A Guided Tour of
Analytical Mechanics

with animations in MAPLE and MATHEMATICA

Rouben Rostamian
Department of Mathematics and Statistics
UMBC
rostamian@umbc.edu

December 1, 2014
## Contents

**Preface** vii

1 An introduction through examples 1
   1.1 The simple pendulum à la Newton ................................. 1
   1.2 The simple pendulum à la Euler .................................. 3
   1.3 The simple pendulum à la Lagrange .............................. 3
   1.4 The double pendulum .............................................. 4
   Exercises .............................................................. 6

2 Work and potential energy 9
   Exercises .............................................................. 12

3 A single particle in a conservative force field 13
   3.1 The principle of conservation of energy ......................... 13
   3.2 The scalar case ................................................... 14
   3.3 Stability .......................................................... 16
   3.4 The phase portrait of a simple pendulum ....................... 16
   Exercises .............................................................. 17

4 The Kapitsa pendulum 19
   4.1 The inverted pendulum .............................................. 19
   4.2 Averaging out the fast oscillations ............................. 19
   4.3 Stability analysis ............................................... 22
   Exercises .............................................................. 23

5 Calculus of variations 25
   5.1 Introduction ...................................................... 25
      5.1.1 A straight line is the shortest path ....................... 25
   5.2 The brachistochrone .............................................. 25
   5.3 Mathematical preliminaries ...................................... 27
      5.3.1 Basic lemmas ................................................. 27
      5.3.2 The variation ................................................. 28
   5.4 The central problem of the calculus of variations ............ 30
   5.5 The invariance of Euler’s equation under change of coordinates ... 32
   5.6 The solution of Problem 5.1 .................................... 33
   5.7 The solution of Problem 5.1 in polar coordinates ............. 34
   5.8 The solution of Problem 5.2 .................................... 35
   5.9 A variational problem in two unknowns ........................... 36
   5.10 Lagrange multipliers ............................................. 37
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.11</td>
<td>Calculus of variations with pointwise constraints</td>
<td>38</td>
</tr>
<tr>
<td>5.12</td>
<td>Calculus of variations with integral constraints</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>43</td>
</tr>
<tr>
<td>6</td>
<td>Lagrangian mechanics</td>
<td>45</td>
</tr>
<tr>
<td>6.1</td>
<td>Newtonian mechanics</td>
<td>45</td>
</tr>
<tr>
<td>6.2</td>
<td>Holonomic constraints</td>
<td>46</td>
</tr>
<tr>
<td>6.3</td>
<td>Generalized coordinates</td>
<td>49</td>
</tr>
<tr>
<td>6.4</td>
<td>Virtual displacements, virtual work, and generalized force</td>
<td>50</td>
</tr>
<tr>
<td>6.5</td>
<td>External versus reaction forces</td>
<td>52</td>
</tr>
<tr>
<td>6.6</td>
<td>The equations of motion for a holonomic system</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>55</td>
</tr>
<tr>
<td>7</td>
<td>Angular velocity</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>58</td>
</tr>
<tr>
<td>8</td>
<td>The moment of inertia tensor</td>
<td>59</td>
</tr>
<tr>
<td>8.1</td>
<td>A brief introduction to tensor algebra</td>
<td>59</td>
</tr>
<tr>
<td>8.1.1</td>
<td>Tensor algebra</td>
<td>59</td>
</tr>
<tr>
<td>8.1.2</td>
<td>Connection with $\mathbb{R}^3$ and $3 \times 3$ matrices</td>
<td>61</td>
</tr>
<tr>
<td>8.1.3</td>
<td>Symmetric tensors</td>
<td>63</td>
</tr>
<tr>
<td>8.2</td>
<td>The moment of inertia tensor</td>
<td>64</td>
</tr>
<tr>
<td>8.3</td>
<td>Translation of the origin</td>
<td>65</td>
</tr>
<tr>
<td>8.4</td>
<td>The principal moments of inertia</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>67</td>
</tr>
<tr>
<td>9</td>
<td>Constraint reactions</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>73</td>
</tr>
<tr>
<td>10</td>
<td>The Gibbs-Appell formulation of dynamics</td>
<td>75</td>
</tr>
<tr>
<td>10.1</td>
<td>Gibbs-Appell according to Lurie [9]</td>
<td>75</td>
</tr>
<tr>
<td>10.1.1</td>
<td>Acceleration in generalized coordinates</td>
<td>75</td>
</tr>
<tr>
<td>10.1.2</td>
<td>Ideal constraints and the fundamental equation of dynamics</td>
<td>75</td>
</tr>
<tr>
<td>10.1.3</td>
<td>Virtual work and generalized force</td>
<td>76</td>
</tr>
<tr>
<td>10.1.4</td>
<td>Constraints</td>
<td>77</td>
</tr>
<tr>
<td>10.1.5</td>
<td>Virtual displacements</td>
<td>78</td>
</tr>
<tr>
<td>10.1.6</td>
<td>Back to the fundamental equation: Part 1</td>
<td>78</td>
</tr>
<tr>
<td>10.1.7</td>
<td>Back to the fundamental equation: Part 2</td>
<td>79</td>
</tr>
<tr>
<td>10.1.8</td>
<td>The Gibbs-Appell equations of motion</td>
<td>79</td>
</tr>
<tr>
<td>10.1.9</td>
<td>Quasi-velocities</td>
<td>80</td>
</tr>
<tr>
<td>10.1.10</td>
<td>Appell’s equations of motion in terms of quasi-velocities</td>
<td>81</td>
</tr>
<tr>
<td>10.2</td>
<td>Gibbs-Appell according to Gantmacher [10]</td>
<td>81</td>
</tr>
<tr>
<td>10.2.1</td>
<td>Pseudocoordinates</td>
<td>82</td>
</tr>
<tr>
<td>10.2.2</td>
<td>Work and generalized forces</td>
<td>83</td>
</tr>
<tr>
<td>10.2.3</td>
<td>Newton’s equations in pseudocoordinates</td>
<td>84</td>
</tr>
<tr>
<td>10.2.4</td>
<td>The energy of the acceleration</td>
<td>84</td>
</tr>
<tr>
<td>10.3</td>
<td>A modification noted by Desloge</td>
<td>85</td>
</tr>
<tr>
<td>10.4</td>
<td>The simple pendulum via Gibbs-Appell</td>
<td>86</td>
</tr>
<tr>
<td>10.5</td>
<td>The Čaplygin sleigh</td>
<td>86</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>10.6</td>
<td>The Čaplygin sleigh revisited</td>
<td>88</td>
</tr>
<tr>
<td>10.7</td>
<td>The problem from page 63 of Gantmacher</td>
<td>89</td>
</tr>
<tr>
<td>10.8</td>
<td>Rigid body dynamics</td>
<td>92</td>
</tr>
<tr>
<td>10.8.1</td>
<td>Three frames of reference</td>
<td>93</td>
</tr>
<tr>
<td>10.8.2</td>
<td>The energy of acceleration for a rigid body</td>
<td>93</td>
</tr>
<tr>
<td>10.9</td>
<td>The rolling coin</td>
<td>94</td>
</tr>
<tr>
<td>10.9.1</td>
<td>The three frames</td>
<td>94</td>
</tr>
<tr>
<td>10.9.2</td>
<td>The angular velocity</td>
<td>94</td>
</tr>
<tr>
<td>10.9.3</td>
<td>The no-slip constraint</td>
<td>97</td>
</tr>
<tr>
<td>10.9.4</td>
<td>The acceleration of the coin’s center</td>
<td>97</td>
</tr>
<tr>
<td>10.9.5</td>
<td>The rotational acceleration</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>100</td>
</tr>
<tr>
<td>11</td>
<td>Quaternions</td>
<td>101</td>
</tr>
<tr>
<td>11.1</td>
<td>The quaternion algebra</td>
<td>101</td>
</tr>
<tr>
<td>11.2</td>
<td>The geometry of the quaternions</td>
<td>102</td>
</tr>
<tr>
<td>11.2.1</td>
<td>The reflection operator</td>
<td>102</td>
</tr>
<tr>
<td>11.2.2</td>
<td>The rotation operator</td>
<td>103</td>
</tr>
<tr>
<td>11.3</td>
<td>Angular velocity</td>
<td>106</td>
</tr>
<tr>
<td>11.4</td>
<td>A differential equation for the quaternion rotation</td>
<td>107</td>
</tr>
<tr>
<td>11.5</td>
<td>Unbalanced ball rolling on a horizontal plane</td>
<td>108</td>
</tr>
<tr>
<td>11.5.1</td>
<td>The no-slip condition</td>
<td>108</td>
</tr>
<tr>
<td>11.5.2</td>
<td>The Gibbs function and the equations of motion</td>
<td>109</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>110</td>
</tr>
<tr>
<td>A</td>
<td>Maple basics</td>
<td>113</td>
</tr>
<tr>
<td>A.1</td>
<td>Configuring MAPLE</td>
<td>113</td>
</tr>
<tr>
<td>A.2</td>
<td>The execution group</td>
<td>113</td>
</tr>
<tr>
<td>A.3</td>
<td>MAPLE key bindings</td>
<td>114</td>
</tr>
<tr>
<td>A.4</td>
<td>Expression sequences, lists, and sets</td>
<td>114</td>
</tr>
<tr>
<td>A.5</td>
<td>Selecting and removing subsets</td>
<td>115</td>
</tr>
<tr>
<td>A.6</td>
<td>Solving equations symbolically</td>
<td>116</td>
</tr>
<tr>
<td>A.7</td>
<td>Solving equations numerically</td>
<td>116</td>
</tr>
<tr>
<td>A.8</td>
<td>The eval() function</td>
<td>117</td>
</tr>
<tr>
<td>A.9</td>
<td>Expressions and functions</td>
<td>117</td>
</tr>
<tr>
<td>A.10</td>
<td>Vectors and Matrices</td>
<td>118</td>
</tr>
<tr>
<td>A.11</td>
<td>Differentiation</td>
<td>119</td>
</tr>
<tr>
<td>A.12</td>
<td>Solving differential equations</td>
<td>120</td>
</tr>
<tr>
<td>A.12.1</td>
<td>Solving differential equations symbolically</td>
<td>120</td>
</tr>
<tr>
<td>A.12.2</td>
<td>Solving differential equations numerically</td>
<td>121</td>
</tr>
<tr>
<td>A.13</td>
<td>Plotting</td>
<td>121</td>
</tr>
<tr>
<td>A.13.1</td>
<td>Plotting a single function</td>
<td>121</td>
</tr>
<tr>
<td>A.13.2</td>
<td>Plotting multiple function together</td>
<td>122</td>
</tr>
<tr>
<td>A.13.3</td>
<td>Parametric plot</td>
<td>122</td>
</tr>
<tr>
<td>A.13.4</td>
<td>Plotting points and more</td>
<td>123</td>
</tr>
<tr>
<td>A.13.5</td>
<td>Overlaying multiple plots</td>
<td>123</td>
</tr>
<tr>
<td>A.13.6</td>
<td>Reflecting a plot</td>
<td>124</td>
</tr>
<tr>
<td>A.14</td>
<td>The Euler–Lagrange equations</td>
<td>125</td>
</tr>
<tr>
<td>A.15</td>
<td>The animation of a simple pendulum</td>
<td>125</td>
</tr>
</tbody>
</table>
Preface

By “solving a problem” I mean performing all the steps laid out below:

1. Select configuration parameters.

2. Define the position vectors $r_1, r_2, \ldots$ of the point masses in terms of the generalized coordinates $q_1, q_2, \ldots$.

3. Compute the velocities of the point masses:

$$v_i = \dot{r}_i = \sum_j \left( \frac{\partial r_i}{\partial q_j} \dot{q}_j \right), \quad i = 1, 2, \ldots$$

4. Compute the kinetic energy $T = \frac{1}{2} \sum_i m_i \|v_i\|^2$, the potential energy $V$, and the Lagrangian $L = T - V$.

5. Form the equations of motion (a system of second order differential equations (DEs)) in the unknowns $q_1(t), q_2(t), \ldots$:

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_j} \right) = \frac{\partial L}{\partial q_j}, \quad j = 1, 2, \ldots$$

If done by hand, this step would be the most labor-intensive part of the calculations. The calculations can get unbearably complex and can easily lead to formulas that fill more than one page. Fortunately we can relegate the tedious computations to a computer algebra system (CAS) such as MAPLE or MATHEMATICA. I use MAPLE in my own work, therefore I will use that for the purposes of this class. I believe that MATHEMATICA has the equivalent capabilities, and you are welcome to use it instead, if you so prefer. I have, however, no experience with MATHEMATICA, therefore I cannot help you there.

6. Solve the system of DEs. Except for a few special cases, such system are generally not solvable in terms of elementary function. One solves them numerically with the help of specialized software. MAPLE and MATHEMATICA offer that functionality as well.

The software replaces the continues time variable $t$ by a closely spaced “time ticks” $t_0, t_1, t_2, \ldots$ which span the time interval of interest, say $[0, T]$, and then it applies some rather sophisticated numerical algorithms to evaluate the unknowns $q_1(t), q_2(t), \ldots$ at those time ticks. The result may be presented as:

(a) a table of numbers; but that’s not very illuminating, so it’s rarely done that way;
(b) as a set of plots of $q_j$ versus $t$. This is the most common way. Both MAPLE and MATHEMATICA can do this easily; or

(c) as a computer animation, which is the most “user friendly” choice but which takes some work—and a certain amount of know-how—to produce. I will show you how to do this in MAPLE.
Chapter 1

An introduction through examples

This chapter introduces some of basic ideas involved in Lagrangian formulation of dynamics through examples. You will need to take some of the statements and formulas for granted, since they won’t be formally introduced until several chapters later. The idea here is to acquire some “gut feeling” for the subject which will help motivate some of the abstract concepts to come.

1.1 The simple pendulum à la Newton

A pendulum, specifically a simple pendulum, is a massless rigid rod of fixed length $\ell$, one end of which is attached to, and can swing about, an immobile pivot, and to the other end of which is attached a point of mass $m$, called the bob.\(^1\) The force of gravity tends to pull the pendulum down so that to bring the free end to the lowest possible position, called the pendulum’s stable equilibrium configuration. A pendulum can stay motionless in the stable equilibrium configuration forever. If disturbed slightly away from the equilibrium, however, it will oscillate back and forth about it, indefinitely in principle if there are no frictional/dissipative effects. Figure 1.1 shows a simple pendulum at a generic position where the rod makes an angle $\phi$ relative to the vertical.

The pendulum may also be balanced in an inverted position, obtained by turning it upward about the pivot by 180 degrees (remember that the connecting rod is rigid.) That position, which admittedly is difficult to achieve in practice, is called the pendulum’s unstable equilibrium configuration. A pendulum can stay motionless in the unstable equilibrium configuration forever, in principle. If disturbed slightly away from that equilibrium, however, it will move away from it in general.

The stable and unstable equilibria are the only possible equilibrium position of a simple pendulum. The pendulum cannot stay motionless at an angle, say at 45 degrees, relative to the vertical.

A pendulum’s initial condition completely determine its future motion. I am assuming here that the only external action on the pendulum is the force of gravity. The initial condition consist of a pair of data items, one being the initial angle that the rod makes relative to the vertical, and the other is the initial velocity which is the bob is set into motion.

As a specific instance, consider the case where the rot’s initial angle is zero, and the

\(^1\)The pendulum of a grandfather clock is a reasonably good example of such a pendulum, albeit the rod is not massless, and the mass attached to the end of it is not literally a point mass.
Chapter 1. An introduction through examples

bob’s initial velocity is small. Then the pendulum will oscillate back and forth about the stable configuration, similar to what we see in a grandfather clock. If the initial velocity is slightly larger, the pendulum will undergo wider oscillations. If the initial velocity is large enough, the pendulum will not oscillate at all. It will swing about pivot, reach the unstable equilibrium position, go past it, fall down from the other side, and return to its initial position, having made a complete 360 degree rotation about the pivot. At this point the pendulum finds itself in the same condition that it had at the initial time, therefore it will repeat what it did the first time around. In the absence of energy dissipating factors, the rotations about the pivot will continue indefinitely.

To make a mathematical model of the pendulum, we introduce the Cartesian coordinates $x$ $y$ with the origin at the pendulum’s pivot, and the $y$ axis pointing down. We also introduce the unit vectors $i$ and $j$ along the $x$ and $y$ axes, and the unit vectors $e_r$ along the pendulum’s rod and $e_\phi$ which is perpendicular to it, as shown in Figure 1.1. The vectors $e_r$ and $e_\phi$ may be expressed as linear combinations of the vectors $i$ and $j$:

$$e_r = i \sin \phi + j \cos \phi, \quad e_\phi = -i \cos \phi + j \sin \phi.$$ 

Furthermore, let us observe that their time derivatives are related through

$$\dot{e}_r = i \dot{\phi} \cos \phi - j \dot{\phi} \sin \phi = -\dot{\phi} e_\phi, \quad \dot{e}_\phi = i \dot{\phi} \sin \phi + j \dot{\phi} \cos \phi = \dot{\phi} e_r.$$ (1.1)

The bob’s position vector $r(t)$ relative to the origin is $r = \ell e_r$, therefore, the bob’s velocity $v = \dot{r}$ and acceleration $a = \ddot{v}$ may be computed easily with the help of (1.1):

$$v = \dot{r} = (\ell e_r)' = -\ell \dot{\phi} e_\phi, \quad a = \ddot{v} = (-\ell \dot{\phi} e_\phi)' = -\ell \ddot{\phi} e_\phi - \ell \dot{\phi} \dot{e}_\phi = -\ell \ddot{\phi} e_\phi - \ell \dot{\phi}^2 e_r.$$ 

We see that the bob’s acceleration has a component along $e_\phi$ and another along $e_r$.

Newton’s law of motion asserts that $ma = F$, where $F$ is the resultant of all forces acting on the bob. Referring to Figure 1.1 we see that the forces acting on the bob consist of weight $w$ and the tension $-\tau e_r$ along the rod.\(^2\) It follows that

$$m(-\ell \ddot{\phi} e_\phi - \ell \dot{\phi}^2 e_r) = w - \tau e_r.$$ 

\(^2\)The assertion that the force exerted on the bob by the rod lies along the rod requires justification. See the next section for elaboration.
Upon replacing \( w \) with its decomposition \( w = mgj = (mg \cos \varphi)e_r + (mg \sin \varphi)e_\varphi \), and collecting the coefficients of \( e_r \) and \( e_\varphi \), we arrive at

\[
\begin{bmatrix}
ml \dot{\varphi} + mg \sin \varphi \\
ml \dot{\varphi}^2 + mg \cos \varphi - \tau
\end{bmatrix}e_r = 0.
\]

Since \( e_r \) and \( e_\varphi \) are orthogonal, hence linearly independent, each of the expressions in the square brackets is zero. We conclude that

\[
ml \dot{\varphi} + mg \sin \varphi = 0, \quad ml \dot{\varphi}^2 + mg \cos \varphi - \tau = 0. \tag{1.2}
\]

The first equation is a second order differential equation in \( \varphi \). It has a unique solution for any prescribed initial condition pair \( \varphi(0) \) and \( \dot{\varphi}(0) \), although the solution is not expressible in terms of elementary functions. In practice, one solves the equation through a numerical approximation algorithm on a computer. Once the solution is obtained, it may be substituted in the second equation to evaluate the tension \( \tau \) in the pendulum’s rod, should it be of interest.

### 1.2. The simple pendulum à la Euler

In the previous section we assumed, without explanation, that the force within the pendulum’s rod points along the rod; see Figure 1.1 where that force is shown as the vector \(-\tau e_r\).

That assumption seems to be so “obvious” that many textbooks on mechanics and its applications present it without as much as a comment. A close scrutiny, however, shows that the assumption is far from obvious, and in fact, it is not a logical consequence of any of Newton’s laws of motion. Antman [3] presents a critical analysis of this issue and concludes that the proper approach is through an application of Euler’s law of motion, which states that the rate of change of the pendulum’s angular momentum equals the resultant torque applied to it.

### 1.3. The simple pendulum à la Lagrange

In this section we rederive the previous section’s differential equation of motion of simple pendulum through Lagrange’s analytical approach. We no longer need the vectors \( e_r \) and \( e_\varphi \). Instead, we write the bob’s position vector \( r \) directly in terms of its \( i \) and \( j \) components:

\[ r = (\ell \sin \varphi)i + (\ell \cos \varphi)j, \]

and then differentiate to find the velocity

\[ v = \dot{r} = (\ell \dot{\varphi} \cos \varphi)i - (\ell \dot{\varphi} \sin \varphi)j. \]

It follows that that \( ||v||^2 = \ell^2 \dot{\varphi}^2 \).

To proceed further, we introduce a few definitions and assertions whose motivations and explanations will emerge only in subsequent chapters.

- A the kinetic energy \( T \) of a point mass \( m \) moving with velocity \( v \) is \( T = \frac{1}{2}m||v||^2 \).
  - In the case of the pendulum this becomes \( T = \frac{1}{2}ml^2 \dot{\varphi}^2 \).
- The potential energy \( V \) of a point mass \( m \) in a constant gravitational field equals \( mg \bar{h} \) where \( g \) is the acceleration due to gravity, and \( \bar{h} \) is its height above an arbitrarily selected reference point. In the case of the pendulum, the elevation of the bob relative to the lowest point in its path is \( \bar{h} = \ell (1 - \cos \varphi) \), therefore \( V = mg \ell (1 - \cos \varphi) \).
Chapter 1. An introduction through examples

- The Lagrangian $L$ of a mechanical system is the difference between its kinetic and potential energies, that is, $L = T - V$. In the case of the pendulum we have:

$$L(\phi, \dot{\phi}) = \frac{1}{2} m \ell^2 \dot{\phi}^2 - m g \ell (1 - \cos \phi). \quad (1.3)$$

As the notation above indicates, we are viewing the Lagrangian $L$ as a function of two variables $\phi$ and $\dot{\phi}$. It should be emphasized that $\phi$ and $\dot{\phi}$ are considered independent variables here. If you find the notation $\dot{\phi}$ confusing in that regard, consider renaming it to $\omega$, as in

$$L(\phi, \omega) = \frac{1}{2} m \ell^2 \omega^2 - m g \ell (1 - \cos \phi).$$

Now $L$ is a function of two independent variables $\phi$ and $\omega$.

The Lagrangian completely characterizes a mechanical system. It incorporates the system’s parameters, geometry, and physics, all in one neat bundle. Beyond this point the analysis of the system’s motion is pure calculus—or analysis, as Lagrange called it in his *Mécanique Analytique*—with no need to refer to the system’s components and geometry.

According to the theory that Lagrange developed, the equation of motion of a mechanical system whose Lagrangian depends on two variables $\phi$ and $\dot{\phi}$, is given by

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{\phi}} \right) = \frac{\partial L}{\partial \phi}. \quad (1.4)$$

In the case of pendulum we have:

$$\frac{\partial L}{\partial \dot{\phi}} = m \ell^2 \dot{\phi}, \quad \frac{\partial L}{\partial \phi} = -m g \ell \sin \phi,$$

and therefore the equation of motion is

$$(m \ell^2 \dot{\phi})' = -m g \ell \sin \phi,$$

or equivalently,

$$\ddot{\phi} + \frac{g}{\ell} \sin \phi = 0, \quad (1.5)$$

which agrees with the first equation in (1.2). The second of those equations may be obtained through the Lagrangian approach as well, but we will not get into that right now.

### 1.4 The double pendulum

A **double pendulum** is obtained by suspending a second pendulum from the bob of a first pendulum, as shown in the left diagram in Figure 1.2. The double pendulum’s geometric configuration is specified through the two angles $\phi_1$ and $\phi_2$ that the rods make relative to the vertical.

To make a mathematical model of a double pendulum, we follow the ideas sketched in the previous section. Specifically, we introduce the $xy$ Cartesian coordinates and the unit vectors $\mathbf{i}$ and $\mathbf{j}$ as shown in Figure 1.2, and then express the position vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ of the two bobs in terms of their components relative to $\mathbf{i}$ and $\mathbf{j}$:

$$\mathbf{r}_1 = (\ell_1 \sin \phi_1) \mathbf{i} + (\ell_1 \cos \phi_1) \mathbf{j}, \quad \mathbf{r}_2 = \mathbf{r}_1 + (\ell_2 \sin \phi_2) \mathbf{i} + (\ell_2 \cos \phi_2) \mathbf{j}. \quad (1.6)$$
1.4. The double pendulum

We conclude that the double pendulum’s kinetic energy is

\[ v_1 = (\ell_1 \dot{\varphi}_1 \cos \varphi_1)i - (\ell_1 \dot{\varphi}_1 \sin \varphi_1)j, \quad v_2 = v_1 + (\ell_2 \dot{\varphi}_2 \cos \varphi_2)i - (\ell_2 \dot{\varphi}_2 \sin \varphi_2)j. \]

We see that \( ||v_1||^2 = \ell_1^2 \dot{\varphi}_1^2 \). Computing \( ||v_2||^2 \) takes only a little bit more work. We observe that \( v_2 = v_1 + \tilde{v} \), where \( \tilde{v} = (\ell_2 \dot{\varphi}_2 \cos \varphi_2)i - (\ell_2 \dot{\varphi}_2 \sin \varphi_2)j \). Therefore

\[
||v_2||^2 = ||v_1||^2 + ||\tilde{v}||^2 + 2v_1 \cdot \tilde{v} \\
= \ell_1^2 \dot{\varphi}_1^2 + \ell_2^2 \dot{\varphi}_2^2 + 2[\ell_1 \dot{\varphi}_1 \cos \varphi_1 - (\ell_1 \dot{\varphi}_1 \sin \varphi_1)i][\ell_2 \dot{\varphi}_2 \cos \varphi_2 - (\ell_2 \dot{\varphi}_2 \sin \varphi_2)j] \\
= \ell_1^2 \dot{\varphi}_1^2 + \ell_2^2 \dot{\varphi}_2^2 + 2\ell_1 \ell_2 \dot{\varphi}_1 \dot{\varphi}_2 (\cos \varphi_1 \cos \varphi_2 + \sin \varphi_1 \sin \varphi_2) \\
= \ell_1^2 \dot{\varphi}_1^2 + \ell_2^2 \dot{\varphi}_2^2 + 2\ell_1 \ell_2 \dot{\varphi}_1 \dot{\varphi}_2 (\varphi_2 - \varphi_1).
\]

We conclude that the double pendulum’s kinetic energy is

\[
T = \frac{1}{2} m_1 \ell_1^2 \dot{\varphi}_1^2 + \frac{1}{2} m_2 [\ell_1^2 \dot{\varphi}_1^2 + \ell_2^2 \dot{\varphi}_2^2 + 2\ell_1 \ell_2 \dot{\varphi}_1 \dot{\varphi}_2 \cos(\varphi_2 - \varphi_1)] \\
= \frac{1}{2} (m_1 + m_2) \ell_1^2 \dot{\varphi}_1^2 + \frac{1}{2} m_2 \ell_2^2 \dot{\varphi}_2^2 + m_1 \ell_1 \ell_2 \dot{\varphi}_1 \dot{\varphi}_2 \cos(\varphi_2 - \varphi_1).
\]

As to the potential energy, let us recall that a mass’s potential energy in a constant gravitational field is the product of its weight and its elevation above a certain reference point. In the case of a double pendulum, it is easiest to set the reference point at the origin of the coordinates; see Figure 1.2. Then the \( j \) components of the vectors \( r_1 \) and \( r_2 \) provide the elevations of the bobs below the reference point, therefore their elevations above the reference point will require a sign reversal. Referring to (1.6) we see that

\[
V = -m_1 g \cos \varphi_1 - m_2 g [\ell_1 \cos \varphi_1 + \ell_2 \cos \varphi_2] = -(m_1 + m_2) g \cos \varphi_1 - m_2 g \ell_2 \cos \varphi_2.
\]

Thus, the double pendulum’s Lagrangian, \( L = T - V \), takes the form

\[
L(\varphi_1, \varphi_2, \dot{\varphi}_1, \dot{\varphi}_2) = \frac{1}{2} (m_1 + m_2) \ell_1^2 \dot{\varphi}_1^2 + \frac{1}{2} m_2 \ell_2^2 \dot{\varphi}_2^2 + m_2 \ell_1 \ell_2 \dot{\varphi}_1 \dot{\varphi}_2 \cos(\varphi_2 - \varphi_1) \\
+ (m_1 + m_2) g \cos \varphi_1 - m_2 g \ell_2 \cos \varphi_2.
\]

In the previous section’s simple pendulum, the Lagrangian \( L(\varphi, \dot{\varphi}) \) was a function of two variables. In the present case, the Lagrangian \( L(\varphi_1, \varphi_2, \dot{\varphi}_1, \dot{\varphi}_2) \) is a function of four.

