_2 + \vec{I}_2\right)$$

$$= \left(-g\,M_1\hat{e}_r + m_1\omega^2 R\cos\phi\hat{e}_m\right)\times\left(-g\,M_2\hat{e}_r + m_2\omega^2 R\cos\phi\hat{e}_m\right),$$

$$= -g\,\omega^2 R\cos\phi\,\left(M_1 m_2 - M_2 m_1\right)\left(\hat{e}_r \times \hat{e}_m\right). \tag{5.28}$$

Putting the pieces together yields the component of torque causing a rotation,

$$\Rightarrow T_{11} = \frac{-g\,\omega^2 R\cos\phi\, m_1 m_2\left(\alpha_1 - \alpha_2\right)\vec{b}\cdot\left(\hat{e}_r \times \hat{e}_m\right)}{\left|\vec{F}\right|}. \tag{5.29}$$

Since we are working only to first order in $\left(\alpha_1 - \alpha_2\right)$,, we may use

$$\left|\vec{F}\right| \cong g\left(M_1 + M_2\right), \tag{5.30}$$

$$\frac{m_1 m_2}{M_1 + M_2} \cong \mu, \tag{5.31}$$

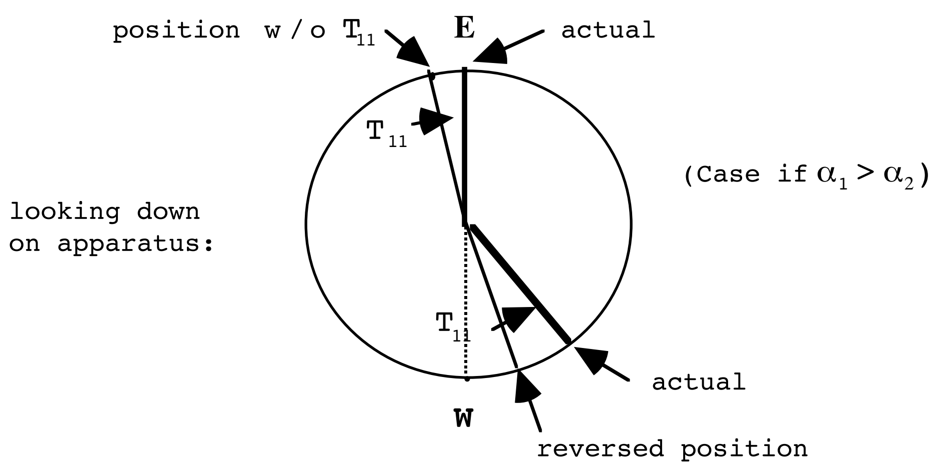
yielding finally

$$T_{11} = -\mu\omega^2 R\cos\phi\,(\alpha_1 - \alpha_2)\,\vec{b}\cdot(\hat{e}_r \times \hat{e}_m), \tag{5.32}$$

where $\vec{b}\cdot(\hat{e}_r \times \hat{e}_m)$ is a max in E–W direction.

Description of the actual measurement:

The torsion balance is brought to equilibrium by turning it until it points in an E – W direction. One can not tell yet if there is a nonzero component of torque, T_{11}, since it is in equilibrium. Now turn the apparatus 180° reversing the sense of **b**. If $\alpha_1 \neq \alpha_2$ there will be a torque which will rotate the rod relative to the frame which supports the balance by an angle twice the (unobserved) initial deflection. See the diagram below.



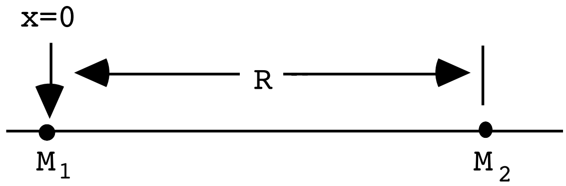
Best to do the expt. at $\phi \sim 45^\circ$ since (with $\vec{b}$ oriented E – W)

$$\cos\phi\left|\vec{b}\cdot\left(\hat{e}_r \times \hat{e}_m\right)\right| = b\sin\phi\cos\phi.$$

Later experiments: Dicke, Roll, Krotkov established equivalence between gravitational and inertial masses, again for specific materials, to about one part in 100,000 million (1 in 10–11!!! See the October '99 **Physics Today** article from Clifford Will for more updated information on gravitational composition dependence experiments. In the literature, the postulate that there is no composition dependence to gravitational forces is called the "weak equivalence principle."

CHAPTER 5 PROBLEMS

1. Consider two point masses which are stationary in some inertial reference frame.



a) Evaluate and sketch as a function of x (measured along the distance between the planets) the gravitational potential, $\Phi(x)$, a test mass would experience.
b) Find the value of x (other than $x = \pm\infty$) for which the test mass feels zero gravitational force. Is this point stable?

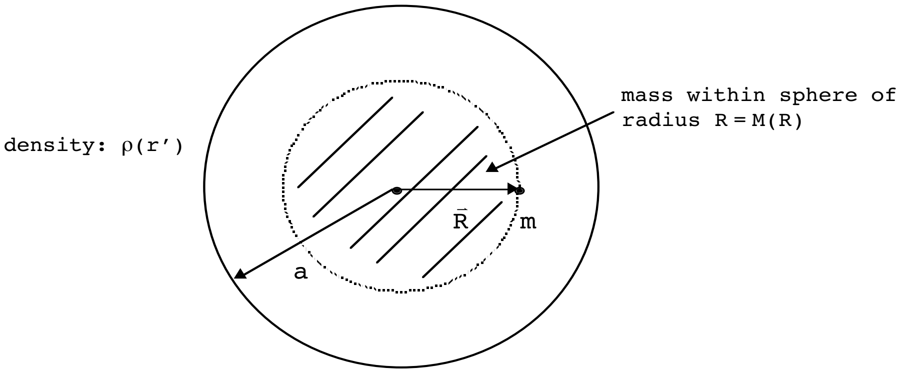
2. Show, for a spherical mass distribution, that the gravitational force on a small test particle of mass m located inside a planet of radius "a" and a distance R from it's center ($a>R$) that

$$\vec{F} = -\frac{GM(R)m}{R^2}\hat{e}_R\ ,$$

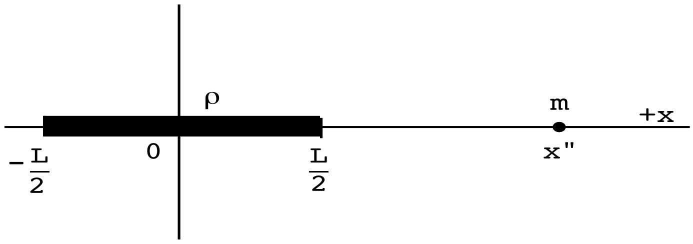
where

$$M(R) \equiv 4\pi\int_0^R dr'\, r'^2\, \rho(r')\ ,$$

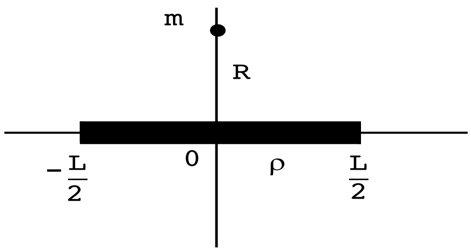
is the total mass within the sphere of radius R. See the following picture of the situation.



3. Find the force of gravity on a small spacecraft of mass m, distance x'' from the middle of a long, straight space station (Babylon 5) of uniform density per unit length, ρ, and length L. Take your origin of coordinates, O, in the middle of the space station, as shown.



4. A small rock of mass m is orbiting the Babylon 5 space station. It orbits in a plane perpendicular to the station, in a circle of radius R away from the space station's midpoint. The station has a length L and a constant mass/length, ρ.



Find the period of the orbit, T, in terms of R, L and ρ. [Hint: Get the gravitational acceleration and set it equal to centripetal acceleration, $\frac{v^2}{R}$. I have a table of integrals.]

5. Scientific Superman digs a hole completely through planet Krypton (which has a spherically symmetric mass distribution) and measures the force of gravity as a function of the distance from Krypton's center, R. He finds

$$\vec{F} = -H(R)\hat{e}_R,$$

where $\hat{e}_R$ is the usual unit vector pointing away from the planet's center. Show that this means Krypton's mass density, $\rho(R)$, is given by

$$\rho(R) = \frac{1}{4\pi Gm}\frac{1}{R^2}\frac{d}{dR}(R^2\, H(R)) \quad .$$

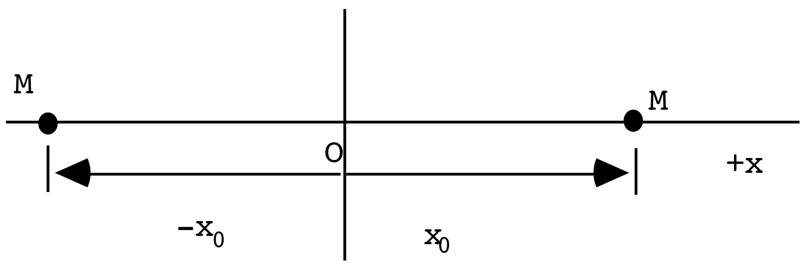
[Hint: Remember that

$$\bar{F} = -\frac{GM(R)m}{R^2}\,\hat{e}_R,$$

where $M(R)$ is the mass contained within a distance R from the planet's center.]

6. Scientific Superman finds that Krypton's mass density, ρ, is a constant in R. He now drops an evil villain down the hole he dug through Krypton. Show that the motion of the evil villain through Krypton is simple harmonic. Find the villain's period of motion in terms of the assumed constant ρ.

7. Two tiny planets of equal mass, M, attract one another and move along the line connecting them (no rotational motion). They are initially stationary and located symmetrical distances x_0 and $-x_0$ from the origin, O, as shown.



Show that the time necessary for the planets to crash into one another is given by

$$T^2 = \frac{\pi^2}{2G\,M}(x_0)^3 \; .$$

[Hint: I had to look up an integral. This result is actually a special case of what is known as Kepler's third law of motion which says that the square of the period of motion is proportional to the cube of an orbit parameter.]

6 CALCULUS OF VARIATIONS

EULER-LAGRANGE EQUATION

In ordinary calculus, we are used to finding stationary values (max's, min's or saddle points) of functions $f(x)$. We now consider the problem of finding a **function** to make an **integral** stationary.

Simplest form:

$$J = \int_{x_1}^{x_2} f\left[y(x), y'(x); x\right] dx .$$

dep. variable ↑ ↑ ↑ indep. variable

$\frac{dy}{dx}$ an indep. dep. variable

Some examples of $f(y,y';x)$ are:

a) $y\sqrt{1 + y'^2}$

b) $\sqrt{\frac{1 + y'^2}{y}}$

c) $a(x)y'^2 + b(x)y^2$

Just an in differential calculus, the vanishing of a first derivative is a necessary, but not sufficient condition for a max or min, so in the calculus of variations one speaks of 1^{st} or 2^{nd} variations of J. Here we will only work with the first variation and rely on geometrical or physical reasoning to decide whether we have found a max, min or inflection point. The point of learning this is the next chapter, where it will help us formulate a new, more general, view of classical dynamics. This chapter, however, is largely mathematics and geometry.

We will often need the following theorem in justifying statements in this Chapter:

If $F(x)$ is a continuous real function for $x_1 \le x \le x_2$ and if

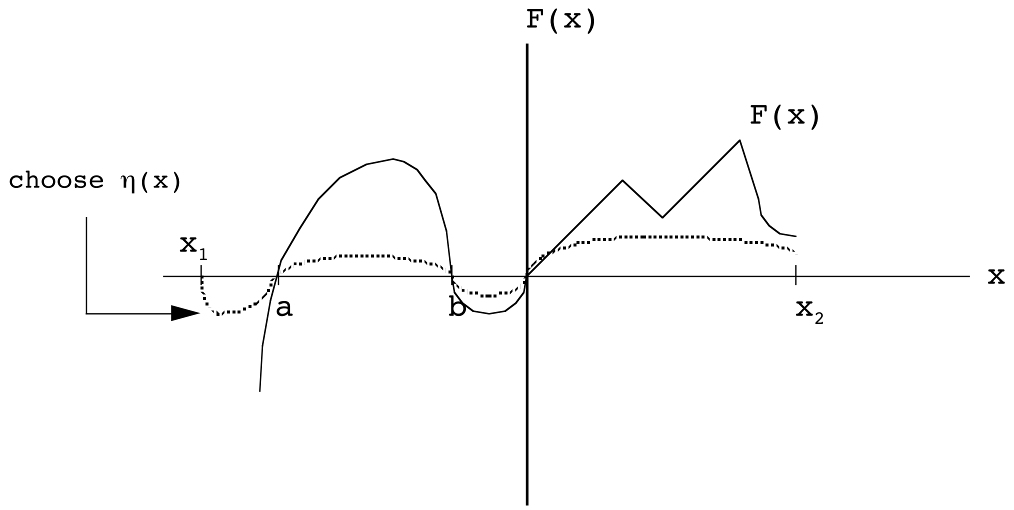
$$\boxed{\int_{x_1}^{x_2} dx\, \eta(x)\, F(x) = 0} \qquad (6.2)$$

for all real, continuous, once-differentiable functions $\eta(x)$ such that

$$\eta(x_1) = \eta(x_2) = 0 \, , \qquad (6.3)$$

then F(x) vanishes identically.

I will not prove it, but here is an "heuristic"[5] argument.



In each region (say $a \le x \le b$ above) we have

$$\int_a^b dx\, \eta(x)\, F(x) > 0\,. \tag{6.4}$$

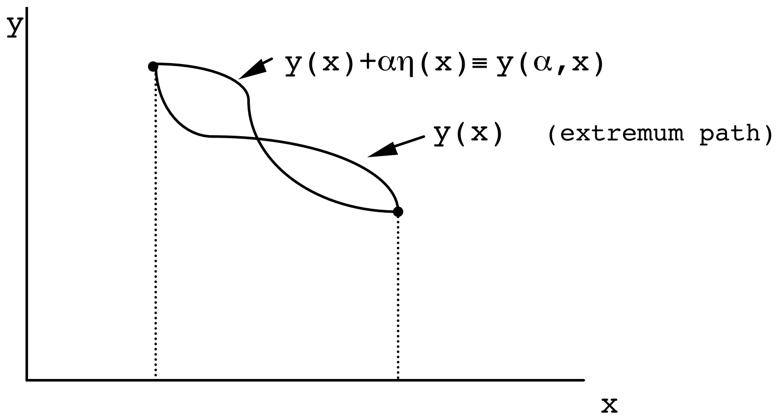
But we have that

$$\int_{x_1}^{x_2} dx\ \ \eta(x)\ \ F(x) = \int_{x_1}^{a} dx\ \ \eta(x)\, F(x) + \int_{a}^{b} dx\ \ \eta(x)\, F(x) + \ldots\,. \tag{6.5}$$

The only way to reconcile (6.2), (6.4) and (6.5) is if $F(x) = 0$ everywhere in the interval $x_1 \le x \le x_2$.

We can now go ahead with the rest of the development of this chapter.

We will need to develop some formalism first. Consider the two paths shown, one of which is assumed to be the extremum path.



$y(\alpha,x)$ is a neighboring or companion function.

$$y(\alpha,x) \equiv y(x) + \alpha\eta(x);\ y'(\alpha,x) = y'(x) + \alpha\eta'(x)\,. \tag{6.6}$$

α is a continuous real parameter and $\eta(x)$ is an arbitrary, once differentiable function for which $\eta(x_1) = \eta(x_2) = 0$. That is, we are keeping the end points fixed. We can then form the comparison integral (assuming $y(x)$ is twice differentiable)

$$J(\alpha) \equiv \int_{x_1}^{x_2} f\left[y(\alpha, x)\,, y'(\alpha, x)\,; x\right] dx\,, \tag{6.7}$$

where $J(0)$ is the extremum sought. The condition is then just

$$\left.\frac{dJ(\alpha)}{d\alpha}\right|_{\alpha=0} = 0\,. \tag{6.8}$$

Using the chain rule (I am using simplified notation),

$$\frac{dJ}{d\alpha}\bigg|_{\alpha=0} \int_{x_1}^{x_2} \overbrace{\left[\frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha} + \frac{\partial f}{\partial y'}\frac{\partial y'}{\partial \alpha}\right]}^{\frac{df}{d\alpha}} dx\,, \tag{6.9}$$

where, using the previous definitions

$$\left.\begin{aligned} \frac{\partial y}{\partial \alpha} &\equiv \eta(x)\,, \\ \frac{\partial y'}{\partial \alpha} &= \frac{d\eta}{dx} \equiv \eta'(x). \end{aligned}\right\} \tag{6.10}$$

We will call objects like $\frac{\partial f}{\partial y}, \frac{\partial f}{\partial y'}$ “functional derivatives” since we are taking derivatives with respect to (unknown) functions, not variables.

We have now

$$\frac{dJ}{d\alpha}\bigg|_{\alpha=0} = \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y}\eta(x) + \frac{\partial f}{\partial y'}\eta'(x)\right] dx\,. \tag{6.11}$$

Let us concentrate on the second term on the right hand side. By integration by parts,

$$\int_{x_1}^{x_2} \frac{\partial f}{\partial y'}\eta'(x)dx = \frac{\partial f}{\partial y'}\eta(x)\bigg|_{x_1}^{x_2} - \int_{x_1}^{x_2} \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\eta(x)dx\,. \tag{6.12}$$

However, since $\eta(x_{1,2}) = 0$ and assuming $\frac{\partial f}{\partial y'}$ is nonsingular at $x_{1,2}$ we have that the boundary term vanishes. Therefore, the stationary condition is

$$\int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right]\eta(x)\,dx = 0. \tag{6.13}$$

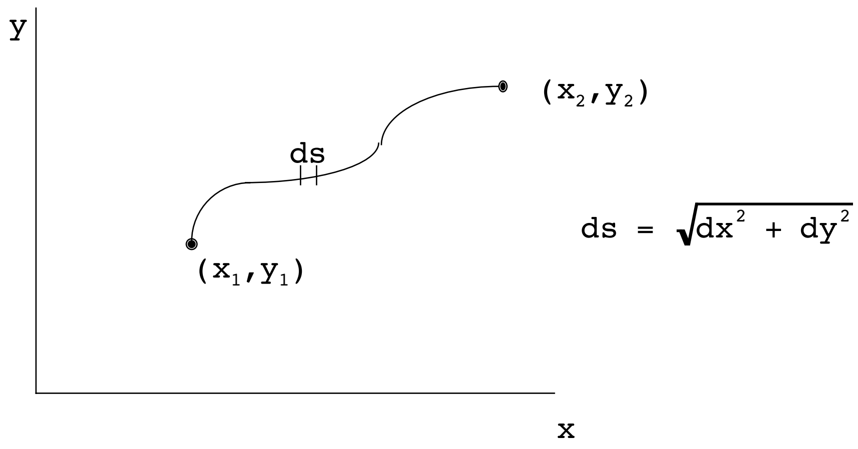
From our initial theorem, we now conclude that

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = 0. \tag{6.14}$$

This is the **Euler-Lagrange equation**. It represents only a necessary condition for an extremum.

Example 1

Consider the problem of finding the shortest distance between two points in a plane.



Solution

The distance along an infinitesimal distance ds along a presumed path can be written

$$ds = \sqrt{dx^2 + dy^2} = \pm\sqrt{1 + y'^2}\, dx\,.$$

We will choose the + sign to go along with the above picture where $x_2 > x_1$. The total length of the curve is

$$L = \int_{x1}^{x2} ds = \int_{x1}^{x2} \underbrace{\sqrt{1 + y'^2}}_{=f}\, dx\,.$$

Notice $\dfrac{\partial f}{\partial y} = 0\,,$, so the E-L $eq^{\underline{n}}$ is simply

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = 0\,,$$

or

$$\frac{\partial f}{\partial y'} = C_1\,,$$

where C_1 is a constant. Therefore

$$C_1^2 = \frac{y'^2}{1 + y'^2}\,,$$

$$\Rightarrow y' = \frac{\pm C_1}{\sqrt{1 - C_1^2}} \equiv C_2\,.$$

This is just the equation of a straight line of slope C_2

$$y = C_2 x + C_3 \ ,$$

where C_3 is another allowed constant. In order to find this specific line, set

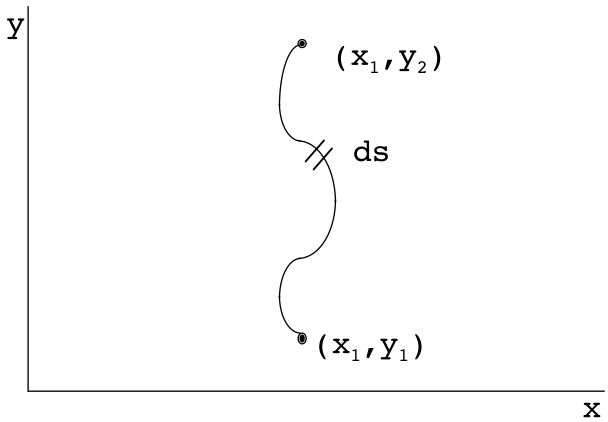
$$y_1 = C_2 x_1 + C_3 \ ,$$

$$y_2 = C_2 x_2 + C_3 \ ,$$

$$\Rightarrow y = \left(\frac{y_2 - y_1}{x_2 - x_1}\right) x + \frac{y_1 x_2 - y_2 x_1}{x_2 - x_1} .$$

Of course, the extremum corresponds to a minimum since there is no maximum distance. The choice of y as the dependent variable and x as the independent one above was completely arbitrary.

In the above example, you may have noticed a slight indiscretion. I overlooked the possibility of certain paths in formulating the integral. Thus, consider the problem of finding the shortest path for the situation:



Our formalism assumes that the stationary path is single-valued and can be written as $y = y(x)$. We can not characterize any path in this case in such a manner. One's first impulse is to simply reformulate the problem so that $x = x(y)$. However, a more general solution to the problem is to consider parametric representations of the paths, $x(t),\ y(t)$, where t is a parameter.[5] We will consider such path representations later, which we will see involve characterizing x and y as dependent variables and t as the independent variable.

Integrated Form of Euler's Equation

It often happens that $\frac{\partial f}{\partial x} = 0$. In this case there is an integrated form of the E-L equation that is easier to use. In applications to mechanics, this integrated form will yield the energy equation rather than Newton's equations. Consider

$$\frac{d}{dx}\left(y'\frac{\partial f}{\partial y'}\right) = y''\frac{\partial f}{\partial y'} + y'\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right). \tag{6.15}$$

However, by the definition of the total derivative

$$\frac{df}{dx} = \underbrace{\frac{\partial f}{\partial x}}_{0} + \frac{\partial f}{\partial y}y' + \frac{\partial f}{\partial y'}y''. \tag{6.16}$$

Thus

$$\frac{d}{dx}\left(y'\frac{\partial f}{\partial y'}\right) = \frac{df}{dx} + y'\left(-\frac{\partial f}{\partial y} + \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right). \tag{6.17}$$

However, if the E-L $\text{eq}^{\underline{n}}$s hold we have

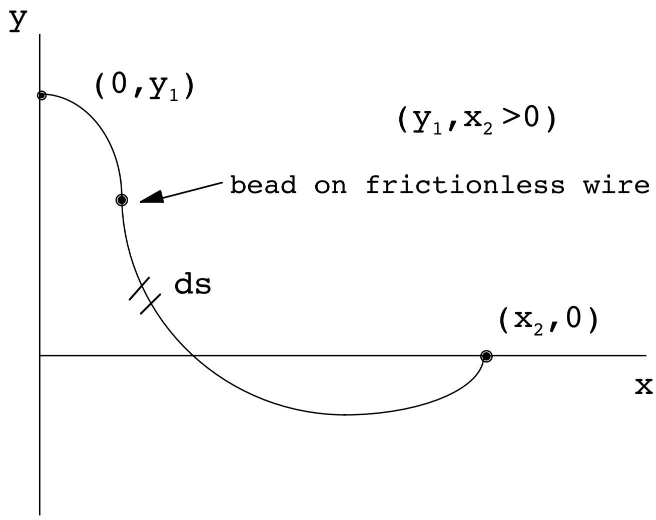
$$\frac{d}{dx}\left(y'\frac{\partial f}{\partial y'}\right) = \frac{df}{dx},$$

$$\Rightarrow \quad f - y'\frac{\partial f}{\partial y'} = \text{const.} \tag{6.18}$$

This yields a first order differential equation whereas the E-L equations give a second order form. The form of this equation when there is more than one dependent variable will be investigated in a HW problem.