Figure 1.2: On the left is a depiction of the physical shape of the double pendulum. On the right we see the pendulum’s mathematical model given by the position vectors \( r_1 \) and \( r_2 \) or the two bobs.
variables. In general, if a mechanical system’s geometric configuration is specified through \( n \) variables \( q_1, \ldots, q_n \), then its Lagrangian is a function of \( 2n \) variables \( q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n \). The equivalent of the single equation of motion (1.4) now is a system of \( n \) Euler–Lagrange equations:

\[
\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) = \frac{\partial L}{\partial q_i}, \quad i = 1, \ldots, n.
\]

The variable \( q_1, \ldots, q_n \) are called the system’s generalized coordinates.

To evaluate these explicitly, we begin by computing

\[
\frac{\partial L}{\partial \dot{\varphi}_1} = (m_1 + m_2) \ell_1^2 \dot{\varphi}_1 + m_2 \ell_1 \ell_2 \dot{\varphi}_2 \cos(\varphi_2 - \varphi_1),
\]

\[
\frac{\partial L}{\partial \dot{\varphi}_2} = m_2 \ell_2^2 \dot{\varphi}_2 + m_2 \ell_1 \ell_2 \dot{\varphi}_1 \cos(\varphi_2 - \varphi_1),
\]

\[
\frac{\partial L}{\partial \varphi_1} = m_2 \ell_1 \ell_2 \dot{\varphi}_1 \dot{\varphi}_2 \sin(\varphi_2 - \varphi_1) - (m_1 + m_2)g \sin \varphi_1,
\]

\[
\frac{\partial L}{\partial \varphi_2} = -m_2 \ell_1 \ell_2 \dot{\varphi}_1 \dot{\varphi}_2 \sin(\varphi_2 - \varphi_1) - m_2 g \ell_2 \sin \varphi_2.
\]

We conclude that the differential equations of motion are

\[
\left[ (m_1 + m_2) \ell_1^2 \ddot{\varphi}_1 + m_2 \ell_1 \ell_2 \ddot{\varphi}_2 \cos(\varphi_2 - \varphi_1) \right] = m_2 \ell_1 \ell_2 \dot{\varphi}_1 \dot{\varphi}_2 \sin(\varphi_2 - \varphi_1) - (m_1 + m_2)g \sin \varphi_1,
\]

\[
\left[ m_2 \ell_2^2 \ddot{\varphi}_2 + m_2 \ell_1 \ell_2 \dot{\varphi}_1 \cos(\varphi_2 - \varphi_1) \right] = -m_2 \ell_1 \ell_2 \dot{\varphi}_1 \dot{\varphi}_2 \sin(\varphi_2 - \varphi_1) - m_2 g \ell_2 \sin \varphi_2.
\]

**Exercises**

1. **Pendulum with a mobile pivot.** Figure 1.3 shows a pendulum whose pivot is allowed to move horizontally without friction. The pivot has mass \( m_1 \) while the bob has mass \( m_2 \). Find the equations of motion of the pendulum.

2. **A spherical pendulum.** The motion of the simple pendulum of length \( \ell \) introduced in this chapter was confined to a single vertical plane, and therefore the pendulum’s bob moved along a circular arc of radius \( \ell \). If off-plane motions are permitted, then the bob will move on a sphere of radius \( \ell \) centered at the pivot. In that setting the pendulum is called a spherical pendulum; see Figure 1.4. Derive the equations of motion of the spherical pendulum.

3. **Bead on a spinning hoop.** A circular wire hoop of radius \( R \) spins about a vertical diameter at a constant angular velocity \( \Omega \). A bead of mass \( m \) can slide without friction along the hoop. The hoop’s radius that connects to the bead makes an angle of \( \varphi(t) \) with respect to the vertical; see Figure 1.5. Find the differential equation satisfied by \( \varphi \).
1.4. A governor mechanism. Figure 1.6 is a schematic drawing of a (simplified) Watt governor which was invented for the automatic control of the speed of steam engines. Our version consists of four massless rigid links of length $\ell$ each, hinged at their ends to form a rhombus. The vertex $O$ remains motionless, while the sleeve at vertex $S$ can slide on the device’s vertical shaft, thereby change the rhomus’s shape. Two balls of mass $m_1$ each are attached to the vertices $A$ and $B$. The sleeve’s mass is $m_2$. The entire assembly rotates at a constant angular speed $\Omega$ about the vertical shaft. Find the differential equation satisfied by the angle $\varphi$ marked on the diagram.

1.5. Two masses on a string. A particle $P$ of mass $m_1$ lies on a smooth horizontal table and is attached to a long, inextensible string which passes through a smooth hole $O$ in the table and hangs down. The other end of the string carries a particle $Q$ of mass $m_2$; see the illustration in Figure 1.7.

The particle $P$ is positioned at the point $(a, 0, 0)$ in the $xyz$ coordinates shown, and given a horizontal initial velocity perpendicular to the $x$ axis. Find the differential equations of motion.

Hint: Let $\rho(t)$ and $\varphi(t)$ be $P$’s position at time $t$ in polar coordinates as seen in Figure 1.7. The equations of motions constitute a system of differential in $\rho(t)$ and $\varphi(t)$. 
Figure 1.5: Bead on a rotating hoop (Exercise 3).

Figure 1.6: A simplified Watt governor (Exercise 4).

Figure 1.7: The point $P$ slides on the table. The point $Q$ moves vertically (Exercise 5).
Chapter 2

Work and potential energy

Work is the product of force and its displacement. To be precise, the infinitesimal work $dW$ performed in displacing a force $F$ by an infinitesimal distance $dr$ is $dW = F \cdot dr$. If the point of the application of the force moves along a path $C$ in space, then the work performed along the path is the line integral

$$W = \int_C F \cdot dr. \quad (2.1)$$

If you are repositioning a massive desk in an office, for example, then work measures the amount of effort exerted by you in performing the task.

Expanding upon the moving of the desk scenario, suppose that you intend to move the desk from a point $A$ to a point $B$. It should be obvious that the amount of work performed will vary, depending on the path along which you move the desk between $A$ and $B$. Chances are that the shortest (straight line) path will require lesser effort than a long path that winds around the office.

There are many interesting and important situations where, unlike the moving of the desk example, the work performed in going from a point $A$ to a point $B$ is independent of the path taken between $A$ and $B$. The most elementary example is the raising or lowering of a weight. To see how it works, set up a Cartesian coordinate system in space so that the $x$ and $y$ axes are horizontal, and the $z$ axis points up. Let $r_a = (x_a, y_a, z_a)$ and $r_b = (x_b, y_b, z_b)$ be the position vectors of the starting and ending points $A$ and $B$, and let $r = (x, y, z)$ be the position vector of a generic point along a path $C(A,B)$ with endpoints $A$ and $B$. Suppose that we move an object of mass $m$ along that path. The force of the object’s weight is $F = (0, 0, -mg)$, where $g$ is the acceleration of gravity. The work performed along the path is

$$W = \int_{C(A,B)} F \cdot dr$$

$$= \int_{C(A,B)} (0, 0, -mg) \cdot (dx, dy, dz) = \int_{C(A,B)} -mg \, dz = -mg(z_b - z_a).$$

We see that the work in moving the weight from $A$ to $B$ is expressed in terms the $z$ coordinates of the endpoints, thus it is the same on all conceivable paths that go from $A$ to $B$.

---

3A position vector of a point $P(x,y,z)$ is the vector $r = (x,y,z)$ that extends from the origin to the point $P$. 
To generalize, consider a (possibly position dependent) force field \( \mathbf{F}(\mathbf{r}) \) with the property that the work performed in going from a given point \( A \) to an arbitrary point \( \mathbf{r} \) in space is independent of the path from \( A \) to \( r \). Let us write \( V(\mathbf{r}) \) for the negative of the value of that integral, that is,

\[
V(\mathbf{r}) = - \int_{C(A, \mathbf{r})} \mathbf{F}(\mathbf{r}') \cdot d\mathbf{r}'.
\] (2.2)

The function \( V \) defined this way is called the potential function, or simply the potential, of the the vector field \( \mathbf{F} \). Equivalently, the vector field \( \mathbf{F} \) is said to be derived from a potential.

In (2.2) I have written \( r' \) for the dummy variable of integration in order to distinguish it from the position vector \( \mathbf{r} \) which designates the path’s endpoint. The minus sign does not have a deep significance; it’s convenient to build it into the definition since it leads to more pleasing forms of general statements, such as the one on conservation of energy.

**Theorem 2.1.** Consider a continuous vector field \( \mathbf{F} \) defined in an open and connected domain \( \mathcal{D} \) in the \( n \)-dimensional space, and suppose that \( \mathbf{F} \) possesses a potential function \( V \) as in (2.2). Then \( V \) is differentiable and \( \mathbf{F}(\mathbf{r}) = -\nabla V(\mathbf{r}) \).

**Proof.** By definition, the gradient \( \nabla V \) of a function \( V \) at a point \( \mathbf{r} \) is the vector with the property that for any unit vector \( \mathbf{e} \), the directional derivative of \( V \) in the direction of \( \mathbf{e} \) is given by

\[
\nabla V(\mathbf{r}) \cdot \mathbf{e} = \lim_{h \to 0} \frac{V(\mathbf{r} + h\mathbf{e}) - V(\mathbf{r})}{h}.
\]

To simplify the discussion, let us write \( P \) and \( Q \) for the points in space corresponding to the position vectors \( \mathbf{r} \) and \( \mathbf{r} + h\mathbf{e} \), as illustrated in Figure 2.1. Pick a path \( C(A, \mathbf{r}) \) to evaluate \( V \) at \( P \), then extend that path as a straight line segment to \( Q \) to evaluate \( V \) at \( Q \). Then the difference \( V(Q) - V(P) \) amounts to an integration along the straight segment \( PQ \):

\[
V(\mathbf{r} + h\mathbf{e}) - V(\mathbf{r}) = V(Q) - V(P) = -\int_0^h \mathbf{F}(\mathbf{r} + \xi\mathbf{e}) \cdot \mathbf{e} \, d\xi.
\]

Then it follows that

\[
\frac{V(\mathbf{r} + h\mathbf{e}) - V(\mathbf{r})}{h} = -\frac{1}{h} \int_0^h \mathbf{F}(\mathbf{r} + \xi\mathbf{e}) \cdot \mathbf{e} \, d\xi = -\mathbf{F}(\mathbf{r} + \xi\mathbf{e}) \cdot \mathbf{e}
\]

for some \( \xi \in (0, h) \), the latter assertion being a consequence of the Mean Value Theorem for integrals; see e.g., Stewart [4].

As \( h \) goes to zero, so does \( \xi \) because \( \xi \in (0, h) \). It follows that

\[
\nabla V(\mathbf{r}) \cdot \mathbf{e} = \lim_{h \to 0} \frac{V(\mathbf{r} + h\mathbf{e}) - V(\mathbf{r})}{h} = -\mathbf{F}(\mathbf{r}) \cdot \mathbf{e}.
\]

Since this holds for every \( \mathbf{e} \), it follows that \( \nabla V(\mathbf{r}) = -\mathbf{F}(\mathbf{r}) \). \( \square \)

**Remark 2.1.** Let is point out the roles that the theorem’s technical assumptions play in the proof:

- The *continuity* of the vector field \( \mathbf{F} \) enters at the point where the Mean Value Theorem is applied. That theorem is not true without continuity.
Figure 2.1: If $V(P)$ is evaluated by integration along the path $C(A, r)$, then $V(Q)$ may be evaluated by integrating along that same path, and then continuing along the straight line segment $PQ$ of length $h$ in the direction $e$.

- The assumption that the domain $\mathcal{D}$ is connected is needed to ensure that a path may be established between any pair of points in $\mathcal{D}$. That’s what enabled us to sketch the curve $C(A, r)$ that connects the points $A$ and $P$ in Figure 2.1.

- The assumption that the domain $\mathcal{D}$ is open means that a ball of positive radius may be placed around any point $P \in \mathcal{D}$ so that the ball lies entirely within $\mathcal{D}$. It’s that property which ensures that moving away from $P$ by a small distance $h$, as we did in Figure 2.1, we land safely on a point $Q$ which lies within $\mathcal{D}$.

Remark 2.2. Adding a constant to the potential function $V$ does not affect the equality $F(r) = \nabla V(r)$. Thus, a vector field’s potential, if it has one, is defined modulo an additive constant.

Example 2.2. Earlier in the this section we observed that the force field corresponding to an object’s weight is $F = (0, 0, -mg)$. We see that $F(r) = -\nabla V(r)$ where $V(x, y, z) = mgz$. We will use $mgz$ as the potential of a weight throughout these notes. Note the effect of the minus sign in (2.2); in its absence the potential of a weight would have been $-mgz$.

Example 2.3. In the previous example we assumed that the acceleration of gravity $g$ is a constant. That’s a good assumption if the changes in height during the motion are small relative to the radius of the Earth. In general, the gravitational force that a point mass $M$ exerts on a point mass $m$ drops as the inverse square of the distance. Specifically, Newton’s law of gravitation says

$$F(r) = -\left(\frac{GMm}{||r||^2}\right)\frac{r}{||r||} \tag{2.3}$$

where $r$ is $m$’s position vector relative to $M$, and $G$ is the universal gravitational constant. The inverse square law is manifested through the $||r||^2$ term that appears in the denominator inside the parentheses. The factor $r/||r||$ is a unit vector that points from $M$ to $m$. It is possible to show (see Exercise ?) that $F$ is derived from a potential.

Theorem 2.4. Suppose the force field $F$ is derived from a potential $V$. Then the work per-
formed in moving the force along any path from a point $r_a$ to $r_b$ is given by
\[ W = V(r_a) - V(r_b). \]  
(2.4)

**Proof.** We have $F(r) = -\nabla V(r)$ therefore

\[ F \cdot dr = -\nabla V(r) \cdot dr = \left( \frac{\partial V}{\partial x_1}, \ldots, \frac{\partial V}{\partial x_n} \right) \cdot \langle dx_1, \ldots, dx_n \rangle = -\left( \frac{\partial V}{\partial x_1} dx_1 + \cdots + \frac{\partial V}{\partial x_n} dx_n \right) = -dV, \]

therefore

\[ W = \int_C F \cdot dr = \int_{r_a}^{r_b} -dV = V(r_a) - V(r_b). \]

\[ \square \]

**Exercises**

2.1. Verify that the gravitational potential field $F(r)$ in (2.3) is derived from the potential

\[ V(r) = \frac{GMm}{||r||}. \]
Chapter 3
A single particle in a conservative force field

3.1 The principle of conservation of energy

Newton’s law of motion, $F = m\ddot{r}$ relates the acceleration $\ddot{r}$ of a point of constant mass $m$ subjected to a force $F$. If the force is derived from a potential $V(r)$, that is, $F = -\nabla V$, then the law of motion takes the form

$$m\ddot{r} = -\nabla V(r). \quad (3.1)$$

Multiplying this through by the velocity $\dot{r}$

$$m\dot{r} \cdot \ddot{r} = -\nabla V(r) \cdot \dot{r},$$

and then rearranging

$$\frac{1}{2} m(\dot{r} \cdot \dot{r}) + (V(r))' = 0,$$

we arrive at

$$\frac{d}{dt} \left( \frac{1}{2} m||\dot{r}||^2 + V(r) \right) = 0,$$

which tells us that the quantity

$$E = \frac{1}{2} m||\dot{r}||^2 + V(r) \quad (3.2)$$

remains constant during the motion. The constant $E$ is called the particle’s mechanical energy (or just the energy for short). The first term on the right-hand side of (3.2) is called the kinetic energy; the second term is called the potential energy. The constancy of $E$ in a motion is called the principle of conservation of energy.

Remark 3.1. The kinetic and potential energies don’t remain constant during the motion; it’s their sum that does. Therefore the reduction of one is accompanied by the increase of the other. It helps to think of this as a conversion of one form of energy to the other. The myriad of motion phenomena encountered in our daily experiences are manifestations of such interplay between the kinetic and potential energies.

Remark 3.2. The conservation of the total energy $E$ is a consequence of the assumption that the force field $F$ is derived from a potential. This should explain the alternative name,
a conservative force field, which is commonly used to refer to a force field derived from a potential.

Remark 3.3. Had we chosen against putting the minus sign in the definition (2.2), the principle of conservation of energy would have stated that the difference between the kinetic and potential energies remains constant, which is not as appealing as saying that their sum remains a constant.

3.2 The scalar case

The rest of this chapter is devoted to a study of the scalar version of equation (3.1), that is,

\[ m\ddot{x} = -V'(x), \]  

(3.3)

where \( x \) is scalar, and \( V' \) is the derivative of a potential \( V \). In addition to the obvious applications in one-dimensional dynamics, this equation occurs in quite a number of other contexts which are far from one-dimensional motions. The equation of motion of a simple pendulum (1.5), for instance, falls in this category, but the motion is certainly not linear. We will more on this in Section 3.4.

The previous section’s statement on conservation of energy, which in the scalar case takes the form

\[ E = \frac{1}{2}m\dot{x}^2 + V(x), \]  

(3.4)

is a first order differential equation in the unknown \( x(t) \), and which may be solved, in principle, through a separation of variables. We have \( \dot{x}^2 = \frac{2}{m}(E - V(x)) \), therefore \( \dot{x} = \pm \sqrt{\frac{2}{m}(E - V(x))} \), and hence

\[ \int \frac{dx}{\sqrt{2(E - V(x))}} = \pm \int dt = \pm t + C. \]

Expect for the most trivial cases, the integral on the left is impossible to evaluate in terms of elementary functions. It is possible, however, to obtain quite an adequate “feel” for the solution, without performing any integration at all, through a phase plane analysis of the equation.

To explain the idea, consider the potential function \( V \) whose graph is shown in Figure 3.1(a). Regard the solution \( x(t) \) of the differential equation (3.3) as the abscissa of a point \( P \) that moves along the horizontal axis in that figure according to the equation’s dynamics. Then the point \( Q \) with coordinates \((x(t), V(x(t))\) slides on the graph of \( V \) accordingly. The length of the line segment \( PQ \) equals the potential energy \( V(x) \). We extend that segment upward to a point \( R \) so the the length of \( QR \) equals the kinetic energy \( \frac{1}{2}m\dot{x}^2 \). Since the sum of the kinetic and potential energies remains a constant \( E \) during the motion, the locus of the point \( R \) is the horizontal line \( V = E \), as marked on the figure.

The point \( Q \) cannot rise above the horizontal line \( V = E \) during the motion because the nonnegative length of the line segment \( QR \) (which equals \( \frac{1}{2}m\dot{x}^2 \)) prevents it. Consequently, the motion of \( Q \) is confined to the graph’s red-colored arc. We refer to that arc as a potential well corresponding to the energy \( E \) and we think of \( Q \) as a point that has fallen into the well and is unable to get out.

At the edges of the potential well the potential energy equals \( E \) and the kinetic energy, and therefore the velocity \( \dot{x} \), are zero. In the interior, where the potential energy
3.2. The scalar case

The scalar case

Figure 3.1: The dynamics of the equation \( \ddot{x} + V'(x) = 0 \) is completely determined by the potential function \( V \). The figure on top shows the graph of \( V(x) \), and an energy well corresponding to an energy level \( E \). The coordinate \( x \) is confined to the energy well shown in red. Since the total energy is conserved, as the potential energy drops below \( E \) within the well, the kinetic energy increases, resulting in the phase portrait shown in the bottom figure.

drops below \( E \), the kinetic energy, and therefore the velocity squared, \( \dot{x}^2 \), are positive. We conclude that as we traverse the potential well from left to right, the velocity begins at zero, increases gradually (in absolute value) to a maximum, then drops back to zero at the rightmost end. The sign of the velocity may be positive or negative since the only information we are getting from Figure 3.1(a) is on the square of the velocity.

This observation leads to the red oval in the diagram shown in Figure 3.1(b). The horizontal axis in that figure is the same as the \( x \) in Figure 3.1(a). The vertical axis is the velocity \( \dot{x} \). Observe that at the leftmost and rightmost points of the oval, which correspond to the extremes of the potential well, the velocity is zero, and in between rises to a maximum (or falls to a minimum), as we expect. The oval is symmetric with respect to the \( x \) axis because a given \( \dot{x}^2 \) yields two velocities \( \pm |\dot{x}| \).

The red oval constructed in the previous discourse depends on the choice of the energy level \( E \). It should be clear that lowering \( E \) slightly will shrink the oval, and raising \( E \) slightly, will expand it. The black curves in Figure 3.1(b) are the result of selecting various values of \( E \).

Figure 3.1(b) is called the phase diagram or phase portrait of the differential equation (3.3). The curves in it are called orbits. An alternative to the geometric construction of the orbits carried out above, we may equally well view them as implicitly defined curves in the \( x-\dot{x} \) plane through the equation (3.4). Varying the parameter \( E \) produces the family of all orbits, some of which are shown in Figure 3.1(b).

Yet another way of viewing the orbits is as level curves of the the surface defined by the function

\[
E(x, \dot{x}) = \frac{1}{2} m \dot{x}^2 + V(x).
\]

Two views of the surface \( E(x, \dot{x}) \) corresponding to the potential \( V(x) \) of Figure 3.1(a) are
Chapter 3. A single particle in a conservative force field

3.3 Stability

We have seen that the dynamics of the equation (3.3) dictate that a hypothetical particle trapped in an energy well cannot escape. The closer the energy level \( E \) is to the bottom of the well, the lesser room there is for the particle to maneuver.\(^4\) In the extreme case when the particle’s energy matches that of the well’s bottom, the particle cannot move at all. We express this by saying that the bottom of the well is a stable equilibrium point of the differential equation (3.3). If the energy level is increases just slightly, the particle will move along an oval around the equilibrium point. Referring to the construction of the phase portrait from the potential \( V \), it should be evident that a local minimum of \( V \) corresponds to a stable equilibrium. Similarly, a local maximum of \( V \) corresponds to a saddle on the energy surface (see Figure 3.2) therefore a local maximum of \( V \) corresponds to an unstable equilibrium.

3.4 The phase portrait of a simple pendulum

The previous section’s treatment of the scalar equation of motion has greater applicability than it may appear on the surface. The variable \( x(t) \) there need not be the coordinate of a moving point along a straight line. For instance, according to (1.5) on page 4, the motion of a simple pendulum is described by the differential equation

\[
\ddot{\varphi} + \frac{g}{\ell} \sin \varphi = 0,
\]  

where \( \varphi = \varphi(t) \) is the angle the the pendulum’s rod makes relative to the downward pointing vertical, as depicted in Figure 3.3 on the left. Although the motion of the pendulum takes places in two dimensions, the equation of motion (3.5) is exactly of the form (3.3) with \( V'(\varphi) = g/\ell \sin \varphi \), hence \( V(\varphi) = g/\ell (1 - \cos \varphi) \), therefore the analysis of the pre-

\(^4\)I am assuming a round-bottomed energy well here. In a flat-bottomed energy well the particle can move around no matter how shallow the well.
The graph of $V(\phi) = \frac{g}{\ell}(1 - \cos \phi)$ (with $g/\ell = 1$) is shown at top right. The corresponding phase portrait is shown at bottom right. The function $V$ has a period of $2\pi$, therefore it would have sufficed to limit the plots to the range $-\pi \leq \phi \leq \pi$. Outside of that range, things repeat by periodicity.

Exercises

3.1. Analyze the stability of the spinning hoop of Exercise 3. Show that the lower equilibrium is stable if $\Omega$ is small, and unstable if it is large. Find the value of $\Omega$ where the transition takes place.

3.2. Plot representative phase portraits for the two cases of the problem above.

---

5As noted earlier, the potential function is defined within an additive arbitrary constant, therefore the "1" in $1 - \cos \phi$ is immaterial; its inclusion makes $V(0) = 0$ which is nice but of no special consequence.
Figure 3.4: Two views of the pendulum’s phase portrait as wrapped into a cylinder to emphasize that the pendulum’s configuration depends on $\phi \mod 2\pi$. 
Chapter 4

The Kapitsa pendulum

4.1 The inverted pendulum

The pendulum in Figure 4.1 consists of a massless rod of length $\ell$, a point mass $m$ as the bob, and a pivot which oscillates vertically according to $y = a \cos \omega t$, where $a$ and $\omega$ are prescribed constants. In Exercise 1 you will show that the equation of motion is

$$\ddot{\varphi} + \frac{g}{\ell} \sin \varphi + \frac{a \omega^2}{\ell} \sin \varphi \cos \omega t = 0, \quad (4.1)$$

where $\varphi$ is the angle of the rod relative to the vertical, as shown.

If the amplitude $a$ of the pivot’s oscillation is small, and if $\omega$ is not too large, then we expect the system to behave similar to an ordinary pendulum, albeit with somewhat jittery oscillations. In particular, the lower equilibrium $\varphi = 0$ would be stable and the upper equilibrium $\varphi = \pi$ would be unstable. A graph of the solution $\varphi(t)$ of the pendulum’s equation of motion with smallish $a$ and $\omega$ is shown on Figure 4.1.

It is the purpose of this chapter to show that as $\omega$ is increased beyond a critical threshold, the pendulum’s behavior changes drastically. Specifically, the lower equilibrium becomes unstable and the upper equilibrium becomes stable. Thus, the pendulum turns around by 180 degrees, points upward, and oscillates about the $\varphi = \pi$ position! Figure 4.2 shows the solution of the pendulum’s equation of motion for a relatively fast $\omega$. Note that the oscillation now take place about the upper equilibrium $\varphi = \pi$. The pendulum is standing upright, pointing up!

4.2 Averaging out the fast oscillations

To explain this interesting phenomenon, introduce a hypothetical “nominal rod” which connects the origin to the bob, and let $\psi$ be its angle relative to the vertical, as shown in the schematic diagram in Figure 4.1. The let $\delta = \varphi - \psi$ be the angle between the real rod and the nominal rod. According to the Law of Sines applied to the triangle shown in the figure, we have:

$$\frac{\sin \psi}{\ell} = \frac{\sin \delta}{a \cos \omega t}$$

that is,

$$\sin \delta = \frac{a}{\ell} \sin \psi \cos \omega t.$$
Figure 4.1: The pivot oscillates vertically according to $a \cos \omega t$ about the pendulum’s nominal pivot. In the schematic diagram on the left the pivot’s displacement is exaggerated; we assume that $a/\ell$ is very small in our computations. When the pendulum’s arm makes a “nominal” angle $\psi$ with the vertical, the angle actually oscillates rapidly in the range $\psi \pm \delta$. The graph on the right is that of the angle $\varphi(t)$ obtained by solving the differential equation (4.1) with parameters $\ell = 1$, $g = 1$, $a = 0.05$, and $\omega = 20$, and initial conditions $\varphi(0) = 5\pi/6$, $\dot{\varphi}(0) = 0$. The oscillation about the lower equilibrium position ($\varphi = 0$) is stable since $\omega^2 < 2g/\ell/a$.