Brachistochrone Problem

As an application of this, we consider a problem which originally gave rise to the calculus of variations, the **brachistochrone**. (Johann Bernoulli posed it in 1696; it was solved by Newton within 12 hours of his receiving it!)



We wish to find the curve that minimizes the time necessary for a bead attached to a wire to go from $(0,y_1)$ to $(x_2,0)$ under the influence of a constant gravitational field. First, we have to construct the integral we wish to vary.

$$dt = \frac{ds}{v}, \quad \leftarrow \text{mg}^{\underline{\text{n}}} \text{ of velocity}$$

$$\Rightarrow T \equiv \int \frac{ds}{v} = \int \frac{\sqrt{dx^2 + dy^2}}{v} = \int_0^{x_2} dx \frac{\sqrt{1 + y'^2}}{v}. \tag{6.19}$$

Conserve energy: $\frac{1}{2} mv^2 + mgy = mgy_1$,

$$\Rightarrow v = \sqrt{2g(y_1 - y)}, \tag{6.20}$$

$$\Rightarrow \quad T = \frac{1}{\sqrt{2g}} \int_0^{x_2} \underbrace{\left(\frac{1 + y'^2}{y_1 - y}\right)^{1/2}}_{=f[y,y';x]} dx. \tag{6.21}$$

Notice $\frac{\partial f}{\partial x} = 0$.. Use 2nd form of the E-L equations:

$$f - y' \frac{\partial f}{\partial y'} = C_1 ,$$

$$\Rightarrow \left(\frac{1 + y'^2}{y_1 - y}\right)^{1/2} - \frac{y'^2}{\left((y_1 - y)(1 + y'^2)\right)^{1/2}} = C_1$$

$$y'^2 = \frac{\frac{1}{C_1^2} - y_1 + y}{y_1 - y},$$

$$\Rightarrow y' = \pm \left(\frac{2a - y_1 + y}{y_1 - y}\right)^{1/2}. \qquad \left(2a \equiv \frac{1}{C_1^2}\right) \tag{6.22}$$

↑ +: positive slope, -: negative slope

Separate variables:

$$dx = \pm \left(\frac{y_1 - y}{2a - y_1 + y}\right)^{1/2} dy,$$

$$= \pm \frac{(y_1 - y)\, dy}{\sqrt{2a(y_1 - y) - (y_1 - y)^2}}. \tag{6.23}$$

Define $z \equiv y_1 - y$,

$$\Rightarrow x = \mp \int_0^{y_1 - y} \frac{z dz}{\sqrt{2az - z^2}}. \tag{6.24}$$

The top sign still implies a positive slope, the bottom means a negative slope on the trajectory, Notice the denominator is only real if

$$0 \le z \le 2a \ ,$$
$$\Rightarrow y_1 - 2a \le y \le y_1 . \tag{6.25}$$

Also notice that y can turn negative (go below the x-axis) if $2a > y_1$. Let us introduce

$$z \equiv a(1 - \cos\theta') , \tag{6.26}$$

which respects the above inequality in z. We then have

$$x = \mp a \int_0^{\theta} d\theta' \overbrace{\frac{\sin\theta'}{|\sin\theta'|}}^{\pm 1} (1 - \cos\theta') . \tag{6.27}$$

Notice the absolute value above and the fact that the upper limit satisfies $y_1 - y = a(1 - \cos\theta)$ If the upper limit, θ, is such that $\theta < \pi$, then $\sin\theta' > 0$ and we choose the root for the negative slope, $\frac{dy}{dx} < 0$ (and positive x):

$$x = a(\theta - \sin\theta). \qquad \theta > \pi$$

Note that if $\theta > \pi$ then $\sin\theta' > 0$ for $\pi < \theta' < \theta$. However, it is at $\theta' = \pi$ that an "undershoot" in the trajectory occurs (see later discussion) and the path must have positive slope, $\frac{dy}{dx} > 0$, which means we have two changes in sign. Thus, requiring $x > 0$ for all θ values results in the representation:

$$\left.\begin{aligned} z = y_1 - y = a(1 - \cos\theta) , \\ x = a(\theta - \sin\theta) . \end{aligned}\right\} \tag{6.28}$$

This is a parametric representation of the path. θ is just a convenient parameter in this problem; however, we will get a physical interpretation of it in a moment.

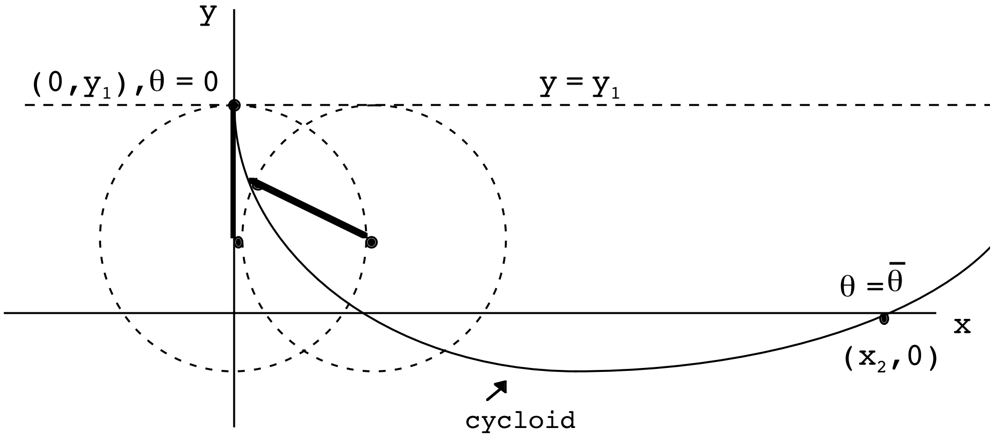
To build in the endpoints, we need to solve

$$\left.\begin{aligned} x_2 = a(\bar\theta - \sin\bar\theta) , \\ y_1 = a(1 - \cos\bar\theta) . \end{aligned}\right\} \tag{6.29}$$

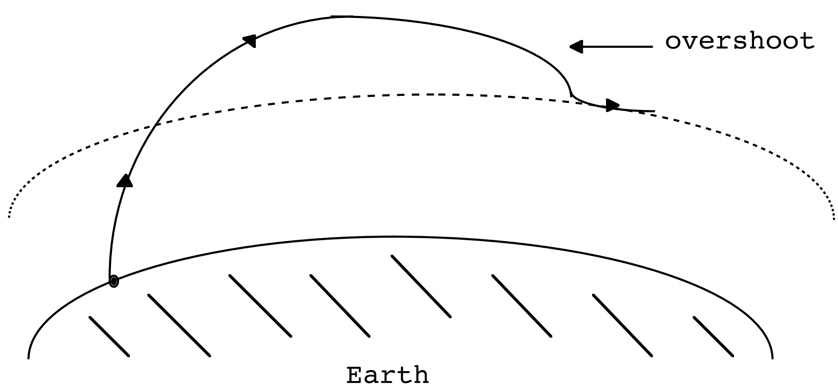
to find $\bar\theta$ (the final value of θ) and "a". One way to do this is to take the ratio,

$$\frac{x_2}{y_1} = \frac{\bar\theta - \sin\bar\theta}{1 - \cos\bar\theta} , \tag{6.30}$$

solving for $\bar{\theta}$ by numerical means and then substitute in either of Eqs.(6.29) to find "a" . Eqs.(6.28) are the equations of a **cycloid**, which is generated by rolling a hoop of radius a without slipping along the underside of the line $y = y_1$.



I mentioned above the possibility of an undershoot along the path, where $y<0$. This is rather nonintuitive behavior, given that we are looking at minimum time paths to travel from one given position to another. In a sense, we have traded a longer path length for a greater average velocity along the path caused by gravity, the extra velocity "kick" occurring on the initial part of the motion. The same sort of phenomenon happens to high thrust rockets launched from Earth into low circular orbits. Then optimal paths often produce an overshoot in the trajectory and actual path rises above the final desired orbital radius, the extra "kick" happening on the final part of the motion. In this case also we are trading a longer path length for a gravity "boost" (free from atmospheric drag) given by the downward inserion into orbit.



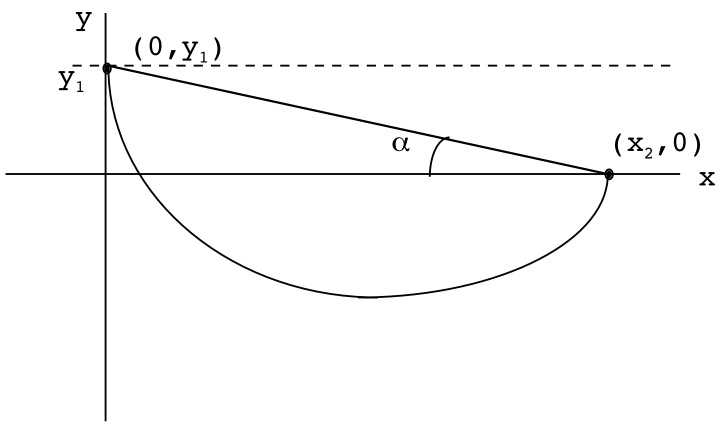
You will examine in a HW problem the precise condition under which an undershoot occurs in the brachistocrone problem. It may be helpful at this point to examine the undershoot cycloid path in more detail.

Example

Compare and contrast the time necessary to reach a point $x_2 >> y_1$ (causing an undershoot) two ways: a) a straight line path and b) a cycloid path.

Solution

The two paths look like the following.



On the straight line path, we have

$$x = \frac{1}{2} a_x t^2,$$

$$a_x = g \sin\alpha \cos\alpha,$$

so the total time is

$$T_s = \left(\frac{2x_2}{g \sin\alpha \cos\alpha}\right)^{1/2},$$

$$\sin\alpha = \frac{y_1}{\left(y_1^2 + x_2^2\right)^{1/2}}, \quad \cos\alpha = \frac{x_2}{\left(y_1^2 + x_2^2\right)^{1/2}}$$

$$T_s = \left(\frac{2\left(x_2^2 + y_1^2\right)}{gy_1}\right)^{1/2} \xrightarrow[x_2 >> y_1]{} x_2 \left(\frac{2}{gy_1}\right)^{1/2}.$$

The time necessary for the cycloid path is given by

$$T_C = \frac{1}{\sqrt{2g}} \int_0^{x_2} dx \left(\frac{1 + y'^2}{y_1 - y}\right)^{1/2}.$$

However

$$y'^2 = \left(\frac{2a - y_1 + y}{y_1 - y}\right),$$

$$\Rightarrow T_C = \sqrt{\frac{a}{g}} \int_0^{x_2} \frac{dx}{y_1 - y}.$$

From (6.28)

$$dx = a(1 - \cos\theta)d\theta,$$
$$(y_1 - y) = a(1 - \cos\theta),$$

$$\Rightarrow T_C = \sqrt{\frac{a}{g}} \int_0^{\bar{\theta}} d\theta.$$

From this we conclude that increments in the parameter θ and time, t, are related by $d\theta = \sqrt{\frac{g}{a}}\,dt$. Since $x_2 >> y_1$, we notice that the cycloid is almost complete, so $\bar{\theta} \simeq 2\pi$ and $x_2 \simeq 2\pi a$,

$$T_C \simeq 2\pi\sqrt{\frac{a}{g}} \simeq \sqrt{\frac{2\pi x_2}{g}}.$$

Thus, we have

$$\frac{T_s}{T_C} \underset{x_2>>y_1}{\longrightarrow} \left(\frac{x_2}{\pi y_1}\right)^{1/2} >> 1.$$

and the cycloid path is clearly superior. Thus, although the undershoot path goes "out of it's way" in dipping vertically, it acquires a larger average x-component of velocity by the extra fall.

The Case of More than One Dependent Variable

Often problems have more than one dependent variable. We can characterize such integral problems in the following form.

$$J = \int_{x_1}^{x_2} f\left[\underbrace{y_1(x), y_2(x), \ldots}_{n \text{ variables}}, y_1'(x), y_2'(x), \ldots; x\right] dx. \tag{6.31}$$

One constructs a set of comparison functions $(i = 1, \ldots, n)$

$$\left.\begin{aligned} y_i(\alpha, x) &\equiv y_i(x) + \alpha\eta_i(x), \\ \eta_i(x_1) &= \eta_i(x_2) = 0. \end{aligned}\right\} \tag{6.32}$$

We proceed as before using the chain rule:

$$\left.\frac{dJ}{d\alpha}\right|_{\alpha=0} = \int_{x_1}^{x_2} dx \sum_{i=1}^{n} \left[\frac{\partial f}{\partial y_i} \underbrace{\frac{\partial y_i}{\partial \alpha}}_{\eta_i(x)} + \frac{\partial f}{\partial y_i'} \underbrace{\frac{\partial y_i'}{\partial \alpha}}_{\eta_i'(x)} \right]. \tag{6.33}$$

Do the integration by parts in the second set of terms:

$$\left.\frac{dJ}{d\alpha}\right|_{\alpha=0} = \int_{x1}^{x2} dx \sum_{i=1}^{n} \left(\frac{\partial f}{\partial y_i} - \frac{d}{dx} \frac{\partial f}{\partial y_i'} \right) \eta_i(x). \tag{6.34}$$

Given that the $\eta_i(x)$ represent n independent arbitrary functions, one can use reasoning similar to the above to conclude,

$$\frac{\partial f}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial f}{\partial y_i'} \right) = 0 \quad , \quad i = 1, \ldots, n. \tag{6.35}$$

Example

Find the shortest path between two points on a plane again using both x and y as dependent variables.

Solution:

We reformulate the problem as

$$ds = \sqrt{\dot{x}^2 + \dot{y}^2}\, dt\,,$$

where $\dot{x} = \frac{dx}{dt}$ and "t" is a parametric variable like θ in the brachistochrone problem. It is not related to time. (There are no dynamics here.)

$$L = \int_{t_1}^{t_2} \underbrace{\sqrt{\dot{x}^2 + \dot{y}^2}}_{=f[\dot{x},\dot{y};t]}\, dt\,.$$

The E-L $eq^{\underline{n}}s$ are

$$\frac{\partial f}{\partial x} - \frac{d}{dt}\left(\frac{\partial f}{\partial \dot{x}}\right) = 0 \Rightarrow \frac{d}{dt}\left(\frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2}}\right) = 0\,,$$

$$\frac{\partial f}{\partial y} - \frac{d}{dt}\left(\frac{\partial f}{\partial \dot{y}}\right) = 0 \Rightarrow \frac{d}{dt}\left(\frac{\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2}}\right) = 0\,.$$

From the above

$$\frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2}} = C_1\,,$$

$$\frac{\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2}} = C_2\,,$$

where $C_{1,2}$ are constants. From combining them we find that $C_1{}^2 + C_2{}^2 = 1$ and therefore that

$$x = (\cos\theta)t + C_3\,,$$

$$y = (\sin\theta)t + C_4\,,$$

is the general parametric solution. This way of doing the problem is as easy or easier than our first solution. By taking $\theta = \frac{\pi}{2}$ we see that the solution mentioned earlier, $x = x_0$, is indeed an extremum path.

The Case of More than One Independent Variable

One can encounter a problem also with several **independent** degrees of freedom:

$$f = f\left[y\,(x_1, x_2), \frac{\partial y}{\partial x_1}, \frac{\partial y}{\partial x_2}; x_1, x_2\right]. \tag{6.36}$$

↑ ↑

partial derivatives

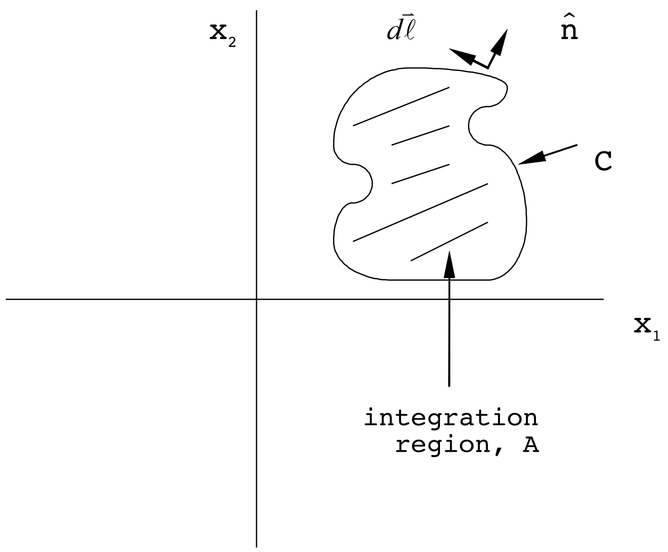
Here, I am showing two independent degrees of freedom, but the argument will be general. Then

$$J = \iint_A dx_1\, dx_2\, f[\ldots]. \tag{6.37}$$

Now introduce

$$y(\alpha, x_1, x_2) \equiv y(x_1, x_2) + \alpha\eta(x_1, x_2), \tag{6.38}$$

where $\eta(x_1,x_2) = 0$ everywhere on the boundary C:



As usual

$$\left.\frac{dJ}{d\alpha}\right|_{\alpha=0} = \iint_A dx_1\, dx_2 \left[\frac{\partial f}{\partial y}\underbrace{\frac{\partial y}{\partial \alpha}}_{=\eta} + \sum_{i=1}^{2} \frac{\partial f}{\partial\left(\frac{\partial y}{\partial x_i}\right)} \underbrace{\frac{\partial\left(\frac{\partial y}{\partial x_i}\right)}{\partial \alpha}}_{=\frac{\partial \eta}{\partial x_i}}\right]. \tag{6.39}$$

This can also be written more compactly as

$$\frac{dJ}{d\alpha}\bigg|_{\alpha=0} = \iint_A dx_1\, dx_2 \left[\frac{\partial f}{\partial y}\eta + \vec{v}\cdot\vec{\nabla}\,\eta\right], \tag{6.40}$$

where $v_1 \equiv \dfrac{\partial f}{\partial\left(\dfrac{\partial y}{\partial x_1}\right)}$, $v_2 \equiv \dfrac{\partial f}{\partial\left(\dfrac{\partial y}{\partial x_2}\right)}$ and $\vec{\nabla}$ is the gradient operator from Ch.1. The two-dimensional form of the Gauss law is

$$\int_A dx_1\, dx_2\; \vec{\nabla}\cdot\vec{v} = \oint_C d\ell\, \vec{v}\cdot\hat{n}, \tag{6.41}$$

where $\hat{n}$ is the outward unit normal and $d\ell$ is an infinitesimal element of the perimeter (see the above figure; this is actually equivalent to Stokes' theorem in two dimensions). This gives the two-dimensional form of integration by parts if we let $\vec{v} \rightarrow \eta\,\vec{v}$,

$$\int_A dx_1\, dx_2\; \vec{v}\cdot\vec{\nabla}\,\eta = -\int_A dx_1\, dx_2\; \eta\,\vec{\nabla}\cdot\vec{v} + \oint_C d\ell\; \eta\,\vec{v}\cdot\hat{n}. \tag{6.41A}$$

Apply (6.41A) to the second term in (6.40) to obtain

$$\frac{dJ}{d\alpha}\bigg|_{\alpha=0} = \iint_A dx_1\, dx_2 \left[\frac{\partial f}{\partial y}\eta - \eta\;\vec{\nabla}\cdot\vec{v}\right], \tag{6.40A}$$

since the function $\eta(x_1,x_2)=0$ everywhere on C. More explicitly,

$$\frac{dJ}{d\alpha}\bigg|_{\alpha=0} = \iint_A dx_1\, dx_2 \left[\frac{\partial f}{\partial y} - \sum_{i=1}^{2}\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial\left(\dfrac{\partial y}{\partial x_i}\right)}\right)\right]\eta. \tag{6.42}$$

Since $\eta(x_1,x_2)$ is arbitrary in A, by setting $\dfrac{dJ}{d\alpha}\bigg|_{\alpha=0} = 0$ and with a suitable generalization of our initial theorem we may conclude

$$\frac{\partial f}{\partial y} - \sum_{i=1}^{2}\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial\left(\dfrac{\partial y}{\partial x_i}\right)}\right) = 0. \tag{6.43}$$

For n independent variables, this argument easily extends to show the upper limit in the second term is n.

Example

Extremize the integral,

$$J = \iint_A dx_1\, dx_2 \underbrace{\left[\left(\frac{\partial y}{\partial x_1}\right)^2 + \left(\frac{\partial y}{\partial x_2}\right)^2\right]}_{=f\left[y,\frac{\partial y}{\partial x_1},\frac{\partial y}{\partial x_2};x_1,x_2\right]}.$$

Solution:

The E-L equation for one independent and two dependent degrees of freedom reads

$$\frac{\partial f}{\partial y} - \frac{\partial}{\partial x_1}\frac{\partial f}{\partial\left(\frac{\partial y}{\partial x_1}\right)} - \frac{\partial}{\partial x_2}\frac{\partial f}{\partial\left(\frac{\partial y}{\partial x_2}\right)} = 0.$$

We have

$$\frac{\partial f}{\partial y} = 0, \qquad \frac{\partial f}{\partial\left(\frac{\partial y}{\partial x_1}\right)} = 2\left(\frac{\partial y}{\partial x_1}\right), \quad \frac{\partial f}{\partial\left(\frac{\partial y}{\partial x_2}\right)} = 2\left(\frac{\partial y}{\partial x_2}\right).$$

We obtain

$$\frac{\partial^2 y}{\partial x_1^2} + \frac{\partial^2 y}{\partial x_2^2} = 0.$$

Notice these terms form the two dimensional Laplacian, so

$$\nabla^2 y = 0,$$

is the E-L equation.

Obviously, one can have cases where a combination of more than one independent as well as dependent degrees of freedom occur in a problem. Then

$$f = f\left[\underbrace{y_1, y_2, \ldots,}_{\text{n dep. variables}} \underbrace{\underbrace{\frac{\partial y_1}{\partial x_1}, \frac{\partial y_1}{\partial x_2}, \ldots,}_{\text{p of them}} \underbrace{\frac{\partial y_2}{\partial x_1}, \frac{\partial y_2}{\partial x_2}, \ldots}_{\text{p of them}}}_{\text{n of these groupings}}; \underbrace{x_1, x_2, \ldots}_{\text{p indep. variables}}\right].$$

In the case of n dependent and p independent degrees of freedom, the E-L equations are simply

$$\frac{\partial f}{\partial y_i} - \sum_{j=1}^{p} \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial\left(\frac{\partial y_i}{\partial x_j}\right)}\right) = 0, \quad i=1,2,...,n. \tag{6.44}$$

In the above, "i" is a free index and "j" is a dummy (summed) index. In applications to mechanics, there is usually only a single independent variable, the time.

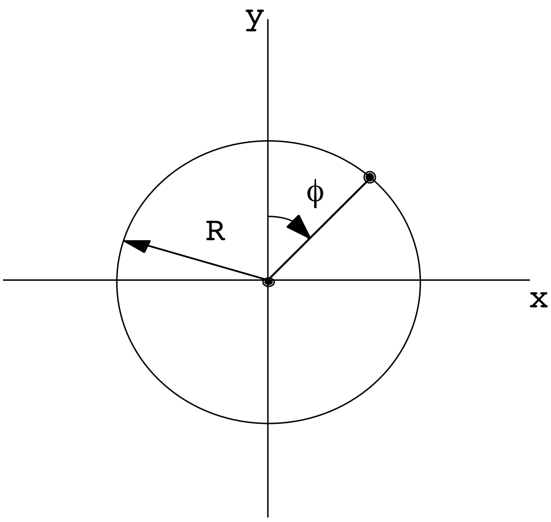
Constraints

It often happens that we wish to minimize the value of some integral in a situation where there are more than one dependent degrees of freedom, which, however can not vary independently. In other words, there is some **constraint** connecting variations of two (or more) dependent variables that appear in f. There are three ways of dealing with this situation:

1. Algebraically eliminate one variable using the constraint condition.
2. Reformulate the problem so that the variations in the dependent variables become truly independent.
3. Use Lagrange multipliers.

Situation 1 occurs when one can actually solve the equation of constraint (when it's linear, for example) and eliminate one or more variables. Thus, the constraint is rather trivially removed.

Situation 2 usually means to change to a new coordinate systems. An example would be the motion of a bead on a hoop.

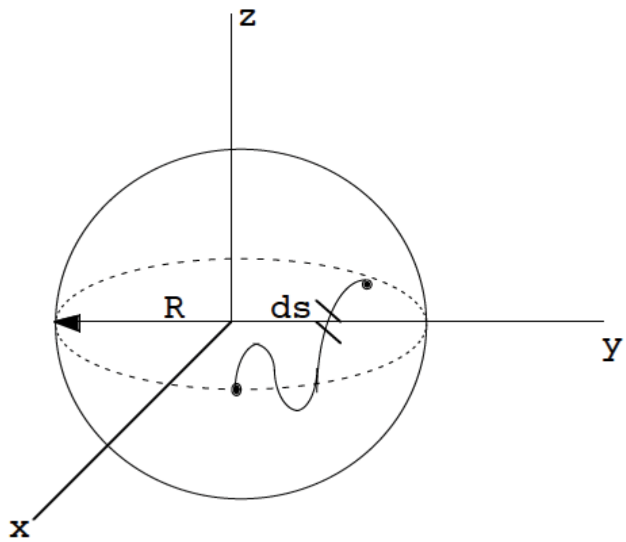


Regarding x and y as dependent variables, it would be foolish to use a rectangular coordinate system since

$$x^2 + y^2 = R^2 ,$$

and variations in x and y are not independent. It would be far superior to change to cylindrical coordinates. Then the only dependent variable is the angular coordinate ϕ, for which there is no constraint. (It is understood that the independent variable in this situation is t, the time.)

As an example of situation 2, in the case of geometry rather than dynamics, let us find the shortest distance between two points on a sphere.



In general it takes 3 coordinates to locate a point in 3 dimensions, but since our point is given to be on the surface of a sphere of radius R, we will need only 2 coordinates, which we will take to be the spherical coordinate angles θ and ϕ. The constraint has been eliminated.

The length element in these coordinates is

$$ds = R\left(d\theta^2 + \sin^2\theta d\phi^2\right)^{1/2} . \tag{6.45}$$

Without loss of generality we can write

$$ds = R\left[\theta'^2 + \sin^2\theta\right]^{1/2} |d\varphi| , \tag{6.46}$$

where $\theta' = \dfrac{d\theta}{d\phi}$. ϕ is now the independent variable, θ is the dependent. Identify

$$f[\theta, \theta' ; \phi] = \left(\theta'^2 + \sin^2\theta\right)^{1/2} . \tag{6.47}$$

Since $\dfrac{\partial f}{\partial \phi} = 0$, we may use the second (integrated) form of the E-L equations:

$$\left(\theta'^2 + \sin^2\theta\right)^{1/2} - \theta'\left(\frac{\theta'}{\left(\theta'^2 + \sin^2\theta\right)^{1/2}}\right) = a ,$$

$$\Rightarrow \sin^2\theta = a\left(\theta'^2 + \sin^2\theta\right)^{1/2} ,$$

$$\Rightarrow \left(\frac{d\phi}{d\theta}\right)^2 = \frac{a^2}{\sin^2\theta\left(\sin^2\theta - a^2\right)} ,$$

$$\Rightarrow d\phi = \pm\frac{d\theta}{\sin\theta\sqrt{\frac{1}{a^2}\sin^2\theta - 1}} . \tag{6.48}$$

We will carry the $\pm$ sign with us to see the effect. We can now write

$$d\phi = \pm \frac{\dfrac{1}{\sin^2\theta}d\theta}{\sqrt{\dfrac{1}{a^2} - \dfrac{1}{\sin^2\theta}}} = \pm \frac{\csc^2\theta d\theta}{\sqrt{\left(\dfrac{1}{a^2} - 1\right) - \cot^2\theta}},$$

$\uparrow$
$\csc^2\theta = 1 + \cot^2\theta$

$$\Rightarrow d\phi = \pm \frac{\beta^{-1}\dfrac{d}{d\theta}\cot\theta}{\sqrt{1 - \beta^{-2}\cot^2\theta}} = \pm\left[\frac{d}{d\theta}\sin^{-1}\left(\frac{\cot\theta}{\beta}\right)\right]d\theta. \tag{6.49}$$

$\uparrow$
$$\frac{d\cot\theta}{d\theta} = \csc^2\theta,\ \beta \equiv \sqrt{\frac{1}{a^2} - 1}.$$

In this form we see that the right hand side is a perfect differential, and thus

$$\phi = \pm\sin^{-1}\left(\frac{\cot\theta}{\beta}\right) + b, \tag{6.50}$$

where "b" is another constant. Notice that there are now two constants, b and β, available to fit two points on the sphere. To make this result look more familiar, take the sine of both sides and use

$$\sin(\phi - b) = \sin\phi\cos b - \sin b\cos\phi,$$

$$\Rightarrow \sin\phi\cos b - \sin b\cos\phi = \pm\frac{\cot\theta}{\beta},$$

$$\Rightarrow R\sin\theta(\sin\phi\cos b - \sin b\cos\phi) = \pm\frac{R\cos\theta}{\beta}.$$

Identify
$$R\cos\theta = z,$$
$$R\sin\theta\cos\phi = x,$$
$$R\sin\theta\sin\phi = y,$$

$$\Rightarrow y\cos b - x\sin b = \pm\frac{z}{\beta}, \tag{6.51}$$

which is the equation of a plane passing through the origin $(x,y,z = 0)$. The sign on the right of (6.51) may be absorbed into a redefinition of β and adds no extra generality in defining the plane. The intersection of the plane and sphere forms a "great circle" connecting the two points on the sphere. There is still a choice of path (long way or short way) which connects the two arbitrary points along this great circle. This is a case where both extreme values are uncovered by the variational principle.

Lagrange Multipliers

Let me begin to explain what is meant by the use of Lagrange multipliers, as in situation 3 above. I'll start in a context of functions rather than integrals. Let us say we wish to extremize a function of more than one variable, such as $f(A,B)$, but subject to some constraint connecting the variables A and B so that their variations are not independent. We assume the constraint can be written as (called "holonomic"; we'll see more of this next chapter)

$$g(A,B) = 0. \tag{6.52}$$

Examples:

1. $A^2 + B^2 = R^2 \Rightarrow g(A,B) = A^2 + B^2 - R^2$

2. $A = CB + D \Rightarrow (A,B) = A - CB - D$

The conventional approach to this problem is to try to eliminate one variable in terms of the other, as in

$$g(A,B) = 0 \Rightarrow f(A,B(A)),$$

and then apply the zero slope necessary condition

$$\frac{df}{dA}\bigg|_{A_0} = 0.$$

The Lagrange method avoids the explicit elimination of variables by using the constraint equation as a subsidiary condition. In the following, it is understood that all variations are being evaluated at the point or points of extremum.

The total variation of f is given by the chain rule,

$$df = \left(\frac{\partial f}{\partial A}\right) dA + \left(\frac{\partial f}{\partial B}\right) dB,$$

$$= \left[\frac{\partial f}{\partial A} + \frac{\partial f}{\partial B}\frac{dB}{dA}\right] dA. \tag{6.53}$$

The extremum condition is

$$df = 0,$$

$$\Rightarrow \frac{\partial f}{\partial A} = -\frac{\partial f}{\partial B}\frac{dB}{dA}.$$

Now from $g(A,B) = 0$ we have

$$0 = dg = \left(\frac{\partial g}{\partial A}\right) dA + \left(\frac{\partial g}{\partial B}\right) dB,$$

$$\Rightarrow \frac{dB}{dA} = -\left(\frac{\partial g}{\partial A}\right)\left(\frac{\partial g}{\partial B}\right)^{-1},$$

$$\Rightarrow \frac{\partial f}{\partial A} = -\frac{\partial f}{\partial B}\left(-\left(\frac{\partial g}{\partial A}\right)\left(\frac{\partial g}{\partial B}\right)^{-1}\right),$$

$$\Rightarrow \frac{\partial f}{\partial A}\left(\frac{\partial g}{\partial A}\right)^{-1}\bigg|_{A_0,B_0} = \frac{\partial f}{\partial B}\left(\frac{\partial g}{\partial B}\right)^{-1}\bigg|_{A_0,B_0} = -\lambda. \tag{6.54}$$

Again, the variations are being carried out at the extremum, and to emphasize this I am indicating this point explicitly as A_0, B_0 in (6.54). I have also made the conventional definition of the Lagrange multiplier, λ, in a variable-symmetric way in the above. It is just a constant. Writing out the two equations contained in (6.54) separately and explicitly stating the constraint, viewed as a subsidiary condition, we have

$$\left.\begin{aligned} \frac{\partial f}{\partial A} + \lambda\frac{\partial g}{\partial A} &= 0, \\ \frac{\partial f}{\partial B} + \lambda\frac{\partial g}{\partial B} &= 0, \\ g(A,B) &= 0, \end{aligned}\right\} \tag{6.55}$$

which is a complete characterization of the problem. Notice we can write these equations a bit more elegantly if we define

$$f^* \equiv f + \lambda g. \tag{6.56}$$

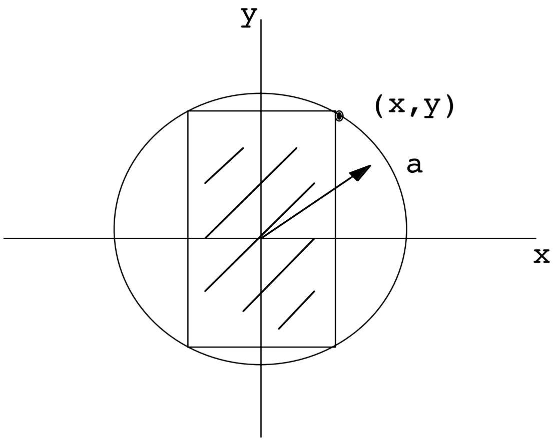
Then the conditions

$$\left.\begin{aligned} \frac{\partial f^*}{\partial A} &= 0, \\ \frac{\partial f^*}{\partial B} &= 0, \\ g(A,B) &= 0. \end{aligned}\right\} \tag{6.57}$$

look like a normal, unconstrained, extremization problem, except for the last explicit constraint.

Example

Find the rectangle of greatest area which can be inscribed in a circle of radius "a."



Solution:

The area is 4xy and the constraint is $x^2 + y^2 = a^2$. Therefore

$$f(x,y) = 4xy,$$

$$g(x,y) = x^2 + y^2 - a^2 = 0.$$

We have

$$\frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x} = 4y + 2\lambda x = 0, \tag{E.1}$$

$$\frac{\partial f}{\partial y} + \lambda \frac{\partial g}{\partial y} = 4x + 2\lambda y = 0. \tag{E.2}$$

$$x^2 + y^2 = a^2. \tag{E.3}$$

Combining (E.1) and (E.2), we find $x^2 = y^2$. Then using (E.3), we obtain $x = y = \frac{a}{\sqrt{2}}$. This just gives a square.

This method can be generalized to m variables, x_i, and n constraints as

$$\frac{\partial f^*(x_1, \ldots, x_m)}{\partial x_i} = 0 \quad , \quad i = 1, \ldots, m \tag{6.58}$$

$$g_j(x_1, x_2, \ldots, x_m) = 0 \quad , \quad j = 1, \ldots, n \tag{6.59}$$

where $f^* = f(x_1, x_2, \ldots, x_m) + \sum_{k=1}^{n} \lambda_k g_k(x_1, \ldots, x_m)$.

How does this general method apply to the situation with integrals and functionals rather than functions and variables? Let us consider the simplest case of 2 dependent and one independent variable (like the x,y position of a particle at time t). We have

$$f = f\left[y_1, y_2, y_1', y_2'; x\right].$$

As usual,

$$\left.\begin{aligned} y_1(\alpha, x) &= y_1(x) + \alpha\eta_1(x), \\ y_2(\alpha, x) &= y_2(x) + \alpha\eta_2(x), \\ \eta_1(x_{1,2}) &= \eta_2(x_{1,2}) = 0. \end{aligned}\right\} \tag{6.60}$$

Then, skipping some hopefully familiar steps, we have

$$\left.\frac{dJ}{d\alpha}\right|_{\alpha=0} = \int_{x_1}^{x_2} dx \left[\left(\frac{\partial f}{\partial y_1} - \frac{d}{dx}\left(\frac{\partial f}{\partial y_1'}\right)\right)\eta_1(x) + \left(\frac{\partial f}{\partial y_2} - \frac{d}{dx}\left(\frac{\partial f}{\partial y_2'}\right)\right)\eta_2(x)\right]. \tag{6.61}$$

But now assume that the variations

$$\frac{\partial y_1}{\partial \alpha} = \eta_1 \;, \quad \frac{\partial y_2}{\partial \alpha} = \eta_2 \;,$$

are no longer independent because of a constraint of the form

$$g(y_1, y_2; x) = 0 \;. \tag{6.62}$$

However, we can write

$$\left.\frac{dJ}{d\alpha}\right|_{\alpha=0} = \int_{x_1}^{x_2} dx \left[\left(\frac{\partial f}{\partial y_1} - \frac{d}{dx}\left(\frac{\partial f}{\partial y_1'}\right)\right) + \left(\frac{\partial f}{\partial y_2} - \frac{d}{dx}\left(\frac{\partial f}{\partial y_2'}\right)\right)\frac{\eta_2}{\eta_1}\right]\eta_1(x). \tag{6.63}$$

Requiring this to vanish and since $\eta_1(x)$ is arbitrary, we have

$$\frac{\partial f}{\partial y_1} - \frac{d}{dx}\left(\frac{\partial f}{\partial y_1'}\right) = -\left(\frac{\partial f}{\partial y_2} - \frac{d}{dx}\left(\frac{\partial f}{\partial y_2'}\right)\right)\frac{\eta_2(x)}{\eta_1(x)}.$$

Since $g(y_1(\alpha,x), y_2(\alpha,x); x) = 0$, we have

$$\frac{dg}{d\alpha} = \frac{\partial g}{\partial y_1}\underbrace{\frac{\partial y_1}{\partial \alpha}}_{=\eta_1} + \frac{\partial g}{\partial y_2}\underbrace{\frac{\partial y_2}{\partial \alpha}}_{=\eta_2} = 0,$$

$$\Rightarrow \quad \frac{\eta_2(x)}{\eta_1(x)} = -\left(\frac{\partial g}{\partial y_1}\right)\left(\frac{\partial g}{\partial y_2}\right)^{-1}. \tag{6.64}$$

Used above to eliminate this ratio, we find

$$\left[\frac{\partial f}{\partial y_1} - \frac{d}{dx}\left(\frac{\partial f}{\partial y_1'}\right)\right]\left(\frac{\partial g}{\partial y_1}\right)^{-1} = \left[\frac{\partial f}{\partial y_2} - \frac{d}{dx}\left(\frac{\partial f}{\partial y_2'}\right)\right]\left(\frac{\partial g}{\partial y_2}\right)^{-1}. \tag{6.65}$$

Set both sides of (6.65) equal to $-\lambda(x)$, an unknown function of x. Our three equations, representing a complete characterization of the extremum problem, are now

$$\left.\begin{aligned} &\frac{\partial f}{\partial y_1} - \frac{d}{dx}\left(\frac{\partial f}{\partial y_1'}\right) + \lambda(x)\frac{\partial g}{\partial y_1} = 0, \\ &\frac{\partial f}{\partial y_2} - \frac{d}{dx}\left(\frac{\partial f}{\partial y_2'}\right) + \lambda(x)\frac{\partial g}{\partial y_2} = 0, \\ &g(y_1, y_2; x) = 0. \end{aligned}\right\} \tag{6.66}$$

This represents 3 equations in 3 unknowns: $y_1, y_2,$ and λ. Again, we may simplify these to

$$\left.\begin{aligned} &\frac{\partial f^*}{\partial y_1} - \frac{d}{dx}\left(\frac{\partial f^*}{\partial y_1'}\right) = 0 \,, \\ &\frac{\partial f^*}{\partial y_2} - \frac{d}{dx}\left(\frac{\partial f^*}{\partial y_2'}\right) = 0 \,, \\ &g(y_1, y_2; x) = 0 \,, \end{aligned}\right\} \tag{6.67}$$

where $f^* = f + \lambda g$. The general case of m dependent variables, $y_i(x)$, and n constraints can be written as

$$\frac{\partial f^*}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial f^*}{\partial y_i'}\right) = 0 \quad , \quad i = 1, \ldots, m \tag{6.68}$$

$$g_j(y_1, \ldots, y_m; x) = 0 \quad , \quad j = 1, \ldots, n \,, \tag{6.69}$$

where

$$f^* \equiv f + \sum_{k=1}^{n} \lambda_k(x)\, g_k \,. \tag{6.70}$$

The above represent $m + n$ equations in $m + n$ unknowns.

We will see in the next Chapter that the generalized E-L equations, with an appropriate choice of f, are equivalent to Newton's force equations on a particle. Therefore, it is not surprising to learn that the Lagrange multipliers, $\lambda_k(x)$, often have the physical interpretation as the forces of constraint. We will leave examples of the use of these equations to the next chapter.

Isoperimetric Problems

I can not leave the subject of the calculus of variations without mention of the subject of so-called isoperimetric problems, even though we will not be using this in the Chapters which follow. Isoperimetric problems are also constraint problems, except the constraint is in the form of another, fixed, integral rather than a relation between dependent variables. In it's simplest form, the problem is to find an extremum of

$$J = \int_{x_1}^{x_2} dx\, f[y, y'; x], \tag{6.71}$$

subject to the integral

$$L = \int_{x_1}^{x_2} dx\, g\left[y, y'; x\right], \tag{6.72}$$

having a fixed value. Typically, L represents a perimeter on a given curve. Here, the use of Lagrange multipliers is also useful. The formalism is quite similar to the use of Lagrange multipliers in the function/variable case studied earlier.

Let us introduce a path varied about the extremum with independent arbitrary functions $\eta(x)$ and $\zeta(x)$ using α_1, and α_2 as parameters. Then define

$$y(\alpha_1, \alpha_2, x) \equiv y(x) + \alpha_1\, \eta(x) + \alpha_2\, \zeta(x), \tag{6.73}$$

where at the endpoints

$$\eta(x_{1,2}) = \zeta(x_{1,2}) = 0, \tag{6.74}$$

as usual. Then J and L become functions of α_1 and α_2,

$$J \to J(\alpha_1, \alpha_2),$$

$$L \to L(\alpha_1, \alpha_2),$$

and the sought extremum is at $\alpha_1 = \alpha_2 = 0$. At this point, the formalism becomes identical to that studied before, with a function of two variables connected by a holonomic constraint between the variables. We can follow the same reasoning as before to arrive at

$$\frac{\partial J}{\partial \alpha_1}\left(\frac{\partial L}{\partial \alpha_1}\right)^{-1}\Bigg|_{\alpha_{1,2}=0} = \frac{\partial J}{\partial \alpha_2}\left(\frac{\partial L}{\partial \alpha_2}\right)^{-1}\Bigg|_{\alpha_{1,2}=0} \equiv -\lambda, \tag{6.75}$$

where I have defined the Lagrange multiplier, λ, which is again a (unknown) constant. Writing part of (6.75) out in full tells us (the other part gives the same information)

$$\frac{\partial J}{\partial \alpha_1} + \lambda \frac{\partial L}{\partial \alpha_1} = 0, \tag{6.76}$$

$$\Rightarrow \int_{x_1}^{x_2} dx \left[\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right) + \lambda\left(\frac{\partial g}{\partial y} - \frac{d}{dx}\left(\frac{\partial g}{\partial y'}\right)\right)\right]\eta(x) = 0. \tag{6.77}$$

Since $\eta(x)$ is arbitrary, we conclude that

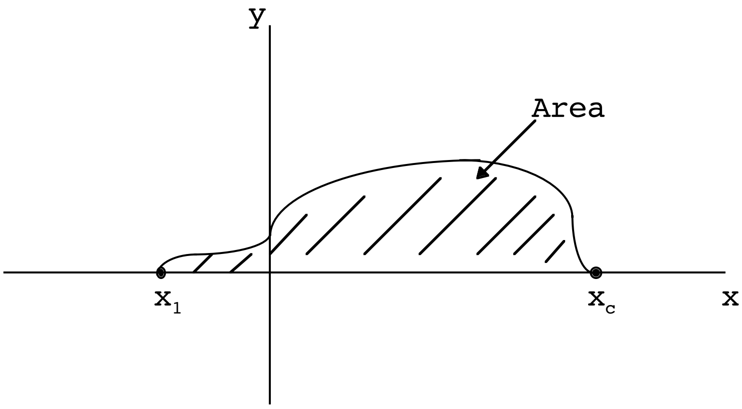
$$\left.\begin{aligned} \frac{\partial f^*}{\partial y} - \frac{d}{dx}\left(\frac{\partial f^*}{\partial y'}\right) = 0, \\ f^* \equiv f + \lambda g. \end{aligned}\right\} \tag{6.78}$$

After $y(x)$ is found, the condition $L = \int_{x_1}^{x_2} dx\, g$ then fixes one of the parameters in the solution. There is an interesting duality in such problems in that if we extremize L and fix the value of J, reversing the roles of f and g we would get the exact same equations as above and thus the same solution.

Thus the constraint in isoperimetric problems is built in simply by adding to f the constraint condition multiplied by λ, and varying f^* in the E-L equations rather than f. The general rule we are seeing is: to build in a constraint, add the constraint condition, multiplied by an unknown multiplier, and vary the result as if there were no constraint!

Example

Find the largest area enclosed between the x-axis and a string of fixed length attached at it's ends to points x_1 and x_2 on the x-axis. ("Dido's problem")



Solution:

We have for the area and the constraint, respectively,

$$J = \int_{x_1}^{x_2} dx\, y\,,$$

$$L = \int_{x_1}^{x_2} dx \sqrt{1 + y'^2}\,.$$

We identify

$$f^* = y + \lambda\sqrt{1 + y'^2}\,,$$

and we write

$$\frac{\partial}{\partial y}\left(y + \lambda\sqrt{1 + y'^2}\right) - \frac{d}{dx}\frac{\partial}{\partial y'}\left(y + \lambda\sqrt{1 + y'^2}\right) = 0\,,$$

$$\Rightarrow \qquad \frac{d}{dx}\left(\frac{y'}{\sqrt{1 + y'^2}}\right) = \frac{1}{\lambda}\,. \qquad \text{(E.4)}$$

(One can also use the integrated form of the E-L equations here since $\frac{\partial f^*}{\partial x} = 0$, but the work necessary either way is almost the same.) Integrating (E.4), we find

$$\frac{y'}{\sqrt{1 + y'^2}} = \frac{x}{\lambda} + C\,,$$

where C is a constant of integration. Now, solving for y', one finds

$$y' = \pm\frac{\frac{x}{\lambda} + C}{\sqrt{1 - \left(\frac{x}{\lambda} + C\right)^2}}\,,$$

where a ± sign from taking the square root is being kept. Separating variables and introducing

$$z \equiv \frac{x}{\lambda} + C\,,$$

we find

$$y = \pm\lambda\int_0^{\frac{x}{\lambda}+C}\frac{z\,dz}{\sqrt{1 - z^2}}\,,$$

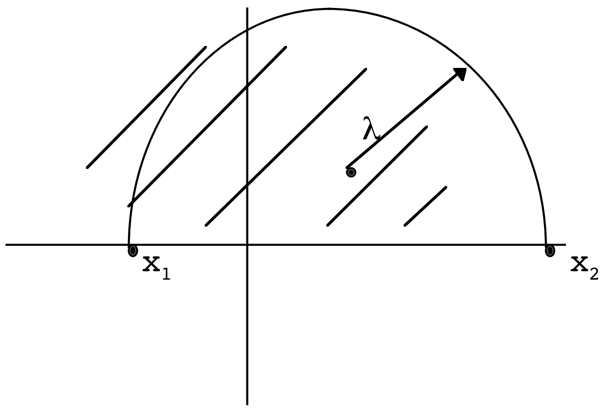
which integrates immediately to

$$y = \pm\lambda\sqrt{1 - \left(\frac{x}{\lambda} + C\right)^2} + K\,,$$

where K is another constant of integration. This can be rearranged, after squaring, to read

$$(y - K)^2 + (x + C\lambda)^2 = \lambda^2\,.$$

Thus, the solution is always a section of a circle. In the case above, it might look like



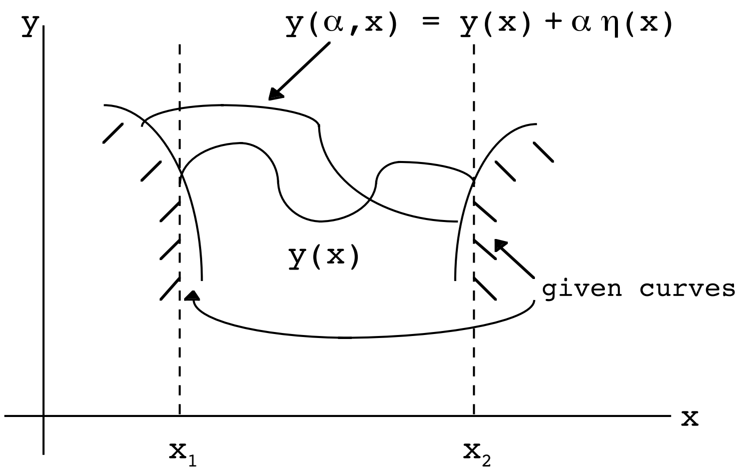
The three constants in the solution, K, C and λ, are fixed by x_1, x_2 and the length of the string, L. Obviously, in this case our extremum corresponds to a maximum. Also, obviously, there is no solution to the problem when $L < |x_2 - x_1|$!