Figure 4.2: The solution of the differential equation (4.1) with parameters $\ell = 1$, $g = 1$, $a = 0.05$, and $\omega = 40$, and initial conditions $\varphi(0) = \pi/6$, $\dot{\varphi}(0) = 0$. The oscillation about the upper equilibrium position ($\varphi = \pi$) is stable since $\omega^2 > 2g/\ell/a$. The figure on the right is an enlarged copy of a portion of the graph on the left. We see that $\varphi(t)$ consists of high-frequency, small-amplitude oscillations riding on a slowly oscillating function.

Then the assumption $a \ll \ell$ implies that $\sin \delta \ll 1$, therefore $\sin \delta \approx \delta$. We conclude that

$$\delta \approx \frac{a}{\ell} \sin \psi \cos \omega t.$$  \hspace{1cm} (4.2)

We are going to need $\delta$’s second derivative soon, so let’s calculate it right now. We
4.2. Averaging out the fast oscillations

have:

\[ \ddot{\delta} \approx \frac{a}{\ell} [ \dot{\psi} \cos \psi \cos \omega t - \omega \sin \psi \sin \omega t ], \]
\[ \ddot{\delta} \approx \frac{a}{\ell} [ \dot{\psi} \cos \psi \cos \omega t - \dot{\psi}^2 \sin \psi \cos \omega t - 2 \omega \dot{\psi} \cos \psi \sin \omega t - \omega^2 \sin \psi \cos \omega t ]. \]

We are interested in high frequency oscillations of the pivot, that is, \( \omega \gg 1 \). Therefore, the term with \( \omega^2 \) in the expression above dominates the rest. We conclude that

\[ \ddot{\delta} \approx -\frac{a \omega^2}{\ell} \sin \psi \cos \omega t. \]

Now we go to the differential equation (4.1) and replace \( \phi \) by \( \psi + \delta \), and replace \( \sin \phi \) with its Taylor series approximation

\[ \sin \phi = \sin(\psi + \delta) \approx \sin \psi + \delta \cos \psi. \]

We get:

\[ \ddot{\psi} + \frac{g}{\ell} [ \sin \psi + \delta \cos \psi ] + \frac{a \omega^2}{\ell} [ \sin \psi + \delta \cos \psi ] \cos \omega t = 0. \]

We multiply out everything and replace \( \ddot{\delta} \) with the expression obtained above, and arrive at

\[ \ddot{\psi} - \frac{a \omega^2}{\ell} \sin \psi \cos \omega t + \frac{g}{\ell} [ \sin \psi + \delta \cos \psi ] + \frac{a \omega^2}{\ell} [ \sin \psi + \delta \cos \psi ] \cos \omega t = 0. \]

The second and fifth terms cancel, leaving us with

\[ \ddot{\psi} + \frac{g}{\ell} [ \sin \psi + \delta \cos \psi ] + \frac{a \omega^2}{\ell} [ \sin \psi + \delta \cos \psi ] \cos \omega t = 0. \]

Then we substitute for \( \delta \) from (4.2):

\[ \ddot{\psi} + \frac{g}{\ell} [ \sin \psi + \frac{a \omega^2}{\ell} \sin \psi \cos \psi \cos \omega t + \frac{a \omega^2}{\ell} \sin \psi \cos \psi \cos^2 \omega t ] = 0. \quad (4.3) \]

In the graphs of \( \phi(t) \) in figures 4.1 and 4.2 we see that the period \( 2\pi/\omega \) of the pivot's oscillations is much smaller than the oscillations of the pendulum itself. Consequently, within one such time period, the value of \( \psi \) and its derivatives are essentially constants. On the basis of this observations, we average (4.3) over one \( 2\pi/\omega \) period, where we regard \( \psi \) as constant. Since

\[ \frac{1}{2\pi/\omega} \int_0^{2\pi/\omega} \cos \omega t \, dt = 0, \quad \frac{1}{2\pi/\omega} \int_0^{2\pi/\omega} \cos^2 \omega t \, dt = \frac{1}{2}, \]

we get

\[ \ddot{\psi} + \frac{g}{\ell} [ \sin \psi + \frac{a^2 \omega^2}{2\ell^2} \sin \psi \cos \psi ] \sin \psi = 0, \]

or the equivalent

\[ \ddot{\psi} + \frac{g}{\ell} [ 1 + \frac{a^2 \omega^2}{2\ell g} \cos \psi ] \sin \psi = 0. \quad (4.4) \]
In comparison with the equation (3.5) of the motion of an ordinary unforced pendulum, the Kapitsa pendulum sees an “effective acceleration of gravity” given by

\[ g \left[ 1 + \frac{a^2 \omega^2}{2\ell g} \cos \psi \right]. \]

If the pendulum’s motion is in the \( 0 < \psi < \pi/2 \) regime, the quantity in the square brackets is greater than 1, therefore vibrating the support is tantamount to increasing the acceleration of gravity.\(^6\) If, however, the pendulum’s motion is in the \( \pi/2 < \psi < \pi \) regime, then the effective acceleration of gravity may become negative if the coefficient of \( \cos \psi \) is sufficiently large. The latter will happen if \( \omega \) is sufficiently large. That’s tantamount to reversing the direction of gravity, which sort of explains why the pendulum turns upright.

### 4.3 Stability analysis

The effective equation of motion (4.4) of the Kapitsa pendulum is of the type (3.3) which was studied in Chapter 3. Comparing the two, we see that

\[ V'(\psi) = \frac{g}{\ell} \sin \psi + \frac{a^2 \omega^2}{2\ell^2} \sin \psi \cos \psi, \quad (4.5) \]

whence

\[ V(\psi) = \frac{g}{\ell} (1 - \cos \psi) + \frac{a^2 \omega^2}{4\ell^2} \sin^2 \psi \quad (4.6) \]

The analysis presented in Chapter 3 is based entirely on the shape of \( V' \)'s graph. Therefore we proceed to analyze the shape.

The equation’s equilibria are the roots of the equation \( V'(\psi) = 0 \). Upon factorizing the equation as

\[ \left[ \frac{g}{\ell} + \frac{a^2 \omega^2}{2\ell^2} \cos \psi \right] \sin \psi \]

we see that the roots are the solutions of

\[ \sin \psi = 0 \quad \text{and} \quad \cos \psi = -\frac{2g \ell}{a^2 \omega^2}. \]

The first equation yields \( \psi = 0 \) and \( \psi = \pi \) as roots. (It suffices to look for roots in the \( 0 \leq \psi \leq \pi \) range.) The second equation yields a root \( \hat{\psi} \) given by

\[ \hat{\psi} = \cos^{-1} \left( -\frac{2g \ell}{a^2 \omega^2} \right) \quad (4.7) \]

if and only if \( \omega^2 > 2g \ell / a^2 \). (If that ratio equals to 1 then the root is \( \pi \), which duplicates what we have already found.)

Table 4.1 lists the critical points of the function \( V(\psi) \), along with the values of \( V \), \( V' \), and \( V'' \) at those points. We see that:

- \( V''(0) > 0 \) regardless of the parameter values, therefore the hanging-down equilibrium, \( \psi = 0 \), is always stable;

\(^6\)Don’t take this literally; the acceleration of gravity in (3.5) is a constant while the effective acceleration of gravity in (4.4) depends on \( \cos \psi \), therefore is not a constant.
Table 4.1: The analysis of the critical points of the function $V(\psi)$ defined in (4.6). The critical point $\hat{\psi}$, defined in (4.7) exists if and only if $\omega^2 > \frac{2g\ell}{a^2}$.

- $V''(\hat{\psi}) < 0$ regardless of the parameter values, therefore the equilibrium $\hat{\psi} = 0$, if it exists, is unstable; and
- if $\omega^2 < \frac{2g\ell}{a^2}$ then $V''(\pi) < 0$, therefore the inverted equilibrium, $\psi = \pi$, is unstable; but if $\omega^2 > \frac{2g\ell}{a^2}$ then $V''(\pi) > 0$, therefore the inverted equilibrium, $\psi = \pi$, is stable.

Figure 4.3 shows the graphs of $V(\psi)$ for two representative cases. The graphs are plotted over the range $[-2\pi, 2\pi]$ to give a clear sense of their nature; only the range $[0, \pi]$ is of true relevance to us.

Exercises

4.1. Derive the equation of motion (4.1) of Kapitsa’s pendulum.
4.2. Horizontally oscillating pivot. Consider a pendulum similar to Kapitsa’s, but whose pivot oscillates horizontally rather than vertically. Derive the equation of motion and do a stability analysis.
Chapter 5
Calculus of variations

5.1 Introduction

5.1.1 A straight line is the shortest path

That a straight line is the shortest path that connects a pair of points in the Euclidean plane seems so obvious that it hardly needs elaboration. It is somewhat surprising, therefore, that when the question is formulated as a minimization problem, the conclusion, though true, is far from immediate. Let us look into the details.

Fix two points \( P_1(x_1, y_1) \) and \( P_2(x_2, y_2) \) in the Cartesian coordinate plane, where \( x_2 > x_1 \), and consider a function \( f \) whose graph goes through the points \( P_1 \) and \( P_2 \). The length of the graph is given by the usual formula from calculus:

\[
\ell(f) = \int_{x_1}^{x_2} \sqrt{1 + f'(x)^2} \, dx.
\]  

(5.1)

The notation \( \ell(f) \) emphasizes that the length depends on the choice of the function \( f \).

Now the question of the path of minimal length between two points may stated as follows:

**Problem 5.1.** Find a function \( f \) whose graph goes through the points \( P_1 \) and \( P_2 \), and which minimizes the length \( \ell(f) \).

*Calculus of variations* is the branch of mathematics that deals with that and similar questions.

5.2 The brachistochrone

Set up a Cartesian coordinate system in a vertical plane, with a horizontal \( x \) axis and a downward pointing \( y \) axis. Fix a point \( P_2(x_2, y_2) \) as shown in Figure 5.1 and consider a smooth function \( f \) whose graph passes through the point \( P_1 \) at the origin and \( P_2 \). Think of the curve shown there as a slide in a playground. A particle slides without friction from \( P_1 \) (with zero initial velocity) and arrives at \( P_2 \). Can we tell how long it takes?

Yes. The particle’s position vector, relative to the origin of the coordinate system is \( r = (x, f(x)) \), therefore its velocity is \( v = \dot{r} = (\dot{x}, f'(x)\dot{x}) \), where a superposed dot indicates the time derivative. It follows that \( ||v||^2 = (1 + f'(x)^2)\dot{x}^2 \), therefore the particle’s
mechanical energy, which is the sum of its kinetic and potential energies, is given by

\[ E = \frac{1}{2} m \left( 1 + f'(x)^2 \right) \dot{x}^2 - mg f(x), \]

where \( m \) is the particle’s mass, and \( g \) is the acceleration of gravity. In particular, since the initial velocity is zero, the mechanical energy at the initial time zero. Furthermore, since there is no friction, mechanical energy is conserved, therefore

\[ \frac{1}{2} m (1 + f'(x)^2) \dot{x}^2 - mg f(x) = 0 \text{ for all } t. \]

This is a first order differential equation for the unknown \( x(t) \). To solve, we isolate \( \dot{x} \):

\[ \dot{x} = \pm \sqrt{\frac{2g f(x)}{1 + f'(x)^2}}. \]

Let us consider the forward motion of the particle only, therefore select the “+” sign. Since \( \dot{x} = dx/dt \), we get

\[ \frac{dx}{dt} = \sqrt{\frac{2g f(x)}{1 + f'(x)^2}}. \]

Separate the variables:

\[ dt = \sqrt{\frac{1 + f'(x)^2}{2g f(x)}}, \]

and integrate between \( t = 0 \) when \( x = 0 \), and \( t = T \) when \( x = x_2 \). We have

\[ \int_0^T dt = \int_{x_0}^{x_2} \frac{1 + f'(x)^2}{2g f(x)} \, dx. \]

We conclude that the time \( T \) to go from \( P_1 \) to \( P_2 \) is given by

\[ T(f) = \int_{x_0}^{x_1} \frac{1 + f'(x)^2}{2g f(x)} \, dx, \quad (5.2) \]

where we have written \( T \) as \( T(f) \) to emphasize the dependence of the elapsed time on the function \( f \), that is, on the slide’s shape. This leads to the following question, known as the brachistochrone problem:
Problem 5.2. Find a function \( f \) whose graph goes through the points \( P_1 \) and \( P_2 \), and which minimizes the travel time \( T(f) \).

In other words, we are looking for the shape of a slide that will let us slide from \( P_1 \) to \( P_2 \) in the fastest possible way.

A historical remark

The brachistochrone problem (from the Greek brachistos + chronos; brachistos = shortest, chronos = time) was proposed by Johann Benoulli in 1696 in *Acta Eruditorum*, one of the earliest European scientific journals. His solution, along with those of several other prominent mathematicians, was published in the journal in the following year. The various techniques developed for analyzing the problem were distilled, decades later, in the capable hands of Euler and Lagrange, into what became known as the calculus of variations.

5.3 Mathematical preliminaries

Before moving on to the analysis of the problems introduced in the previous section, and that of the calculus of variations in general, we collect here a few elementary mathematical concepts and results which facilitates the subsequent treatment. The presentation here follows closely that of Gelfand and Fomin [5].

5.3.1 Basic lemmas

We write \( C[a, b] \) for the set of continuous functions defined on the closed interval \([a, b]\). For any nonnegative integer \( n \), we write \( C^n[a, b] \) for the set of those functions in \( C[a, b] \) whose derivatives of order up to \( n \) are in \( C[a, b] \). In particular, \( C^0[a, b] = C[a, b] \).

Lemma 5.1. If \( \alpha \in C[a, b] \) and if

\[
\int_a^b \alpha(x) h'(x) \, dx = 0 \quad \text{for all } h \in \{ C^1[a, b] : h(a) = h(b) = 0 \},
\]

then \( \alpha \) is constant on \([a, b]\).

*Proof.* Define the constant \( c \) through

\[
\int_a^b [\alpha(x) - c] \, dx = 0,
\]

and let

\[
h(x) = \int_a^x [\alpha(s) - c] \, ds.
\]

Then \( h \) meets the requirements of (5.3). We also have \( h'(x) = \alpha(x) - c \), therefore

\[
\int_a^b [\alpha(x) - c]^2 \, dx = \int_a^b \alpha(x) \, dx - \int_a^b c \alpha(x) \, dx = \int_a^b \alpha(x) \, dx - \int_a^b \alpha(x) \, dx - \int_a^b c h'(x) \, dx = 0,
\]

whence \( \alpha \equiv c \). \( \Box \)
Lemma 5.2. Let $\alpha$ and $\beta$ be in $C[a, b]$ and
\[
\int_a^b \left[ \alpha(x)h(x) + \beta(x)b'(x) \right] dx = 0 \quad \text{for all } h \text{ in } \{ C^1[a, b] : b(a) = b(b) = 0 \}. \tag{5.4}
\]
Then $\beta$ is differentiable, and $\beta'(x) = \alpha(x)$ for all $x$ in $[a, b]$.

Proof. Let
\[
A(x) = \int_a^x \alpha(s) ds,
\]
then calculate as follows, applying integration by parts:
\[
\int_a^b \alpha(x)h(x) dx = \int_a^b A'(x)h(x) dx = A(x)h(x)|_a^b - \int_a^b A(x)h'(x) dx
\]
\[
= - \int_a^b A(x)h'(x) dx.
\]
Therefore (5.4) may be written as
\[
\int_a^b \left[ -A(x) + \beta(x) \right] h'(x) dx = 0 \quad \text{for all } h \text{ in } \{ C^1[a, b] : b(a) = b(b) = 0 \}.
\]
Then from Lemma 5.1 it follows that
\[
-A(x) + \beta(x) \equiv \text{constant on } [a, b],
\]
which implies, in particular, that $\beta \in C^1[a, b]$ since $A \in C^1[a, b]$. Upon differentiation we see that $\beta'(x) = A'(x) = \alpha(x)$. $\square$

5.3.2 The variation

An affine subspace is a subset of a linear space $X$ obtained by translating a subspace of $X$ parallel to itself. The illustration in Figure 5.2 shows a subspace $\mathcal{N}$ of $X$ which has been translated to produce the affine subspace $\mathcal{A}$. Thus, if $x_1$ and $x_2$ are in $\mathcal{A}$, then $x_1 - x_2$ is in $\mathcal{N}$. Equivalently, if $x_2 \in \mathcal{A}$, then every $x \in \mathcal{A}$ is of the form $x = x_2 + h$, where $h \in \mathcal{N}$.

Definition 5.3. Let $X$, $\mathcal{N}$, and $\mathcal{A}$ be as above, and suppose that $X$ is equipped with a norm $\| \cdot \|$. Consider the (generally nonlinear) function $J : \mathcal{A} \to \mathbb{R}$. The variation of $J$ at a point $x \in \mathcal{A}$ is a linear function $f_x : \mathcal{N} \to \mathbb{R}$, if one exists, with the property that
\[
J(x + b) = J(x) + f_x(b) + \epsilon \|b\|, \quad \text{for all } b \in \mathcal{N}, \tag{5.5}
\]
where $\epsilon \to 0$ as $\|b\| \to 0$.

Remark 5.1. The variation of a function is closely related to the derivative. In fact, the two concepts coincide if $\mathcal{N} = X$.

Remark 5.2. The traditional notation for the variation of $J$ is $\delta J$. However the $\delta$ symbol is assigned too many diverse roles in mechanics. I prefer the more innocuous
Figure 5.2: The oval represents a normed space $X$ with $O$ as the origin. The affine subspace $\mathcal{A}$ is obtained by a parallel translation of the subspace $\mathcal{N}$.

notation $f_x$, which has the additional advantage that it makes explicit the point $x$ at which it is evaluated.

**Theorem 5.4.** The variation of a function, if it exists, is unique.

**Proof.** Suppose that $J_x$ and $\hat{J}_x$ are both variations of $J$ at $x$. Then

$$J(x + b) = J(x) + J_x(b) + \epsilon_1 ||b||,$$

$$J(x + b) = J(x) + \hat{J}_x(b) + \epsilon_2 ||b||,$$

for all $b \in \mathcal{N}$. It follows that

$$J_x(b) - \hat{J}_x(b) = (\epsilon_1 - \epsilon_2)||b||$$

for all $b$ in $\mathcal{N}$.

Fix $b \in \mathcal{N}$ arbitrarily, and set $b_n = b/n$. Then we have

$$J_x(b_n) - \hat{J}_x(b_n) = (\epsilon_1 - \epsilon_2)||b_n||,$$

which has two consequences. First, by the linearity of $J_x$ and $\hat{J}_x$, and the homogeneity of the norm we have:

$$J_x(b_n) = \frac{1}{n} J_x(b), \quad \hat{J}_x(b_n) = \frac{1}{n} \hat{J}_x(b), \quad ||b_n|| = \frac{1}{n} ||b||,$$

therefore

$$J_x(b) - \hat{J}_x(b) = (\epsilon_1 - \epsilon_2)||b||.$$

Second, $b/n \to 0$ as $n \to \infty$, therefore $\epsilon_1$ and $\epsilon_2$ go to zero as $n \to \infty$. It follows that $J_x(b) - \hat{J}_x(b) = 0$ for all $b \in \mathcal{N}$, therefore $J_x = \hat{J}_x$. \qed

**Definition 5.5.** Let $J : \mathcal{A} \to \mathbb{R}$ be as in Definition 5.3. We say $J$ is stationary at $x \in \mathcal{A}$ if $J_x(b) = 0$ for all $b \in \mathcal{N}$. 
Chapter 5. Calculus of variations

Note that if \( J \) is stationary at \( x \), then it follows from (5.5) that

\[
\frac{J(x + h) - J(x)}{\|b\|} = \epsilon \to 0, \quad \text{as } b \to 0,
\]

which corresponds to the familiar statement from elementary calculus which says that a function is stationary when its derivative is zero.

5.4 The central problem of the calculus of variations

The minimum length problem of subsection 5.1.1, and the minimum time problem of subsection 5.2, are special cases of the following general problem.

Let \( X = C^1[a, b] \) equipped with the norm

\[
\|y\|_1 = \max_{a \leq x \leq b} \{|y(x)|, |y'(x)|\}, \quad y \in X.
\]  

(5.6)

Define the subspace \( \mathcal{N} \) of \( X \) through

\[
\mathcal{N} = \{ y \in X : y(a) = y(b) = 0 \}
\]

(5.7)

and the affine subspace \( \mathcal{A} \) of \( X \) through

\[
\mathcal{A} = \{ y \in X : y(a) = A, \ y(b) = B \}
\]

(5.8)

where \( A \) and \( B \) are given. Note that \( \mathcal{A} \) is a parallel translation of \( \mathcal{N} \); the difference of any two elements of \( \mathcal{A} \) is in \( \mathcal{N} \).

Let’s say we are given a function \( F : \mathbb{R}^3 \to \mathbb{R} \), which we will assume is twice continuously differentiable. Define the function \( J : \mathcal{A} \to \mathbb{R} \) through

\[
J(y) = \int_a^b F(x, y(x), y'(x)) \, dx, \quad y \in \mathcal{A},
\]

(5.9)

and then introduce the central problem of the calculus of variations:

Problem 5.3. Find \( y \in \mathcal{A} \) so that \( J \) is stationary at \( y \).

Remark 5.3. The function \( J \) is defined over the domain \( \mathcal{A} \) which consists of a set of functions. This is unlike the common functions studied in elementary calculus where function are typically defined over \( \mathbb{R} \) or \( \mathbb{R}^n \). In the modern theory of functions, a function is a mapping from a set (the domain) to another set (the range). The domain and range may be sets of any type, therefore the fact that the domain of \( J \) is a set of functions is of no way special. However, in the historical context in which the calculus of variations arose, a function whose domain was anything other than \( \mathbb{R}^n \) was an oddity, therefore such a function was called a functional, implying a “function-like object”. In modern times there is no particular justification for such a special terminology, nevertheless functions like \( J \) are often called functionals, carrying on the tradition.

Remark 5.4. Section 5.1.1’s curve of minimal length (equation (5.1)) corresponds to

\[
F(x, y, y') = \sqrt{1 + y'^2},
\]
5.4. The central problem of the calculus of variations

and Section 5.2’s brachistochrome problem (equation (5.2)) corresponds to

\[ F(x, y, y') = \sqrt{1 + y'^2}. \]

Note that we are using \( y' \) as a formal argument here—the prime in \( y' \) does not mean the derivative. The two functions \( F \) shown above could have been given equally well as

\[ F(x, y, z) = \sqrt{1 + z^2}, \quad F(x, y, z) = \sqrt{1 + z^2}. \]

but the \( y' \) notation is a useful reminder that the \( y' \) is a placeholder for the actual derivative \( y'(x) \) in the integrand of in \( J \) in (5.9).

**Theorem 5.6.** Consider the function \( J: \mathcal{A} \to \mathbb{R} \) defined in (5.9). If \( J \) is stationary at \( y \in \mathcal{A} \), then

\[ F_y - \frac{d}{dx} F_{y'} = 0, \tag{5.10} \]

where a subscript to \( F \) means the partial derivative of \( F \) with respect to the corresponding argument.

**Proof.** Let us say \( J \) is stationary at \( y \in \mathcal{A} \). From the Taylor series expansion of \( F \) about \( y '\):

\[
F(x, y + h, y' + b') = F(x, y, y') + F_y(x, y, y')h + F_{y'}(x, y, y')hb' \\
+ \frac{1}{2}F_{yy}(x, y, y')h^2 + F_{y'y'}(x, y, y')hb' + \frac{1}{2}F_{y'y'}(x, y, y')b'^2 + \cdots.
\]

Integrating over \([a, b]\) we obtain

\[
J(y + h) = J(y) + \int_a^b \left[ F_y(x, y, y')h + F_{y'}(x, y, y')hb' \right] dt + \epsilon\|h\|_1,
\]

where the norm \( \| \cdot \|_1 \) is defined in (5.6), and \( \epsilon \to 0 \) as \( \|b\|_1 \to 0 \). Then it follows from Definition 5.3 that the variation of \( J \) is given by

\[
J_y(h) = \int_a^b \left[ F_y(x, y, y')h + F_{y'}(x, y, y')hb' \right] dt.
\]

Since \( J \) is stationary at \( y \), we have

\[
\int_a^b \left[ F_y(x, y, y')h + F_{y'}(x, y, y')hb' \right] dt = 0, \quad \text{for all } h \in \mathcal{A}.
\]

Then according to Lemma 5.2, \( F_{y'}(x, y, y') \) is differentiable and \( \frac{d}{dx} F_{y'}(x, y, y') = F_{y'}(x, y, y') \).

\( \square \)

**Remark 5.5.** Equations (5.10) is called the Euler equation corresponding to the function \( J \). In expanded form it looks like this:

\[
\frac{\partial F}{\partial y'}(x, y(y'(x))) - \frac{d}{dx} \left( \frac{\partial F}{\partial y'}(x, y(y'(x))) \right) = 0, \quad x \in [a, b].
\]
This is a second order differential equation in $y$. It is to be solved for $y \in \mathcal{A}$ which supplies the boundary conditions $y(a) = A$ and $y(b) = B$ in accordance with (5.8).

In the frequently occurring special case where the function $F(x, y, y')$ in (5.9) has no explicit dependence on $x$, that is, when $J$ is of the form

$$J = \int_{a}^{b} F(y(x), y'(x)) \, dx, \quad y \in S,$$

(5.11)

then the second order differential (5.10) may be reduced to a first order equation, thus resulting in significant simplification. We state this as the following:

**Theorem 5.7.** If $y$ is a stationary point of $J$ defined in (5.11), then

$$F - y' F_{y'} = C$$

(5.12)

where $C$ is a constant.

**Proof.** To show that the left-hand side of (5.12) is a constant, let us compute its derivative:

$$\frac{d}{dx} \left( F(y(x), y'(x)) - y'(x) F_{y'}(y(x), y'(x)) \right)$$

$$= F_{y}(y(x), y'(x)) y'(x) + F_{y'}(y(x), y'(x)) y''(x)$$

$$- y''(x) F_{y'}(y(x), y'(x)) - y'(x) \frac{d}{dx} \left( F_{y'}(y(x), y'(x)) \right).$$

The second and third terms on the right-hand side cancel, and we are left with

$$\frac{d}{dx} \left( F(y(x), y'(x)) - y'(x) F_{y'}(y(x), y'(x)) \right)$$

$$= \left[ F_{y}(y(x), y'(x)) - \frac{d}{dx} \left( F_{y'}(y(x), y'(x)) \right) \right] y'(x).$$

The right-hand side is zero due to (5.10), completing the proof. \[\square\]

### 5.5 • The invariance of Euler’s equation under change of coordinates

A remarkable property of Euler’s equation is its invariance under change of coordinates in the following sense. Consider a change of variables from $(x, y)$ to $(u, v)$, where

$$x = X(u, v), \quad y = Y(u, v),$$

and suppose that the transformation is smooth and one-to-one, which in particular implies that Jacobian of the transformation is nonzero:

$$\det \begin{pmatrix} X_u & X_v \\ Y_u & Y_v \end{pmatrix} \neq 0.$$
A subscript indicates partial derivative, as in $X_u = \frac{\partial X}{\partial u}$. A curve given by the equation $y = y(x)$ in the $xy$-plane corresponds to a curve $v = v(u)$ in the $uv$-plane. In particular, we have

$$dx = X_u \, du + X_v \, dv, \quad dy = Y_u \, du + Y_v \, dv,$$

which has two immediate consequences. First, we have $dx = (X_u + X_v v') \, du$, and second,

$$y' = \frac{dy}{dx} = \frac{Y_u \, du + Y_v \, dv}{X_u \, du + X_v \, dv} = \frac{Y_u + Y_v \, v'}{X_u + X_v \, v'}.$$

With the aid of these, we change the variables from $x$ and $y$ to $u$ and $v$ in (5.9):

$$\int_a^b F(x, y(x), y'(x)) \, dx = \int_{\tilde{a}}^{\tilde{b}} \tilde{F}(u, v, v') \, du,$$

where $\tilde{a}$ and $\tilde{b}$ are the images of $a$ and $b$ under our transformation. Thus, we set

$$\tilde{F}(u, v, v') = F\left(X(u, v), Y(u, v), \frac{Y_u + Y_v \, v'}{X_u + X_v \, v'}\right)(X_u + X_v \, v'),$$

and define

$$\tilde{f}(v) = \int_{\tilde{a}}^{\tilde{b}} \tilde{F}(u, v, v') \, dv.$$

**Remark 5.6.** The rest of the argument requires the introduction of the concept of variational derivative which we haven’t done. See [5, Section 8] for the rest. But for now, perhaps the following heuristic argument will do.