Variation of the End Points of Integration

The last thing I want to discuss in this Chapter is the subject of varying endpoints in our integrals. Suppose the limits of the integral

$$J = \int_{x_1}^{x_2} dx\, f[y, y'; x], \tag{6.79}$$

were not fixed, but were allowed to vary as well as the "interior" points, as shown.



Obviously, $\eta(x)$ no longer vanishes at the endpoints. Although this seems like an unlikely situation, we will see in the next Chapter that endpoint variations yield conservation laws. Note that the infinitesimal endpoint variations occur on given curves at the two boundaries of the integral.

Let us evaluate the variation in J. We have[6]

$$\left.\frac{dJ}{d\alpha}\right|_{\alpha=0} = \frac{d}{d\alpha}\int_{x_1(\alpha)}^{x_2(\alpha)} dx\, f\left[y(\alpha, x), y'(\alpha, x); x\right], \tag{6.80}$$

where we are making the endpoints change with α, as in the above picture, and where it is understood the right hand side terms are evaluated at $\alpha = 0$. We find (making use of the Leibnitz rule since there is a dependence in $x_{1,2}$)

$$\left.\frac{dJ}{d\alpha}\right|_{\alpha=0} = f\left.\frac{dx}{d\alpha}\right|_{x_1}^{x_2} + \int_{x_1}^{x_2} dx\,\frac{df}{d\alpha}, \tag{6.81}$$

The second term on the right in (6.81) is the usual term we get when the endpoints are fixed. We have

$$\frac{df}{d\alpha} = \frac{\partial f}{\partial y}\underbrace{\frac{\partial y}{\partial \alpha}}_{\eta} + \frac{\partial f}{\partial y'}\underbrace{\frac{\partial y'}{\partial \alpha}}_{\eta'} \tag{6.82}$$

and doing an integration by parts on the term involving η' yields

$$\left.\frac{dJ}{d\alpha}\right|_{\alpha=0} = \left.\left(f\frac{dx}{d\alpha} + \eta\frac{\partial f}{\partial y'}\right)\right|_{x_1}^{x_2} + \int_{x_1}^{x_2} dx\left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right]\eta(x)\,. \tag{6.83}$$

We now require

$$\left.\frac{dJ}{d\alpha}\right|_{a=0} = 0\,, \tag{6.84}$$

and assume that the E-L equations are satisfied in the interior so that the integral in (6.83) vanishes. This results in the endpoint condition:

$$\left.\left(f\frac{dx}{d\alpha} + \eta\frac{\partial f}{\partial y'}\right)\right|_{x_1}^{x_2} = 0\,. \tag{6.85}$$

The endpoint requirement looks arbitrary because h appears there. However, the value of η at the endpoints is actually determined through the given endpoint curves,

$$\left.\frac{dy(\alpha,x)}{d\alpha}\right|_{\alpha=0} = \left.\left(\frac{dy(x)}{dx}\frac{dx}{d\alpha} + \eta + \alpha\frac{d\eta}{dx}\frac{dx}{d\alpha}\right)\right|_{\alpha=0}, \tag{6.86}$$

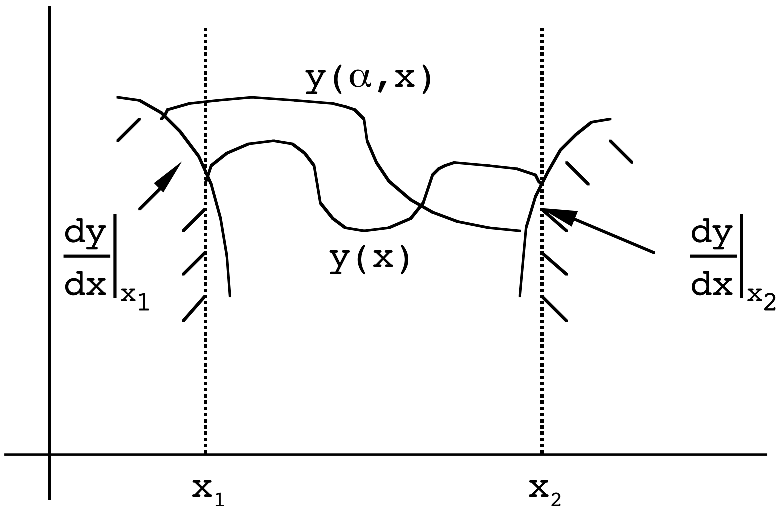
so that

$$\eta = \frac{dy}{d\alpha} - y'\frac{dx}{d\alpha}\,. \tag{6.87}$$

Multiplying by da, the endpoint requirement above assumes the form

$$\left.\left[\left(f - y'\frac{\partial f}{\partial y'}\right)dx + \frac{\partial f}{\partial y'}dy\right]\right|_{x_1}^{x_2} = 0\,. \tag{6.88}$$

This is called the "transversality condition." The slope at the endpoints determines the ratio of the variations of dy to dx; see the figure below.



In the case of more than one dependent variable, the second and third terms in (6.88) become summed over all the dependent variables, y_i. For example, when there are two dependent variables, y_1 and y_2, we obtain

$$\left[(f - \sum_{i=1}^{2} y_i' \frac{\partial f}{\partial y_i'})dx + \sum_{i=1}^{2} \frac{\partial f}{\partial y_i'} dy_i'\right]\Bigg|_{x_1}^{x_2} = 0 \tag{6.88a}$$

There are many special cases and re-writings of (6.88) that can be considered. In cases where the first point is fixed and only x_2 is varied, we have

$$\left[\left(f - y' \frac{\partial f}{\partial y'}\right) dx + \frac{\partial f}{\partial y'} dy\right]\Bigg|_{x_2} = 0\,. \tag{6.89}$$

If in addition one has $dx=0$ at x_2, the condition becomes simply

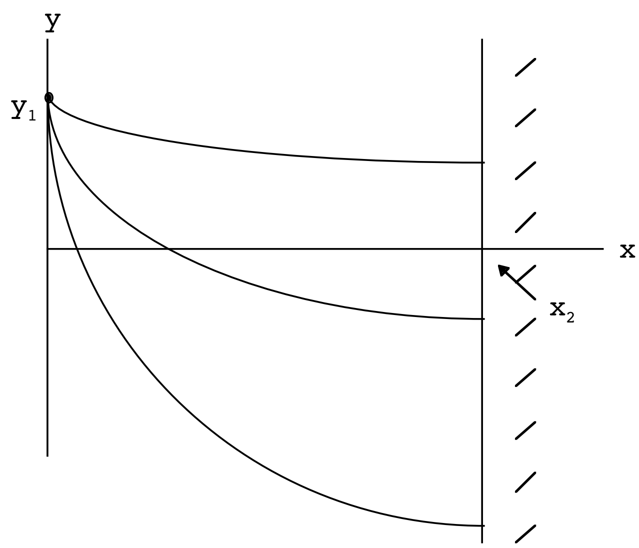
$$\frac{\partial f}{\partial y'}\Bigg|_{x_2} = 0\,, \tag{6.90}$$

and if $dy=0$ at x_2,

$$\left(f - y' \frac{\partial f}{\partial y'}\right)\Bigg|_{x_2} = 0\,. \tag{6.91}$$

Example

Find the path of least time for a sliding bead from the point $(0,y_1)$ to any point along the line $x=x_2$.



Solution:

First of all, we know that the fastest path from the given point to any given point on $x=x_2$ is a cycloid. The question is, which of these cycloids yields the smallest time to arrive at $x=x_2$? Since a single endpoint, x_2, is being varied for which $dx=0$, the endpoint condition is just (6.90) above. We have from our earlier example that

$$f = \left(\frac{1 + y'^2}{y_1 - y}\right)^{1/2},$$

and we obtain

$$\frac{\partial f}{\partial y'} = \frac{y'}{\sqrt{(y_1 - y)(1 + y'^2)}}\bigg|_{x_2} = 0,$$
$$\Rightarrow y'(x_2) = 0.$$

Therefore, the path which has zero slope at $x=x_2$ is the winner. Paths above this one have a shorter distance to go but acquire a smaller a smaller average velocity. Paths below this have greater magnitude of velocity by dipping lower in y, but have longer path lengths. Here, an "undershoot" solution never occurs, or rather, is always on the verge of occurring.

CHAPTER 6 PROBLEMS

1. Let's say we had an integral of the form

$$J = = \int_{x_1}^{x_2} dx\, f[y'', y', y; x].$$

Find the equation that f must satisfy for J to be an extremum. Only consider those paths $y(\alpha,x) = y(x)+\alpha\,\eta(x)$ such that

$$\eta(x_1) = \eta(x_2) = 0,$$

and

$$\eta'(x_1) = \eta'(x_2) = 0.$$

(assume $\eta(x)$ is twice differentiable.)

2. For $\frac{\partial f}{\partial x} = 0$, show that the integrated form of the Euler-Lagrange equation becomes (C is a constant)

$$f - \sum_{i=1}^{n} y_i' \frac{\partial f}{\partial y_i'} = C,$$

when f is now a function of **n dependent variables**,

$$f = f[y_1, y_2, \ldots, y_n, y_1', y_2' \ldots y_n' ; x].$$

[Hint: Proceed as in the text, but use the chain rule.]

3. Given

$$J = \int_{x_1}^{x_2} dx\ f[y, y'; x],$$

show that the Euler-Lagrange equation may also be written as

$$y''\left(\frac{\partial^2 f}{\partial y'^2}\right) + y'\left(\frac{\partial^2 f}{\partial y \partial y'}\right) + \left(\frac{\partial^2 f}{\partial x \partial y'}\right) - \left(\frac{\partial f}{\partial y}\right) = 0.$$

4. a) Extremize the integral

$$J = \int_0^1 dx\ [ay'^2 - by^2],$$

where a and b are positive constants, to find $y(x)$ such that $y(0)=0$, $y(1)=1$.

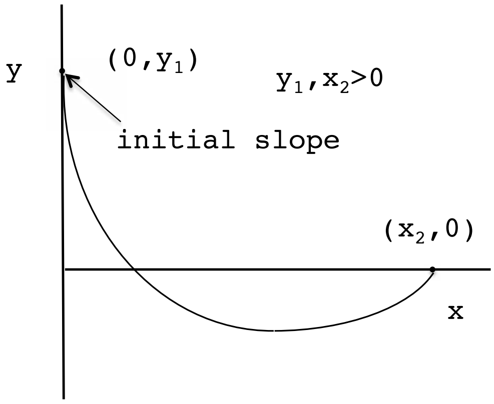
b) Find the value of the integral J using the stationary path $y(x)$ found in a).

5. Find the $y(x)$ that gives a minimum value of

$$I = \frac{\int_0^1 dx(y')^2}{\int_0^1 dx y^2},$$

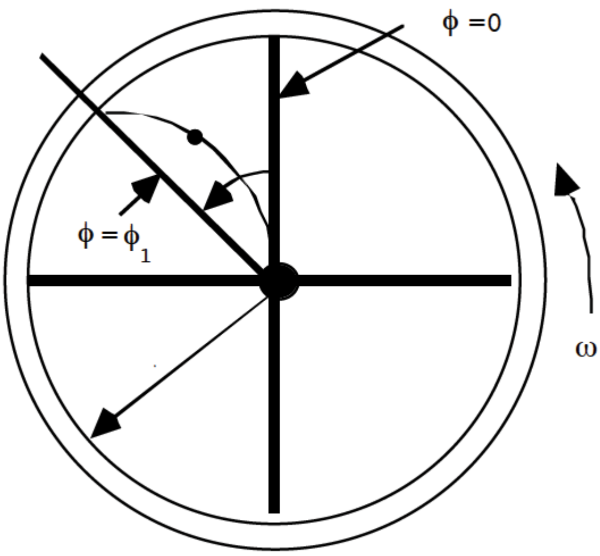
with $y(0)=y(1)=0$. [From Ch.2 problems, "Variational Mechanics for Engineers" by M. W. Wilcox (my Father). Hint: What would you do if you took the derivative of $f(x) = \frac{g(x)}{h(x)}$?]

6. a) What is the initial slope, $\frac{dy}{dx}$, of the brachistochrone solution?



b) Under what conditions on the ratio $\frac{x_2}{y_1}$ will the cycloid **dip below** (undershoot) the point at $(x_2,0)$ before reaching the point $(x_2,0)$? [Hint: The lowest point on the cycloid occurs when $\theta = \pi$ and an undershoot will occur if the y value is negative there.]

7. A space station which is rotating at an angular velocity ω about an axis through it's central hub has a inter-spoke mail delivery system in which packages are sent down enclosed frictionless tubes from a central hub to various stations along the rim of the station.



All packages start from rest at the hub along $\phi = 0, r = a$, then curve along a path to reach their final destination at $\phi=\phi_1, r=R$. The connection between speed along the path, v, and r is $v=\omega r$. (This comes from energy conservation since kinetic energy is $\frac{1}{2}mv^2$ and potential energy is $\frac{1}{2}m\omega^2 r^2$.)

a) Show that the time for delivery can be written (use cylindrical coordinates; $r' \equiv \frac{dr}{d\phi}$)

$$T = \frac{1}{\omega}\int_0^{\phi_1} d\phi \frac{\sqrt{r'^2 + r^2}}{r}.$$

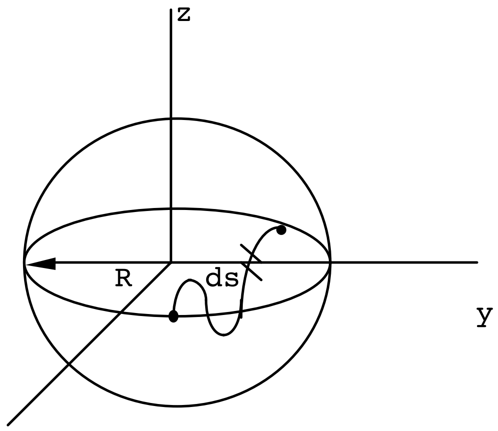
b) Minimize the integral in (a) and show the path is a logarithmic spiral.

8. Extremize the 3 dimensional integral $(\phi = \phi(\vec{x}))$,

$$J = \int\int\int dx_1 dx_2 dx_3 \left[\frac{1}{2}(\vec{\nabla}\phi)^2 - \rho(\vec{x})\,\phi\right],$$

where $\rho(\bar{x})$ is a given function and $\bar{\nabla}$ is the three dimensional gradient operator. Find the partial differential equation for ϕ that results.

9. Find the path of shortest distance between two points on a sphere of radius R as in the text.



This time treat θ as a dependent variable and ϕ the independent variable.

10. a) Re-do the sphere as in the text, using θ as a dependent variable and ϕ the independent variable. However, use the usual (nonintegrated) Euler-Lagrange equations and show that one obtains

$$\frac{\cos\theta}{\sin\theta} - \frac{d}{d\phi}\left(\theta'\frac{1}{\sin^2\theta}\right) = 0,$$

where $\theta' = \dfrac{d\theta}{d\phi}$.

b) Show that

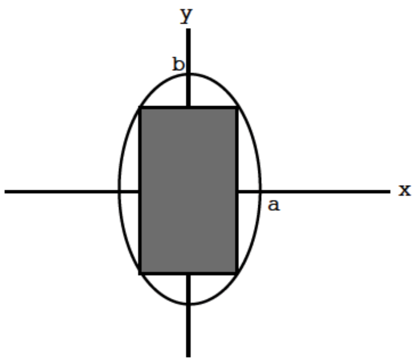
$$\frac{d}{d\phi}\left(\frac{\cos\theta}{\sin\theta}\right) = -\,\theta'\frac{1}{\sin^2\theta}$$

and use this to solve the equation in (a).

11. One more time with the sphere. Using both θ and ϕ as dependent variables, formulate this situation as a parametic problem. It is enough to simply write down the appropriate extreemum conditions in this case. Combine the two E-L equations and show that one obtains the same equation given in prob.10(a) above.

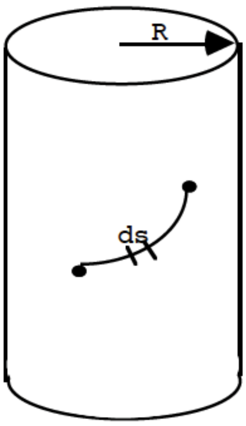
12. Use the method of Lagrange multipliers to find the area of the largest rectangle, with sides parallel to the major and minor axes, which can lie inscribed in the ellipse:

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1.$$

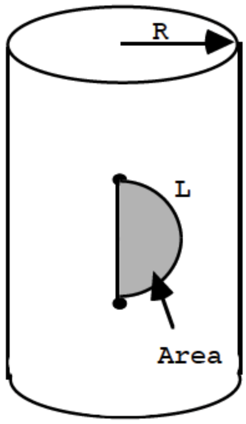


13. Find the cubical of greatest volume which can be inscribed in a sphere of radius R given that $x=2y$ where $2x$ and $2y$ are lengths of two sides of the cubical. Formulate this as a Lagrange multipier problem.

14. a) Find the equation of the shorest curve between two points on a cylinder of radius R using cylindrical coordinates. [$ds^2 = R^2 d\phi^2 + dz^2$.]

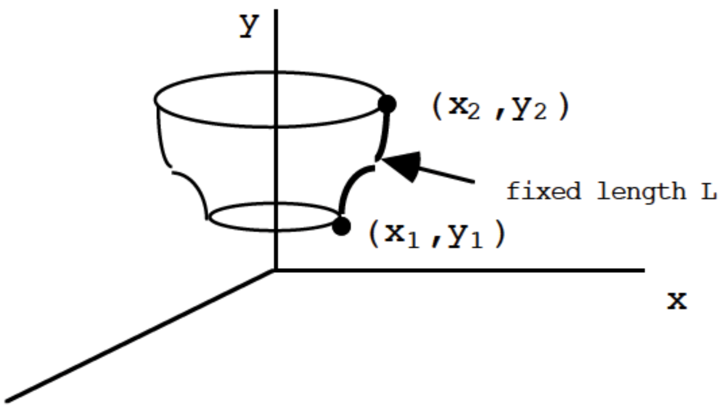


b) Let us say we are given an unstreachable length of string, L. Find the equation of the line on the cylinder's surface which encloses the greatest area on the cylinder's surface between the string and a vertical line on the cylinder. See the below:



Ans.: $(z-K)^2 + \frac{R^2}{\lambda^2}(\lambda\phi - C)^2 = \frac{R^2}{\lambda^2}$, where C, K and λ are constants. (This is just the equation of a portion of a circle pasted on the surface of the cylinder.)

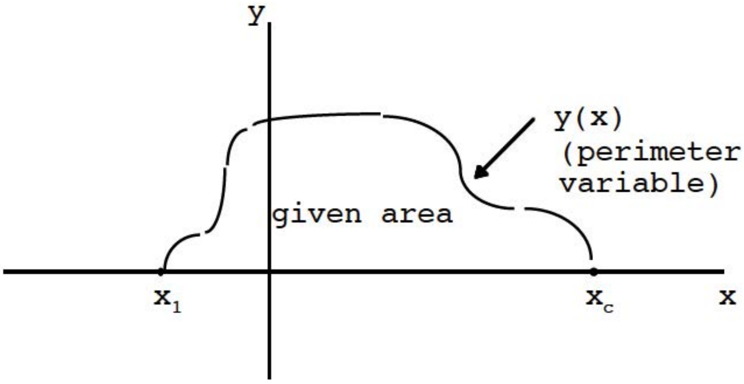
15. An economical mathematician wishes to make a lampshade with as small an area as possible by rotating a given curve in the $y-x$ plane passing through (x_1,y_1) and (x_2,y_2)around the y-axis. A wire form is used to construct this curve, which can be bent but has a fixed length, L.



Find the form of the equation, $y(x)$, which describes this lampshade. [Hint: Set this problem up as a constraint problem using a Lagrange multiplier. Make sure you tell me how the Lagrange multiplier is determined.]

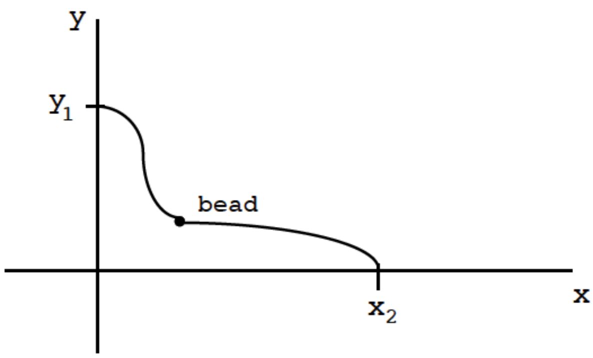
16. Prove the following:

The shortest perimeter of an arbitrary path enclosing a fixed **area** between the x-axis and a curve $y(x)$, as shown, is a portion of a circle.



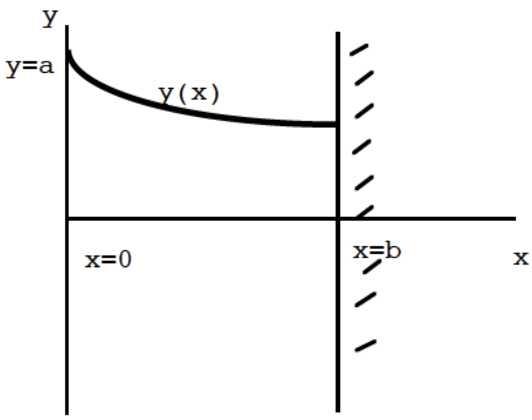
[Hint: This may not be as difficult as you may think.]

17. A wire of **fixed length** L is used to connect the points shown. A bead slides without friction along the wire.



We wish to fashion the wire's shape to construct a path leading from $(0,y_1)$ to $(x_2,0)$ which uses minimum time, but with fixed length. We of course have $L > \sqrt{y_1^2 + x_2^2}$. Set this problem up as a constraint problem using a Lagrange multiplier. **It is not necessary to attempt to solve the resulting differential equation for** $y(x)$.

18. Another mathematician now wishes to construct a lampshade of least area for the surface generated by revolving about the x-axis a line, $y(x)$, which connects the point at $x=0, y=a$ to **anywhere** on the fixed wall at $x=b$, as shown.



Show that the equation is

$$y(x) = C\cosh\left(\frac{x-b}{C}\right),$$

where C (the value of $y(b)$ is determined by

$$a = C\cosh\left(\frac{b}{C}\right).$$

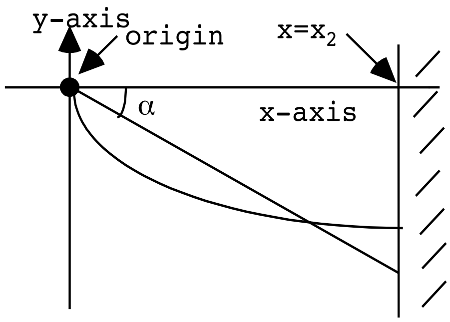
[Hint: You must consider endpoint variations at $x=b$.]

19. Going back to the space station mail problem, we found that the time for delivery was (using cylindrical coordinates; $r' \equiv \frac{dr}{d\phi}$)

$$T = \frac{1}{\omega}\int_0^{\phi_1} d\phi\, \frac{\sqrt{r'^2 + r^2}}{r}$$

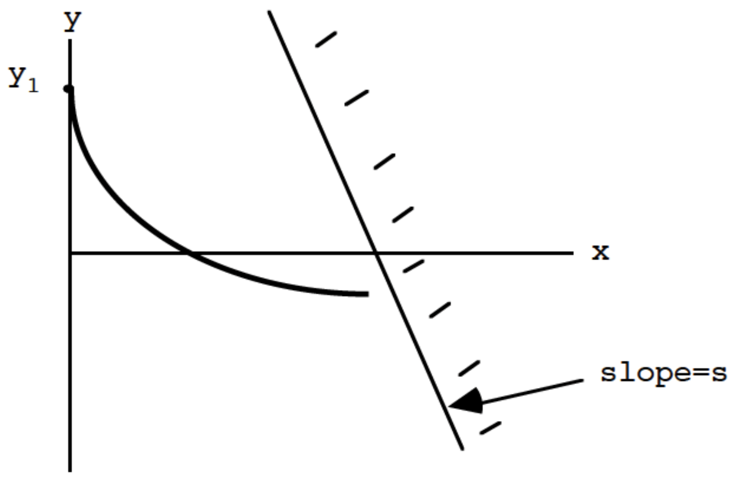
a) Minimize the time subject to the condition that the length of the tube, L, is fixed. Simplify the implied differential equation as much as possible. (There is no need to attempt to solve the resulting differential equation, which is rather difficult.)
b) Given the solution to 7(b), show that the fastest mail path to the outside ring is the shortest path by considering endpoint variations of the final location.

20. Find the angle, α, of the straight-line path for which a sliding bead will pass from the origin to the wall at $x = x_2$ the fastest. Note we have a constant gravitational field downward in the diagram.



Find the equation of the brachistichrone which has zero slope at the wall ($\left.\frac{dy}{dx}\right|_{x=x_2} = 0$) and show that the bead's time of flight on this path is less than the result found above.

21. Find the endpoint condition on the cycloid which provides the fastest descent from a fixed point $(0,y_1)$ on the y-axis to any point on a line of slope s, as shown.



[This generalizes the last example of the text. Answer: the fastest cycloid always approaches perpendicularly to the given line.]

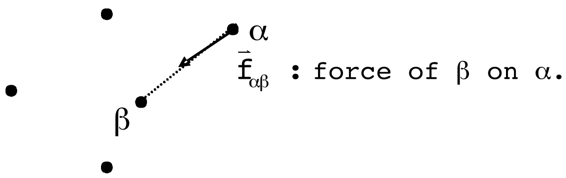
7 LAGRANGIAN AND HAMILTONIAN MECHANICS

THE ACTION AND HAMILTON'S PRINCIPLE

For an n-particle system, Newton's equations are

$$m_\alpha \frac{d^2\vec{x}_\alpha}{dt^2} = \vec{f}_\alpha . \qquad (3n \text{ eq}^{\underline{n}}\text{s}) \tag{7.1}$$

Imagine (no external forces):



Total force on α:

$$\vec{f}_\alpha = \sum_{\beta \neq \alpha} \vec{f}_{\alpha\beta}.$$

‖ no self-force (7.2)

Assume that $\vec{f}_{\alpha\beta}$ results from a two-body potential

$$\vec{f}_{\alpha\beta} = -\vec{\nabla}_\alpha \bar{U}_{\alpha\beta}, \qquad (\text{no } \alpha \text{ sum!}) \tag{7.3}$$

$$\bar{U}_{\alpha\beta} \equiv \bar{U}_{\alpha\beta}\left(\left|\vec{x}_\alpha - \vec{x}_\beta\right|\right). \tag{7.4}$$

Notice that (7.4) implies $\bar{U}_{\alpha\beta} = \bar{U}_{\beta\alpha}$, which gives Newton's third law from (7.3). Eq.(7.4) also means that the force between the particles is along the line connecting them. This is not the most general assumption.

Explicitly, the notation in the following is

$$\vec{x}_\alpha \rightarrow x_{\alpha i},$$

particle no.↑ ↑coordinate no.

$$\Rightarrow m_\alpha \ddot{x}_{\alpha i} + \frac{\partial}{\partial x_{\alpha i}}\left(\sum_{\beta \neq \alpha} \bar{U}_{\alpha\beta}\right) = 0. \tag{7.5}$$

Another way of writing this equation will now be explained. The total kinetic and potential energies are

$$T \equiv \frac{1}{2}\sum_{\alpha, i} m_\alpha \dot{x}_{\alpha i}^2, \tag{7.6}$$

$$U \equiv \sum_{\alpha < \beta} \bar{U}_{\alpha\beta}\left(\left|\vec{x}_\alpha - \vec{x}_\beta\right|\right). \tag{7.7}$$

↑ summing over both α, β such that $\alpha < \beta$.

The reason for the $\alpha < \beta$ restriction in the sum in (7.7) is to count a given interaction between two different particles only once.

Form $L \equiv T - U$. Notice

$$\frac{\partial L}{\partial x_{\alpha i}} = -\frac{\partial U}{\partial x_{\alpha i}} = -\frac{\partial}{\partial x_{\alpha i}}\left(\sum_{\gamma < \beta} \bar{U}_{\gamma\beta}\left(\left|\vec{x}_\gamma - \vec{x}_\beta\right|\right)\right). \tag{7.8}$$

Beginning to look like the second term in (7.5).

Take the special case of 3 bodies $(\alpha = 1,2,3)$ to see another way of writing (7.8). We have

$$\sum_{\gamma<\beta} \overline{U}_{\gamma\beta} = \underbrace{\overline{U}_{13} + \overline{U}_{23}}_{\gamma<3} + \underbrace{\overline{U}_{12}}_{\gamma<2} . \tag{7.9}$$

Let's now consider the derivative of this with respect to x_{1i}:

$$\begin{aligned} \frac{\partial}{\partial x_{1i}} \sum_{\gamma<\beta} \overline{U}_{\gamma\beta} &= \frac{\partial}{\partial x_{1i}} \left(\overline{U}_{13} + \overline{U}_{23} + \overline{U}_{12}\right) \\ &= \frac{\partial}{\partial x_{1i}} \left(\overline{U}_{13} + \overline{U}_{12}\right) \\ &= \frac{\partial}{\partial x_{1i}} \sum_{\beta\neq 1} \overline{U}_{1\beta} . \end{aligned} \tag{7.10}$$

Do the same thing for $\alpha = 2$ (to see a necessary property for the rewriting we will do)

$$\begin{aligned} \frac{\partial}{\partial x_{2i}} \sum_{\gamma<\beta} \overline{U}_{\gamma\beta} &= \frac{\partial}{\partial x_{2i}} \left(\overline{U}_{13} + \overline{U}_{23} + \overline{U}_{12}\right) , \\ &= \frac{\partial}{\partial x_{2i}} \left(\overline{U}_{23} + \overline{U}_{12}\right) \\ &= \frac{\partial}{\partial x_{2i}} \left(\overline{U}_{23} + \overline{U}_{21}\right) \quad \Big\} \; \overline{U}_{\gamma\beta} = \overline{U}_{\beta\gamma} \text{ necessary} \\ &= \frac{\partial}{\partial x_{2i}} \sum_{\beta\neq 2} \overline{U}_{2\beta} . \end{aligned} \tag{7.11}$$

Generalization for an n-body system:

$$\frac{\partial}{\partial x_{\alpha i}} \sum_{\gamma<\beta} \overline{U}_{\gamma\beta} = \frac{\partial}{\partial x_{\alpha i}} \sum_{\beta\neq\alpha} \overline{U}_{\alpha\beta} . \tag{7.12}$$

Using this above, we get

$$\frac{\partial L}{\partial x_{\alpha i}} = -\frac{\partial}{\partial x_{\alpha i}} \left(\sum_{\beta\neq\alpha} \overline{U}_{\alpha\beta}\right) . \tag{7.13}$$

What's more,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}_{\alpha i}}\right) = \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{x}_{\alpha i}}\right) , \tag{7.14}$$

and

$$\frac{\partial}{\partial \dot{x}_{\alpha i}} \left(\frac{1}{2} \sum_{\beta,j} m_\beta \dot{x}_{\beta j}^2\right) = m_\alpha \dot{x}_{\alpha i} ,$$

$$\Rightarrow \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{x}_{\alpha i}}\right) = \frac{d}{dt}\left(m_\alpha \dot{x}_{\alpha i}\right) = m_\alpha \ddot{x}_{\alpha i} . \tag{7.16}$$

Now, comparing with (7.5), we have

$$\Rightarrow \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}_{\alpha i}}\right) - \frac{\partial L}{\partial x_{\alpha i}} = 0. \quad \text{(still 3n equations)} \tag{7.17}$$

Reversing our usual logic, this now implies that the integral

$$S \equiv \int_{t_1}^{t_2} dt\, L\left(x_{\alpha i}, \dot{x}_{\alpha i}; t\right), \tag{7.18}$$

("the action") is an extremum, or stationary, when the equations (7.17) hold. "L" is called the Lagrangian. Eq.(7.18) represents a vast simplification of the situation. All information is contained in a single, scalar, quantity. This gives rise to Hamilton's principle:

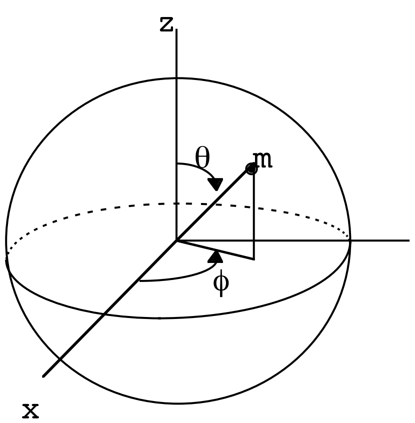
> Of all possible paths along which a dynamical system moves, the actual path is that which makes the action,
>
> $$S = \int_{t_1}^{t_2} dt\,(T - U),$$
>
> stationary.

Notice we have not proven this principle, we have simply constructed the particular form of the action for a special type of system.

Generalized Coordinates

Let's now generalize the systems we are describing by introducing new "generalized" coordinates, q_i, which uniquely specify the positions of the particles. These coordinates are assumed to automatically respect the possible constrained motions of the system. Examples are:

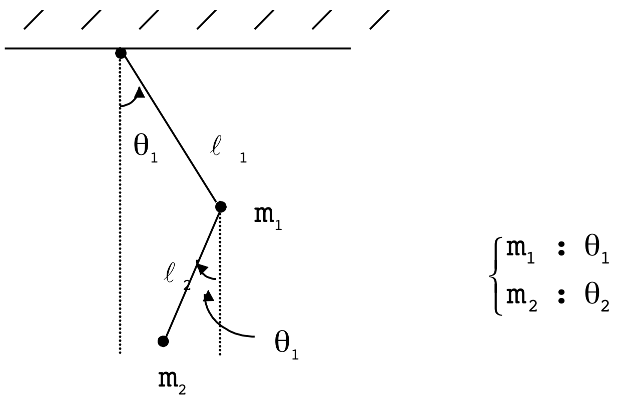
1.



radius = R

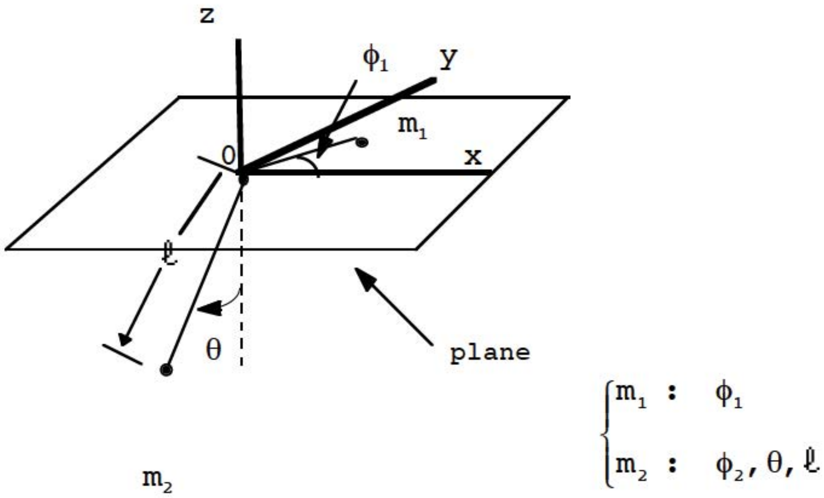
θ,φ (spherical coordinates angles)

2.



planar double pendulum (ℓ_1, ℓ_2 fixed)

3.



String of length fixed L with masses m_1, m_2 at ends slipping through a hole in a plane (ϕ_2 not shown).

Notice that some generalized coordinates have units of length while others do not. Notice also the choice is not unique. In writing the generalized coordinates as q_i, I am dropping the double indices on particle number and coordinate axis. In general there are fewer qi variables than xaj variables because of the elimination of constraints among the $x_{\alpha j}$. Our new "phase space" for such a system is specified by the generalized positions, qi, and the generalized velocities, q_i. If there are m constraints, then the number of generalized coordinates, s, is $3n - m$. (For solids the counting goes like $s = 6n - m$ since three coordinates and three angles are needed to specify a body's position and orientation.) If there are no remaining constraints among the remaining generalized variables, they are called a **proper** set of coordinates.

Let us assume we know how to write the kinetic and potential energies, which are simply scalar quantities under coordinate changes, of the system in terms of the q_i and $\dot{q}_i$ That is, we know the transformation equations from the $x_{\alpha i}$ to the q_j variables. Symbolically, we write

$$x_{\alpha i} = x_{\alpha i}(q_j, t), \tag{7.19a}$$

which implies

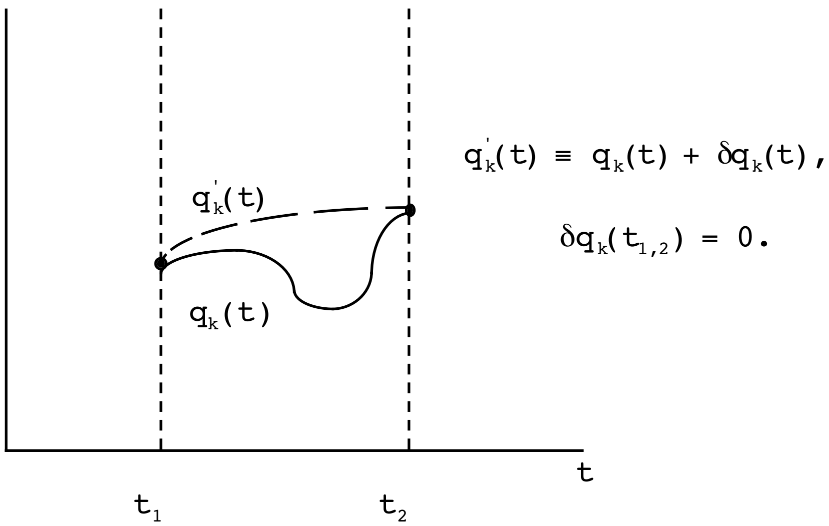
$$\dot{x}_{\alpha i} = x_{\alpha i}(q_j, \dot{q}_j, t). \tag{7.19b}$$

Notice that the time can explicitly enter the transformation. We assume that T depends only on $x_{\alpha i}$ and $\dot{x}_{\alpha i}$ and U depends only on $x_{\alpha i}$, so therefore $T(\dot{x}_{\alpha i})=T(q_i, \dot{q}_i, t)$ and $U(x_{\alpha i})=U(q_i, t)$and we may form

$$S = \int_{t_1}^{t_2} dt\, L(q_i, \dot{q}_i; t). \tag{7.20}$$

By extremizing this expression, we will find the form of Newton's equations in the new, generalized coordinates. To perform this variation and to introduce a new notation, let us again reverse our logic and extremize the action in the form (7.20). This new notation is easier to use and more elegant than the notation in the last Chapter.

Let us imagine varying a particular coordinate $q_k(t)$ around the stationary path, as shown below. (Note the prime on $q_k^{'}(t)$ does NOT represent a derivative!)



$\delta q_k(t)$ is imagined to be an infinitesimal, arbitrary variation about the extremum solution, $q_k(t)$. This "frees" us from the clumsy use of α and η in the previous chapter. The connection with the previous notation is

$$\delta(\ \ ) \rightarrow d\alpha \frac{\partial}{\partial\alpha}(\ \ )\Big|_{\alpha=0}. \tag{7.21}$$

Also

$$\frac{dq_k'(t)}{dt} = \frac{dq_k(t)}{dt} + \frac{d}{dt}\delta q_k(t). \tag{7.22}$$

Now

$$\delta\left(\frac{dq_k}{dt}\right) = \left(\frac{dq_k}{dt}\right)' - \frac{dq_k}{dt} = \frac{dq_k'}{dt'} - \frac{dq_k}{dt}. \tag{7.23}$$

However $t' = t$ in the variation, so we have

$$\delta\left(\frac{dq_k}{dt}\right) = \frac{d}{dt}\left(q_k' - q_k\right) = \frac{d}{dt}\delta q_k. \tag{7.24}$$

This means we can interchange the "δ" and the "$\frac{d}{dt}$" operations. This makes sense since the variation defined in (7.21) is independent of t.

It's important to realize that the varied paths in general do not conserve energy or other quantities. Think of a particle subject to a potential, $U(x)$. The varied paths change the relationship between x and $\dot{x}$ and, therefore, between kinetic and potential energies. Thus, the varied paths generally are not physically possible in classical mechanics. However there is a formulation of quantum mechanics which uses a quantity called the path integral. In this approach, particles are seen to actually experience these "unphysical" paths, the actual motion being a sum of the wave amplitudes associated with each of the paths; the path integral represents this sum. One can think of the motion as a tube of possible paths surrounding the classical path picked out by the E-L equations. This tube becomes thinner in the limit that the motion becomes more classical, i.e. where quantum numbers become large and the discrete nature of the quantum world becomes almost continuous. The path integral approach to quantum mechanics provides a beautiful physical picture which connects to and generalizes the ideas of classical mechanics.

Under these variations, we have the change in the Lagrangian (sums from 1 to S)

$$\delta L = \sum_k \left[\left(\frac{\partial L}{\partial q_k}\right) \delta q_k + \left(\frac{\partial L}{\partial \dot{q}_k}\right) \underbrace{\delta \dot{q}_k}_{\frac{d}{dt}\delta q_k} \right], \qquad (7.25)$$

$\uparrow$ from (7.24)

$$\Rightarrow \delta S = \int_{t_1}^{t_2} dt \sum_k \left[\frac{\partial L}{\partial q_k} \delta q_k + \frac{\partial L}{\partial \dot{q}_k} \frac{d}{dt} \delta q_k \right]. \qquad (7.26)$$

Do an integration by parts on the second set of terms on the right:

$$\int_{t_1}^{t_2} dt \sum_k \frac{\partial L}{\partial \dot{q}_k} \frac{d}{dt} \delta q_k = \sum_k \frac{\partial L}{\partial \dot{q}_k} \underbrace{\delta q_k}_{0} \Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} dt \sum_k \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_k}\right) \delta q_k . \qquad (7.27)$$

If the δq_k are independent and arbitrary, when we demand

$\delta S = 0$. we obtain

$$\frac{\partial L}{\partial q_k} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) = 0. \qquad (7.28)$$

One can also show (7.28) follows directly from the original set, (7.5), and the transformation equations, (7.19).

Examples of the Formalism

Some examples at this point would be helpful.

Example 1

Set up the simple harmonic oscillator in the Lagrangian formalism.

Solution:

The simple harmonic oscillator has kinetic and potential energies, given by

$$T = \frac{1}{2}m\dot{x}^2 \quad , \quad U = \frac{1}{2}kx^2 ,$$
$$\Rightarrow L = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2 .$$

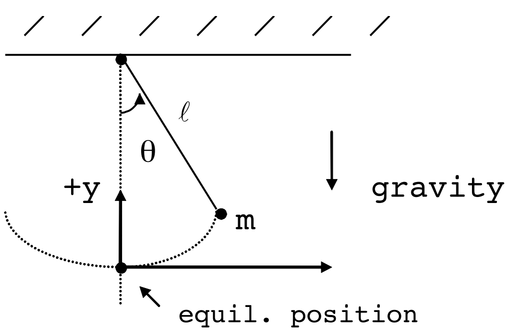
The E-L equation is

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = 0 ,$$
$$\Rightarrow -kx - \frac{d}{dt}\left(m\dot{x}\right) = 0 ,$$
$$\Rightarrow \quad \ddot{x} + \frac{k}{m}x = 0 ,$$

which is the usual equation of motion.

Example 2

Do the undamped plane pendulum in the Lagrangian formalism. As the generalized coordinate use an angle.



Solution:

In rectangular coordinates the kinetic and potential energies are

$$T = \frac{1}{2} m \left(\dot{x}^2 + \dot{y}^2\right),$$

$$U = mgy.$$

The transformation equations are

$$x = \ell \sin \theta ,$$
$$y = \ell(1 - \cos\theta).$$

Thus

$$\dot{x}^2 + \dot{y}^2 = \ell^2 \left(\cos^2\theta + \sin^2\theta\right)\dot{\theta}^2 = \ell^2\dot{\theta}^2 ,$$

and

$$L = \frac{1}{2} m\ell^2\dot{\theta}^2 - mg\ell\left(1 - \cos\theta\right).$$

The generalized E-L equation is now

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

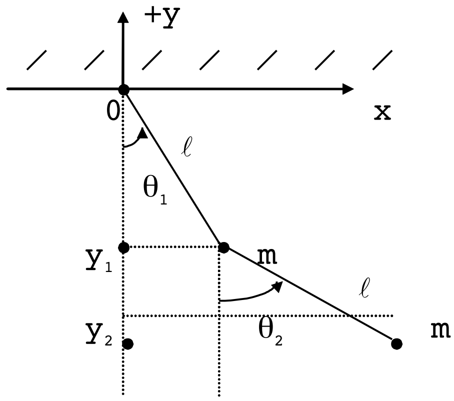
$$\Rightarrow -mg\ell \sin\theta - \frac{d}{dt}\left(m\ell^2\dot{\theta}\right) = 0$$

$$\Rightarrow \ddot{\theta} + \frac{g}{\ell}\sin\theta = 0.$$

This nonlinear equation is also quite familiar from Ch.4.