Review Section 5.4 to verify for yourself that neither the statement of the problem (5.9), nor its solution (5.10) refer to any particular coordinate system. Therefore the $x$ and $y$ notation notwithstanding, the entire treatment is independent of a choice of coordinates. For instance, if $F(\theta, \rho, \rho')$ is constructed in polar coordinates $(\theta, \rho)$, then the stationary point is given by the differential equation

$$F_{\rho} - \frac{d}{d\theta} F_{\rho'} = 0.$$

This is an absolutely crucial observation in the development of the theory of Lagrangian mechanics.

### 5.6 The solution of Problem 5.1

To illustrate the previous section’s general development, let us apply it to solve Problem 5.1. In view of (5.1), we set $F(x, y, y') = \sqrt{1 + y'^2}$, therefore

$$\frac{\partial F}{\partial y} = 0, \quad \frac{\partial F}{\partial y'} = \frac{y'}{\sqrt{1 + y'^2}},$$

whence

$$\frac{d}{dx} \left( \frac{\partial F}{\partial y'} \right)_{(x, y, y'(x))} = \frac{d}{dx} \left( \frac{y'(x)}{\sqrt{1 + y'(x)^2}} \right).$$
Chapter 5. Calculus of variations

and the differential equation (5.10) takes the form

\[ \frac{d}{dx} \left( \frac{y'(x)}{\sqrt{1 + y'(x)^2}} \right) = 0. \]

Solving this is quite straightforward. The derivative of the parenthesized quantity is zero, therefore the parenthesized quantity is a constant, say \( c \):

\[ \frac{y'(x)}{\sqrt{1 + y'(x)^2}} = c. \]

It follows that \( y'(x) = c \sqrt{1 + y'(x)^2} \), therefore \( y'(x)^2 = c^2(1 + y'(x)^2) \), whence \( (1 - c^2)y'(x)^2 = c^2 \), which implies that \( y'(x) \) is constant, therefore the graph of \( y(x) \) is a straight line, as asserted.

**Remark 5.7.** Since \( F = \sqrt{1 + y'^2} \) has no explicit dependence on \( x \) in this case, we could have solved the problem by applying the alternative formula (5.12) which would have lead to

\[ \sqrt{1 + y'^2} - y'y' = C. \]

This simplifies to \( C \sqrt{1 + y'^2} = -1 \), whence we conclude that \( y' \) is a constant, as before.

### 5.7 The solution of Problem 5.1 in polar coordinates

In Section 5.5 we noted that the Euler equations are independent of the choice of coordinates. To illustrate this, let us solve Problem 5.1, that is, determining the shortest path between two points, in polar coordinates.

Thus, let \( \rho(\theta) \) be a function in the polar coordinates \((\theta, \rho)\) so that its graphs passes through the given points \( P_1(\theta_1, \rho_1) \) and \( P_2(\theta_2, \rho_2) \). In calculus it is shown that the path’s length is obtained through the formula

\[ \ell(\rho) = \int_{\theta_1}^{\theta_2} \sqrt{\rho(\theta)^2 + \rho'(\theta)^2} \, d\theta, \]

therefore \( F(\theta, \rho, \rho') = \sqrt{\rho^2 + \rho'^2} \). Note that \( F \) has no explicit dependence on \( \theta \). From here we get

\[ \frac{\partial F}{\partial \rho'} = \frac{\rho'}{\sqrt{\rho^2 + \rho'^2}}, \]

therefore the Euler equation (5.12) takes the form

\[ \sqrt{\rho^2 + \rho'^2} - \frac{\rho'^2}{\sqrt{\rho^2 + \rho'^2}} = c. \]

Multiplying the equation through by \( \sqrt{\rho^2 + \rho'^2} \) and simplifying, we get \( c \sqrt{\rho^2 + \rho'^2} = \rho^2 \), whence \( c(\rho^2 + \rho'^2) = \rho^4 \), and consequently, \( c \rho' = \rho \sqrt{\rho^2 - c^2} \), which may be solved through separation of variables:

\[ \int \frac{c \rho}{\rho \sqrt{\rho^2 - c^2}} \, d\rho = \int d\theta = \theta + k, \]
where \( k \) is the constant of integration.

The integral on the left-hand side is may be evaluated with the help of the change of variables \( \rho = c/u \). We have \( d\rho = -c/u^2 \, du \), therefore

\[
\int \frac{c \, d\rho}{\rho \sqrt{\rho^2 - c^2}} = \int \frac{c(-d\rho/u)}{\rho \sqrt{(c/u)^2 - c^2}} = -\int \frac{du}{\sqrt{1-u^2}} = -\cos^{-1} u = -\cos^{-1} \frac{c}{\rho}.
\]

We conclude that the solution of the differential equation is \( \frac{c}{\rho} = \theta + k \), that is,

\[
\frac{c}{\rho} = \cos(-\theta - k) = \cos(\theta + k) = \cos k \cos \theta - \sin k \sin \theta.
\]

Multiplying through by \( \rho \) noting that \( x = \rho \cos \theta \) and \( y - \rho \sin \theta \), where \( x \) and \( y \) are the Cartesian coordinates corresponding the the polar \( (\theta, \rho) \), the solution reduces to

\[
c = x \cos k + y \sin k,
\]

which is the equation of a straight line.

### 5.8 The solution of Problem 5.2

In view of (5.2), we set

\[
F(x, y, y') = \sqrt{\frac{1+y^2}{y}}.
\]

I have dropped the factor of \( 2g \) in the denominator a constant multiplier does not affect the location of a function's minimum. Note that \( F \) does not have an explicit dependence on \( x \), therefore the special Euler formula (5.12) yields,

\[
\sqrt{\frac{1+y^2}{y}} - \frac{y^2}{\sqrt{y(1+y^2)}} = C,
\]

which simplifies to

\[
y(1+y^2) = \frac{1}{C^2},
\]

Introducing the alternative constant \( a = 1/C^2 \), and then solving for \( y' \) we obtain

\[
y' = \sqrt{a - yy}.
\]

Upon separating the variables and integrating we get

\[
\int \sqrt{\frac{y}{a-y}} \, dy = \int dx = x + b. \quad (*)
\]

The integral on the left-hand side may be evaluated with the help of a change of variables of the form \( y = a \sin^2 \theta = \frac{1}{2} (1 - \cos 2\theta) \). With hindsight, the change of variables \( y = a \sin \frac{\theta}{2} \) produces a better-looking result, and that's what we will use. We have \( dy = a \sin \frac{\theta}{2} \cos \frac{\theta}{2} \, d\theta \), and

\[
\int \sqrt{\frac{y}{a-y}} \, dy = \int \sqrt{\frac{a \sin^2 \frac{\theta}{2}}{a - a \sin^2 \frac{\theta}{2}}} a \sin \frac{\theta}{2} \cos \frac{\theta}{2} \, d\theta
\]

\[
= a \int \sin^2 \frac{\theta}{2} \, d\theta = \frac{a}{2} \int (1-\cos \theta) \, d\theta = \frac{a}{2} (\theta - \sin \theta).
\]
From (\ref{eq:5.9}) we conclude that \( \theta - \sin \theta = x + b \). Thus, we have arrived at the general solution of our Euler equation in parametric form
\[
\begin{align*}
x &= \frac{a}{2}(\theta - \sin \theta) - b, \\
y &= \frac{a}{2}(1 - \cos \theta).
\end{align*}
\]
The constant \( a \) and \( b \) may be determined from the problem’s auxiliary conditions. Referring to Figure 5.1, we see that the solution curve is to go through the points \((0, 0)\) and \((x_2, y_2)\). The first of these conditions is fulfilled by taking \( b = 0 \). The second condition requires that
\[
\begin{align*}
x_2 &= \frac{a}{2}(\hat{\theta} - \sin \hat{\theta}), \\
y_2 &= \frac{a}{2}(1 - \cos \hat{\theta}),
\end{align*}
\]
where \( \hat{\theta} \) is the value of \( \theta \) at the point \( P_2 \). Upon solving this system for the unknowns \( a \) and \( \hat{\theta} \), we arrive at the solution of the brachistochrone problem, expressed parametrically as follows:
\[
\begin{align*}
x &= \frac{a}{2}(\theta - \sin \theta), \\
y &= \frac{a}{2}(1 - \cos \theta).
\end{align*}
\]
You will recognize that these are the parametric equations of a cycloid. Therefore the fastest playground slide is a cycloid that goes through the slide’s start and finish points.

5.9 • A variational problem in two unknowns

This section needs to be completely rewritten!

The variational problem defined in equation (5.9) involves a single unknown function \( y = y(x) \). The following generalization involves two unknowns, \( x(t) \) and \( y(t) \):
\[
J(x, y) = \int_{t_1}^{t_2} F(t, x(t), y(t), \dot{x}(t), \dot{y}(t)) dt,
\]
where a superposed dot indicates a derivative with respect to \( t \).

Consider the set \( \mathcal{X} \) of all smooth functions from \((t_1, t_2)\) to \( \mathbb{R} \) such that \( x(t_1) = a_1, \ y(t_1) = b_1, \ x(t_2) = a_2, \ y(t_2) = b_2 \), where \( a_1, b_1, a_2, b_2 \) are given.

Problem 5.4. Find functions \( x \in \mathcal{X} \) and \( y \in \mathcal{X} \) which minimize \( J(x, y) \).

Following the technique introduced in Section 5.4, it is not too hard to prove the following theorem whose proof will be left as an exercise:

Theorem 5.8. Consider the function \( f : \mathcal{X} \times \mathcal{X} \to \mathbb{R} \) defined in (5.13). If \( J \) achieves a minimum at \((x, y) \in \mathcal{X} \times \mathcal{X} \), then \( x \) and \( y \) satisfy the following system of second order differential equation
\[
\begin{align*}
F_x - \frac{d}{dt}F_{\dot{x}} &= 0, \\
F_y - \frac{d}{dt}F_{\dot{y}} &= 0.
\end{align*}
\]
5.10 • Lagrange multipliers

This section does not really necessary; consider removing it.

This section reviews the Lagrange multipliers technique which is usually covered in a multivariable calculus course. If you are familiar with the concept, have a quick look through this section nevertheless, since the methodology developed here carries in a parallel fashion to the subsequent sections.

Problem 5.5. Let the functions \( f : \mathbb{R}^3 \rightarrow \mathbb{R} \) and \( g : \mathbb{R}^3 \rightarrow \mathbb{R} \) be given. Find the minimum of \( f \) subject to \( g(x, y, z) = 0 \).

Geometrically, we think of \( g(x, y, z) = 0 \) as a surface \( S \) in \( \mathbb{R}^3 \), and we are asking for the minimum of \( f \) on that surface.

Solution. Let’s say \( f \) achieves a minimum at a point \( P_0(x_0, y_0, z_0) \in S \). The rest of the solution falls into two parts, which analyze the properties of \( g \) and \( f \) at \( P_0 \).

Part 1: The analysis of \( g \)

Let \( C \) be an arbitrary curve in \( S \) which passes through \( P_0 \), and let’s write \( (x_0 + \alpha(s), y_0 + \beta(s), z_0 + \gamma(s)) \) for the parametric representation of \( C \). The requirement that \( C \) passes through \( P_0 \) implies that \( \alpha(0) = \beta(0) = \gamma(0) = 0 \). The requirement that \( C \) lies within \( S \) implies that \( g(x_0 + \alpha(s), y_0 + \beta(s), z_0 + \gamma(s)) = 0 \) for all \( s \), and consequently

\[
\frac{\partial g}{\partial x} \alpha'(s) + \frac{\partial g}{\partial y} \beta'(s) + \frac{\partial g}{\partial z} \gamma'(s) = 0, \quad \text{for all } s,
\]

or equivalently,

\[
\left( \frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}, \frac{\partial g}{\partial z} \right) \cdot (\alpha'(s), \beta'(s), \gamma'(s)) = 0.
\]

This says that in particular at the point \( P_0 \) where \( s = 0 \) we have

\[
\nabla g \bigg|_{P_0} \perp (\alpha'(0), \beta'(0), \gamma'(0)).
\]

(5.15)

The vector \( (\alpha'(0), \beta'(0), \gamma'(0)) \) is tangent to the curve \( C \) at the point \( P_0 \). Since \( C \) is arbitrary, the vector \( (\alpha'(0), \beta'(0), \gamma'(0)) \) can be anything. We conclude that \( \nabla g \bigg|_{P_0} \) is perpendicular to all tangent vectors of all curves that pass through \( P_0 \), therefore it is perpendicular to the surface \( S \) at \( P_0 \).

Part 2: The analysis of \( f \)

We introduce the function

\[
\phi(s) = f(x_0 + \alpha(s), y_0 + \beta(s), z_0 + \gamma(s)) = 0
\]

which expresses the values taken on by the function \( f \) on the curve \( C \). Since \( f \) achieves a minimum at \( P_0 \) on \( S \), the \( \phi(s) \) achieves a minimum at \( s = 0 \), and therefore \( \phi'(0) = 0 \). We have

\[
\phi'(s) = \left( \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right) \cdot (\alpha'(s), \beta'(s), \gamma'(s)),
\]
therefore
\[ \nabla f \big|_{P_0} \perp \langle x'(0), y'(0), \gamma'(0) \rangle. \]  

(5.16)

We conclude that \( \nabla f \big|_{P_0} \) is perpendicular to all tangent vectors of all curves that pass through \( P_0 \), therefore it is perpendicular to the surface \( S \) at \( P_0 \).

### Part 3: Conclusion

We have seen that \( \nabla g \big|_{P_0} \) and \( \nabla f \big|_{P_0} \) are both perpendicular to the surface \( S \) at \( P_0 \). It follows that those two vectors are parallel, that is, there exists a scalar \( \lambda \) so that
\[ \nabla f = \lambda \nabla g \]  

at \( P_0 \).

The solution of Problem 5.5 then reduces to solving the system of 4 equations
\[
\frac{\partial f}{\partial x} = \frac{\lambda}{9} \frac{\partial g}{\partial x}, \quad \frac{\partial f}{\partial y} = \frac{\lambda}{2} \frac{\partial g}{\partial y}, \quad \frac{\partial f}{\partial z} = \frac{\lambda}{4} \frac{\partial g}{\partial z}, \quad g(x, y, z) = 0,
\]
for the four unknowns \( x, y, z, \lambda \).

**Remark 5.8.** The statement and conclusion of Problem 5.5 generalizes to \( \mathbb{R}^n \) in the obvious way.

**Example 5.9.** Find the point on the ellipse \( \frac{x^2}{9} + \frac{y^2}{4} = 1 \) which is closest to the point \((a, b)\).

**Solution.** We let \( f(x, y) = (x-a)^2 + (y-b)^2 \) be the square of the distance from any point \((x, y)\) to the point \((a, b)\); and \( g(x, y) = x^2/9 + y^2/4 - 1 \). Then the problem may be restated as: minimize \( f \) subject to \( g(x, y) = 0 \). Therefore we need to solve the nonlinear system of equations
\[
2(x-a) = \frac{2}{9} \lambda x, \quad 2(y-b) = \frac{1}{2} \lambda y, \quad \frac{x^2}{9} + \frac{y^2}{4} = 1.
\]

There is no simple solution to this system; the usual method of elimination leads to a fourth degree polynomial equation. If \( a \) and \( b \) are given as numbers, then we may solve the system numerically.

### 5.11 Calculus of variations with pointwise constraints

**This section needs to be completely rewritten!**

Gelfand & Fomin’s approach calls for the introduction of the variational derivative

Here is a variant of Problem 5.4 with an added constraint.

Consider the set \( \mathcal{X} \) of all smooth functions from \((t_1, t_2)\) to \( \mathbb{R} \) such that \( x(t_1) = a_1, y(t_1) = b_1, x(t_2) = a_2, y(t_2) = b_2 \), where \( a_1, b_1, a_2, b_2 \) are given.

**Problem 5.6.** The functions \( L : \mathbb{R}^3 \to \mathbb{R} \) and \( G : \mathbb{R}^2 \to \mathbb{R} \) are given. We define
\[ J(x, y) = \int_{t_1}^{t_2} L(t, x(t), y(t), \dot{x}(t), \dot{y}(t)) dt. \]

(5.17)

Find functions \( x \in \mathcal{X} \) and \( y \in \mathcal{X} \) which minimize \( J(x, y) \) subject to \( G(x, y) = 0 \).
5.11. Calculus of variations with pointwise constraints

Solution. Let $\alpha$ and $\beta$ be a pair of arbitrary smooth scalar-valued functions defined on some interval $(-\delta, \delta)$, and such that $\alpha(0) = \beta(0) = 0$.

Let $\xi$ and $\eta$ be a pair of arbitrary smooth scalar-valued functions defined on the interval $(t_1, t_2)$ which vanish at $t_1$ and $t_2$.

Suppose the constrained minimum of $J$ is achieved at $x \in \mathcal{X}$ and $y \in \mathcal{X}$. Then, if $\alpha$, $\beta$, $\xi$, and $\eta$ are such that

$$G(x(t) + \alpha(s) \xi(t), y(t) + \beta(s) \eta(t)) = 0, \quad -\delta \leq s \leq \delta, \quad t_1 \leq t \leq t_2,$$

we have

$$J(x + \alpha \xi, y + \beta \eta) \geq J(x, y).$$

Upon differentiation (5.18) with respect to $s$ and then setting $s$ to zero we get

$$\frac{\partial G}{\partial x} \xi(t) \alpha'(0) + \frac{\partial G}{\partial y} \eta(t) \beta'(0) = 0.$$ 

Now integrate the result

$$\alpha'(0) \int_{t_1}^{t_2} \frac{\partial G}{\partial x} \xi(t) dt + \beta'(0) \int_{t_1}^{t_2} \frac{\partial G}{\partial y} \eta(t) dt = 0.$$

We conclude that the vectors $\langle \alpha'(0), \beta'(0) \rangle$ and

$$\left\langle \int_{t_1}^{t_2} \frac{\partial G}{\partial x} \xi(t) dt, \int_{t_1}^{t_2} \frac{\partial G}{\partial y} \eta(t) dt \right\rangle$$

are orthogonal.

Next, let us introduce $\varphi : (-\delta, \delta) \to \mathbb{R}$ via

$$\varphi(s) = J(x + \alpha \xi, y + \beta \eta) = \int_{t_1}^{t_2} L(t, x(t) + \alpha(s) \xi(t), y(t) + \beta(s) \eta(t), \dot{x}(t) + \alpha(s) \dot{\xi}(t), \dot{y}(t) + \beta(s) \dot{\eta}(t)) dt.$$

Then the previous inequality takes the form $\varphi(s) \geq \varphi(0)$, which indicates that $\varphi$ achieves a minimum at $0$. We conclude that $\varphi'(0) = 0$. We proceed, therefore, to calculate $\varphi'(s)$.

We have:

$$\varphi'(s) = \int_{t_1}^{t_2} \left[ \frac{\partial L}{\partial x} \xi(t) \alpha'(s) + \frac{\partial L}{\partial y} \eta(t) \beta'(s) + \frac{\partial L}{\partial \dot{x}} \dot{\xi}(t) \alpha'(s) + \frac{\partial L}{\partial \dot{y}} \dot{\eta}(t) \beta'(s) \right] dt$$

$$= \alpha'(s) \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial x} \xi(t) + \frac{\partial L}{\partial \dot{x}} \dot{\xi}(t) \right) dt + \beta'(s) \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial y} \eta(t) + \frac{\partial L}{\partial \dot{y}} \dot{\eta}(t) \right) dt.$$ 

Then from $\varphi'(0) = 0$ we get that the vectors $\langle \alpha'(0), \beta'(0) \rangle$ and

$$\left\langle \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial x} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} \right) \xi(t) dt, \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial y} - \frac{d}{dt} \frac{\partial L}{\partial \dot{y}} \right) \eta(t) dt \right\rangle$$

are orthogonal.
Since the vectors (5.19) and (5.20) are both orthogonal to \((\alpha'(0), \beta'(0))\), and since the latter is arbitrary, we conclude that the two former vectors are parallel, therefore, there exists a function \(\lambda(t)\) so that
\[
\frac{d}{dt} \frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = \lambda \frac{\partial G}{\partial x}, \quad \frac{d}{dt} \frac{\partial L}{\partial \dot{y}} - \frac{\partial L}{\partial y} = \lambda \frac{\partial G}{\partial y}.
\]
(5.21)

**Example 5.10.** Find the curve of shortest length in \(\mathbb{R}^3\) that lies on the cylinder \(x^2 + y^2 = 1\) and connects the points ??.

**Solution.** Let the shortest curve be parametrized as \(\langle x(t), y(t), z(t) \rangle, \ t_1 \leq t \leq t_2\). The curve’s length is
\[
\int_{t_1}^{t_2} \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} \, dt.
\]
Therefore we let
\[
L(t, x(t), y(t), z(t), \dot{x}(t), \dot{y}(t), \dot{z}(t)) = \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}, \quad G(x, y, z) = x^2 + y^2 - 1,
\]
and apply the equations (5.21). We have
\[
\frac{\partial L}{\partial \dot{x}} = \frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}}, \quad \frac{\partial L}{\partial x} = 0, \quad \frac{\partial G}{\partial x} = 2x.
\]
These lead to the system of equations
\[
\left(\frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}}\right) = 2\lambda x, \quad \left(\frac{\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}}\right) = 2\lambda y, \quad \left(\frac{\dot{z}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}}\right) = 0,
\]
\[
x^2 + y^2 = 1.
\]
If the curve is parametrized by the arc length \(s\), then \(\dot{x}^2 + \dot{y}^2 + \dot{z}^2 = 1\), therefore the equations simplify to
\[
\ddot{x} = 2\lambda x, \quad \ddot{y} = 2\lambda y, \quad \ddot{z} = 0, \quad x^2 + y^2 = 1, \quad (5.22)
\]
where a superposed dot means differentiation with respect to \(s\). The \(z\) equation integrates readily to \(z(s) = as + b\).

To continue, let us differentiate the constraint equation twice. We get \(x\dot{x} + y\dot{y} = 0\) and \(\dot{x}^2 + x\ddot{x} + \dot{y}^2 + y\ddot{y} = 0\). For the second derivatives we substitute from (5.22) to get
\[
\dot{x}^2 + 2\lambda x^2 + \dot{y}^2 + 2\lambda y^2 = 0,
\]
which we rearrange into
\[
2\lambda (x^2 + y^2) + \dot{x}^2 + \dot{y}^2 = 0.
\]
This further simplifies by observing that:
1. \( x^2 + y^2 = 1 \) (the constraint equation);
2. \( \dot{x}^2 + \dot{y}^2 + \dot{z}^2 = 1 \) (the arc length parametrization) therefore \( \dot{x}^2 + \dot{y}^2 = 1 - \dot{z}^2 \); and
3. \( \dot{z} = a \).

We conclude that \( 2\lambda = 1 - a^2 \), therefore
\[
\lambda = \frac{1}{2}(1 - a^2).
\]

It just happens that in this case \( \lambda \) is a constant; in general we expect \( \lambda \) to be a function.

Going back to the differential equations in (5.22), we have
\[
\ddot{x} = (1 - a^2)x, \quad \ddot{y} = (1 - a^2)y,
\]
which yields the solution
\[
\begin{align*}
x(s) &= \cos(\sqrt{1 - a^2}s + c), \\
y(s) &= \sin(\sqrt{1 - a^2}s + c), \\
z(s) &= as + b,
\end{align*}
\]
which we recognize as the parametric equations of a helix.

5.12 • Calculus of variations with integral constraints

This section needs to be completely rewritten!

Gelfand & Fomin’s approach calls for the introduction of the variational derivative

Consider the set \( \mathcal{X} \) of all smooth functions from \((t_1, t_2)\) to \( \mathbb{R} \) such that \( x(t_1) = a_1 \), \( x(t_2) = a_2 \), where \( a_1 \) and \( a_2 \) are given. Similarly, let \( \mathcal{Y} \) be the set of all smooth functions from \((t_1, t_2)\) to \( \mathbb{R} \) such that \( y(t_1) = b_1 \), \( y(t_2) = b_2 \), where \( b_1 \) and \( b_2 \) are given.

**Problem 5.7.** The functions \( L : \mathbb{R}^5 \to \mathbb{R} \) and \( G : \mathbb{R}^5 \to \mathbb{R} \) are given. We define
\[
J(x, y) = \int_{t_1}^{t_2} L(t, x(t), y(t), \dot{x}(t), \dot{y}(t)) \, dt. \tag{5.23}
\]

Find the functions \( x \in \mathcal{X} \) and \( y \in \mathcal{Y} \) which minimize \( J(x, y) \) subject to
\[
\int_{t_1}^{t_2} G(t, x(t), y(t), \dot{x}(t), \dot{y}(t)) \, dt = 0.
\]

With a technique similar to those in the previous sections, it is not difficult to see that the solution of the problem is given by
\[
\frac{d}{dt} \frac{\partial}{\partial \dot{x}} (L + \lambda G) = \frac{\partial}{\partial x} (L + \lambda G), \quad \frac{d}{dt} \frac{\partial}{\partial \dot{y}} (L + \lambda G) = \frac{\partial}{\partial y} (L + \lambda G), \tag{5.24}
\]
where \( \lambda \) is a constant to be determined.
Example 5.11. Find the closed curve of a prescribed length $\ell$ in the plane which encloses the largest possible area.

Solution. Suppose the curve is given in parametric form as $(x(t), y(t))$, $0 \leq t \leq 1$, with $x(0) = x(1)$ and $y(0) = y(1)$. The area enclosed by the curve, according to Green’s Theorem, is

$$\frac{1}{2} \int_0^1 (x \dot{y} - y \dot{x}) \, dt$$

and the curve’s length is

$$\int_0^1 \sqrt{\dot{x}^2 + \dot{y}^2} \, dt.$$ (5.26)

Therefore we wish to maximize (5.25) subject to (5.26). We apply the equation (5.24) to solve this. We have:

$$L(t, x(t), y(t), \dot{x}(t), \dot{y}(t)) = \frac{1}{2}(x \dot{y} - y \dot{x})$$

and

$$G(t, x(t), y(t), \dot{x}(t), \dot{y}(t)) = \sqrt{\dot{x}^2 + \dot{y}^2},$$

therefore

$$L + \lambda G = \frac{1}{2}(x \dot{y} - y \dot{x}) + \lambda \sqrt{\dot{x}^2 + \dot{y}^2},$$

whence

$$\frac{\partial}{\partial \dot{x}}(L + \lambda G) = -\frac{1}{2} \dot{y} + \frac{\lambda \ddot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2}}, \quad \frac{\partial}{\partial \dot{y}}(L + \lambda G) = \frac{1}{2} \ddot{y} + \frac{\lambda \ddot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2}}.$$ (5.24)

If we pick the arc length, $s$, for parametrization, we get $\dot{x}^2 + \dot{y}^2 = 1$, therefore equations (5.24) lead to

$$\lambda \ddot{x} = \dot{y}, \quad \lambda \ddot{y} = -\dot{x}.$$ (5.24)

Integrating once we obtain

$$\lambda \dot{x} = y + c_1, \quad \lambda \dot{y} = -x + c_2.$$ (5.24)

We solve the resulting system of ODEs and arrive at

$$x(s) = A \sin \frac{1}{\lambda} s + B \cos \frac{1}{\lambda} s + c_3, \quad y(s) = A \cos \frac{1}{\lambda} s - B \sin \frac{1}{\lambda} s - c_3,$$ (5.24)

which we recognize as a parametric equation of a circle.
Exercises

5.1. **Minimal surface.** Find the function \( y(x), x_1 \leq x \leq x_2 \), whose graph passes through the points \((x_1, y_1)\) and \((x_2, y_2)\) in the Cartesian plane, and the surface of revolution obtained by rotating the graph about the \( x \) axis has the least area. See Figure 5.3.