Example 3

Set up the Lagrangian equations of motion for the double pendulum. For simplicity, assume $m_1 = m_2 = m$ as well as $\ell_1 = \ell_2 = \ell$



Solution:

The kinetic and potential energies are

$$T = \frac{1}{2} m\left(\dot{x}_1^2 + \dot{y}_1^2\right) + \frac{1}{2} m\left(\dot{x}_2^2 + \dot{y}_2^2\right),$$

$$U = mg\left(y_1 + y_2\right).$$

The transformation equations are

$$x_1 = \ell \sin\theta_1 \quad , \quad y_1 = -\ell \cos\theta_1 ,$$

$$x_2 = \ell\left(\sin\theta_1 + \sin\theta_2\right),$$

$$y_2 = -\ell\left(\cos\theta_1 + \cos\theta_2\right).$$

I'll leave it to you to show in a HW problem, that the kinetic and potential energies now become

$$T = m\ell^2\dot{\theta}_1^2 + \frac{1}{2} m\ell^2\dot{\theta}_2^2 + m\ell^2\dot{\theta}_1\dot{\theta}_2 \cos\left(\theta_1 - \theta_2\right),$$

$$U = -mg\ell\left(2\cos\theta_1 + \cos\theta_2\right),$$

Needed for the E-L equations:

$$\frac{\partial L}{\partial \theta_1} = -m\ell^2\dot{\theta}_1\dot{\theta}_2 \sin\left(\theta_1 - \theta_2\right) - 2mg\ell \sin\theta_1 ,$$

$$\frac{\partial L}{\partial \theta_2} = m\ell^2\dot{\theta}_1\dot{\theta}_2 \sin\left(\theta_1 - \theta_2\right) - mg\ell \sin\theta_2 \,,$$

$$\frac{\partial L}{\partial \dot{\theta}_1} = 2m\ell^2\dot{\theta}_1 + m\ell^2\dot{\theta}_2 \cos\left(\theta_1 - \theta_2\right),$$

$$\frac{\partial L}{\partial \dot{\theta}_2} = m\ell^2\dot{\theta}_2 + m\ell^2\dot{\theta}_1 \cos\left(\theta_1 - \theta_2\right).$$

The equations of motion are now

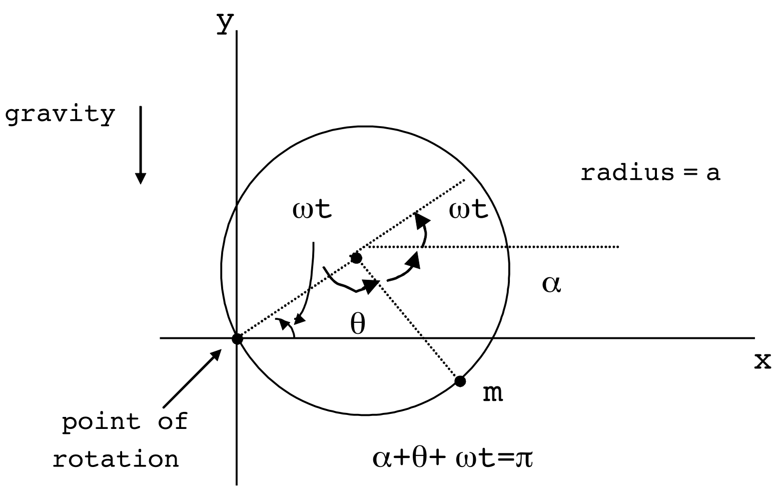
$$\theta_1 : \ddot{\theta}_1 + \frac{1}{2}\ddot{\theta}_2 \cos\left(\theta_1 - \theta_2\right) + \frac{1}{2}\dot{\theta}_2^2 \sin\left(\theta_1 - \theta_2\right) + \frac{g}{\ell}\sin\theta_1 = 0\,,$$

$$\theta_2 : \ddot{\theta}_2 + \ddot{\theta}_1 \cos\left(\theta_1 - \theta_2\right) + \dot{\theta}_1^2 \sin\left(\theta_1 - \theta_2\right) + \frac{g}{\ell}\sin\theta_2 = 0\,.$$

Do you think you could have derived these equations directly without using a Lagrangian?

Example 4

Find the equations of motion for a bead of mass m free to move on a hoop of radius a, rotating at constant angular velocity ω, as shown below. (Assume the mass can move past the point of rotation.)



Solution:

We have

$$T = \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2\right),$$

$$U = mgy\,.$$

The transformation equations are

$$x = a\cos\omega t - a\cos(\theta + \omega t)\ ,$$
$$y = a\sin\omega t - a\sin(\theta + \omega t)\ .$$

Notice in this case, unlike the other examples, the transformation equations explicitly involve the time. Using these, I find

$$T = \frac{1}{2}ma^2\omega^2 + \frac{1}{2}ma^2\left(\dot\theta + \omega\right)^2 - ma^2\omega\left(\dot\theta + \omega\right)\cos\theta\,,$$

$$U = mga\sin\omega t - mga\sin(\theta + \omega t)\,.$$

The E-L equation of motion becomes

$$\ddot\theta - \omega^2\sin\theta - \frac{g}{a}\cos(\theta + \omega t) = 0\,.$$

Notice, if we define $\phi = \theta + \pi$ and do a rotation in the plane perpendicular to the gravitational field, we obtain ($g = 0$ effectively)

$$\ddot\phi + \omega^2\sin\phi = 0\,,$$

which says the bead moves like a pendulum about the far end of the hoop.

Two Points about Lagrangian Methods

Let me make several points about Lagrangian methods before we go on. First, the Lagrangian of a system is not unique. Consider a new Lagrangian, L' given by

$$L' = L + \frac{d}{dt} G(q_i, t), \tag{7.29}$$

in one dimension. The action associated with L' is

$$S' = S + \int_{t_1}^{t_2} \underbrace{dt \frac{d}{dt} G(q_i, t)}_{G(q_i,t)\Big|_1^2}. \tag{7.30}$$

The quantity $G(q_i, t)\Big|_1^2$ is fixed in the variation because $(q_i)_{1,2}$ and $t_{1,2}$ are not being changed at the endpoints. Therefore $\delta S' = \delta S$ and the equations of motion (= Newton's equations = E-L equations) for L' are the same as those with L.

Another important point is that Hamilton's principle can be used for systems which have potentials very different from Eq. (7.4) above and which are subject to specific types of constraints. The crucial assumption in the demonstration that Newton's equations, (7.1), can be written in the E-L form (7.17) is that potentials exist from which the non-constraint forces may be derived and which obey Newton's third law. Such potentials can be time-dependent and even velocity-dependent. The form of the relation between such potentials and the forces in this case is

$$F_{\alpha i} = -\frac{\partial U}{\partial x_{\alpha i}} + \frac{d}{dt}\left(\frac{\partial U}{\partial \dot{x}_{\alpha i}}\right). \tag{7.31}$$

Even more generally, there are also E-L- type equations, but not a Lagrangian, for systems which do not possess potentials. These are called D'Alembert's equations. We will pass over a description of these types of systems.

Types of Constraints

The general form of constraint relations in rectangular coordinates is ($j=1,...,m$ for m constraints)

$$\sum_{\alpha,i} \omega^j_{\alpha i}(x,t)\,dx_{\alpha i} + \omega^j_t(x,t)\,dt = 0\,,$$

$$\Rightarrow \sum_{\alpha,i} \omega^j_{\alpha i}(x,t)\,\dot{x}_{\alpha i} + \omega^j_t(x,t) = 0\,.$$

Then we may classify constraints as follows:

(1). "Holonomic": $\omega^j_{\alpha i} = \dfrac{\partial f^j}{\partial x_{\alpha i}}\,,\ \omega^j_t = \dfrac{\partial f^j}{\partial t}$. Then since

$$\frac{df^j}{dt} = \sum_{\alpha i} \frac{\partial f^j}{\partial x_{\alpha i}}\dot{x}_{\alpha i} + \frac{\partial f^j}{\partial t}\,,$$

we may write these simply as

$g^j(x_{\alpha i}, t) = f^j(x_{\alpha i}, t) - \text{const.} = 0$.

Also called "integrable" constraints.

(2). "Non-holonomic": All other cases. Also called "Non-integrable".

As pointed out above, a Lagrangian exists and the E-L equations, Eqs.(7.17), hold for systems which have a potentials of the form (7.31). Systems which have holonomic constraints and generalized forces derivable from potentials, $U(q_i, q_i)$, have E-L equations have the form (7.28). Examples of constraint relations for the three examples which began the Generalized Coordinates section are:

1. Particle on a sphere

   1. $g_1(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - R^2 = 0\,.$

      $n = 1,\quad m = 1$

   $s = 3n - m = 2$ generalized coordinates (θ, ϕ)

2. Plane double pendulum

   1. $g_1 = x_{13} = 0$

   2. $g_2 = x_{23} = 0$

   3. $g_3 = x_1^2 + x_2^2 - \ell^2 = 0$

   4. $g_4 = (x_{21} - x_{11})^2 + (x_{22} - x_{12})^2 - \ell^2 = 0$

$$n = 2, \quad m = 4$$

$s = 3n - m = 2$ generalized coordinates (θ_1, θ_2)

3. String of length L through a plane

1. $g_1 = x_{13} = 0$

2. $g_2 = \sqrt{x_{21}^2 + x_{22}^2 + x_{23}^2} + \sqrt{x_{11}^2 + x_{12}^2} - L = 0$

$$n = 2, \; m = 2$$

$s = 3n - m = 4$ generalized coordinates $(\phi_1, \phi_2, \theta, \ell)$

All of these are holonomic constraints. A further independent classification is whether the constraints involve the time ("rheonomic") or are time independent ("scleronomic"). It would take us too far afield to further discuss these classifications in any further detail. I suggest "Principles of Mechanics" by Synge and Griffth for a deeper investigation of these ideas.

Although I stated the holonomic constraint condition above in terms of rectangular coordinates, it is also possible to state such conditions between generalized coordinates and to use the constraint formalism of the last chapter. The E–L equations with explicit generalized, holonomic constraints (from (6.68)–(6.70)) can be written as

$$\left.\begin{aligned} &\frac{\partial L^*}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial L^*}{\partial \dot{q}_i}\right) = 0 \quad , && (i = 1, \dots, s) \\ &g_j(q_i, t) = 0, && (j = 1, \dots, m) \end{aligned}\right\} \tag{7.32}$$

where

$$L^* \equiv T - U^* \; , \; U^* \equiv U + \sum_{j=1}^{m} \lambda_j g_j \; . \tag{7.33}$$

I will call U^* the "effective potential". Using the connection between potential and force, we see that if one can relate q_i to an instantaneous rectangular coordinate x_i by a scale factor s such that $x_i = q_i s$, then the actual force of constraint is

$$F_i = s^{-1} \Im_i \; ,$$

where

$$\Im_i \equiv -\sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial q_i} \; , \tag{7.34}$$

The constraint formalism is especially useful in revealing the explicit forces of constraint.

Example 5

Find the effective force of constraint in the radial direction for the pendulum in Example 2 above.

Solution:

We will write the kinetic and potential energies here as

$$T = \frac{1}{2} m r^2 \dot{\theta}^2 , \tag{E.1}$$

$$U = mg\left(\ell - r\cos\theta\right), \tag{E.2}$$

where r and θ are the generalized coordinates and the constraint as (ℓ = constant)

$$g = r - \ell = 0 .$$

(Notice how I write the potential in (E.2). If r is changed, this changes only the pendulum's length, not the zero of potential.) The only new equation here is (note $g = g(r)$, so the constraint does not affect the θ-equation)

$$\frac{\partial L}{\partial r} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right) + \Im_r = 0 ,$$

where I have used $L^* = T - U^*$ and identified $\Im_r$ from (7.34). Thus (using $r = \ell$ only after the partial derivative is taken)

$$\Im_r = -mg\cos\theta - m\ell\dot{\theta}^2 .$$

Notice the overall minus sign: the first term counteracts gravity, the second counteracts the centrifugal force. This is a general result, applicable even if the pendulum were driven. In our case, we know that $E = T + U$ is a constant. Thus

$$E = \frac{1}{2} m\ell^2\dot{\theta}^2 + mg\ell\left(1 - \cos\theta\right).$$

Substituting for $\dot{\theta}^2$ in $\Im_r$, we find

$$\Im_r = mg\left(2 - 3\cos\theta\right) - \frac{2E}{\ell} .$$

Now since

$$E = mg\ell(1-\cos\theta_0),$$

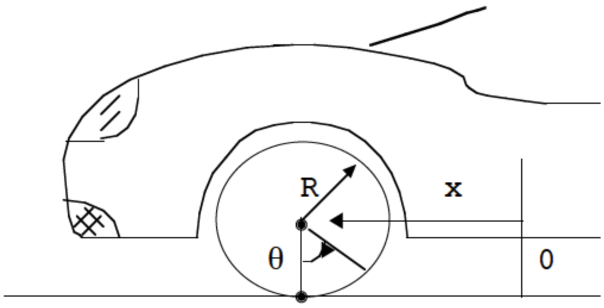
where θ_0 is the maximum value of θ, we can write

$$\Im_r = mg\left(2\cos\theta_0 - 3\cos\theta\right).$$

This is easily seen to be the actual constraining force (due to the pendulum attachment) on the mass in the radial direction. Notice $\Im_r = -mg$ for $\theta = \theta_0 = 0$, as it should.

Example 6

A car accelerates (or decelerates) at an instantaneous rate, $\ddot{x}$. Find the generalized force on a tire at the point of contact with the road if there is no slippage. What does this generalized force represent?



Solution:

For this problem, all we need know is the kinetic energy of the tire,

$$T = \frac{1}{2}I\dot{\theta}^2 + \frac{1}{2}m\dot{x}^2, \tag{E.3}$$

(I is the moment of inertia of the tire) and the fact that the tire rolls without slipping on the pavement:

$$x = R\theta \tag{E.4}$$

In writing (E.3) I am anticipating the result that a rigid body's kinetic energy can always be written as translational plus rotational energies. The E-L equation in θ is just

$$\frac{\partial L}{\partial \theta} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) + \Im_\theta = 0$$

$$\Rightarrow \Im_\theta = \frac{I\ddot{x}}{R}, \tag{E.5}$$

where I have used the constraint equation, (E.4), after the partials are taken. Clearly, the generalized force, $\Im_\theta$, is just the torque exerted on the wheel, the force being given by $R^{-1}\Im_\theta$, R being the scale factor, S. This simple result is applicable to other situations once the instantaneous acceleration $\ddot{x}$ is solved for (often by solving additional E-L equations).

Endpoint Invariance: Multiparticle Conservation Laws

I had mentioned in the last Chapter that the endpoint variation considerations there would be useful in uncovering conserved quantities. There, we were assuming invariance of the action, S, under endpoint variations and finding the consequences. Here we will need to directly show this invariance as the first step. This will be done for a model system.

First, let us examine two types of endpoint variations. Then, from (6.83) and (6.87) we can write for the case of s generalized coordinates, q_i (using our new "δ" notation),

$$\delta S = \delta_i S + \delta_e S, \tag{7.35}$$

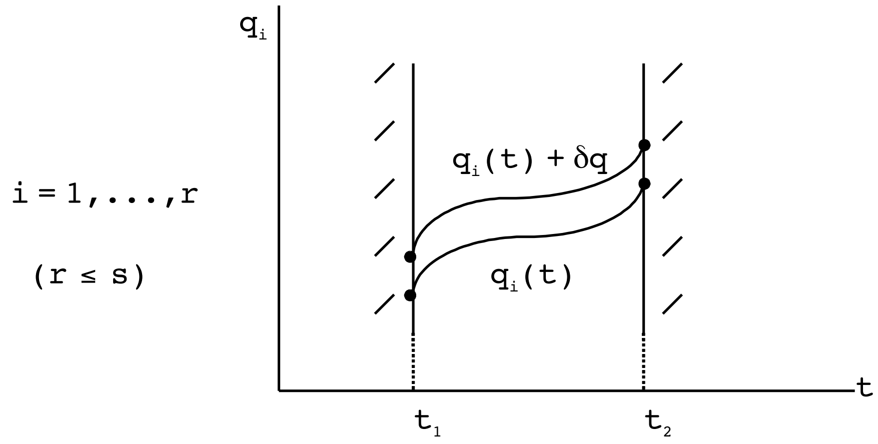
$$\delta_i S = \int_{t_1}^{t_2} dt \sum_{i=1}^{s} \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i} \right) \right) \delta q_i(t), \tag{7.36}$$

$$\delta_e S = \left[\left(L - \sum_{i=1}^{s} \dot{q}_i \frac{\partial L}{\partial \dot{q}_i}\right)\delta t + \sum_{i=1}^{s} \frac{\partial L}{\partial \dot{q}_i}\delta q_i\right]\Bigg|_{t_1}^{t_2} . \tag{7.37}$$

"$\delta_i S$" denotes the interior contribution and "$\delta_e S$" represents the endpoint part. We will assume that the E-L equations hold in the interior. Thus, although $\delta q_i(t) \neq 0$ one has $\delta_i S=0$ from (7.36).

Let us look at two specific types of varied paths.

Case 1:



In this case we are looking at the special case of variations, at endpoints as well as interior points,

$$\left.\begin{array}{l} \delta q_i = \delta q\ ,\quad (i = 1,\ldots,r) \\ \delta t = 0\ , \end{array}\right\} \tag{7.38}$$

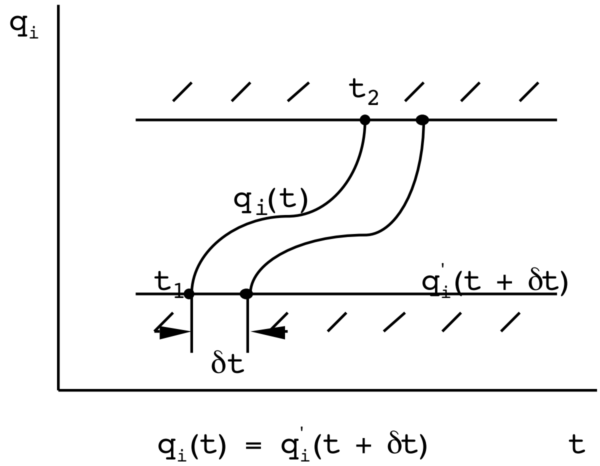
for a subset ($r \le s$) of the original s coordinates. δq is just a constant, infinitesimal parameter. Then, for this type of variation

$$\delta_e^{(1)} S = \delta q \sum_{i=1}^{r} \left(\frac{\partial L}{\partial \dot{q}_i}\right)\Bigg|_{t_1}^{t_2} . \tag{7.39}$$

The second type of variation is illustrated below.

Case 2:

$i = 1, \ldots, s$



In this case at the endpoints,

$$\left.\begin{aligned} &\delta q_i = 0, \quad (i = 1, \ldots, s) \\ &\delta t \neq 0. \end{aligned}\right\} \tag{7.40}$$

Therefore,

$$\delta_e^{(2)} S = \delta t \left(L - \sum_{i=1}^{s} \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} \right) \Bigg|_{t_1}^{t_2}. \tag{7.41}$$

Now let now apply these considerations to our model Lagrangian from the beginning of this Chapter. It was $(\alpha = 1, \ldots, n;\ i = 1, 2, 3)$

$$L\left(x_{\alpha i}, \dot{x}_{\alpha i}\right) = T - U, \tag{7.42}$$

$$U = \sum_{\alpha < \beta} \bar{U}_{\alpha\beta}\left(\left\| \vec{x}_\alpha - \vec{x}_\beta \right\|\right), \tag{7.43}$$

$$T = \frac{1}{2} \sum_{i,\alpha} m_\alpha\, \dot{x}_{\alpha i}^2, \tag{7.44}$$

We are assuming the E-L equations for this system hold. We notice that L and therefore the E-L equations do not change under the translation of coordinates,

$$\left.\begin{aligned} &x_{\alpha i} \to x_{\alpha i} + \delta x_i \quad (\text{all } \alpha) \\ &\qquad t \to t, \end{aligned}\right\} \tag{7.45}$$

where the δx_i are three arbitrary, infinitesimal, time-independent parameters. This is a "Case 1" variation with $\delta x_{\alpha i} = \delta x_i$, $\delta t = 0$. Moreover, the total variation in the action,

$$\delta^{(1)} S = \int_{t_1}^{t_2} dt \left\{ L\left(x_{\alpha i} + \delta x_i, \dot{x}_{\alpha i}\right) - L\left(x_{\alpha i}, \dot{x}_{\alpha i}\right) \right\}, \tag{7.46}$$

vanishes by the invariance of L. We thus find

$$\delta^{(1)}S = \delta_e^{(1)}S = 0 \,, \tag{7.47}$$

where

$$\delta_e^{(1)}S = \sum_i \delta x_i \sum_{\alpha=1}^{n} \left(\frac{\partial L}{\partial \dot{x}_{\alpha i}} \right) \Bigg|_{t_1}^{t_2} . \tag{7.48}$$

Since the δx_i are arbitrary, this means

$$\sum_\alpha \frac{\partial L}{\partial \dot{x}_{\alpha i}} \bigg|_{t_1} = \sum_\alpha \frac{\partial L}{\partial \dot{x}_{\alpha i}} \bigg|_{t_2} , \tag{7.49}$$

or

$$P_i\Big|_{t_1} = P_i\Big|_{t_2} , \tag{7.50}$$

where

$$P_i \equiv \sum_{\alpha=1}^{n} m_\alpha \dot{x}_{\alpha i} \,, \tag{7.51}$$

is the system's total linear momentum. Linear momentum is conserved as a consequence of invariance of L under translations.

Now let us write this same Lagrangian in cylindrical coordinates (ρ, θ, z) The kinetic energy is

$$T = \frac{1}{2} \sum_\alpha m_\alpha \left(\dot{\rho}_\alpha^2 + \rho_\alpha^2 \dot{\theta}_\alpha^2 + \dot{z}_\alpha^2 \right) . \tag{7.52}$$

This time we notice that L is unchanged under the rotation,

$$\left.\begin{aligned} \theta_\alpha &\to \theta_\alpha + \delta\theta \,, \qquad (\text{all } \alpha) \\ t &\to t \,, \end{aligned}\right\} \tag{7.53}$$

where $\delta\theta$ is a constant parameter, while keeping ρ_α and z_α fixed. This is again a "Case 1" variation with $\delta\theta_\alpha = \delta\theta$, $\delta t = 0$. As before, the total variation vanishes as a consequence of the invariance of L, showing that $\delta_e^{(1)}S = 0$ The consequence this time is that

$$\sum_\alpha \frac{\partial L}{\partial \dot{\theta}_\alpha} \bigg|_{t_1} = \sum_\alpha \frac{\partial L}{\partial \dot{\theta}_\alpha} \bigg|_{t_2} , \tag{7.54}$$

or

$$\sum_\alpha m_\alpha \rho_\alpha^2 \dot{\theta}_\alpha \Big|_{t_1} = \sum_\alpha m_\alpha \rho_\alpha^2 \dot{\theta}_\alpha \Big|_{t_2} , \tag{7.55}$$

expressing conservation of total angular momentum about a particular (arbitrary) axis. Angular momentum is conserved as a consequence of invariance of L under rotations. Although I used a specific Lagrangian in this argument, any Lagrangian which is invariant under spatial translations will conserve total linear momentum, and any Lagrangian invariant under rotations will conserve total angular momentum.

Finally, we notice that the model Lagrangian (7.44) is also unchanged under the substitutions

$$\left.\begin{aligned} x_{\alpha i} &\to x_{\alpha i}, \\ t &\to t + \delta t. \end{aligned}\right\} \Rightarrow \begin{array}{c} x_{\alpha i}(t) = x'_{\alpha i}(t + \delta t), \\ (\text{all } \alpha, i) \end{array} \tag{7.56}$$

where dt is again a constant, infinitesimal parameter. this is a "Case 2" variation with $\delta x_{\alpha i} = 0$. The total variation in S is,

$$\begin{aligned} \delta S &= \int_{t_1+\delta t}^{t_2+\delta t} dt' L\left(x'_{\alpha i}, \dot{x}'_{\alpha i}, t'\right) - \int_{t_1}^{t_2} dt\, L\left(x_{\alpha i}, \dot{x}_{\alpha i}, t\right) \\ &= \int_{t_1}^{t_2} dt \left\{L\left(x'_{\alpha i}\left(t + \delta t\right), \dot{x}'_{\alpha i}\left(t + \delta t\right), t + \delta t\right) - L\left(x_{\alpha i}(t), \dot{x}_{\alpha i}(t), t\right)\right\}, \\ &= \delta t \int_{t_1}^{t_2} dt \frac{\partial L}{\partial t}, \end{aligned} \tag{7.57}$$

which vanishes under (7.56) above as long as the time does not appear explicitly in L. This means

$$\delta^{(2)} S = \delta_e^{(2)} S = 0, \tag{7.58}$$

and therefore

$$\left(L - \sum_{i,\alpha} \dot{x}_{\alpha i} \frac{\partial L}{\partial \dot{x}_{\alpha i}}\right)\Bigg|_{t_1} = \left(L - \sum_{i,\alpha} \dot{x}_{\alpha i} \frac{\partial L}{\partial \dot{x}_{\alpha i}}\right)\Bigg|_{t_2}. \tag{7.59}$$

we define the "Hamiltonian" as

$$H \equiv \sum_{i,\alpha} \dot{x}_{\alpha i} \frac{\partial L}{\partial \dot{x}_{\alpha i}} - L. \tag{7.60}$$

The above implies that H is conserved:

$$\frac{dH}{dt} = 0. \tag{7.61}$$

In our case, not surprisingly

$$\frac{\partial L}{\partial \dot{x}_{\alpha i}} = m_\alpha \dot{x}_{\alpha i}, \tag{7.62}$$

$$\Rightarrow H = 2T - (T - U) = T + U \tag{7.63}$$

(It can also happen that $\frac{dH}{dt} = 0$ but that $H \neq T + U$.) Again, although I have used a model Lagrangian, any L which is invariant under time translations ($\frac{\partial L}{\partial t} = 0$) will have a conserved H. Notice that (7.59) says

$$\sum_{i,\alpha} \dot{x}_{\alpha i} \frac{\partial L}{\partial \dot{x}_{\alpha i}} - L = \text{const.}. \tag{7.64}$$

which is just a statement of the integrated form of the E-L equations, with a change in notation from the last Chapter. The condition for this to hold was $\frac{\partial L}{\partial t} = 0$, consistent with my observations above.

We may write the endpoint variation for a given Lagrangian, $L(q_i, \dot{q}_i, t)$, as $(i = 1,\ldots,s)$

$$\delta_e S = \left[-H\delta t + \sum_i p_i \delta q_i\right], \tag{7.65}$$

where the general definitions are

$$H \equiv \sum_i p_i \dot{q}_i - L, \tag{7.66}$$

$$p_i \equiv \frac{\partial L}{\partial q_i}. \tag{7.67}$$

(Note: the p_i are generalized "canonical" momenta, which may not correspond to actual, physical momenta.) We may repeat the argument in generalized coordinates to show:

"Case 1": variation

$$\text{If } \frac{\partial L}{\partial q_i} = 0 \quad \Rightarrow \quad p_i = \text{const.}, \tag{7.68}$$

and that

"Case 2": variation

$$\text{If } \frac{\partial L}{\partial t} = 0 \quad \Rightarrow \quad H = \text{const.} \tag{7.69}$$

These statements may also be shown directly from the E-L equations (representing Newton's equations) without invoking the idea of endpoint variations of S.

Consequences of Scale Invariance

Before I go on, I would like to point out that there may be other types of invariances in situations that can give us information about the system without actually solving the dynamical equations. Let's take the example

$$U(\bar{r}) = Cr^n \quad , \quad (r = |\bar{x}|) \tag{7.70}$$

representing a particle in an external potential. Let's let ("scaling")

$$\bar{r} \to \alpha\bar{r}. \qquad (\alpha \text{ a real number}) \tag{7.71}$$

then

$$U(\alpha\bar{r}) = \alpha^n U(\bar{r}). \tag{7.72}$$

Let's try to make T, the kinetic energy, transform similarly.

$$T = \frac{1}{2} m\dot{\bar{r}}^2 . \tag{7.73}$$

$$\bar{r} \to \alpha\bar{r}, t \to \alpha^{1-n/2} t \Rightarrow T \to \alpha^2 \alpha^{-2+n} T.$$

Therefore $L = T - U$

$$L \to \alpha^n L . \tag{7.74}$$

Since L is changed only by an overall constant, the equations of motion from

$$\delta \int L\, dt = 0 \tag{7.75}$$

will be unchanged! Let's look at three cases to see what specifically this implies. (Works for any portion of a trajectory)

(i) Kepler problem, $n = -1$.

$$\frac{\tau'}{\tau} = \alpha^{3/2} = \left(\frac{a'}{a}\right)^{3/2}$$

$$\Rightarrow \quad \tau^2 \propto a^3 , \qquad \text{Kepler's } 3^{\underline{rd}} \text{ law!}$$

(ii) Free fall, $n = 1$.

$$\frac{\tau'}{\tau} = \alpha^{1/2} = \left(\frac{h'}{h}\right)^{1/2}$$

$$\Rightarrow \tau \propto \sqrt{h}. \qquad \left(h = \frac{1}{2}at^2 \Rightarrow t = \sqrt{\frac{2h}{a}}.\right) \quad \surd$$

(iii) Simple harmonic oscillator, $n = 2$.

$$\frac{\tau'}{\tau} = \alpha^0 = 1$$

$$\Rightarrow \tau \text{ independent of amplitude. } \surd$$

When Does $H=T+U$?

I mentioned above that H is generally not just $T + U$. Under which conditions does the equality hold? We have

$$\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial T}{\partial \dot{q}_i} - \frac{\partial U}{\partial \dot{q}_i}. \tag{7.76}$$

Now given

$$T = \frac{1}{2}\sum_{\alpha,i} m_\alpha \dot{x}_{\alpha i}^2 , \tag{7.77}$$

$$x_{\alpha i} = x_{\alpha i}\left(q_j , t\right) , \tag{7.78}$$

we have

$$\dot{x}_{\alpha i} = \sum_j \frac{\partial x_{\alpha i}}{\partial q_j}\dot{q}_j + \frac{\partial x_{\alpha i}}{\partial t} . \tag{7.79}$$

Then the kinetic energy has the general form,

$$T = \sum_{j,k} a_{jk}\left(q_i , t\right)\dot{q}_j\dot{q}_k + \sum_j b_j\left(q_i , t\right)\dot{q}_j + c\left(q_i , t\right) , \tag{7.80}$$

$$a_{jk} \equiv \frac{1}{2}\sum_{\alpha,i} m_\alpha \frac{\partial x_{\alpha i}}{\partial q_j}\frac{\partial x_{\alpha i}}{\partial q_k} , \tag{7.81}$$

$$b_j \equiv \sum_{\alpha,i} m_\alpha \frac{\partial x_{\alpha i}}{\partial q_j}\frac{\partial x_{\alpha i}}{\partial t} , \tag{7.82}$$

$$c \equiv \frac{1}{2}\sum_{\alpha,i} m_\alpha \left(\frac{\partial x_{\alpha i}}{\partial t}\right)^2 . \tag{7.83}$$

Note that $a_{jk} = a_{kj}$ Now we have

$$\frac{\partial T}{\partial \dot{q}_i} = \sum_{j,k} a_{jk}\left(\delta_{ji}\,\dot{q}_k + \dot{q}_j\,\delta_{ki}\right) + b_i ,$$

$$= 2\sum_k a_{ik}\,\dot{q}_k + b_i , \tag{7.84}$$

which means that,

$$\sum_i \frac{\partial T}{\partial \dot{q}_i}\dot{q}_i = 2\sum_{j,k} a_{jk}\,\dot{q}_k\dot{q}_j + \sum_j b_j\dot{q}_j ,$$

and using (7.80),

$$\sum_i \frac{\partial T}{\partial \dot{q}_i}\dot{q}_i = 2T - \sum_j b_j\,\dot{q}_j - 2c . \tag{7.85}$$

Finally, therefore from (7.66),

$$H = \underbrace{2T - (T - U)}_{T + U} - \sum_i \left(\frac{\partial U}{\partial \dot{q}_i} + b_i\right)\dot{q}_i - 2c \tag{7.86}$$

Thus, the conditions under which $H = T + U$ are quite different from that for $H = const$. A sufficient set of conditions for $H = T + U$ is seen to be $\frac{\partial x_{\alpha i}}{\partial t} = 0$ and $\frac{\partial U}{\partial \dot{q}_i} = 0$

Investigation into the Meaning of $\frac{dE}{dt} = 0$ (Optional)

One can separately inquire as to the conditions giving $\frac{dE}{dt} = 0$, where $E \equiv T + U$ For convenience in the following, I will assume that $\frac{\partial x_{\alpha i}}{\partial t} = 0$. We have

$$E = \sum_{jk} a_{jk}\,\dot{q}_j\dot{q}_k + U\left(q_i, \dot{q}_i, t\right),$$

"Term (1)"

"Term (1)"

$$\Rightarrow \frac{dE}{dt} = 2\sum_{j,k} a_{jk}\,\dot{q}_j\ddot{q}_k + \sum_{j,k} \frac{\partial a_{jk}}{\partial q_l}\,\dot{q}_j\dot{q}_k\dot{q}_l + \frac{dU}{dt}. \tag{7.88}$$

However, the E-L equation is explicitly given by

"Term (2)" "Term (3)"

$$\frac{\partial\left(T-U\right)}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial\left(T-U\right)}{\partial \dot{q}_i}\right) = 0 \tag{7.89}$$

"Term (2)": $$\frac{\partial T}{\partial q_i} = \sum_{j,k}\frac{\partial a_{jk}}{\partial q_i}\,\dot{q}_j\dot{q}_k \tag{7.90}$$

"Term (3)": $$-\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_i}\right) = -\frac{d}{dt}\left(2\sum_k a_{ik}\dot{q}_k\right)$$

$$= -2\sum_k\left(\dot{a}_{ik}\dot{q}_k + a_{ik}\ddot{q}_k\right)$$

$$= -2\sum_{k,\ell}\frac{\partial a_{ik}}{\partial q_l}\,\dot{q}_l\dot{q}_k - 2\sum_k a_{ik}\,\ddot{q}_k \tag{7.91}$$

Substituting in (7.89) gives,

"Term (4)"

$$\sum_{j,k}\frac{\partial a_{jk}}{\partial q_i}\,\dot{q}_j\dot{q}_k - \frac{\partial U}{\partial q_i} - 2\sum_{k,l}\frac{\partial a_{ik}}{\partial q_\ell}\,\dot{q}_\ell\dot{q}_k - 2\sum_k a_{ik}\,\ddot{q}_k + \frac{d}{dt}\left(\frac{\partial U}{\partial \dot{q}_i}\right) = 0 \tag{7.92}$$

Isolating Term (4) above, multiplying by $\dot{q}_i$ throughout and summing on I gives:

$$2\sum_{i,k} a_{ik}\,\ddot{q}_k\dot{q}_i = -\sum_{i,j,k}\frac{\partial a_{jk}}{\partial q_i}\,\dot{q}_j\dot{q}_k\dot{q}_i - \sum_i\left(\frac{\partial U}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial U}{\partial \dot{q}_i}\right)\right)\dot{q}_i. \tag{7.93}$$

This is just an alternate form for Term (1) above. Substituting, (7.93) into (7.88), we now have

$$\frac{dE}{dt} = -\sum_i \left(\frac{\partial U}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial U}{\partial \dot{q}_i}\right)\right)\dot{q}_i + \frac{dU}{dt}. \tag{7.94}$$

The chain rule on $\frac{dU}{dt}$ gives

$$\frac{dU}{dt} = \frac{\partial U}{\partial t} + \sum_i \frac{\partial U}{\partial q_i}\dot{q}_i + \sum_i \frac{\partial U}{\partial \dot{q}_i}\ddot{q}_i \tag{7.95}$$

resulting in our final form,

$$\frac{dE}{dt} = \frac{\partial U}{\partial t} + \frac{d}{dt}\left(\sum_i \frac{\partial U}{\partial \dot{q}_i}\dot{q}_i\right) \tag{7.96}$$

(If we had not assumed $\frac{\partial x_{\alpha i}}{\partial t} = 0$, there would be additional non-canceling terms on the right hand side of (7.96).) So, the condition for $E = \text{const.}$ is much less general than the condition for $H = \text{const.}$ $\left(\frac{\partial L}{\partial t} = 0\right)$ A sufficient set of conditions for $\frac{dE}{dt} = 0$ are $\frac{\partial x_{\alpha i}}{\partial t} = 0$, $\frac{\partial U}{\partial t} = 0$ and $\frac{\partial U}{\partial \dot{q}_i} = 0$. Not surprisingly, this represents a union of the sufficient conditions for $H = T + U$ $\left(\frac{\partial x_{\alpha i}}{\partial t} = 0, \quad \frac{\partial U}{\partial \dot{q}_i} = 0\right)$ and those for H = const. $\left(\frac{\partial L}{\partial t} = 0 \Rightarrow \frac{\partial U}{\partial t} = 0 \text{ since } \frac{\partial x_{\alpha i}}{\partial t} = 0\right)$. H is undoubtedly a more useful quantity in mechanics than E.

Hamilton's Equations

There is another formulation of mechanics which also comes from Hamilton's principle. Lagrange's equations are second order in time derivatives. But mathematically, we know that a single second order differential equation can always be replaced by two first order equations. This is a hint that another formulation of mechanics exists which yields first order differential equations.

Let's say we have found a "proper" (unconstrained) set of generalized coordinates, q_i. The key to the new formulation is the replacement of the $q_i(t)$ as the dependent generalized variables, with a set which has twice as many members:

Lagrangian	Hamiltonian
$L(\underbrace{q_i, \dot{q}_i}; t)$	$H(\underbrace{q_i, p_i}; t)$
related by	independently varied
$\delta\dot{q}_i = \frac{d}{dt}\delta q_i$	$\left(p_i \equiv \frac{\partial L}{\partial \dot{q}_i}\right)$

As pointed out above, the "canonical momenta" are in general not "real" momenta. The reasons are:

1. The q_i are not generally rectangular coordinates.
2. The potential, U, may be velocity dependent. Then, even if we use a rectangular coordinate, x, and $T = \frac{1}{2}m\dot{x}^2$,

$$p_x = \frac{\partial L}{\partial \dot{x}} = m\dot{x} - \frac{\partial U}{\partial \dot{x}}. \tag{7.97}$$

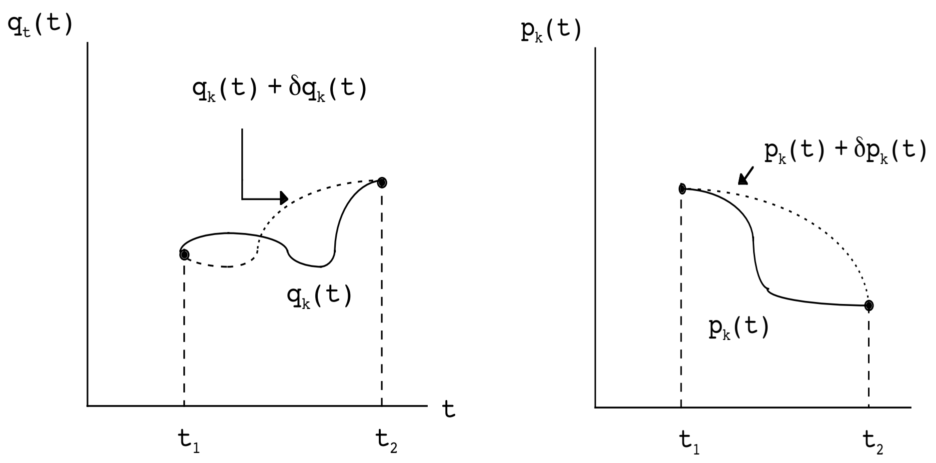
Keeping these facts in mind, the Hamiltonian is

$$H = \sum_k p_k \dot{q}_k - L,$$

where $\dot{q}_k$ is considered a function of q_k and p_k. Solving for L in terms of H, the action is

$$S = \int_{t_1}^{t_2} dt \left[\sum_k p_k \dot{q}_k - H\left(p_k, q_k, t\right)\right]. \tag{7.98}$$

We will look for the stationary path, this time by varying p_k, q_k simultaneously and independently.



$\delta q_k(t_{1,2}) = 0$ $\delta p_k(t_{1,2}) = 0$

Variation of the action gives

$$\delta S = \int_{t_1}^{t_2} dt \sum_k \left[\delta p_k \dot{q}_k + p_k \, \delta\dot{q}_k - \frac{\partial H}{\partial p_k} \delta p_k - \frac{\partial H}{\partial q_k} \delta q_k \right]. \tag{7.99}$$

Now

$$\delta\dot{q}_k = \frac{d}{dt} \delta q_k \, , \tag{7.100}$$

as before (it is q_k and p_k which vary independently, not q_k and $\dot{q}_k$), so

$$\int_{t_1}^{t_2} dt \, p_k \, \delta\dot{q}_k = p_k \underbrace{\delta q_k}_{0} \Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} dt \, \dot{p}_k \, \delta q_k \, ,$$

$$\Rightarrow \delta S = \int_{t_1}^{t_2} dt \sum_k \left[\delta p_k \left(\dot{q}_k - \frac{\partial H}{\partial p_k} \right) - \delta q_k \left(\dot{p}_k + \frac{\partial H}{\partial q_k} \right) \right] \tag{7.101}$$

We require

$$\delta S = 0 \tag{7.102}$$

which results in $k = 1,\ldots,s$

$$\dot{q}_k = \frac{\partial H}{\partial p_k}, \quad \text{(gives } p_k, \dot{q}_k \text{ connection: } \dot{q}_k = \dot{q}_k(p_i, q_i)) \tag{7.103}$$

$$\dot{p}_k = -\frac{\partial H}{\partial q_k}, \quad \text{(dynamical part)} \tag{7.104}$$

These $2s$ set of first order differential equations are called **Hamilton's equations**. They replace the Lagrange equations, which may be written as

$$\dot{p}_k = \frac{\partial L}{\partial q_k} \tag{7.105}$$

Hamilton's formulation is not as useful, practically speaking, as the E-L formulation. However, they are just as important, if not more so, theoretically speaking. Hamiltonian methods permeate all active research areas in theoretical physics and help provide a formalism bridge from classical mechanics to quantum mechanics.

First of all, let's make sure that Hamilton's equations imply the E-L equation, (7.105), above. Notice that the first equation, (7.103), reduces to an identity

$$\dot{q}_k = \frac{\partial}{\partial p_k}\left(\sum_i p_i \dot{q}_i - L(q_i, \dot{q}_i, t)\right)$$

$$= \dot{q}_k + \sum_i p_i \frac{\partial \dot{q}_i}{\partial p_k} - \sum_i \underbrace{\frac{\partial L}{\partial \dot{q}_i}}_{p_i} \frac{\partial \dot{q}_i}{\partial p_k},$$

↖ cancel ↗

$$= \dot{q}_k \qquad \surd \tag{7.106}$$

The second equation, (7.104), now gives

$$\dot{p}_k = -\frac{\partial}{\partial q_k}\left(\sum_i p_i\dot{q}_i - L(q_i, \dot{q}_i, t)\right)$$
$$= -\sum_i p_i \frac{\partial \dot{q}_i}{\partial q_k} + \frac{\partial L}{\partial q_k} + \sum_i \underbrace{\frac{\partial L}{\partial \dot{q}_i}}_{p_i} \frac{\partial \dot{q}_i}{\partial q_k},$$

↖ cancel ↗

$$\Rightarrow \dot{p}_k = \frac{\partial L}{\partial q_k}. \qquad \surd \tag{7.107}$$

In this form, it is particularly convenient to calculate $\frac{dH}{dt}$ directly. We find

$$\frac{dH}{dt} = \sum_i \left(\frac{\partial H}{\partial q_i}\dot{q}_i + \frac{\partial H}{\partial p_i}\dot{p}_i\right) + \frac{\partial H}{\partial t} \tag{7.108}$$

Using Hamilton's equations for $\dot{q}_i$ and $\dot{p}_i$ now shows that the two terms in the first sum cancel, and we have

$$\frac{dH}{dt} = \frac{\partial H}{\partial t}. \tag{7.109}$$

Also notice, using the definition of H and the E-L equations,

$$\frac{dH}{dt} = \sum_i (\dot{p}_i\dot{q}_i + p_i\ddot{q}_i)$$

↖ cancel ↘ ↖ cancel ↘

$$-\sum_i \left(\frac{\partial L}{\partial q_i}\dot{q}_i + \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i\right) - \frac{\partial L}{\partial t}. \tag{7.110}$$

$$\Rightarrow \frac{dH}{dt} = -\frac{\partial L}{\partial t}. \tag{7.111}$$

Thus, if $H \neq H(t)$ (or equivalently) $L \neq L(t)$) then $H = \text{const}$.

Here is a cookbook summary of Hamilton's method:

Step 1: Find $T(q_i, \dot{q}_i, t)$ and $U(q_i, \dot{q}_i, t)$ and form $L(q_i, \dot{q}_i, t) = T - U$ in the usual way.

Step 2: Solve for the generalized momenta from

$$p_j = \frac{\partial L}{\partial \dot{q}_j}.$$ (terminology: p_j is "canonically conjugate" to q_j.)

Step 3: Form H from

$$H = \sum_i p_i \dot{q}_i - L(q_i, \dot{q}_i, t),$$

remembering that $\dot{q}_i = \dot{q}_i(q_j, p_j, t)$ in general.

Step 4: Write down Hamilton's equations,

$$\dot{q}_k = \frac{\partial H}{\partial p_k}, \qquad \text{(definition)}$$

$$\dot{p}_k = -\frac{\partial H}{\partial q_k}. \qquad \text{(dynamics)}$$

There is a shortcut possible in this procedure. If one can establish that $H = T + U$ (see (7.86) above), then we need only eliminate the $\dot{q}_i$ variables in terms of the p_i and q_i when constructing H. Notice that if a particular coordinate, q_k, does not appear explicitly in H, we have

$$\dot{p}_k = -\frac{\partial H}{\partial q_k} = 0, \qquad (7.112)$$

which implies the momentum conjugate to qk is conserved. (The same conclusion also holds in Lagrangian mechanics when $L \neq L(q_k)$ Such coordinates are called "cyclic."

Let's look at two examples of this process.

Example 7

Formulate the undamped pendulum in Hamiltonian mechanics.

Solution:

We will use the angle, θ, as the generalized coordinate as before. I have written down the transformation equations between x, y and θ in Example 2 above. These equations do not explicitly involve the time. In addition, the gravitational potential is velocity independent, which means from (7.86) that

$$H = T + U,$$

$$H = \frac{1}{2} m\ell^2\dot{\theta}^2 + mg\ell(1 - \cos\theta).$$

$\dot{\theta}$ is not an appropriate variable to have in H. We eliminate it by the definition,

$$p_\theta \equiv \frac{\partial L}{\partial \dot{\theta}} = m\ell^2\dot{\theta},$$

$$\Rightarrow H = \frac{p_\theta^2}{2m\ell^2} + mg\ell(1 - \cos\theta).$$

Hamilton's equations are now

$$\dot{\theta} = \frac{\partial H}{\partial p_\theta} = \frac{p_\theta}{m\ell^2}, \text{ (definition)}$$

and

$$\dot{p}_\theta = -\frac{\partial H}{\partial \theta} = -mg\ell\sin\theta.$$

Of course we can recover the correct second order equation,

$$\ddot{\theta} = \frac{\dot{p}_\theta}{m\ell^2} = -\frac{g}{\ell}\sin\theta,$$

by combining the two. You can see the Lagrangian technique is more direct in this case.

Example 8

Given the (nonrelativistic) Lagrangian for a particle (charge q, mass m) in an external electromagnetic field (using Gaussian units),

$$L = T - U$$

$$T = \frac{1}{2}m\sum_i \dot{x}_i^2, \quad U = q\phi(\vec{x}, t) - \frac{q}{c}\sum_{i=1}^{3}\dot{x}_i A_i\left(\vec{x}, t\right),$$

($\phi(\vec{x}, t)$: "scalar potential", $\vec{A}(\vec{x}, t)$: "vector potential"). Derive the Hamiltonian and the Hamiltonian equations of motion. Is $H = T + U$? Is H a constant in time?