5.2. Derive the Euler equations (5.14). *Hint:* Suppose \((x, y)\) is a minimizing pair. Then for any pair of scalars \( \alpha, \beta \in \mathbb{R} \), and any pair of functions \( \xi, \eta \) which vanish at \( t_1 \) and \( t_2 \), we have:

\[
J(x + \alpha \xi, y + \beta \eta) \geq J(x, y).
\]

For fixed \( \xi \) and \( \eta \), let

\[
\varphi(\alpha, \beta) = J(x + \alpha \xi, y + \beta \eta).
\]

Argue that \( \varphi \) achieves a minimum at \((0, 0)\), then proceed as on the proof of Theorem 5.6.

5.3. **The geodesic.** Find the differential equations corresponding to the shortest curve lying on the surface \( z = f(x, y) \) in \( \mathbb{R}^3 \), which connects the points \( P_1(x_1, y_1, f(x_1, y_1)) \) and \( P_2(x_2, y_2, f(x_2, y_2)) \). Such a curve is called a geodesic. *Hint:* Any such curve may be represented parametrically as

\[
r(t) = (u(t), v(t), f(u(t), v(t))), \quad 0 \leq t \leq 1,
\]

such that

\[
u(0) = x_1, \quad v(0) = y_1, \quad u(1) = x_2, \quad v(1) = y_2.
\]

Then we have

\[
dr(t) = \left( \dot{u}, \dot{v}, \frac{\partial f}{\partial u} \dot{u} + \frac{\partial f}{\partial v} \dot{v} \right) dt.
\]

Let \( ds \) be the infinitesimal curve length corresponding to the increment \( dt \). We have

\[
ds^2 = ||dr(t)||^2 = \left[ \dot{u}^2 + \dot{v}^2 + \left( \frac{\partial f}{\partial u} \dot{u} + \frac{\partial f}{\partial v} \dot{v} \right)^2 \right] dt^2,
\]

Figure 5.3: The graph of \( y(x) \) is revolved about the \( x \) axis (left) to produce a surface of revolution (right) (Exercise 1).
therefore the curve’s length is

\[ L(u, v, u', v') = \int_0^1 \left\{ \left[ 1 + \left( \frac{\partial f}{\partial u} \right)^2 \right] u'^2 + \left[ 1 + \left( \frac{\partial f}{\partial v} \right)^2 \right] v'^2 + 2 \frac{\partial f}{\partial u} \frac{\partial f}{\partial v} u' v' \right\}^{1/2} \, dt. \]

5.4. Apply the previous exercise’s general result to the special case where \( z = f(x, y) = \sqrt{1 - x^2} \), that is, the surface is (a part of) a circular cylinder. Conclude that geodesics are helixes.
Chapter 6
Lagrangian mechanics

6.1 Newtonian mechanics

Let \( r(t) \) be the position vector at time \( t \) of a particle (point mass) of constant mass \( m \) moving in the three-dimensional space under the influence of a force \( f(t) \). According to Newton, the equation of motion is \( m \ddot{r} = f \), where, to simplify the notation, I have written \( r \) and \( f \) for \( r(t) \) and \( f(t) \). A superimposed dot on a variable indicates the time derivative of that variable. Thus, \( \dot{r} \) is the particle’s velocity and \( \ddot{r} \) is its acceleration.

The motion of a collection of \( N \) particles is given as a set \( N \) vectorial equations

\[
m_k \ddot{r}_k = f_k, \quad k = 1, 2, \ldots, N,
\]

where \( m_k \) is the mass of the \( k \)th particle, \( r_k \) is its position vector, and \( f_k \) is the resultant of all forces acting on \( m_k \).

Example 6.1. Consider an idealized “dumbbell” consisting of two particles of masses \( m_1 \) and \( m_2 \), connected with a rigid massless rod, as shown in Figure 6.1(a). In the free flight of the dumbbell, as when it is tossed up in the air, the force exerted on \( m_1 \) is the resultant of the (known) weight vector \( m_1 g \) and the (unknown) push/pull \( f_{12} \) the rod.

That is, \( f_1 = m_1 g + f_{12} \).

Let us write \( r_k = \langle r_{k,1}, r_{k,2}, r_{k,3} \rangle \) and \( f_k = \langle f_{k,1}, f_{k,2}, f_{k,3} \rangle \) for the Cartesian representations of the vectors \( r_k \) and \( f_k \). Then the \( N \) vector equations above may equivalently be viewed as \( 3N \) scalar equations

\[
m_k \ddot{r}_{k,j} = f_{k,j}, \quad j = 1, 2, 3, \quad k = 1, 2, \ldots, N.
\]

The following obvious trick flattens the doubly-indexed variables into singly-index quantities and results in a significant algebraic simplification. We introduce the vectors

\[
x = \langle \underbrace{r_{1,1}, r_{1,2}, r_{1,3}}_{r_1}, \underbrace{r_{2,1}, r_{2,2}, r_{2,3}}_{r_2}, \ldots, \underbrace{r_{N,1}, r_{N,2}, r_{N,3}}_{r_N} \rangle, \tag{6.3a}
\]

\[
f = \langle \underbrace{f_{1,1}, f_{1,2}, f_{1,3}}_{f_1}, \underbrace{f_{2,1}, f_{2,2}, f_{2,3}}_{f_2}, \ldots, \underbrace{f_{N,1}, f_{N,2}, f_{N,3}}_{f_N} \rangle, \tag{6.3b}
\]

\[
m = \langle \underbrace{m_1, m_1, m_1}_{m_1}, \underbrace{m_2, m_2, m_2}_{m_2}, \ldots, \underbrace{m_N, m_N, m_N}_{m_N} \rangle. \tag{6.3c}
\]
and write (6.2) as

\[ m_i \ddot{x}_i = f_i, \quad i = 1, 2, \ldots, 3N. \] (6.4)

The change from (6.2) to (6.4) may seem merely cosmetic, but it entails a major change of philosophy and opens the doors to Lagrangian mechanics, as we shall see. Specifically, we view (6.4) as the differential equation of a motion of a point \( x \) in \( \mathbb{R}^{3N} \). According to (6.3a), knowing the position of the single point \( x \in \mathbb{R}^{3N} \) is equivalent to knowing the positions of the \( N \) points \( r_1, r_2, \ldots, r_N \) in the (physical) three-dimensional space. Thus, the study of the motion of a system of \( N \) points in the three-dimensional space is equivalent to the study of the motion of a single point in the abstract \( \mathbb{R}^{3N} \). Specifying an \( x \) in the \( \mathbb{R}^{3N} \) amounts to specifying the geometrical configuration of the particle system.

**Definition 6.2.** The \( 3N \)-dimensional space introduced above is called the mechanical system’s Cartesian configuration space. In analogy with Newton’s equations of motion, the vectors \( x, f, m \) defined in (6.3) are called the position, the force, and the mass of the single abstract “particle” moving in the configuration space.

**Remark 6.1.** Although it is tempting to think of the equation of motion (6.4) as a generalization of Newton’s equation \( m \ddot{x} = f \) to \( \mathbb{R}^{3N} \), the analogy is imperfect. The true generalization would have been

\[ m_i \ddot{x}_i = f_i, \quad i = 1, 2, \ldots, 3N, \]

involving only a single \( m \). In contrast, (6.4)’s fictitious “particle” exhibits different masses along different coordinate directions.

### 6.2 Holonomic constraints

The motion of a particle in the three-dimensional physical space traces a curve, as in the arc of a thrown ball, or the orbit of a planet. The motion of \( N \) particle then traces \( N \) curves in the three-dimensional space. The position \( x \) in configuration space, defined in (6.3a), merges the coordinates of the \( N \) particles into one, therefore the motion of the entire \( N \)-particle system appears as a single curve in the configuration space. We call that
curve the system’s orbit in the configuration space. When there is no risk of confusion, we will simply call it the orbit.

If there are no impediments in placing the particles independently in arbitrary positions in space, then the orbit of the system of $N$ particles may reach any point in the configuration space—all is needed is the application of an appropriate force to get there. If, however, the relative movements of the points are constrained, as in the dumbbell of Figure 6.1(a), only a subset of the configuration space may be reached.

**Example 6.3.** Let $r_1 = (r_{1,1}, r_{1,2}, r_{1,3})$ and $r_2 = (r_{2,1}, r_{2,2}, r_{2,3})$ be the position vectors of the dumbbell of Figure 6.1(a). Then, according to (6.3a) the position vector $x \in \mathbb{R}^6$ is given by

$$x = (r_{1,1}, r_{1,2}, r_{1,3}, r_{2,1}, r_{2,2}, r_{2,3}),$$

The constraint of the fixed length $\ell$ of the connecting rod is expressed as $||r_1 - r_2|| = \ell$, or more explicitly, as

$$(r_{1,1} - r_{2,1})^2 + (r_{1,2} - r_{2,2})^2 + (r_{1,3} - r_{2,3})^2 = \ell^2,$$

that is,

$$(x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_3 - x_4)^2 - \ell^2 = 0. \quad (6.5)$$

This defines a 5-dimensional “surface”—a manifold is the technical term—embedded in $\mathbb{R}^6$. The point orbit cannot roam arbitrarily in $\mathbb{R}^6$; it is constrained to stay on that manifold.

**Example 6.4.** Figure 6.1(b) shows three point masses connected with three massless rigid rods, and thus forming a rigid triangle. The position vectors $r_i, i = 1, 2, 3$, of the masses are constrained through the three constraint equations

$$||r_1 - r_2|| = \ell_{12}, \quad ||r_2 - r_3|| = \ell_{23}, \quad ||r_3 - r_1|| = \ell_{31},$$

which, in terms of the extended variable

$$x = (r_{1,1}, r_{1,2}, r_{1,3}, r_{2,1}, r_{2,2}, r_{2,3}, r_{3,1}, r_{3,2}, r_{3,3}),$$

$$= (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9),$$

take on the form

$$(x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_3 - x_4)^2 = \ell_{12}^2,$$

$$(x_4 - x_5)^2 + (x_5 - x_6)^2 + (x_7 - x_8)^2 = \ell_{23}^2,$$

$$(x_7 - x_1)^2 + (x_8 - x_2)^2 + (x_9 - x_3)^2 = \ell_{31}^2.$$

These confine the triangle’s orbit in the configuration space to a 6-dimensional manifold embedded in $\mathbb{R}^9$.

**Example 6.5.** Reconsider the previous example with a added twist. Suppose that the triangle’s rods are equipped with remote-controlled motors with may be activated to vary the rods’ lengths as desired during the flight. The previous example’s constraint equations take the form

$$(x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_3 - x_4)^2 = \ell_{12}(t)^2,$$

$$(x_4 - x_5)^2 + (x_5 - x_6)^2 + (x_7 - x_8)^2 = \ell_{23}(t)^2,$$

$$(x_7 - x_1)^2 + (x_8 - x_2)^2 + (x_9 - x_3)^2 = \ell_{31}(t)^2,$$
where \( \ell_{12}(t) \), \( \ell_{23}(t) \), and \( \ell_{31}(t) \) are given. The manifold \( \mathcal{M} \) in this case is a 6-dimensional manifold embedded in \( \mathbb{R}^9 \) whose shape changes with time.

**Example 6.6.** Recall the bead on the rotating hoop of Exercise 3 on page 6. With the obvious choice of the \( x,y,z \) coordinates, the position vector of the bead is

\[
\mathbf{r} = (R \sin \varphi \cos \Omega t, R \sin \varphi \sin \Omega t, R \cos \varphi).
\]

The manifold \( \mathcal{M} \) in this case is a spinning hoop itself. Geometrically it is a one-dimensional spinning object (a circle) embedded in \( \mathbb{R}^3 \). It is given by the pair of equations

\[
x \sin \Omega t = y \cos \Omega t, \\
x^2 + y^2 + z^2 = a^2.
\]

The first equation is that of a plane that contains the \( z \) axis and spins about it with an angular velocity of \( \Omega \). The second equation is that of a sphere of radius \( a \) centered at the origin. The intersection of the two objects is the spinning hoop.

In general, a system of \( N \) particles subject to \( M \) constraint equations of the form

\[
\varphi_i(x,t) = 0, \quad i = 1, 2, \ldots, M, \tag{6.6}
\]

where \( \varphi_i : \mathbb{R}^{3N} \times \mathbb{R} \to \mathbb{R}, i = 1, 2, \ldots, M \). These define a \((3N - M)\)-dimensional manifold \( \mathcal{M} \) embedded in \( \mathbb{R}^{3N} \). The system’s possible orbits are confined to lie in that manifold. Generally \( \mathcal{M} \) may move/deform with time, as it was the case in Examples 6.5 and 6.6. However, if the equations (6.6) are independent of time, as it was the case in Examples 6.3 and 6.4, then \( \mathcal{M} \) remains unchanged during the motion. That corresponds to a set of constraints of the form

\[
\varphi_i(x) = 0, \quad i = 1, 2, \ldots, M, \tag{6.6'}
\]

Constraints of type (6.6) are not the most general. Some very interesting mechanical systems impose constraints on the velocity, \( \dot{x} \), as in \( \varphi_i(x, \dot{x}, t) = 0 \). The rolling of a coin on the floor, for instance, has a constraint that depends on \( \dot{x} \), therefore (6.6) is inadequate for that purpose.

**Definition 6.7.** Constraints of the type (6.6) are called holonomic. All other types of constraints are called nonholonomic.

**Definition 6.8.** A mechanical system whose only constraints are of the holonomic type is called a holonomic system.

We will begin our study of Lagrangian dynamics with holonomic systems. Nonholonomic constraints will be brought up in the later chapters.

**Remark 6.2.** You may be interested to know that holonomic constraints of type (6.6) are called rheonomic while those of type (6.6') are called scleronomic. I prefer to call them with the more user-friendly terms “time-dependent” and “time-independent” instead.

**Remark 6.3.** The term holonomic was introduced by Hertz in [6]:

§123. A material system between whose possible positions all conceivable continuous motions are also possible motions is called a holonomous system. The term means that such a system obeys integral (\( \delta \lambda, \lambda \)) laws (\( \delta \mu, \mu \)), whereas material systems in general obey only differential conditions.
Admittedly that definition is rather vague, but its meaning is clarified further down:

§132. When from the differential equations of a material system an equal number of finite equations between the coordinates of the system can be deduced, the system is holonomous.

By “finite equations” he means algebraic, as opposed to differential, equations.

### 6.3 Generalized coordinates

In a holonomic system of \( N \) particles subject to \( M \) holonomic constraints, the \( 3N \) Cartesian components of the position vector \( \mathbf{x} \) are not quite suitable for the analysis of motion— they cannot serve as independent variables since they are interrelated through the \( M \) constraint equations (6.6). A much better approach is to parametrize the \( n = (3N - M) \)-dimensional configuration manifold \( \mathcal{M} \) through a suitably chosen \( n \) independent variables \( q_1, q_2, \ldots, q_n \), called the system’s generalized coordinates. The parametrization is certainly not unique, however in practice there often is an “obvious” choice. We write \( \mathbf{q} \) when we wish to refer to the \( n \) variables \( q_1, q_2, \ldots, q_n \) collectively.

The parameters \( \mathbf{q} \) form a (generally curvilinear) coordinate system on \( \mathcal{M} \). Since the motion’s orbit lies in \( \mathcal{M} \), the system’s state as a function of time may be expressed in terms of \( \mathbf{q}(t) \). The purpose of analytical mechanics is to express Newton’s equations of motion (6.4) in terms of the generalized coordinates \( \mathbf{q} \).

**Remark 6.4.** A familiar example curvilinear coordinates, albeit not directly related to mechanics, is the system of addressing locations on the surface of the Earth through their longitude \( \lambda \) and latitude \( \varphi \). In this context, \( \mathcal{M} \) is the Earth’s surface, and \( \lambda \) and \( \varphi \) are the coordinates \( q_1 \) and \( q_2 \).

Any \( \mathbf{q} \) identifies a point on the manifold \( \mathcal{M} \). Since \( \mathcal{M} \) is embedded in \( \mathbb{R}^{3N} \), it also identifies a point \( \mathbf{x} \in \mathbb{R}^{3N} \). That is, the system’s configuration vector \( \mathbf{x} \) is a function of \( \mathbf{q} \). We write this as \( \mathbf{x} = \mathbf{x}(\mathbf{q}, t) \), or in components:

\[
x_i = x_i(\mathbf{q}, t), \quad i = 1, 2, \ldots, 3N.
\] (6.7)

The \( t \) in these equations accounts for the possible motion/deformation of the manifold related to time-dependent constraints (6.6). In the case of time-independent constraints (6.6'), \( \mathcal{M} \) is independent of time, and (6.7) reduces to

\[
x_i = x_i(\mathbf{q}), \quad i = 1, 2, \ldots, 3N.
\] (6.7')

Differentiating the \( \mathbf{q} \) to \( \mathbf{x} \) mapping of (6.7) with respect to time, we obtain an expression for the velocities in terms of generalized coordinates:

\[
\dot{x}_i = \sum_k \frac{\partial x_i(\mathbf{q}, t)}{\partial q_k} \dot{q}_k + \frac{\partial x_i(\mathbf{q}, t)}{\partial t}.
\] (6.8)

[The last term will be absent in the case of (6.7').] Let us observe that although the position \( x_i \) is a function of \( \mathbf{q} \) and \( t \) only, the velocity \( \dot{x}_i \) is a function of \( \mathbf{q}, \dot{\mathbf{q}}, \) and \( t \). Let’s record this here for future reference:

\[
\dot{x}_i = \dot{x}_i(\mathbf{q}, \dot{\mathbf{q}}, t), \quad i = 1, 2, \ldots, 3N.
\] (6.9)

The following theorem establishes a couple of very useful mathematical identities:
Theorem 6.9. Let $x_i$ and $\dot{x}_i$ be as in (6.7) and (6.9). Then for any $i = 1, 2, \ldots, 3N$ and $j = 1, 2, \ldots, M$, we have:

$$\frac{\partial \dot{x}_i(q, \dot{q}, t)}{\partial q_j} = \frac{\partial x_i(q, t)}{\partial q_j}. \quad (6.10)$$

$$\frac{\partial \dot{x}_i(q, \dot{q}, t)}{\partial q_j} = \frac{d}{dt} \left( \frac{\partial x_i(q, t)}{\partial q_j} \right). \quad (6.11)$$

Proof. The assertion (6.10) is an immediate consequence of (6.8). As to (6.11), it’s a matter of differentiating (6.8) with respect to $q_j$ and then exchanging the differentiation order in the resulting second order partial derivatives:

$$\frac{\partial \dot{x}_i(q, \dot{q}, t)}{\partial q_j} = \sum_k \frac{\partial^2 x_i(q, t)}{\partial q_j \partial q_k} \dot{q}_k + \frac{\partial^2 x_i(q, t)}{\partial q_j \partial t} \dot{q}_k$$

$$= \sum_k \frac{\partial}{\partial q_k} \left( \frac{\partial x_i(q, t)}{\partial q_j} \right) \dot{q}_k + \frac{\partial}{\partial t} \left( \frac{\partial x_i(q, t)}{\partial q_j} \right)$$

$$= \frac{d}{dt} \left( \frac{\partial x_i(q, t)}{\partial q_j} \right).$$

6.4 Virtual displacements, virtual work, and generalized force

Figure 6.2 depicts a representation of the orbit of a system of $N$ particles on a manifold $\mathcal{M}$ embedded in $\mathbb{R}^{3N}$. Pick an arbitrary point, let’s say $\hat{x}$, of the orbit and then consider the tangent at that point to the manifold. We explore that tangent through infinitesimal excursions away from $\hat{x}$. Such excursions are called virtual displacements and commonly written as $\delta x$. I should emphasize that we are viewing the whole picture as a fossil frozen in time. The excursions have nothing to do with the system’s motion which will continue along the predetermined orbit once we unfreeze the time. The “$\delta x$” notation is used to distinguish between virtual displacements and the actual differential of the motion $d\mathbf{x}$.

The obvious way of producing a virtual displacement is through incrementing the generalized coordinates $q$. A change in $q$ amounts to a displacement within the manifold $\mathcal{M}$. Therefore the differential $\delta q$ is a displacement within $\mathcal{M}$’s tangent. In view of (6.7), we have:

$$\delta x_i = \sum_{j=1}^{n} \frac{\partial x_i}{\partial q_j} \delta q_j. \quad (6.12)$$

Let $f$ be the force vector, see (6.3b), at the point $\hat{x}$. Under a virtual displacement $\delta x$, the force performs a work $\delta W$, called virtual work, given by

$$\delta W = f \cdot \delta x = \sum_{i=1}^{3N} f_i \delta x_i = \sum_{i=1}^{3N} \left( \sum_{j=1}^{n} \frac{\partial x_i}{\partial q_j} \delta q_j \right) = \sum_{i=1}^{3N} \left( \sum_{j=1}^{n} f_i \frac{\partial x_i}{\partial q_j} \right) \delta q_j = \sum_{j=1}^{n} Q_j \delta q_j.$$

Letting

$$Q_j = \sum_{i=1}^{3N} f_i \frac{\partial x_i}{\partial q_j}, \quad j = 1, 2, \ldots, n, \quad (6.13)$$
the virtual work is now expressed as

$$\delta W = f \cdot \delta x = Q \cdot \delta q.$$ 

The vector $Q$ is called the \textit{generalized force} at $\dot{x}$. The component $Q_j$ is called the component of the generalized force along the generalized coordinate $q_j$.

\textbf{Example 6.10.} Consider the simple pendulum of Figure 1.1. The position vector $r = \langle \ell \sin \varphi, \ell \cos \varphi \rangle$, therefore the vector $x$ (see (6.3a)) is

$$x = \langle x_1, x_2 \rangle = \langle \ell \sin \varphi, \ell \cos \varphi \rangle,$$

and the constraint is $x_1^2 + x_2^2 - \ell^2$, therefore the configuration manifold $\mathcal{M}$ coincides with the circle swept by the pendulum’s bob, embedded in the configuration space $\mathbb{R}^2$. The angle $\varphi$ plays the role of the generalized coordinate in this case; any value of $\varphi$ identifies a point on $\mathcal{M}$. Let us write $\varphi$ and $Q_\varphi$ instead of $q_1$ and $Q_1$ for clarity. The force vector is $\langle 0, mg \rangle$. We compute the generalized force by applying (6.13):

$$Q_\varphi = f_1 \frac{\partial x_1}{\partial \varphi} + f_2 \frac{\partial x_2}{\partial \varphi} = 0 \times (\ell \cos \varphi) + m g \times (-\ell \sin \varphi) = -m g \ell \sin \varphi.$$

Observe that $Q_\varphi$ turns out to be equal to the moment of the weight vector $f$ about the pendulum’s pivot.

\textbf{Example 6.11.} Consider the double-pendulum of Figure 1.2. The position vectors of its two masses are given by

$$r_1 = \langle \ell_1 \sin \varphi, \ell_1 \cos \varphi \rangle, \quad r_2 = r_1 + \langle \ell_2 \sin \psi, \ell_2 \cos \psi \rangle,$$
therefore the vectors \( x \) and \( f \) (see (6.3)) are
\[
x = (x_1, x_2, x_3, x_4)
= (l_1 \sin \varphi, l_1 \cos \varphi l_1 \sin \varphi + l_2 \sin \psi, l_1 \cos \varphi + l_2 \cos \psi),
\]
\[
f = (f_1, f_2, f_3, f_4)
= (0, m_1 g, 0, m_2 g).
\]
The configuration space is \( \mathbb{R}^4 \) in this case. The two constraints
\[
(x_1 - x_2)^2 = \ell_1^2, \quad (x_3 - x_4)^2 = \ell_2^2
\]
result in a two-dimensional configuration manifold \( \mathcal{M} \) embedded in \( \mathbb{R}^4 \). The angles \( \varphi \) and \( \psi \) serve as generalized coordinates on \( \mathcal{M} \). Let us write \( \varphi \) and \( \psi \) for the generalized coordinates instead of the generic \( q_1 \) and \( q_2 \), for clarity. We write \( Q_\varphi \) and \( Q_\psi \) for the corresponding generalized forces. forces instead of \( Q_1 \) and \( Q_2 \). By applying (6.13) we get
\[
Q_\varphi = f_1 \frac{\partial x_1}{\partial \varphi} + f_2 \frac{\partial x_2}{\partial \varphi} + f_3 \frac{\partial x_3}{\partial \varphi} + f_4 \frac{\partial x_4}{\partial \varphi} = -m_1 g \ell_1 \sin \varphi - m_2 g \ell_2 \sin \psi,
\]
\[
Q_\psi = f_1 \frac{\partial x_1}{\partial \psi} + f_2 \frac{\partial x_2}{\partial \psi} + f_3 \frac{\partial x_3}{\partial \psi} + f_4 \frac{\partial x_4}{\partial \psi} = -m_2 g \ell_2 \sin \psi.
\]

### 6.5 External versus reaction forces

In equation (6.1) the force \( f_k \) applied to particle \( k \) is the resultant of all forces acting on that particle. For instance, in the triangular system of Figure 6.1(b), forces applied to \( m_1 \) consist of \( m_1 g + f_{11} + f_{13} \). The first term is the gravitational force applied to \( m_1 \), that is its weight, which is known. We call it an external force. The other two are generated dynamically within the rods, and are unknowns to be determined. We call then internal forces or more frequently, constraint reactions because they arise due to the unchanging lengths of the rods.

The constraint reactions get eliminated in the Lagrangian formulation as we shall see. Their elimination reduces the problem’s unknowns, and hence simplifies the equations of motion significantly. In anticipation of that development, we write the total force \( f_k \) in (6.1) as \( f_k + f_k' \), where, with some abuse of notations, we have recycled the notation \( f_k \) to signify the external forces only, and \( f' \) the internal forces, applied to the particle \( k \). After flattening the vectors in accordance with (6.3), equation (6.4) takes on the form
\[
m_i \ddot{x}_i = f_i + f_i', \quad i = 1, 2, \ldots, 3N.
\]

The argument that leads to the elimination of the constraint reactions proceeds as follows. The orbit of (6.14) lies in the constraint manifold \( \mathcal{M} \) in \( \mathbb{R}^{3N} \). External forces applied to the particles push and pull the point \( x \) in a direction tangent to \( \mathcal{M} \). But what keeps \( x \) from flying away from \( \mathcal{M} \)? The manifold holds it back, that’s what! If, for example, \( x \) speeds over a round protrusion on the manifold, centrifugal forces will tend to pull it away from the manifold. The manifold, however, exerts just the right amount of opposite force, the constraint reaction, which holds \( x \) attached to \( \mathcal{M} \). In the physical space, that is \( \mathbb{R}^{3} \), the manifold’s reactions manifests itself as the forces that develop in the system’s interconnecting links, such as \( f_{13} \) and \( f_{13} \) noted above.