Solution:

Notice that the potential is velocity-dependent, so we can not write $H = T + U$ as in the last example. The canonical momenta are given by

$$p_j = \frac{\partial L}{\partial \dot{x}_j} = m\dot{x}_j + \frac{q}{c}A_j$$

We now construct H as

$$H = \sum_i p_i \dot{x}_i - L,$$

$$= \sum_i p_i \left(\frac{p_i}{m} - \frac{q}{cm} A_i\right) - \frac{1}{2} m \sum_i \left(\frac{p_i}{m} - \frac{q}{cm} A_i\right)^2$$

$$+ q\phi - \frac{q}{c} \sum_i A_i \left(\frac{p_i}{m} - \frac{q}{cm} A_i\right).$$

After the algebra settles down, we find

$$H = \frac{1}{2m} \sum_i \left(p_i - \frac{q}{c} A_i\right)^2 + q\phi,$$

which confirms that $H \neq T + U$. Hamilton's equations are

$$\dot{x}_k = \frac{\partial H}{\partial p_k} = \frac{1}{m}\left(p_k - \frac{q}{c} A_k\right), \quad \text{(definition)}$$

and

$$\dot{p}_k = -\frac{\partial H}{\partial x_k} = \frac{q}{mc} \sum_i \left(p_i - \frac{q}{c} A_i\right) \frac{\partial A_i}{\partial x_k} - q \frac{\partial \varphi}{\partial x_k}.$$

These equations are a disguised form of the familiar Lorentz force equation for charged particles. To see this, form

$$m\ddot{x}_k = \dot{p}_k - \frac{q}{c} \dot{A}_k,$$

$$= \frac{q}{mc} \sum_i \overbrace{\left(p_i - \frac{q}{c} A_i\right)}^{m\dot{x}_i} \frac{\partial A_i}{\partial x_k} - q \frac{\partial \varphi}{\partial x_k}$$

$$- \frac{q}{c}\left(\frac{\partial A_k}{\partial t} + \sum_i \frac{\partial A_k}{\partial x_i} \dot{x}_i\right),$$

$$= q\left(-\nabla_k \varphi - \frac{1}{c}\frac{\partial A_k}{\partial t}\right) + \frac{q}{c} \sum_i \dot{x}_i \left(\nabla_k A_i - \nabla_i A_k\right).$$

Using the definition of the curl from Ch.1,

$$\left[\dot{\vec{x}} \times \left(\vec{\nabla} \times \vec{A}\right)\right]_i = \sum_{j,k} \varepsilon_{ijk}\, \dot{x}_j \left(\vec{\nabla} \times \vec{A}\right)_k$$

$$= \sum_{\substack{j,k\\ \ell,m}} \underbrace{\varepsilon_{ijk}\, \varepsilon_{k\ell m}}\, \dot{x}_j\, \nabla_\ell\, A_m\,,$$

$$\underset{k\ \text{sum}}{\rightarrow} \left(\delta_{i\ell}\delta_{jm} - \delta_{im}\delta_{j\ell}\right)$$

$$= \sum_j \dot{x}_j \left(\nabla_i A_j - \nabla_j A_i\right),$$

$$\Rightarrow m\ddot{x}_k = q\, E_k + \frac{q}{c}\left(\dot{\vec{x}} \times \vec{B}\right)_k, \text{ (Lorentz force law)}$$

where $E_k \equiv -\nabla_k \phi - \dfrac{1}{c}\dfrac{\partial A_k}{\partial t}$, (electric field)

$$B_k \equiv \left(\vec{\nabla} \times \vec{A}\right)_k. \text{ (magnetic field)}$$

Notice that $\dfrac{dH}{dt} \neq$ constant if ϕ and/or $\vec{A}$ have explicit time dependence.

Holonomic Constraints in Hamiltonian Formalism

Remember how we handled holonomic constraints in the Lagrangian formalism $k=1,\ldots,s$

$$\frac{\partial L^*}{\partial q_k} - \frac{d}{dt}\left(\frac{\partial L^*}{\partial \dot{q}_k}\right) = 0, \tag{7.113}$$

$$\begin{cases} L^* = T - U^*, \\ U^* = U - \sum_{i=1}^{m} \lambda_i g_i, \end{cases} \tag{7.114}$$

where the constraints were

$$g_i\left(q_j, t\right) = 0. \qquad (i = 1, \ldots, m) \tag{7.115}$$

The incorporation of the constraints as a modification of the original potential gives us a hint as to how to proceed in the Hamiltonian method. A natural hypothesis is that the constraint Hamiltonian equations are just $(k = 1,\ldots,s)$

$$\dot{q}_k = \frac{\partial H^*}{\partial p_k}, \tag{7.116}$$

$$\dot{p}_k = -\frac{\partial H^*}{\partial q_k}, \tag{7.117}$$

where

$$H^* \equiv H + \sum_{i=1}^{m} \lambda_i g_i ,$$

$$g_j(q_i, t) = 0. \tag{7.118}$$

(7.116), (7.117) and (7.118) represent $2s + m$ equations in $2s + m$ unknowns. We can easily verify that this formulation is correct by reducing this set to the equivalent Euler-Lagrange equations. We need only repeat the steps leading to (7.107) above. There are essentially no new aspects of this reduction beyond what was seen above. I'll leave it to you to show that these reduce to the modified **E-L** equations

$$\dot{p}_k = \frac{\partial L^*}{\partial q_k}, \tag{7.119}$$

where

$$p_k = \frac{\partial L^*}{\partial \dot{q}_k} = \frac{\partial L}{\partial \dot{q}_k}. \tag{7.120}$$

CHAPTER 7 PROBLEMS

1. Consider a system of two masses in one dimension, one (m_1) attached to a wall with a spring (spring constant k_1), the other mass (m_2) attached to the first mass by another spring (spring constant k_2).

   a) Formulate the Lagrangian of this system.
   b) Write down Lagrange's equations and show that one obtains the earlier Eqs. (3.62) of the notes.

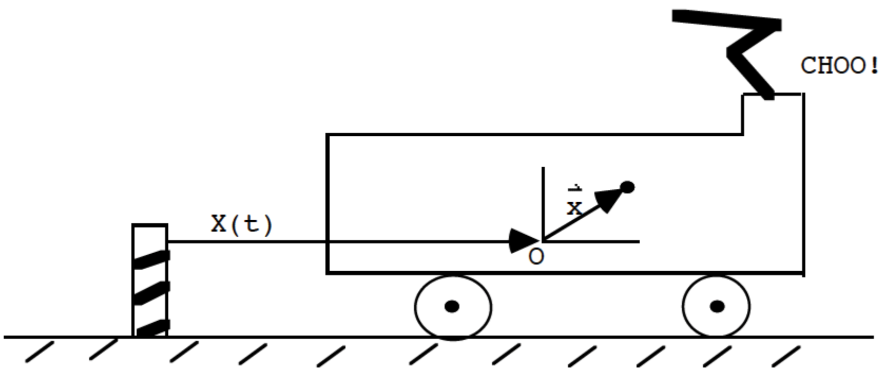
2. a) In a given inertial system ("the laboratory") the Lagrangian of an unconstrained particle is given by

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

   Write down the Euler-Lagrange equations of motion.

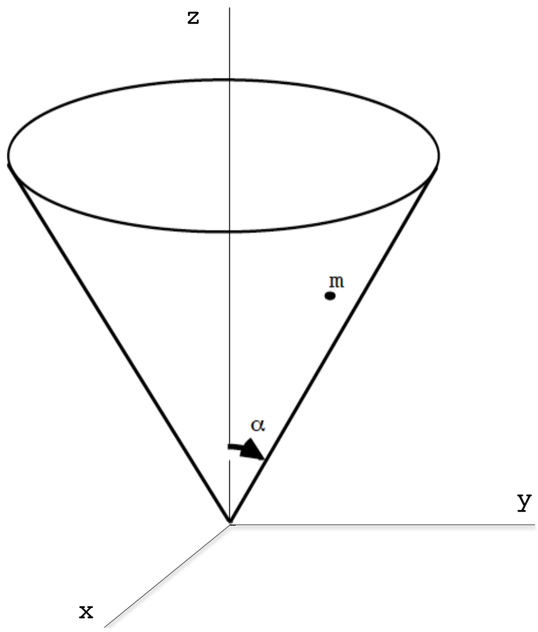
   b) The laboratory in (a) is inside a train, moving at a constant rate with respect to a fixed point on the Earth,

$$X(t) = vt, \; v = \text{const.}$$



   What is the Lagrangian of the particle, L in the Earth's frame of reference? Show that although $L' \neq L$, one obtains the same Euler-Lagrange equations as in (a).

3. A small particle of mass m is constrained to move on the surface of a cone with opening angle α as shown uner the influence of gravity in the z-direction.

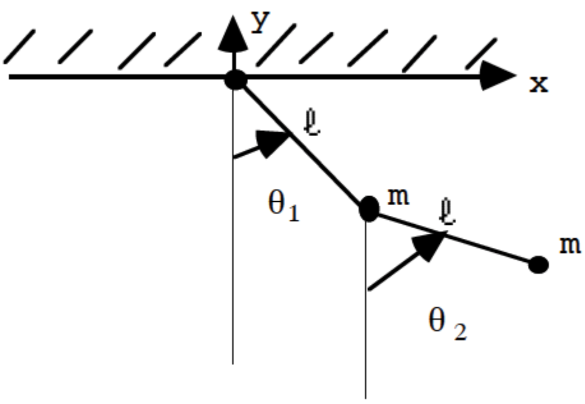


Find the Lagrangian and Lagrange's equations in some set of unconstrained variables. Show that these equations may be solved.

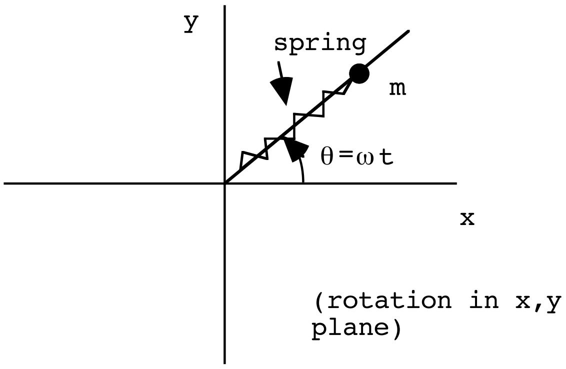
4. Show, using the transformation from (x_1,y_1) and (x_2,y_2) to θ_1 and θ_2 (see figure), that the kinetic and potential energies of the double pendulum can be written as

$$T = m\ell^2\dot{\theta}_1^2 + \frac{m}{2}\ell^2\dot{\theta}_2^2 + m\ell^2\dot{\theta}_1\dot{\theta}_2 \cos(\theta_1-\theta_2),$$

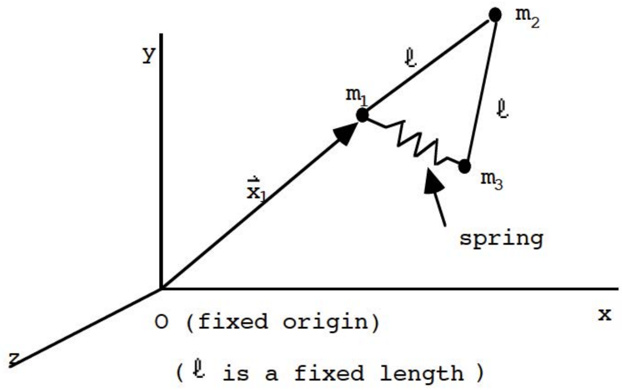
$$U = -mg\ell(2\cos\theta_1 + \cos\theta_2).$$



5. System 1:



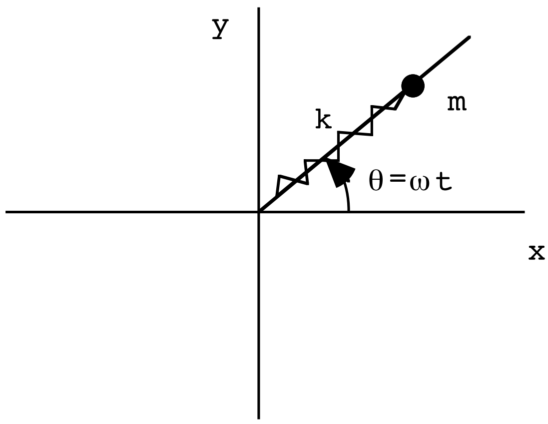
System 2:



(Note: all the masses above are point masses.) For the two systems depicted above:

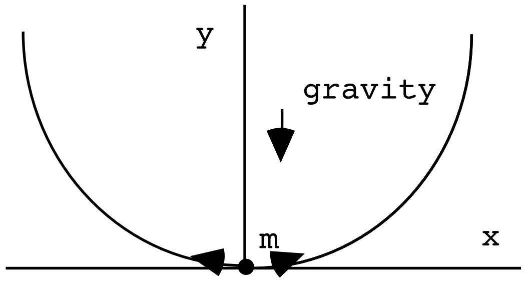
a) Explicitly write down all the constraints in the form $g_j(x_i,t)=0$ for both systems.
b) Find a set of generalized coordinates to describe each system. Explain your choice.

6. A point mass m is constrained to move on a thin wire while being rotated at a constant angular velocity. It is attached to the center of rotation by a massless spring with spring constant **k** and unstretched length ℓ_0. There is no gravitational potential energy in this problem.



a) Find the equilibrium distance of the mass from the center of rotation.
b) Write down the Euler-Lagrange equations for this system. Show that the motion is simple harmonic. What is the frequency of oscillations?

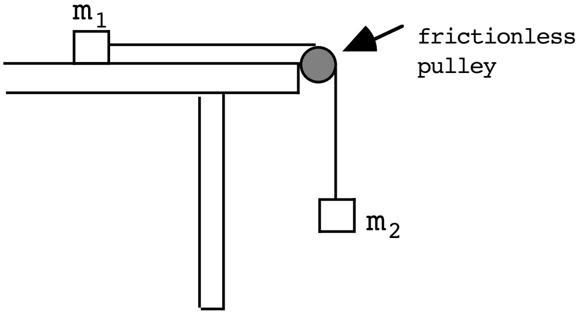
7. (Harkening back to prob.17, Ch.2) A particle of mass m on a frictionless wire follows the trajectory $y = Kx^2$ (K is a constant) in the xy plane under the influence of gravity.



a) Write the Lagrangian of the system, find the E-L equation and find the period of motion for small oscillations. [Hint: Drop terms that are small in the E-L equation and get it in harmonic oscillator form.]
b) By introducing a Lagrange multiplier, find the force of the mass on the wire in the y-direction. (The answer is:

$$-m\left(\frac{2K\dot{x}^2 + g}{1 + 2K^2x^2}\right).)$$

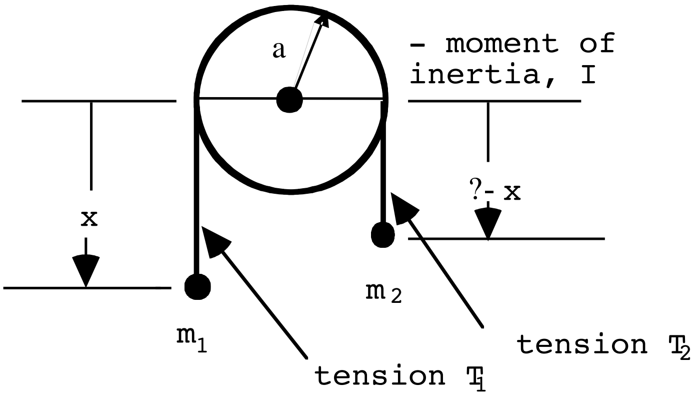
8. A particle of mass m_1 on a frictionless table is attached to another of mass m_2 by a rope. The mass m_2 is hanging off the table as shown ($m_1 \neq m_2$ in general).



Formulate this as a Lagrangian constraint problem.

a) Find the kinetic, potential energies and the Lagrangian equations of motion.
b) Find the acceleration of mass 2 and the tension in the rope.

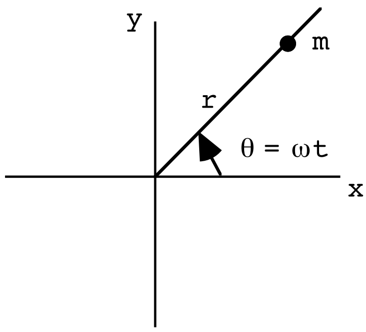
9. Consider:



The pulley is free to rotate either clockwise or counterclockwise. It's moment of inertia is I.

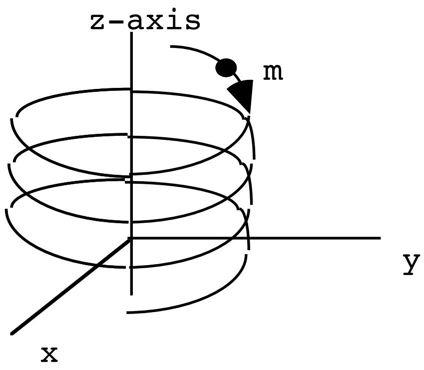
a) Set up this problem with an undetermined Lagrange multiplier to determine $\ddot{x}$ the acceleration of mass 1.
b) The Lagrange mutiplier will turn out to be related to the tension in the rope; it's exact meaning depends on how you formulate the problem. Actually, since the pulley is massive, the tension in the rope on the m_1 and m_2 sides are different. Find the tension on both sides of the rope, T_1 and T_2

10. Consider a mass m constrained to move on a wire that is being rotated at a constant angular velocity as shown (no gravity in this problem!):



Find the force of the wire on the mass. [Hint: Introduce a Lagrange multiplier, λ associated with the constraint.]

11. A particle of mass m moves under the influence of gravity along the spiral $z = k\theta$ $r = const.$ where k is a constant and z is the vertical axis.



a) Obtain the Hamiltonian equations of motion (use cylindrical coordinates $r,\ \theta,\ z.$).

b) Finish this equation: $\ddot{z} = ??$

12. Often the KE is a quadratic function of the generalized coordinate,

$$T = f(q)\dot{q}^2,$$

where $f(q)$ is some function of the generalized coordinate, q. Often, too, the PE does not depend on $\dot{q}$:

$$U = U(q,t).$$

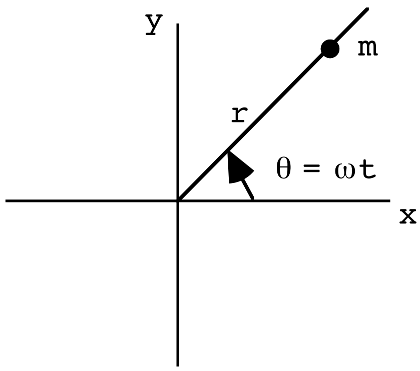
Construct the Hamiltonian, H. Under these circumstances is H equal to the total energy? Is it a constant of the motion? Explain.

13. Given the general form of the kinetic and potential energies of a system of particles ("α" is the particle label, "i" is the axis label),

$$T = \frac{1}{2}\sum_{\alpha,i} m_{\alpha i}\dot{x}_{\alpha i}^2, \quad U = U(x_{\alpha i}, \dot{x}_{\alpha i}),$$

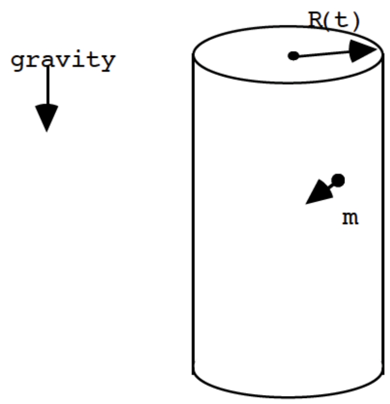
(Notice the potential energy is velocity dependent) construct the form of the Hamiltonian. Is this H conserved? Is it equal to $T+U$ in general?

14. Consider a mass m constrained to move on a wire which is being rotated at a constant angular velocity as shown (no gravity in this problem!):



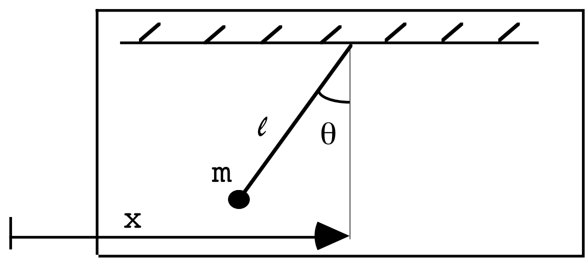
a) Find $L = L(r, \dot{r}, t)$ and construct $H = H(p_r, r, t)$.
b) Is H a constant of the motion? Is it equal to $T+U$?

15. Consider a point mass, m, constrained to moving on the surface of a hollow cylinder of radius R under the influence of gravity (no friction). The radius $R = R(t)$, is a given function of the time.



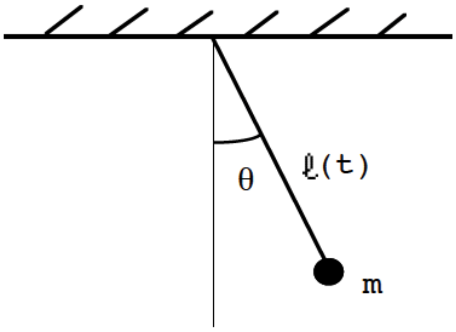
a) Find the Lagranian and Lagrange's equations of motion in cylindrical coordinates.
b) Construct the Hamiltonian. Is H a constant in time? Is $H = T+U$?

16. A particle of mass m subject to gravity is suspended from a simple pendulum of fixed length ℓ as shown. The whole system is placed in a box and accelerated with acceleration "a" along the x direction.



a) Choosing θ as the generalized coordinate (be careful!), write down the kinetic, potential energies of the system. Show that they are explicit functions of the time.
b) Construct the system's Hamiltonian. (You do not have to find the equations of motion.) Is $H = T+U$? Is H a constant of the motion? Explain your answers.

17. A point mass m acted on by gravity is suspended from above by a wire whose length is a given function of time: $\ell(t)$.

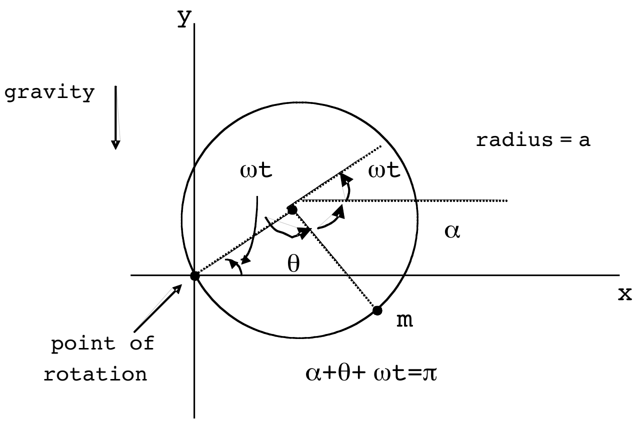


a) Find the Hamiltonian for the system and Hamilton's equations. Eliminate p_θ and show that the equation of motion for q is:

$$\ddot{\theta} + 2\frac{\dot{\ell}}{\ell(t)}\dot{\theta} + \frac{g}{\ell(t)}\sin\theta = 0.$$

b) Is the Hamiltonian a constant of the motion? Is it equal to the total energy, $T + U$? Explain your answers.

18. Remember the bead of mass m attached to a rotating hoop (in the $x\ y$ plane) of radius 'a' (Example 4) from Ch.7.



Treat this as a Hamiltonian problem now.

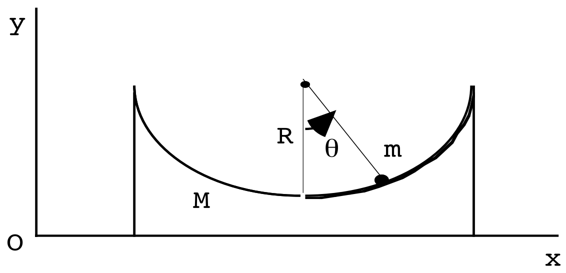
a) Find the Hamiltonian in the appropriate variables.
b) Is $H =$ const.? Is $H = T + U$? Explain your answers.

19. Let's think back to one of the points made in Chapter 4. There I stated that systems which possess a Hamiltonian have phase spaces for which the volume neither expands nor contracts. This was called Liouville's theorem. Show that for Hamiltonian systems,

$$\vec{\nabla} \cdot \vec{v} = 0,$$

where $\vec{v}$ is the generalized "velocity" ($= \frac{d\vec{x}}{dt}$) and $\vec{\nabla}$ is a generalized gradient, thus proving Liouville's theorem for these systems. [Hint: Set $x_i = q_i$ for $i=1,...,n$ and $x_i = p_i$ for $i=n+1,...,2n$ and use Hamilton's equations for x and p.]

20. (A challenging problem.) A particle of mass m slides on the top of a hollowed-out half-cylinder of radius R and mass M. The half-cylinder moves frictionlessly on a horizontal table, as shown. (The total momentum of the system is given to be zero.)



a) By introducing appropriate generalized coordinates, find the Lagrangian of the system and Lagrange's equations.
b) Find the frequency of small oscillations ($\theta \ll 1$) of m.

ENDNOTES

1 We are assuming that λ^{-1} exists, which we know requires det$(\lambda) \neq 0$; we will establish this momentarily.

2 Here it is necessary to use the Leibnitz rule for differentiation.

3 There is even another way to state this. Using the curl concept from Ch.1, $\vec{\nabla} \times \vec{F} = 0$ is another way of specifying a conservative force field.

4 Agrees with the result in Goldstein, p.525, when variables are changed appropriately.

5 For more on parametric representations, see "Calculus of Variations" by Weinstock, ps. 34-36.

6 Even this is not the most general type of variation! One could also alter the meaning of the dependent variable, x, such that $x \rightarrow x + \delta x(x)$ along the path. In the application in the next Chapter, x will be the time variable and there will be nothing gained by considering such scale change variations.