The crucial observation that leads to the elimination of the constraint reactions from the equations of motion is that the constraint reaction is orthogonal to the constraint manifold. If it weren’t, then it would have a component tangent to the manifold, which will
then perform work during the motion. But such a behavior is uncharacteristic of a passive constraint surface, so we disallow it.

Since a virtual displacement $\delta x$ is tangent to the constraint manifold (see Figure 6.2), the orthogonality of the reaction force $f'$ to $M$ is expressed naturally as $f' \cdot \delta x = 0$ for all virtual displacements $\delta x$, or in expanded form

$$\sum_{i=1}^{3N} f'_i \delta x_i = 0 \quad \text{for all virtual displacements} \delta x.$$  

We note that, however, that due to (6.12)

$$\sum_{i=1}^{3N} f'_i \delta x_i = \sum_{i=1}^{3N} f'_i \left( \sum_{j=1}^{n} \frac{\partial x_i}{\partial q_j} \delta q_j \right) = \sum_{j=1}^{n} \left( \sum_{i=1}^{3N} f'_i \frac{\partial x_i}{\partial q_j} \right) \delta q_j,$$

therefore

$$\sum_{j=1}^{n} \left( \sum_{i=1}^{3N} f'_i \frac{\partial x_i}{\partial q_j} \right) \delta q_j = 0 \quad \text{for all virtual displacements} \delta q,$$

from which it follows that

$$\sum_{i=1}^{3N} f'_i \frac{\partial x_i}{\partial q_j} = 0, \quad j = 1, 2, \ldots, n. \quad (6.15)$$

### 6.6 The equations of motion for a holonomic system

At this point, the motion of a system consisting of $N$ point masses and $M$ holonomic constraints has been encapsulated into the $3N + M$ equations (6.14) and (6.6) in the unknowns $x_i$ and $f'_i$. It is the goal of this section to re-express the equations of motions as a system of only $n = 3N - M$ differential equations for the $n$ generalized coordinates $q$ as the unknowns. We begin with multiplying the equation (6.14) by $\partial x_i / \partial q_j$ and summing over $i$:

$$\sum_{i=1}^{3N} m_i \dot{x}_i \frac{\partial x_i}{\partial q_j} = \sum_{i=1}^{3N} f'_i \frac{\partial x_i}{\partial q_j} + \sum_{i=1}^{3N} f'_i \frac{\partial x_i}{\partial q_j}.$$  

The second summation on the right-hand side is zero due to (6.15). The first summation on the right-hand side is the generalized force $Q_j$; see (6.13). Therefore obtain

$$\sum_{i=1}^{3N} m_i \ddot{x}_i \frac{\partial x_i}{\partial q_j} = Q_j, \quad j = 1, 2, \ldots, n. \quad (6.16)$$

To simplify the left-hand side, we begin with a preliminary preparation. The kinetic energy of the system is

$$\hat{T}(\dot{x}) = \sum_{k=1}^{3N} \frac{1}{2} m_k \dot{x}_k^2.$$  

I have written $\hat{T}$ rather than the usual $T$ for a reason which will become obvious shortly. Now observe that for any $i$

$$\frac{\partial \hat{T}(\dot{x})}{\partial \dot{x}_i} = \frac{\partial}{\partial \dot{x}_i} \sum_{k=1}^{3N} \frac{1}{2} m_k \dot{x}_k^2 = m_i \ddot{x}_i,$$
therefore
\[ \frac{d}{dt} \frac{\partial \tilde{T}(\dot{x})}{\partial \dot{x}_i} = m \ddot{x}_i. \]

Then, the left-hand side of (6.16) may be calculated as
\[
\sum_{i=1}^{3N} m_i \ddot{x}_i \frac{\partial x_i}{\partial q_j} = \sum_{i=1}^{3N} \frac{d}{dt} \left[ \sum_{i=1}^{3N} \frac{\partial \tilde{T}(\dot{x})}{\partial \dot{x}_i} \frac{\partial x_i}{\partial q_j} \right] - \sum_{i=1}^{3N} \frac{\partial \tilde{T}(\dot{x})}{\partial \dot{x}_i} \frac{d}{dt} \left( \frac{\partial x_i}{\partial q_j} \right),
\]
where in the last step we have used the differentiation formula \( u'v = (uv)' - u'v' \). Now apply (6.10) to the first summation on the right-hand side, and apply (6.11) to the second summation, to get
\[
\sum_{i=1}^{3N} m_i \ddot{x}_i \frac{\partial x_i}{\partial q_j} = \sum_{i=1}^{3N} \frac{d}{dt} \left[ \sum_{i=1}^{3N} \frac{\partial \tilde{T}(\dot{x})}{\partial \dot{x}_i} \frac{\partial x_i}{\partial q_j} \right] - \sum_{i=1}^{3N} \frac{\partial \tilde{T}(\dot{x})}{\partial \dot{x}_i} \frac{d}{dt} \left( \frac{\partial x_i}{\partial q_j} \right).
\]

Let us recall that the Cartesian velocity components \( \dot{x} \) and the generalized velocity components \( \dot{q} \) are related through (6.8). Therefore the kinetic energy, which we have taken to be a function of \( \dot{x} \), may equally well be expressed as a function of \( q, \dot{q}, \) and \( t \). We write the latter as \( T \) to distinguish it from the former \( \tilde{T} \):
\[ \tilde{T}(\dot{x}) = T(q, \dot{q}, t), \]
and then note that by the chain rule
\[ \frac{\partial T(q, \dot{q}, t)}{\partial \dot{q}_j} = \sum_{i=1}^{3N} \frac{\partial \tilde{T}(\dot{x})}{\partial \dot{x}_i} \frac{\partial x_i}{\partial q_j} \quad \text{and} \quad \frac{\partial T(q, \dot{q}, t)}{\partial q_j} = \sum_{i=1}^{3N} \frac{\partial \tilde{T}(\dot{x})}{\partial \dot{x}_i} \frac{\partial x_i}{\partial q_j}. \]
Consequently
\[ \sum_{i=1}^{3N} m_i \ddot{x}_i \frac{\partial x_i}{\partial q_j} = \frac{d}{dt} \left[ \frac{\partial T(q, \dot{q}, t)}{\partial \dot{q}_j} \right] \frac{\partial T(q, \dot{q}, t)}{\partial q_j}, \]
and therefore (6.16) takes the form
\[ \frac{d}{dt} \left[ \frac{\partial T(q, \dot{q}, t)}{\partial \dot{q}_j} \right] \frac{\partial T(q, \dot{q}, t)}{\partial q_j} = Q_j, \quad j = 1, 2, \ldots, n. \quad (6.17) \]
These \( n \) second order differential equations in the \( n \) unknowns \( q_1, q_2, \ldots, q_n \) are called Lagrange's equation of motion for a holonomic system.

In particular, if the external forces \( Q \) are derived from a potential, that is, if there exists a scalar function function \( V(q, t) \) so that \( Q_j = -\frac{\partial V}{\partial q_j} \), then Lagrange's equation of motion take on the form
\[ \frac{d}{dt} \left[ \frac{\partial T(q, \dot{q}, t)}{\partial \dot{q}_j} \right] \frac{\partial T(q, \dot{q}, t)}{\partial q_j} = -\frac{\partial V(q, t)}{\partial q_j}, \quad j = 1, 2, \ldots, n. \]
In that case we define the system’s Lagrangian as

\[ L(q, \dot{q}, t) = T(q, \dot{q}, t) - V(q, t) \]

and observe that since \( V \) does not depend on \( q \), we have \( \partial L/\partial \dot{q}_j = \partial T/\partial \dot{q}_j \), therefore the equations of motion collapse to

\[ \frac{d}{dt} \left[ \frac{\partial L(q, \dot{q}, t)}{\partial \dot{q}_j} \right] - \frac{\partial L(q, \dot{q}, t)}{\partial q_j} = 0, \quad j = 1, 2, \ldots, n. \]  

(6.18)

**Exercises**

6.1. Find the generalized forces \( Q_\varphi \) and \( Q_\theta \) in the spherical pendulum of Figure 1.4 (page 7).

6.2. A massless hoop of radius \( a \) rolls without slipping on a horizontal line, while remaining in a vertical plane. A particle of mass \( m \) is firmly attached to the hoop, as seen in Figure 6.3. Use the angle \( \varphi \) of the mass’s radius relative to the vertical as generalized coordinate. Find the generalized force \( Q_\varphi \).

6.3. In the previous problem use the distance \( x \) travelled by the contact point (see the figure) as generalized coordinate. Find the generalized force \( Q_x \).

6.4. Suppose the external forces \( f_i \) in (6.14) are derived from a potential, that is, there exists a scalar-valued function \( \tilde{V}(x, t) \) such that \( f_i = -\partial \tilde{V}(x, t)/\partial x \) for \( i = 1, 2, \ldots, 3N \). Let \( V(q, t) = \tilde{V}(x, t) \) be the representation of the potential as a function of generalized coordinates. Show that the generalized forces \( Q_j \) are derived from the potential \( V \), that is, \( Q_j = \partial V(q, t)/\partial q_j \), for \( j = 1, 2, \ldots, n \).
Chapter 7
Angular velocity

When you hurl a rock in the air, or toss a Frisbee, the object spins in general as it moves. Associated with the motion is a vector \( \omega(t) \), called the object’s angular velocity vector, (or just angular velocity for short,) whose direction and magnitude at any instant of time \( t \) indicate orientation and the rate of rotation. It is the goal of this section to make the definition of the angular velocity precise.

To that end, let \( \{b_1, b_2, b_3\} \) be an orthonormal set of vectors which is firmly affixed to the spinning object, therefore moves with it. The choice of the letter \( b \) is this notation is to remind us that we are dealing with a body coordinate system.

We apply the general formula (8.7) (page 61) to express the rate of change, \( \dot{b}_i \), of the vector \( b_i \) in the basis \( \{b_1, b_2, b_3\} \):

\[
\dot{b}_i = \sum_{j=1}^{3}(b_i \cdot b_j)b_j.
\]

Let \( a_{ij} = b_i \cdot b_j \) be the matrix of the coefficients. We claim that the matrix, let’s call it \( A \), is skew-symmetric. Indeed, we have \( b_i \cdot b_j = \delta_{ij} \), where \( \delta_{ij} \) is the Kronecker delta defined in (8.6) (page 61). Upon differentiating this we get \( \dot{b}_i \cdot b_j + b_i \cdot \dot{b}_j = 0 \), whence \( a_{ij} + a_{ji} = 0 \), which proves \( A \) is skew-symmetric.

Since the general form of a \( 3 \times 3 \) skew-symmetric matrix is

\[
A = \begin{pmatrix}
0 & \omega_3 & -\omega_2 \\
-\omega_3 & 0 & \omega_1 \\
\omega_2 & -\omega_1 & 0
\end{pmatrix},
\]

therefore

\[
\dot{b}_1 = \omega_3 b_2 - \omega_2 b_3, \quad (7.2a)
\]
\[
\dot{b}_2 = \omega_1 b_3 - \omega_3 b_1, \quad (7.2b)
\]
\[
\dot{b}_3 = \omega_2 b_1 - \omega_1 b_2. \quad (7.2c)
\]

The coefficients \( \omega_1, \omega_2, \omega_3 \) measure the rate of change of the triad \( \{b_1, b_2, b_3\} \), and consequently, the rate of change of orientation of the body to which the triad is attached. The angular velocity vector, defined as

\[
\omega = \omega_1 b_1 + \omega_2 b_2 + \omega_3 b_3
\]

(7.3)
encapsulates that rate of change.

**Remark 7.1.** Actually calling $\omega$ a vector is premature. A vector, as defined in Section 8.1 is a physical object in the sense that it is independent of any coordinate system which may be used in defining it. Here $\omega$ has been defined in terms of its components on the $\{b_1, b_2, b_3\}$ triad. But the triad is not an integral part of the rotating body. Does $\omega$ survive if the triad goes away? Does $\omega$ remain the same if that triad is replaced by another? It turns out that the answer to both of those question is in the positive. Indeed, $\omega$ is not an artifact of the arbitrarily chosen triad. It is an intrinsic property of the body's motion. The presence of a coordinate triad affixed to the body is immaterial. The proof of this claim is not too hard but it takes us deeper into tensor analysis, so I will skip it for now.

**Remark 7.2.** We leave it as an exercise to show that differentiating (7.3) and applying the equations (7.2), results in

$$\omega = \dot{\omega}_1 b_1 + \dot{\omega}_2 b_2 + \dot{\omega}_3 b_3.$$  \hspace{1cm} (7.4)

**Remark 7.3.** On the one hand, in (7.1) we have $a_{23} = \omega_1$. On the other hand, we have the $a_{ij} = \dot{b}_i \cdot b_j$ by definition. It follows that $\omega_1 = \dot{b}_2 \cdot b_3$. Similar expressions are obtained for $\omega_2$ and $\omega_3$. Let’s make a record of this:

$$\omega_1 = \dot{b}_2 \cdot b_3, \quad \omega_2 = \dot{b}_3 \cdot b_1, \quad \omega_3 = \dot{b}_1 \cdot b_2.$$  \hspace{1cm} (7.5)

**Remark 7.4.** From the definition of $\omega$ in (7.3) it follows that

$$\omega \times b_1 = (\omega_1 b_1 + \omega_2 b_2 + \omega_3 b_3) \times b_1 = -\omega_2 b_3 + \omega_3 b_2.$$ 

Therefore the equations (7.2) may be written as

$$\dot{b}_1 = \omega \times b_1, \quad \dot{b}_2 = \omega \times b_2, \quad \dot{b}_3 = \omega \times b_3.$$  \hspace{1cm} (7.6)

**Remark 7.5.** Consider points $O$ and $P$ in the body, and let $\vec{r} = \overrightarrow{OP}$. Then $r = \alpha_1 b_1 + \alpha_2 b_2 + \alpha_3 b_3$, where $\alpha_1, \alpha_2, \alpha_3$ are constants. The velocity $v$ of $P$ relative to $O$ is given by $v = \dot{\vec{r}}$. Let us observe that

$$v = \dot{r} = \alpha_1 \dot{b}_1 + \alpha_2 \dot{b}_2 + \alpha_3 \dot{b}_3$$

$$= \alpha_1 (\omega \times b_1) + \alpha_2 (\omega \times b_2) + \alpha_3 (\omega \times b_3) = \omega \times (\alpha_1 b_1 + \alpha_2 b_2 + \alpha_3 b_3) = \omega \times r.$$ 

The formula

$$v = \omega \times r$$  \hspace{1cm} (7.7)

is used throughout the rest of this book.

## Exercises

7.1. Derive (7.4).
Chapter 8

The moment of inertia tensor

8.1 A brief introduction to tensor algebra

In much of our working and thinking, we tend to blur the distinction between a vector as a pointed arrow in space $\mathbb{R}^3$, and an $n$-tuple of numbers $(x_1, x_2, \ldots, x_n)$. Yet, the two concepts are drastically different. The former would have been quite a natural object to Euclid in 300 BC, but the latter would have been very foreign to him—for at the time they lay two millennia in the future.

The blurring between the two ways of looking at vectors is harmless much of the time, but there are places where a strict distinction between the two views is crucially important. Tensor algebra makes a bridge between the two and goes substantially beyond. Tensor algebra, along with tensor analysis, are indispensable tools in differential geometry, continuum mechanics, and general relativity, and perhaps other disciplines as well. For a general exposition of tensor algebra and tensor analysis, see [7].

Tensor algebra is not a sine qua non of rigid body mechanics, therefore only rarely it is brought into play. There is, however, the ubiquitous use of the term moment of inertia tensor, which hints tantalizingly to a connection to tensors behind the scenes, however the connection is only rarely brought out. It is the purpose of this section and the next to introduce the minimal tensor algebra which explains the “tensor” in the “moment of inertia tensor”.

8.1.1 Tensor algebra

Vectors, dot product, cross product. In the three-dimensional Euclidean space fix a point called the origin, and consider the set $\mathcal{V}$ of all vectors (in the sense of pointed arrows) whose tails are attached to the origin. We define the sum $x + y$ of two vectors $x$ and $y$ in $\mathcal{V}$ through the parallelogram rule, that is, we form a parallelogram based on the vectors, and consider the parallelogram’s diagonal as their sum as illustrated in Figure 8.1. Multiplying a vector by a number stretches/shrinks the vector’s length by the magnitude of that number. If the number is negative, the resulting vectors flips, that is, it points in the opposite of the original’s direction. See Figure 8.1 for an illustration.

The length of the vector $x$ is written $||x||$. The dot product $x \cdot y$ of a pair of vectors $x$ and $y$ in $\mathcal{V}$ is a number defined by

$$x \cdot y = ||x|| ||y|| \cos \theta,$$

where $\theta$ is the angle between the vectors; see Figure 8.2. Let us note that (a) the dot
product is commutative, that is \( x \cdot y = y \cdot x \); (b) if the vectors are orthogonal to each other, that is, \( \theta = \pi/2 \), then \( x \cdot y = 0 \); and (c) a vector’s length may be expressed as a dot product: \( ||x||^2 = x \cdot x \).

The cross product of \( x \times y \) of a pair of vectors \( x \) and \( y \) in \( \mathcal{V} \) is the vector

\[
x \times y = (||x||||y|| \sin \theta) n,
\]

where \( 0 \leq \theta \leq \pi \) is the angle between the vectors, and \( n \) is a unit vector which is perpendicular to the plane of the vectors \( x \) and \( y \), and is oriented according to the right-hand rule. The latter means that if you align your right hand’s thumb with \( n \), then a rotation by an angle \( \theta \) in the direction pointed at by your fingers will take the vector \( x \) to \( y \). It follows that the cross product is anticommutative, which means that \( x \times y = -y \times x \).

**Tensors.** A function \( L : \mathcal{V} \to \mathcal{V} \) is said to be linear if

\[
L(x + y) = L(x) + L(y) \quad \text{for all } x, y \in \mathcal{V},
\]

\[
L(\alpha x) = \alpha L(x) \quad \text{for all } x \in \mathcal{V}, \alpha \in \mathbb{R}.
\]

A linear function from \( \mathcal{V} \) to \( \mathcal{V} \) is called a tensor on \( \mathcal{V} \), or just tensor, for short. It is customary to write \( Lx \) instead of \( L(x) \) when \( L \) is a tensor, as we will do from now on. The identity tensor, \( I \), is the tensor with the property

\[
Ix = x \quad \text{for all } x \in \mathcal{V}.
\]

The sum \( L_1 + L_2 \) of tensors \( L_1 \) and \( L_2 \) is a tensor defined by

\[
(L_1 + L_2)x = L_1x + L_2x \quad \text{for all } x \in \mathcal{V}.
\]
The product $\alpha L$ of a tensor $L$ and a number $\alpha \in \mathbb{R}$ is a tensor defined by

$$(\alpha L)x = \alpha (Lx) \quad \text{for all } x \in \mathcal{V}.$$ (8.3)

The set of tensors on $\mathcal{V}$, equipped with the operations defined in (8.2) and (8.3), is a vector space in the sense of an abstract vector space (not to be confused with vectors of the pointed-arrow kind).

**The dyadic product.** The dyadic product $a \otimes b$ of vectors $a$ and $b$ in $\mathcal{V}$ is a tensor on $\mathcal{V}$ defined by

$$(a \otimes b)x = (b \cdot x)a \quad \text{for all } x \in \mathcal{V}.$$ (8.4)

**Lemma 8.1.** For any pair of vectors $a, b \in \mathcal{V}$ we have

$$||a \times b||^2 = a \cdot (||b||^2 I - b \otimes b)a,$$ (8.5)

where $I$ is the identity tensor.

**Proof.** Let $\theta$ be the angle between the vectors $a$ and $b$. We have:

$$||a \times b||^2 = ||a||^2 ||b||^2 \sin^2 \theta = ||a||^2 ||b||^2 (1 - \cos^2 \theta) = ||a||^2 ||b||^2 - ||a||^2 ||b||^2 \cos^2 \theta = ||b||^2 (a \cdot a - (a \cdot b)^2) = a \cdot [||b||^2 a - (a \cdot b)b] = a \cdot [||b||^2 a - (b \otimes b)a] = a \cdot (||b||^2 I - b \otimes b)a.$$

[Box]

**8.1.2 Connection with $\mathbb{R}^3$ and $3 \times 3$ matrices**

Let $\{e_1, e_2, e_3\}$ be an orthonormal set in $\mathcal{V}$. This means that the three vectors are mutually perpendicular, and each is of length one. Thus

$$e_i \cdot e_j = \delta_{ij} = \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{if } i = j. \end{cases}$$ (8.6)

The symbol $\delta_{ij}$ defined above is called the Kronecker delta.

In Exercise 1 you will show that an orthonormal set is linearly independent. If follows that the orthonormal set $\{e_1, e_2, e_3\}$ forms a basis for the three-dimensional vector space $\mathcal{V}$. Consequently, any $x \in \mathcal{V}$ may be expressed as a linear combination of the form

$$x = c_1 e_1 + c_2 e_2 + c_3 e_3,$$

Dot-multiplying through by $e_1$ and taking into account the orthonormality of basis vectors, we get $x \cdot e_1 = c_1$. Repeating with $e_2$ and $e_3$ we conclude that $c_i = x \cdot e_i$, $i = 1, 2, 3$. Thus we have established the identity

$$x = (x \cdot e_1)e_1 + (x \cdot e_2)e_2 + (x \cdot e_3)e_3, \quad \text{for all } x \in \mathcal{V}.$$ (8.7)
For each $i$, the coefficient $x \cdot e_i$ is called the component of the vector $x$ along $e_i$.

**Remark 8.1.** Let us state emphatically that the components $x \cdot e_i$ are not properties of the vector $x$ at all! If you replace one basis with another which is rotated in an arbitrary manner relative to the first, then the components of $x$ will be different in general, while $x$ has not been touched. Even if you remove the basis altogether, the vector $x$ will happily continue to exist. Having said all that, it is sometimes useful to work with $\langle c_1, c_2, c_3 \rangle \in \mathbb{R}^3$ as sort of an “avatar” of $x \in \mathcal{V}$, as long you remain cognizant of what it is.

**Remark 8.2.** Applying the definition of the diatic product (8.4), the identity (8.7) may be written in the equivalent form

$$x = (e_1 \otimes e_1) x + (e_2 \otimes e_2) x + (e_3 \otimes e_3) x,$$

which has at least two implications. First, upon factoring the $x$ on the right-hand side, we see that

$$I = e_1 \otimes e_1 + e_2 \otimes e_2 + e_3 \otimes e_3.$$

Second, $e_i \otimes e_i$ acting on any vector $x$ produces the projection of $x$ in the direction $e_i$.

We noted earlier that the set of tensors on $\mathcal{V}$, equipped with the operations defined in (8.2) and (8.3), is a vector space of its own. The following lemma shows how to construct a basis for that vector space.

**Lemma 8.2.** Let $\{e_1, e_2, e_3\}$ be an orthonormal basis in $\mathcal{V}$. Then $\{e_i \otimes e_j\}_{i,j=1}^{3}$ is a basis for the space of tensors on $\mathcal{V}$.

**Proof.** We will show that any tensor $L$ is a linear combination of the nine dyadic products $\{e_i \otimes e_j\}_{i,j=1}^{3}$. Toward that end, pick an arbitrary $x \in \mathcal{V}$,

$$x = \sum_{i=1}^{3} (x \cdot e_i) e_i,$$

and let $y = Lx$. By the linearity of $L$ we have

$$y = Lx = \sum_{i=1}^{3} (x \cdot e_i) Le_i,$$

it follows that

$$y \cdot e_j = \sum_{i=1}^{3} (x \cdot e_i) (Le_i \cdot e_j) = \sum_{i=1}^{3} (Le_i \cdot e_j) (x \cdot e_i),$$

and consequently

$$y = \sum_{j=1}^{3} (y \cdot e_j) e_j$$

$$= \sum_{j=1}^{3} \sum_{i=1}^{3} (Le_i \cdot e_j) (x \cdot e_i) e_j$$

$$= \sum_{j=1}^{3} \sum_{i=1}^{3} (Le_i \cdot e_j) (e_i \otimes e_j) x.$$
8.1. A brief introduction to tensor algebra

where we have made use of the definition (8.4). Since $y = Lx$, this tells us that

\[ Lx = 3 \sum_{i=1}^{3} \sum_{j=1}^{3} (Le_i \cdot e_j)(e_i \otimes e_j)x \quad \text{for all } x \in \mathcal{V}, \]

therefore

\[ L = 3 \sum_{i=1}^{3} \sum_{j=1}^{3} (Le_i \cdot e_j)(e_i \otimes e_j), \tag{8.8} \]

which shows that $L$ is a linear combination of the nine dyadic products $\{e_i \otimes e_j\}_{i,j=1}^{3}$. In other words, the set of vectors $\{e_i \otimes e_j\}_{i,j=1}^{3}$ spans the set of all tensors. To show that the set is a basis, it remains to show that it is linearly independent. You will do that in Exercise 2.

Remark 8.3. The coefficients in (8.8) are called the components of $L$ in the basis $\{e_i \otimes e_j\}_{i,j=1}^{3}$. Letting $\ell_{ij} = Le_i \cdot e_j$, we may write (8.8) in the abbreviated form

\[ L = 3 \sum_{i=1}^{3} \sum_{j=1}^{3} \ell_{ij} e_i \otimes e_j. \]

At times it is useful to identify the tensor $L$ with the $3 \times 3$ matrix with components $\ell_{ij}$ but the caveats of Remark 8.1 apply equally well here. The components $\ell_{ij}$ are artifacts of the choice of the basis vectors. The tensor is an intrinsic property of the system and will continue to exist even when the basis vectors are obliterated.

Remark 8.4. In view of the one-to-one correspondence between $3 \times 3$ matrices and tensors on $\mathcal{V}$ noted above, every property or theorem in matrix algebra finds a counterpart in tensor analysis. For instance, a nonzero vector $x$ is called and eigenvector of the tensor $L$ if $Lx = \lambda x$ for some $\lambda \in \mathbb{R}$. The coefficient $\lambda$ is the eigenvalue corresponding to that eigenvector.

8.1.3 Symmetric tensors

A tensor $L$ is said to be symmetric if

\[ a \cdot (Lb) = (La) \cdot b \quad \text{for all } a, b \in \mathcal{V}. \]

It is easy to show that if $L$ is symmetric, then the matrix $\ell_{ij}$ of its components (see Remark 8.3) is symmetric, that is $\ell_{ij} = \ell_{ji}$ for $i, j = 1, 2, 3$.

We know from matrix analysis that the eigenvalues of a symmetric matrix are real, and its eigenvector may selected as an orthonormal set. This, along with what is known as diagonalization of matrices in linear algebra, lead to the spectral decomposition theorem in tensor algebra:

**Theorem 8.3 (Spectral decomposition, cf. [7, p. 137]).** The eigenvalues $\lambda_1, \lambda_2, \lambda_3$ of a symmetric tensor $L$ are real, and the corresponding eigenvectors, $g_1, g_2, g_3$ may be selected as an orthonormal set. Furthermore,

\[ L = \lambda_1 g_1 \otimes g_1 + \lambda_2 g_2 \otimes g_2 + \lambda_3 g_3 \otimes g_3. \tag{8.9} \]

\[ ^7 \text{say something about complex fields} \]
Remark 8.5. According to (8.8), a tensor is a linear combination of the nine basis elements \( \{ e_i \otimes e_j \}_{i,j=1}^3 \) in general. In contrast, (8.9) presents \( L \) as a linear combination of only three special elements \( \{ g_j \otimes g_j \}_{j=1}^3 \) constructed from \( L \)'s eigenvectors.

8.2 The moment of inertia tensor

Consider a set of particles of masses \( m_i, i = 1, 2, \ldots, N \), which are thoroughly interconnected through massless rigid rods, so that the entire assembly forms a rigid object. Suppose that the object rotates with angular velocity \( \omega \) about a fixed axis passing through the origin. Figure 8.3 depicts the vector \( \omega \), and a representative particle of mass \( m_i \) and position vector \( r_i \). Then the particle’s velocity is \( v_i = \omega \times r_i \), therefore the kinetic energy of the \( N \)-particle system is given by

\[
T = \sum_{i=1}^{N} \frac{1}{2} m_i \|v_i\|^2 = \sum_{i=1}^{N} \frac{1}{2} m_i \|\omega \times r_i\|^2,
\]

which according to Lemma 8.1 is equivalent to

\[
T = \sum_{i=1}^{N} \frac{1}{2} m_i \omega \cdot (\|r_i\|^2 I - r_i \otimes r_i) \omega
\]

\[
= \frac{1}{2} \omega \cdot \left( \sum_{i=1}^{N} m_i (\|r_i\|^2 I - r_i \otimes r_i) \right) \omega.
\]

We introduce the tensor

\[
J = \sum_{i=1}^{N} m_i (\|r_i\|^2 I - r_i \otimes r_i),
\]

(8.10)

whereby the kinetic energy takes the simple form

\[
T = \frac{1}{2} \omega \cdot J \omega.
\]

(8.11)

The tensor \( J \) is called the moment of inertia tensor of the \( N \)-particle system. Let us note that \( J \) is independent of \( \omega \), so the orientation of the axis about which the system rotates, or the speed of rotation, is immaterial. However, \( J \) does depend on the choice of the origin of the vectors—changing the origin will affect the position vectors \( r_i \), therefore...
the the tensor $\mathcal{J}$. The change, however, obeys a simple translation rule which we will develop in the next section. For now let us observe another aspect of (8.10). The fact that the system under consideration consists of $N$ rigidly connected point masses is hardly of particular importance. A general rigid solid may be approximated by a union a large number of tiny parts, as one does in the theory of integration, and then pass to the limit as the number of the parts goes to infinity, and the sizes of the individual parts go to zero, while maintaining a fix mass for the aggregate.

To be specific, let $B$ be the solid object, $dm$ be the differential mass of the "part" of $B$ indicated by the position vector $r$ relative to some origin $o$. Then the obvious extension of (8.10) takes the following form for the solid’s moment of inertia:

$$\mathcal{J} = \int_B (||r||^2 I - r \otimes r) dm.$$  

(8.12)

If $\rho(r)$ is the density of the body at the position $r$, then $dm = \rho \, dV$, where $V$ is the volume element, and the formula above takes the form

$$\mathcal{J} = \int_B \rho(r) (||r||^2 I - r \otimes r) dV.$$  

(8.13)

### 8.3 Translation of the origin

As noted above, the moment of inertia tensor of a rigid body depends on the choice of the origin of the vectors. To see how a translation of the origin affects the tensor, let $\mathcal{J}_o$ and $\mathcal{J}_{o'}$ be the moment of inertia tensors of a rigid body $B$ relative to two origins $o$ and $o'$, respectively. According to (8.12) we have:

$$\mathcal{J}_o = \int_B (||r||^2 I - r \otimes r) dm, \quad \mathcal{J}_{o'} = \int_B (||r'||^2 I - r' \otimes r') dm,$$

where $r$ and $r'$ are the position vectors of a generic point $p \in B$ relative to $o$ and $o'$, as seen in Figure 8.4.

**Theorem 8.4.** Let $c \in B$ the the body’s center of mass, and let us write $r'_c$ for the position vector of $c$ relative to $o'$. Furthermore, let $\tau = o' - o$. Then we have:

$$\mathcal{J}_o = \mathcal{J}_{o'} + m \left[ \left( 2r'_c \cdot \tau + ||\tau||^2 \right) I - r'_c \otimes \tau - \tau \otimes r'_c - \tau \otimes \tau \right].$$  

(8.14)

where $m$ is the body’s mass.

**Proof.** Referring to Figure 8.4 we have $\tau = r' + \tau$, therefore

$$\mathcal{J}_o = \int_B (||r||^2 I - r \otimes r) dm
= \int_B (||r' + \tau||^2 I - (r' + \tau) \otimes (r' + \tau)) dm
= \int_B \left[ (||r'||^2 + 2r' \cdot \tau + ||\tau||^2) I - r' \otimes r' - \tau \otimes r' - \tau \otimes \tau \right] dm
= \int_B (||r'||^2 I - r' \otimes r') dm + \int_B \left[ (2r' \cdot \tau + ||\tau||^2) I - r' \otimes \tau - \tau \otimes r' - \tau \otimes \tau \right] dm.$$
The first integral on the right-hand side equals $J_{o'}$. The second integral may be simplified by noting that
\[ m = \int_B d m, \quad r'_{c} = \frac{1}{m} \int_B r' \, d m, \]
and consequently $\int_B r' \, d m = mr'_{c}$.

**Corollary 8.5.** Let $c \in B$ be the body's center of mass as before, and let $J_{c}$ be the moment of inertia tensor relative to $c$. Then the moment of inertia tensor $J_{o}$ of $B$ relative to any origin $o$ is given by
\[ J_{o} = J_{c} + m\left[ ||r_{c}||^2 I - r_{c} \otimes r_{c} \right]. \quad (8.15) \]

**Proof.** Apply (8.14) with $o'$ set to $c$. Then $r'_{c} = 0$, and the formula reduces to
\[ J_{o} = J_{c} + m\left[ ||\tau||^2 I - \tau \otimes \tau \right]. \]
Then the observation that $\tau = o' - o = c - o = r_{c}$ completes the proof.

**Remark 8.6.** The moment of inertia tensor of a rigid system of $N$ point masses is given in (8.10). In particular, the moment of inertia tensor of a single point of mass $m$ at a position $r$ relative to an origin $o$ is
\[ m\left[ ||r||^2 I - r \otimes r \right]. \]

Therefore the translation formula (8.15) may be interpreted as saying that the moment of inertia of a body relative to any point $o$ equals the moment of inertia relative to its center of mass, plus the moment of inertia relative to $o$ of a fictitious point of mass $m$ situated at the center of mass.

**Remark 8.7.** The moment of inertia tensor $J_{c}$ of a rigid body relative to its center of mass is an intrinsic property of the body, just like its total mass or its center of mass are.
The total mass is a scalar, the center of mass is expressed through its position vector, and the moment of inertia $\mathcal{J}_c$ is a tensor.

## 8.4 The principal moments of inertia

In the exercises you will verify that a moment of inertia tensor is a symmetric tensor, that is:

$$m \cdot \mathcal{J} n = n \cdot \mathcal{J} m, \quad \text{for all } m, n \in \mathcal{V}.$$  

Then according to the Spectral Decomposition Theorem (page 63) $\mathcal{J}$ admits a representation of the form (8.9). In that context, the eigenvalues $\lambda_1, \lambda_2, \lambda_3$ of $\mathcal{J}$ are called the object's principal moments of inertia, and the eigenvectors $\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3$ are called the principal axes of the moment of inertia.

Tables of the principal moments of inertia of many common geometric solids are available in books and websites. Wikipedia has a page for it at:


**Example 8.6.** Consider a solid circular cylinder of radius $r$, length $\ell$, and total mass $m$. Assume that the mass is distributed uniformly, that is, the density is constant. The principal axes of the moment of inertia tensor $\mathcal{J}_c$ relative to the center of mass are: $\mathbf{g}_1$ along the cylinder’s axis; $\mathbf{g}_2$ and $\mathbf{g}_3$ arbitrary unit vectors so that $\{\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3\}$ forms an orthonormal set. The corresponding principal moments of inertia are $\lambda_1 = \lambda_2 = \frac{1}{12} m (3r^2 + \ell^2)$, $\lambda_3 = \frac{1}{2} m r^2$. Therefore

$$\mathcal{J}_c = \frac{1}{12} m (3r^2 + \ell^2) \mathbf{g}_1 \otimes \mathbf{g}_1 + \frac{1}{12} m (3r^2 + \ell^2) \mathbf{g}_2 \otimes \mathbf{g}_2 + \frac{1}{2} m r^2 \mathbf{g}_3 \otimes \mathbf{g}_3. \quad (8.16)$$

The limiting case of $r = 0$ is interesting in its own right. Such an object is called a slender rod of length $\ell$ and mass $m$. In that case we have:

$$\mathcal{J}_c = \frac{1}{12} m \ell^2 \mathbf{g}_1 \otimes \mathbf{g}_1 + \frac{1}{12} m \ell^2 \mathbf{g}_2 \otimes \mathbf{g}_2. \quad (8.17)$$

The principal moments of inertia $\mathcal{J}$ of a slender rod relative to one of its endpoints may be calculated from the translation formula (8.15).

---

**Exercises**

8.1. Show that an orthonormal set of vectors is linearly independent.

8.2. Complete the proof of Lemma 8.2 by showing that the set $\{e_i \otimes e_j\}_{i,j=1}^3$ is linearly independent. *Hint:* It suffices to show that $\sum_{i=1}^3 \sum_{j=1}^3 c_{ij} (e_i \otimes e_j) = 0$ implies that every $c_{ij}$ is zero. Begin by applying each side of that equality to $e_k$.

8.3. Show that if the tensor $L$ is symmetric, then the matrix $\ell_{ij}$ of its components (see Remark 8.3) is symmetric, that is $\ell_{ij} = \ell_{ji}$ for $i, j = 1, 2, 3$.

8.4. Use (8.17) in conjunction with the translation formula (8.15) to calculate the moment of inertia tensor of a slender rod relative to one of its endpoints.

8.5. Look up the principal moments of inertia of a circular hoop in Wikipedia relative to the hoop’s center. Then apply the translation formula (8.15) to calculate the moment of inertia tensor of the hoop relative to a point on its rim.
Chapter 9

Constraint reactions

In Chapter 6 we considered a mechanical system consisting of $N$ point masses subject to $M$ holonomic constraints (6.6), and derived the equation of motion (6.17), where the generalized forces $\{Q_j\}_{j=1}^n$ are related to the externally applied forces $\{f_i\}_{i=1}^N$ through (6.13). Here $n = 3N - M$.

In the derivation of the equations we assumed that the generalized coordinates were independent of each other. This was used in the derivation of equation (6.15) where we assumed that the virtual displacements $\delta q$ were arbitrary. Among other this, this resulted in the elimination of the internal reaction forces $f_i'$ from the equations of motion.

Suppose, however, that we are interested in finding out the reaction forces. It is the goal of this chapter to explain how. The key idea is to forgo the assumption of independence of the generalized coordinates.

Specifically, we assume that the $n$ generalized coordinates $q_1, q_2, \ldots, q_n$ are greater than the minimum necessary to specify the system’s configuration. This implies that one or more relationships exists among the $q_j$’s. Suppose that they are $m$ such relationships:

$$a_{11} dq_1 + a_{12} dq_2 + \cdots + a_{1n} dq_n + a_{1t} dt = 0,$$

$$a_{21} dq_1 + a_{22} dq_2 + \cdots + a_{2n} dq_n + a_{2t} dt = 0,$$

$$\cdots$$

$$a_{m1} dq_1 + a_{m2} dq_2 + \cdots + a_{mn} dq_n + a_{mt} dt = 0,$$

where each $a_{ij}$ and $a_{ij}$ can be a given function of $q$ and $t$. For convenience, we write these set of constraints in the compact form

$$\sum_{k=1}^n a_{jk}(q,t) dq_k + a_{lt}(q,t) dt = 0, \quad l = 1, 2, \ldots, m, \quad (9.1)$$

**Remark 9.1.** By dividing (9.2) through by $dt$, we see that

$$\sum_{k=1}^n a_{jk}(q,t) \dot{q}_k + a_{lt}(q,t) = 0, \quad l = 1, 2, \ldots, m, \quad (9.3)$$

that is, the equations impose restrictions on the system’s velocities.

Repeating (TODO) the calculations of Chapter 6, we arrive at the following equations
of motion:
\[ \frac{d}{dt} \left[ \frac{\partial L(q, \dot{q}, t)}{\partial \dot{q}_j} \right] - \frac{\partial L(q, \dot{q}, t)}{\partial q_j} = \sum_{k=1}^{m} a_{kj} \lambda_k, \quad j = 1, 2, \ldots, n. \tag{9.4} \]

These, together with (9.3), form a system of \( n + m \) equations in the \( m + n \) unknowns \( q_1, q_2, \ldots, q_n, \lambda_1, \lambda_2, \ldots, \lambda_m \). The summation that appear on the right-hand side gives the generalized reaction forces
\[ Q'_j = \sum_{k=1}^{m} a_{kj} \lambda_k. \tag{9.5} \]

Once the coefficients \( \lambda_k \) have been computed, we may use (9.5) in conjunction with
\[ Q'_j = \sum_{i=1}^{3N} f'_i \frac{\partial x_i}{\partial q_j} \tag{9.6} \]
to determine the reaction force components \( f'_i \).

**Example 9.1.** Let us revisit the simple pendulum of Figure 1.1 in page 2, and calculate the force within its connecting rod.

In Section 1.3 we derived the equation of motion (1.5) in terms of the angle \( \varphi \) which served as the generalized coordinate. The position vector of the pendulum’s bob relative to the suspension point was expressed as \( r = \ell e_r \), in that context, where \( \ell \) is the length of the pendulum’s rod.

Here we change the setting of the problem as follows. We express the configuration of the pendulum in terms of not one, but two generalized coordinates: \( \varphi \), which is the pendulum’s angle as before; and \( \rho \) which is the rod’s length and which is viewed as a variable. We impose the constraint \( \rho = \ell \) retroactively to recover the physical model.

Referring to Figure 1.1 we have \( r = \rho e_r \), therefore the bob’s velocity is given by
\[ v = \dot{r} = \rho \dot{e}_r + \rho \dot{e}_\varphi = \rho \dot{e}_r - \rho \dot{\varphi} e_\varphi, \]
where we have made use of (1.1). The kinetic and potential energies are
\[ T = \frac{1}{2} m (\dot{\rho}^2 + \rho^2 \dot{\varphi}^2), \quad V = -mg \rho \cos \varphi, \]
which leads to the lagrangian
\[ L = T - V = \frac{1}{2} m (\dot{\rho}^2 + \rho^2 \dot{\varphi}^2) + mg \rho \cos \varphi. \]

We calculate
\[ \frac{\partial L}{\partial \dot{\rho}} = m \dot{\rho}, \quad \frac{\partial L}{\partial \rho} = m \rho \dot{\varphi}^2 + mg \cos \varphi \quad \frac{\partial L}{\partial \dot{\varphi}} = m \rho^2 \dot{\varphi}, \quad \frac{\partial L}{\partial \varphi} = -mg \rho \sin \varphi. \]

The constraint of inextensibility \( \rho = \ell \) implies that \( d\rho = 0 \), which we write as (1)d\rho + (0)d\varphi + (0)d\ell = 0 \) to conform to the general template (9.1). It follows that \( a_{11} = 1, a_{12} = 0, a_{1\ell} = 0 \), therefore the equations (9.4) take the form
\[ \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{\rho}} \right) - \frac{\partial L}{\partial \rho} = a_{11} \lambda_1, \]
\[ \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{\varphi}} \right) - \frac{\partial L}{\partial \varphi} = a_{12} \lambda_1, \]
that is

\[
(m \dot{\rho})' - (m \rho \dot{\phi}^2 + mg \cos \phi) = \lambda_1,
\]

\[
(m \rho^2 \ddot{\phi})' + mg \rho \sin \phi = 0,
\]

or in expanded form:

\[
m \ddot{\rho} - (m \rho \dot{\phi}^2 + mg \cos \phi) = \lambda_1,
\]

\[
m \rho^2 \ddot{\phi} + 2m \rho \dot{\phi} \dot{\phi}' + mg \rho \sin \phi = 0.
\]

Applying the constraint \( \rho = \ell \) reduces these to

\[
-ml \dot{\phi}^2 - mg \cos \phi = \lambda_1,
\]

\[
ml^2 \ddot{\phi} + m g \ell \sin \phi = 0.
\]

The second equation is the usual equation of motion of a simple pendulum. In principle, we may plug its solution, \( \phi(t) \), into the first equation to find \( \lambda_1 \), but we don’t do it that way. Instead, we compute the generalized reaction forces \( Q'_r \) and \( Q'_\phi \) from (9.5):

\[
Q'_r = (1) \lambda_1 = -ml \dot{\phi}^2 - mg \cos \phi, \quad Q'_\phi = (0) \lambda_1 = 0.
\]

Then, we apply (9.6) to translate these into the physical components of the forces. Toward that end, let us observe that \( r = \rho e_r \), therefore

\[
\frac{\partial r}{\partial \rho} = e_r, \quad \frac{\partial r}{\partial \phi} = \rho \frac{\partial e_r}{\partial \phi} = \rho e_\phi,
\]

therefore (9.6) reads

\[
Q'_r = f' \cdot \frac{\partial r}{\partial \rho} = f' \cdot e_r, \quad Q'_\phi = f' \cdot \frac{\partial r}{\partial \phi} = f' \cdot \rho e_\phi,
\]

whence

\[
f' \cdot e_r = -(ml \dot{\phi}^2 + mg \cos \phi) \quad f' \cdot e_\phi = 0.
\]

Since \( f' = (f' \cdot e_r) e_r + (f' \cdot e_\phi) e_\phi \), we conclude that

\[
f' = -(ml \dot{\phi}^2 + mg \cos \phi) e_r.
\]

This tells us that the constraint reaction force \( f' \) lies in the direction of the pendulum’s rod, and its magnitude equals the sum of the centrifugal force \( ml \dot{\phi}^2 \) and the component \( mg \cos \phi \) of the bob’s weight along the rod.

**Example 9.2.** Consider a the rolling hoop of Exercise (2) of Chapter 6. Find the force at contact point between the hoop and the plane.

Referring to Figure 6.3, we use the hoop’s rotation angle \( \phi \) and the horizontal translation \( x \) of its center as a generalized coordinates. The no-slip condition imposes the constraint \( x = a \phi \). We will use \( \phi \) and \( x \) as overdetermined generalized coordinates subject to \( x = a \phi \), that is \( dx - a \, d \phi = 0 \), therefore according to the template (9.1), \( a_{11} = 1, a_{12} = -a \).
Chapter 9. Constraint reactions

The position vector of the mass $m$ is $r = (x + a \cos \varphi, a + a \cos \varphi)$, therefore the velocity is $v = \dot{r} = (\dot{x} + a \dot{\varphi} \cos \varphi, -a \dot{\varphi} \sin \varphi)$. It follows that $||v||^2 = \dot{x}^2 + 2a\dot{x} \dot{\varphi} \cos \varphi + a^2 \dot{\varphi}^2$, therefore the kinetic and potential energies are

$$T = \frac{1}{2} m (\dot{x}^2 + 2a\dot{x} \dot{\varphi} \cos \varphi + a^2 \dot{\varphi}^2), \quad V = mga(1 + \cos \varphi).$$

Therefore

$$L = \frac{1}{2} m (\dot{x}^2 + 2a\dot{x} \dot{\varphi} \cos \varphi + a^2 \dot{\varphi}^2) - mga(1 + \cos \varphi).$$

and

$$\frac{\partial L}{\partial \dot{x}} = m(\dot{x} + a \dot{\varphi} \cos \varphi), \quad \frac{\partial L}{\partial x} = 0,$$

$$\frac{\partial L}{\partial \dot{\varphi}} = m(a \dot{x} \cos \varphi + a^2 \dot{\varphi}), \quad \frac{\partial L}{\partial \varphi} = m(-a \dot{x} \dot{\varphi} \sin \varphi + ga \sin \varphi).$$

Then (9.4) takes the form

$$m(\dot{x} + a \dot{\varphi} \cos \varphi) = a_{11} \lambda_1 = \lambda_1,$$

$$m(a \dot{x} \cos \varphi + a^2 \dot{\varphi}) - m(-a \dot{x} \dot{\varphi} \sin \varphi + ga \sin \varphi) = a_{12} \lambda_1 = -a \lambda_1.$$

Now substitute $x = a \varphi$ and simplify:

$$-ma(1 + \cos \varphi) \dot{\varphi} + ma \dot{\varphi}^2 \sin \varphi = \lambda_1,$$

$$ma^2(1 + \cos \varphi) \dot{\varphi} + mga \sin \varphi = -a \lambda_1. \quad (9.7)$$

To eliminate $\lambda_1$, divide the second equation through by $a$ and add the result to the first equation. We get

$$2ma(1 + \cos \varphi) \dot{\varphi} - ma\left(\frac{a}{a} + \varphi^2\right) \sin \varphi = 0.$$

This is the hoop’s equation of motion. To compute the constraint force, solve this for $\ddot{\varphi}$:

$$\ddot{\varphi} = a\left(\frac{\varphi}{a} + \varphi^2\right) \sin \varphi \cdot 2a(1 + \cos \varphi)$$

and substitute the result in the first of (9.7). We get:

$$\lambda_1 = \frac{1}{2} ma(\dot{\varphi}^2 - \frac{a}{a}) \sin \varphi. \quad (9.8)$$

Now we compute the generalized forces $Q'_x$ and $Q'_\varphi$ from (9.5):

$$Q'_x = a_{11} \lambda_1 = \lambda_1, \quad Q'_\varphi = a_{12} \lambda_1 = -a \lambda_1.$$

Then, we apply (9.6) to translate these into the physical components of the forces. Toward that end, let us observe that

$$\frac{\partial r}{\partial x} = (1, 0), \quad \frac{\partial r}{\partial \varphi} = (a \cos \varphi, -a \sin \varphi),$$
Figure 9.1: whatever

therefore, letting \( f = (f_x, f_y) \), we get

\[
Q_x' = f \cdot \frac{\partial r}{\partial x} = f_x \\
Q_\phi' = f \cdot \frac{\partial r}{\partial \phi} = af_x \cos \phi - af_y \sin \phi.
\]

It follows that

\[
f_x = \lambda_1, \quad af_x \cos \phi - af_y \sin \phi = -a \lambda_1.
\]

We solve this as a system of two equations in the two unknowns \( f_x \) and \( f_y \):

\[
f_x = \lambda_1, \quad f_y = \frac{1 + \cos \phi}{\sin \phi} \lambda_1.
\]

Upon substitution for \( \lambda_1 \) from (9.8) we conclude that

\[
f_x = \frac{1}{2} ma (\dot{\phi}^2 - \frac{g}{a}) \sin \phi,
\]

\[
f_y = \frac{1}{2} ma (\dot{\phi}^2 - \frac{g}{a}) (1 + \cos \phi),
\]

whence

\[
f = \frac{1}{2} ma (\dot{\phi}^2 - \frac{g}{a})(\sin \phi, 1 + \cos \phi).
\]

This result has a significant mechanical/geometric interpretation. Refer to Figure 9.1. The vector \( a (\sin \phi, 1 + \cos \phi) \) extends from the contact point \( p \) to the mass \( m \), therefore the constraint force \( f' \) is parallel to that vector.

### Exercises

9.1. The hoop of Example 9.2 rolls, without slipping, down an incline which makes an angle \( \alpha \) with respect to the horizontal. See Figure 9.2. Use the angle \( \phi \) as the generalized coordinate. Derive the equation of motion. \textit{Hint}: (a) Express the vectors \( i' \) and \( j' \) in terms of \( i, j \), and the angle \( \alpha \); (b) Express the position vector \( r \) of the mass \( m \) in terms of the basis vectors \( i' \) and \( j' \); (c) Apply the results of (a) and (b) to express \( r' \) in terms of \( i \) and \( j \).

9.2. In the the previous exercise, find the contact force between the hoop and incline.
Figure 9.2: Hoop rolling on an incline. The hoop's initial position is shown in gray (Exercise 1).
Chapter 10

The Gibbs-Appell formulation of dynamics

The goal of this chapter is to derive the Gibbs-Appell equations of motion and then show a few applications. For now, I have two versions of the proof here, one based on the presentation in Lurie \cite{9} and the other based on the presentation in Gantmacher \cite{10}. The part based on Lurie is not quite finished; the one based on Gantmacher is pretty complete.

Due to the reliance on two different presentations, the notation in the illustrations/examples is very inconsistent. At one point I will rewrite this chapter by merging the two presentations into one, and introduce a consistent notation.

Some of the material from the previous chapters is repeated for the sake of making this chapter somewhat self-contained.

10.1 Gibbs-Appell according to Lurie \cite{9}

The goal of this (lengthy) section is to derive the Gibbs-Appell equations of motion (10.27).

10.1.1 Acceleration in generalized coordinates

Consider the dynamics of \( N \) point masses, whose position vectors relative to an origin \( o \) are \( \mathbf{r}_i, \ i = 1, \ldots, N \).

Suppose that the system’s configuration may be specified through \( n \) independent generalized coordinates \( \mathbf{q} = (q_1, \ldots, q_n) \). Thus, \( \mathbf{r}_i = \mathbf{r}_i(\mathbf{q}, t) \), and the velocities are given by

\[
\mathbf{v}_i = \dot{\mathbf{r}}_i = \sum_{s=1}^{n} \frac{\partial \mathbf{r}_i}{\partial q_s} \dot{q}_s + \frac{\partial \mathbf{r}_i}{\partial t}, \quad i = 1, 2, \ldots, N.
\]  

(10.1)

We differentiate the velocities to calculate the accelerations \( \mathbf{w}_i \):

\[
\mathbf{w}_i = \ddot{\mathbf{r}}_i = \sum_{s=1}^{n} \frac{\partial \mathbf{v}_i}{\partial q_s} \dot{q}_s + \sum_{s=1}^{n} \sum_{k=1}^{n} \frac{\partial^2 \mathbf{r}_i}{\partial q_s \partial q_k} \dot{q}_s \dot{q}_k + 2 \sum_{s=1}^{n} \frac{\partial^2 \mathbf{r}_i}{\partial q_s \partial t} \dot{q}_s + \frac{\partial^2 \mathbf{r}_i}{\partial t^2}.
\]

(10.2)

10.1.2 Ideal constraints and the fundamental equation of dynamics

Continuing the previous subsection’s analysis of the motion of \( N \) particles, let \( \mathbf{F}_i + \mathbf{R}_i \) be the totality of forces acting on the particle \( i \), where \( \mathbf{F}_i \) is the resultant of the (known) externally applied forces and \( \mathbf{R}_i \) is the resultant of the (unknown) internal forces/reactions.
Newton’s law of motion states that
\[ m_i \ddot{w}_i = F_i + R_i, \quad i = 1, \ldots, N. \quad (10.3) \]

We assume that the system’s constraints are ideal, that is, they do not perform work in any virtual displacement. (See page 271 of Lurie’s book for a discussion.) Thus, for all virtual displacements \( \delta r_i \) we have
\[ \sum_{i=1}^{N} R_i \cdot \delta r_i = 0. \quad (10.4) \]

As a consequence, multiplying (10.3) by \( \delta r_i \) and summing over \( i \) eliminates the the reaction forces \( R_i \):
\[ \sum_{i=1}^{N} m_i \dot{w}_i \cdot \delta r_i = \sum_{i=1}^{N} F_i \cdot \delta r_i. \quad (10.5) \]

Lagrange called this the fundamental equation of dynamics.

### 10.1.3 Virtual work and generalized force

The virtual displacement \( \delta r_i \) of particle \( i \) is related to the virtual displacement \( \delta q \) of the generalized coordinates through
\[ \delta r_i = \sum_{s=1}^{n} \frac{\partial r_i}{\partial q_s} \delta q_s. \]

The expression on the right-hand side of (10.5) is the virtual work performed by all external forces \( F_i \) under the virtual displacements \( \delta r_i \). It may be expressed as
\[ \sum_{i=1}^{N} F_i \cdot \delta r_i = \sum_{i=1}^{N} F_i \left( \sum_{s=1}^{n} \frac{\partial r_i}{\partial q_s} \delta q_s \right) = \sum_{i=1}^{n} \left( \sum_{i=1}^{N} F_i \frac{\partial r_i}{\partial q_s} \right) \delta q_s = \sum_{i=1}^{n} Q_i \delta q_s, \]
where we have let
\[ Q_i = \sum_{i=1}^{N} F_i \frac{\partial r_i}{\partial q_s}. \quad (10.6) \]

\( Q_i \) is called the generalized force corresponding to the generalized coordinate \( q_i \).

TODO: It can be shown that
\[ \frac{\partial r_i}{\partial q_i} = \frac{\partial \dot{r}_i}{\partial q_i} = \frac{\partial \ddot{r}_i}{\partial q_i}. \]

In formulating the Gibbs-Appell equations, it works better if we replace (10.6) with
\[ Q_i = \sum_{i=1}^{N} F_i \frac{\partial \dot{r}_i}{\partial q_i}. \quad (10.7) \]
10.1.4 • Constraints

Suppose that the \( n \) generalized coordinates are related through \( l \) (generally nonholonomic) constraints

\[
\sum_{i=1}^{n} a_{ki}(q, t) \dot{q}_i + a_k(q, t) = 0, \quad k = 1, \ldots, l, \tag{10.8}
\]

or equivalently,

\[
\sum_{i=1}^{n} a_{ki}(q, t) \frac{dq_i}{dt} + a_k(q, t) dt = 0, \quad k = 1, \ldots, l. \tag{10.9}
\]

Consequently, virtual displacements \( \delta q \) satisfy

\[
\sum_{i=1}^{n} a_{ki}(q, t) \delta q_i = 0, \quad k = 1, \ldots, l. \tag{10.10}
\]

We assume that the \( l \times n \) coefficient matrix \( a_{ki} \) is full-rank, therefore (10.8) may be solved for \( l \) of the generalized velocities in terms of the rest. Thus:

\[
\dot{q}_r = \sum_{i=1}^{n-l} b_{r,i+1}(q, t) \dot{q}_{i+1} + b_r(q, t), \quad r = 1, \ldots, l. \tag{10.11}
\]

Similarly, (10.10) may be solved for \( l \) of the virtual displacements in terms of the rest:

\[
\delta q_r = \sum_{i=1}^{n-l} b_{r,i+1}(q, t) \delta q_{i+1}, \quad r = 1, \ldots, l. \tag{10.12}
\]

Differentiating (10.11) we obtain an expression for the accelerations:

\[
\ddot{q}_r = \sum_{i=1}^{n-l} b_{r,i+1}(q, t) \ddot{q}_{i+1} + \sum_{i=1}^{n-l} b_{r,i+1}(q, t) \dot{q}_{i+1} + \sum_{i=1}^{n-l} b_r(q, t), \quad r = 1, \ldots, l. \tag{10.13}
\]

For reasons which will become clear shortly, the terms which involve no generalized accelerations in our calculations are irrelevant, therefore, to simplify the notation, we write the above as

\[
\ddot{q}_r = \sum_{i=1}^{n-l} b_{r,i+1}(q, t) \ddot{q}_{i+1} + \cdots. \tag{10.13}
\]

From here on, the ellipsis “\( \cdots \)" in an equation indicates additive terms that involve no generalized accelerations.

Now we apply (10.13) to eliminate the accelerations \( \ddot{q}_r, \ r = 1, \ldots, l \) from (10.2). We have:

\[
\mathbf{w}_i = \sum_{k=1}^{n} \frac{\partial \mathbf{r}_i}{\partial \dot{q}_k} \ddot{q}_k + \cdots
\]

\[
= \sum_{r=1}^{l} \frac{\partial \mathbf{r}_i}{\partial q_r} \ddot{q}_r + \sum_{i=1}^{n-l} \frac{\partial \mathbf{r}_i}{\partial \dot{q}_{i+1}} \ddot{q}_{i+1} + \cdots
\]

\[
= \sum_{r=1}^{l} \frac{\partial \mathbf{r}_i}{\partial q_r} \left( \sum_{i=1}^{n-l} b_{r,i+1}(q, t) \ddot{q}_{i+1} + \cdots \right) + \sum_{i=1}^{n-l} \frac{\partial \mathbf{r}_i}{\partial \dot{q}_{i+1}} \ddot{q}_{i+1} + \cdots
\]

\[
= \sum_{i=1}^{n-l} \left( \frac{\partial \mathbf{r}_i}{\partial \ddot{q}_{i+1}} + \sum_{r=1}^{l} \frac{\partial \mathbf{r}_i}{\partial q_r} b_{r,i+1}(q, t) \right) \ddot{q}_{i+1} + \cdots.
\]
Upon introducing the notation
\[ c_{i,l+s} = \frac{\partial r_i}{\partial q_{l+s}} + \sum_{r=1}^{l} \frac{\partial r_i}{\partial q_r} b_{r,l+s}(q, t), \quad i = 1, \ldots, N, \quad s = 1, \ldots, n-l, \] (10.14)
the acceleration takes the form
\[ w_i = \sum_{s=1}^{n-l} c_{i,l+s} \ddot{q}_{l+s} + \cdots, \quad i = 1, \ldots, N. \]

Since the terms hidden under the ellipsis do not involve the generalized accelerations, we conclude that
\[ \frac{\partial w_i}{\partial \ddot{q}_{l+s}} = c_{i,l+s}. \] (10.15)

### 10.1.5 Virtual displacements

According to (10.12), a virtual displacement \( \delta r_i \) is given by
\[ \delta r_i = \sum_{k=1}^{n} \frac{\partial r_i}{\partial q_k} \delta q_k = \sum_{r=1}^{l} \frac{\partial r_i}{\partial q_r} \delta q_r + \sum_{s=1}^{n-l} \frac{\partial r_i}{\partial q_{l+s}} \delta q_{l+s} \]
\[ = \sum_{r=1}^{l} \frac{\partial r_i}{\partial q_r} \left( \sum_{s=1}^{n-l} b_{r,l+s}(q, t) \delta q_{l+s} \right) + \sum_{s=1}^{n-l} \frac{\partial r_i}{\partial q_{l+s}} \delta q_{l+s} \]
\[ = \sum_{s=1}^{n-l} \left( \frac{\partial r_i}{\partial q_{l+s}} + \sum_{r=1}^{l} \frac{\partial r_i}{\partial q_r} b_{r,l+s}(q, t) \right) \delta q_{l+s} \]
\[ = \sum_{s=1}^{n-l} c_{i,l+s} \delta q_{l+s}, \]

where \( c_{i,l+s} \) is as in (10.14). Then, in view of (10.15) we conclude that
\[ \delta r_i = \sum_{s=1}^{n-l} \frac{\partial w_i}{\partial q_{l+s}} \delta q_{l+s}, \quad i = 1, \ldots, N. \] (10.17)

### 10.1.6 Back to the fundamental equation: Part 1

We use (10.17) to reformulate the fundamental equation of dynamics (10.5):
\[ \sum_{i=1}^{N} m_i w_i \cdot \delta r_i = \sum_{i=1}^{N} m_i w_i \cdot \sum_{s=1}^{n-l} \frac{\partial w_i}{\partial q_{l+s}} \delta q_{l+s}, \]
\[ = \sum_{s=1}^{n-l} \left( \sum_{i=1}^{N} m_i w_i \cdot \frac{\partial w_i}{\partial q_{l+s}} \right) \delta q_{l+s}. \]
This motivates the introduction of a quantity \( \mathcal{G} \) known as the *Gibbs function* or the *energy of acceleration*:
\[ \mathcal{G} = \frac{1}{2} \sum_{i=1}^{N} m_i w_i \cdot w_i. \]
We observe that
\[ \frac{\partial \mathbf{\Phi}}{\partial \dot{q}_{l+s}} = \sum_{i=1}^{N} m_i w_i \cdot \frac{\partial w_i}{\partial \dot{q}_{l+s}}, \]
whence
\[ \sum_{i=1}^{N} m_i w_i \cdot \delta r_i = \sum_{i=1}^{N} \frac{\partial \mathbf{\Phi}}{\partial \dot{q}_{l+s}} \delta q_{l+s}. \quad (10.18) \]

Let us note that we consider \( \mathbf{\Phi} \) as a function of the form \( \mathbf{\Phi}(q, \dot{q}_{l+1}, \ldots, \dot{q}_n, t) \), therefore the nonholonomic constraints (10.8) are automatically accounted for.

### 10.1.7 Back to the fundamental equation: Part 2

In the previous section we reformulated the left-hand side of the fundamental equation of dynamics (10.5). In this section we reformulate its right-hand side. According to (10.16) we have:
\[ \sum_{i=1}^{n-l} F_i \cdot \delta r_i = \sum_{i=1}^{n-l} F_i \cdot \left( \sum_{r=1}^{l} \frac{\partial r_i}{\partial q_r} b_{r,l+s}(q, t) \right) \delta q_{l+s} = \sum_{i=1}^{n-l} \tilde{Q}_{l+s} \delta q_{l+s}, \quad (10.19) \]
where we have let
\[ \tilde{Q}_{l+s} = \sum_{i=1}^{N} F_i \cdot c_{i,l+s}, \quad s = 1, \ldots, n-l. \]

Upon substituting for \( c_{i,l+s} \) from its definition in (10.14), we see that
\[ \tilde{Q}_{l+s} = \sum_{i=1}^{N} F_i \left( \frac{\partial r_i}{\partial q_{l+s}} + \sum_{r=1}^{l} \frac{\partial r_i}{\partial q_r} b_{r,l+s}(q, t) \right) + \sum_{i=1}^{n-l} \frac{\partial r_i}{\partial q_{l+s}} b_{r,l+s}(q, t), \]
which, according to (10.6) reduces to
\[ \tilde{Q}_{l+s} = Q_{l+s} + \sum_{r=1}^{l} Q_r b_{r,l+s}(q, t), \quad s = 1, \ldots, n-l. \]

### 10.1.8 The Gibbs-Appell equations of motion

Substituting (10.18) and (10.19) in the fundamental equation of dynamics (10.5), we arrive at
\[ \sum_{i=1}^{n-l} \frac{\partial \mathbf{\Phi}}{\partial \dot{q}_{l+s}} \delta q_{l+s} = \sum_{i=1}^{n-l} \tilde{Q}_{l+s} \delta q_{l+s}, \]
that is,
\[ \sum_{i=1}^{n-l} \left( \frac{\partial \mathbf{\Phi}}{\partial \dot{q}_{l+s}} \delta q_{l+s} - \tilde{Q}_{l+s} \right) \delta q_{l+s}. \]

Since the variations \( \delta q_{l+s} \) are independent, we conclude that
\[ \frac{\partial \mathbf{\Phi}}{\partial \dot{q}_{l+s}} = \tilde{Q}_{l+s}, \quad s = 1, \ldots, n-l. \quad (10.20) \]
These are the Gibbs-Appell equation of motion.

The $n-l$ second order differential equations (10.20), along with the $l$ constraint equations (10.8) form a system of $n$ differential equation in the $n$ unknowns $q_1(t), \ldots, q_n(t)$.

### 10.1.9 Quasi-velocities

In a mechanical system with generalized coordinates $q = \langle q_1, \ldots, q_n \rangle$, expressions of the type

$$\omega_s = \sum_{k=1}^{n} a_{sk}(q, t) \dot{q}_k + a_s(q, t), \quad s = 1, \ldots, n$$  \hspace{1cm} (10.21)

are called quasi-velocities. At times it is simpler to model a system in terms of suitably defined quasi-velocities $\omega_1, \ldots, \omega_n$ rather than generalized coordinates $q_1, \ldots, q_n$. If the number of naturally occurring quasi-velocities is $n' < n$, then we extend their number to $n$ by defining the rest through $\omega_s = \dot{q}_s, s = n' + 1, \ldots, n$.

We assume that the quasi-velocities are defined so that the coefficient matrix $a_{sk}$ is nonsingular. It follows that (10.21) may be solved for $\dot{q}$’s in terms of $\omega$’s:

$$\dot{q}_r = \sum_{s=1}^{n} b_{rs}(q, t) \omega_s + b_r(q, t), \quad s = 1, \ldots, n.$$  

Additionally, with each equation in (10.21) we associate a differential as follows

$$d\pi_s = \sum_{k=1}^{n} a_{sk}(q, t) d\dot{q}_k + a_s(q, t) dt, \quad s = 1, \ldots, n.$$  \hspace{1cm} (10.22)

Comparing (10.21) and (10.22) we see that

$$d\pi_s = \omega_s dt.$$  \hspace{1cm} (10.23)

Remark: There is no reason to expect the expression on the right to be an exact differential, therefore although $d\pi_s$ makes sense as defined, it should not be assumed that it is the differential of a function $\pi_s$.

For any arbitrary function $\varphi(q, t)$ we have:

$$d\varphi = \sum_{r=1}^{n} \frac{\partial \varphi}{\partial q_r} d\dot{q}_r + \frac{\partial \varphi}{\partial t} dt$$

$$= \sum_{r=1}^{n} \frac{\partial \varphi}{\partial q_r} \dot{q}_r dt + \frac{\partial \varphi}{\partial t} dt$$

$$= \sum_{r=1}^{n} \frac{\partial \varphi}{\partial q_r} \left( \sum_{s=1}^{n} b_{rs}(q, t) \omega_s + b_r(q, t) \right) dt + \frac{\partial \varphi}{\partial t} dt$$

$$= \sum_{s=1}^{n} \left( \sum_{r=1}^{n} b_{rs}(q, t) \frac{\partial \varphi}{\partial q_r} \right) d\pi_s + \sum_{r=1}^{n} b_r(q, t) \frac{\partial \varphi}{\partial q_r} dt + \frac{\partial \varphi}{\partial t} dt.$$  

We introduce the purely symbolic notation

$$\frac{\partial \varphi}{\partial \pi_s} = \sum_{r=1}^{n} b_{rs}(q, t) \frac{\partial \varphi}{\partial q_r},$$  \hspace{1cm} (10.24)
whereupon the previous calculation leads to

$$d\varphi = \sum_{i=1}^{n} \frac{\partial \varphi}{\partial \pi_i} d\pi_i + \left( \sum_{i=1}^{n} b_i(q, t) \frac{\partial \varphi}{\partial q_i} + \frac{\partial \varphi}{\partial t} \right) dt.$$  

(10.25)

Applying (10.24) to \( r_i \) we get:

$$\frac{\partial r_i}{\partial \pi_s} = \sum_{r=1}^{n} b_{rs}(q, t) \frac{\partial r_i}{\partial q_r},$$

and therefore we obtain an expression for the velocity vector in terms of the quasi-coordinates:

$$v_i = \dot{r}_i = \text{TODO} \ldots$$

10.1.10 • Appell's equations of motion in terms of quasi-velocities

The main ingredients of Appell's equations of motion (10.20) are the generalized accelerations \( \dot{q}_s \). Those equations, however, are rarely used in that form. We will find that expressing the equations in terms of quasi-velocities, defined below, leads to major simplification.

Thus, given the \( l \) constraints (10.8), define the quasi-velocities

$$\omega_k = \sum_{s=1}^{n} a_{ks}(q, t) \dot{q}_s + a_k(q, t), \quad k = 1, \ldots, l,$$

(10.26)

therefore the constraint equations take the form

$$\omega_k = 0, \quad k = 1, \ldots, l.$$  

We solve (10.26) for the \( \dot{q} \) in terms of the quasi-velocities. Since the first \( l \) of the quasi-velocities are zero, each \( \dot{q}_s \) is a function of \( \omega_{i+s}, \quad s = 1, \ldots, n-l \). Then we express the Gibbs function \( \Phi \) as a function of

$$q_1, \ldots, q_n, \quad \omega_{1+s}, \ldots, \omega_n, \quad \omega_{1+s}, \ldots, \omega_n.$$  

The rest of the derivation is horrendous, so we skip forward...

Then, it turns out that the Gibbs-Appell equation of motion take the form

$$\frac{\partial \Phi}{\partial \omega_{i+s}} = Q_{i+s}, \quad s = 1, \ldots, l.$$  

(10.27)

Despite the complexity of the derivation, equations (10.27) are much easier to apply than the original (10.20). In particular, these result in first order differential equations in quasi-velocities, as apposed to the second order differential equations which we were getting before.

10.2 • Gibbs-Appell according to Gantmacher [10]

The goal of this (lengthy) section is to derive the Gibbs-Appell equations of motion (10.47)
10.2.1 Pseudocoordinates

Consider the motion of $N$ particles of masses $m_{\nu}$ and position vectors $r_{\nu}$ subject to $d$ algebraic and $g$ differential constraints. Utilizing the $d$ algebraic constraints, we introduce $m = 3N - d$ generalized independent coordinates $q = (q_1, \ldots, q_m)$, and express the system’s configures in them as

$$r_{\nu} = r_{\nu}(q, t), \quad \nu = 1, \ldots, N. \quad (10.28)$$

From this it follows that

$$\dot{r}_{\nu} = \sum_{i=1}^{m} \frac{\partial r_{\nu}}{\partial q_i} \dot{q}_i + \frac{\partial r_{\nu}}{\partial t}, \quad \nu = 1, \ldots, N, \quad (10.29)$$

and

$$\delta r_{\nu} = \sum_{i=1}^{m} \frac{\partial r_{\nu}}{\partial q_i} \delta q_i, \quad \nu = 1, \ldots, N. \quad (10.30)$$

The vectors $r_{\nu}$ and $\dot{r}_{\nu}$ should satisfy the $g$ differential constraints:

$$\sum_{\nu=1}^{m} l_{\beta \nu}(r_{\nu}, t) \cdot \dot{r}_{\nu} + D_{\beta}(r_{\nu}, t) = 0, \quad \beta = 1, \ldots, g. \quad (10.31)$$

Substituting for $r_{\nu}$ and $\dot{r}_{\nu}$ from (10.28) and (10.29), the constraint equations take the form

$$\sum_{i=1}^{m} A_{\beta i}(q, t) \dot{q}_i + A_{\beta}(q, t) = 0, \quad \beta = 1, \ldots, g. \quad (10.32)$$

Thus, the generalized coordinates $q_1, \ldots, q_m$ can take on arbitrary values, but the generalized velocities $\dot{q}_1, \ldots, \dot{q}_m$ are constrained by (10.32).

Assuming that the $g$ constraints in (10.32) are independent, that is, the matrix $A_{\beta i}$ has full rank, we may solve for $g$ of the velocities $\dot{q}_1, \ldots, \dot{q}_m$ in terms of the remaining $n = m - g = 3N - d - g$, the number $n$ being the system’s degrees of freedom. Thus, the velocities $\dot{q}_1, \ldots, \dot{q}_n$ may take on arbitrary values, while the remaining $g$ velocities are determined from (10.32).

In practice, instead of the $n$ generalized velocities noted above, it is often preferable to use $n$ linear combination of the generalized velocities, such as

$$\pi_s = \sum_{i=1}^{m} f_{is} \dot{q}_i, \quad i = 1, \ldots, m. \quad (10.33)$$

The quantities $\pi_s$ defined here are called pseudovelocities. The notation $\pi_s$ is entirely pro forma—there is no requirement that the right-hand side of (10.33) be and exact derivative, therefore although the notation $\pi_s$ is well-defined, there is no function $\pi_s$ which it is the derivative of. Nevertheless, we will have occasions to refer to the un-dotted symbol $\pi_s$, which has no predefine meaning, but we will define it in an appropriate and consistent way. The symbol $\pi_s$ is called a pseudocoordinate.

We impose an invertibility requirement on the definitions (10.33) as follows. Consider the system of $g + n = m$ equation obtained as the union of the equations (10.32) and (10.33), and view it as a system of $m$ linear equaitons in the $m$ unknowns $\dot{q}_1, \ldots, \dot{q}_m$. We require that system to be invertible. Thus, the $m$ generalized velocities may be expressed in terms of the $n$ pseudovelocities:

$$\dot{q}_i = \sum_{s=1}^{n} h_{si}(q, t) \pi_s + h_i(q, t), \quad i = 1, \ldots, m. \quad (10.34)$$
Thus, any set of \( m \) generalized velocities that satisfy the motion constraints (10.32) define a corresponding set of pseudovelocities \( \pi_s \), and conversely, any set of \( n \) arbitrary pseudovelocities define a set of generalized velocities that satisfy the motion constraints. The important point here is that there is no constraint on the pseudovelocities.

Let us note that due to the constrain (10.32) on generalized velocities, the generalized displacement are subject to the constraints

\[
\sum_{i=1}^{m} A_{\beta i} \delta q_i = 0, \quad \beta = 1, \ldots, g. \tag{10.35}
\]

Then in view of the equations (10.33) also introduce the notation

\[
\delta \pi_s = \sum_{i=1}^{m} A_{\beta i} f_{i\beta} \delta q_i, \quad s = 1, \ldots, n. \tag{10.36}
\]

From what it has been said, the union of the equations (10.35) and (10.36) is invertible and the inverse has the form

\[
\delta q_i = \sum_{s=1}^{n} h_{is} \delta \pi_s, \quad i = 1, \ldots, m. \tag{10.37}
\]

Thus, as argued before, the expressions \( \delta \pi_s \) may take on arbitrary values. The corresponding \( q_i \) obtained from (10.37) will automatically satisfy the constraints (10.35).

### 10.2.2 Work and generalized forces

Let us compute the work done by the external forces in a virtual displacement:

\[
\delta W = \sum_{v=1}^{N} F_v \cdot \delta r_v = \sum_{i=1}^{m} \sum_{v=1}^{N} F_v \cdot \left( \sum_{i=1}^{m} h_{vi} Q_i \frac{\partial r_v}{\partial q_i} \right) \delta q_i = \sum_{i=1}^{m} Q_i \delta q_i, \tag{10.38}
\]

where

\[
Q_i = \sum_{v=1}^{N} F_v \cdot \frac{\partial r_v}{\partial q_i}, \quad i = 1, \ldots, m. \tag{10.39}
\]

Thus, we have obtained an expression for the virtual work in terms of the virtual displacements \( \delta q_i \). Utilizing (10.37), we may express the virtual work in terms of the \( \delta \pi_s \):

\[
\delta W = \sum_{i=1}^{m} Q_i \sum_{s=1}^{n} h_{is} \delta \pi_s = \sum_{s=1}^{n} \left( \sum_{i=1}^{m} h_{is} Q_i \right) \delta \pi_s = \sum_{s=1}^{n} \Pi_s \delta \pi_s, \tag{10.40}
\]

where we have let

\[
\Pi_s = \sum_{i=1}^{m} h_{is} Q_i, \quad s = 1, \ldots, n. \tag{10.41}
\]

The \( \Pi_s \) defined above are called the generalized forces corresponding to the pseudocoordinate \( \pi_s, s = 1, \ldots, n \).
10.2.3 Newton’s equations in pseudocoordinates

Equations (10.29) express the particle velocities in terms of the generalized velocities. Substituting for the latter from (10.34) we obtain the particle velocities in terms of the pseudovelocities:

\[ \dot{r}_\nu = \sum_{s=1}^{n} e_{\nu s}(q,t) \pi_s + e_{\nu}(q,t), \quad \nu = 1, \ldots, N. \tag{10.42} \]

Then it follows that

\[ \delta r_\nu = \sum_{s=1}^{n} e_{\nu s}(q,t) \delta \pi_s, \quad \nu = 1, \ldots, N, \tag{10.43} \]

and

\[ \ddot{r}_\nu = \sum_{s=1}^{n} e_{\nu s}(q,t) \ddot{\pi}_s + \cdots, \quad \nu = 1, \ldots, N. \]

where the ellipsis indicate terms that are free of the pseudoaccelerations \( \ddot{\pi}_s \), \( s = 1, \ldots, n \).

For future reference, let us note that

\[ \frac{\partial \ddot{r}_\nu}{\partial \ddot{\pi}_s} = e_{\nu s}. \tag{10.44} \]

Consider the fundamental equations of dynamics:

\[ \delta W - \sum_{\nu=1}^{N} m_\nu \ddot{r}_\nu \cdot \delta r_\nu = 0. \]

Substituting from (10.40) and (10.43) this takes the form

\[ \sum_{s=1}^{n} \Pi_s \delta \pi_s - \sum_{s=1}^{n} m_\nu \ddot{r}_\nu \cdot \left( \sum_{s=1}^{n} e_{\nu s}(q,t) \delta \pi_s \right) = 0, \]

that is

\[ \sum_{s=1}^{n} \Pi_s \delta \pi_s - \sum_{s=1}^{n} \left( m_\nu \ddot{r}_\nu \cdot e_{\nu s}(q,t) \right) \delta \pi_s = 0, \]

and finally

\[ \sum_{s=1}^{n} \left[ \Pi_s - \sum_{s=1}^{n} \left( m_\nu \ddot{r}_\nu \cdot e_{\nu s}(q,t) \right) \right] \delta \pi_s = 0. \]

Since \( \delta \pi_s \) are unconstrained, we conclude that

\[ \sum_{s=1}^{N} m_\nu \ddot{r}_\nu \cdot e_{\nu s}(q,t) = \Pi_s, \quad s = 1, \ldots, n. \tag{10.45} \]

10.2.4 The energy of the acceleration

We introduce the “energy of acceleration” which is defined as

\[ U = U(q, \dot{q}, \ddot{q}) = \frac{1}{2} \sum_{\nu=1}^{N} m_\nu ||\dot{r}_\nu||^2. \tag{10.46} \]
Then, in view of (10.44), we see that
\[ \frac{\partial U}{\partial \ddot{\pi}_s} = \sum_{v=1}^{N} m_v \ddot{r}_v \cdot \frac{\partial \ddot{r}_v}{\partial \ddot{\pi}_s} = \sum_{v=1}^{N} m_v \ddot{r}_v \cdot e_{v_s}, \quad s = 1, \ldots, n, \]
which, in conjunction with (10.45) leads to
\[ \frac{\partial U}{\partial \ddot{\pi}_s} = \Pi_s, \quad s = 1, \ldots, n, \quad (10.47) \]
This are Appell's equations of motion.

Upon a close examination, equations (10.47) are a system of \( n \) first order differential equations in the \( n \) unknowns \( \dot{\pi}_s, s = 1, \ldots, n \). Note that there are no such things and \( \pi_s \), so the equations are indeed first order in \( \dot{\pi}_s \).

The equations are not complete, however, since they also involve the generalized co-ordinates \( q_i, i = 1, \ldots, n \). To complete the system, we append to it the \( g \) equations (10.32) and (10.33), as in:
\[ \frac{\partial U}{\partial \dot{\pi}_s} = \Pi_s, \quad s = 1, \ldots, n, \]
\[ \sum_{i=1}^{m} A_{\beta i}(q, t) \dot{q}_i + \dot{A}_{\beta}(q, t) = 0, \quad \beta = 1, \ldots, g, \]
\[ \dot{\pi}_s = \sum_{i=1}^{m} f_{s i} \dot{q}_i, \quad i = 1, \ldots, n. \]
This combined set consists of \( 2n + g \) equations. Let us note that \( n + g = m \), therefore we have a system of \( m + n \) first order differential equations in the \( m + n \) unknowns \( \{q_i\}_{i=1}^{m} \) and \( \{\ddot{\pi}_s\}_{s=1}^{n} \).

### 10.3 A modification noted by Desloge

The generalized forces \( \Pi_s \) in the Gibbs-Appell equations of motion (10.47) are computed from (10.41), where, in turn, the \( Q_i \) are computed from (10.39). Desloge [2, 11] notes that if we define
\[ R = \sum_{v=1}^{N} F_v \cdot \ddot{r}_v, \quad (10.48) \]
and express the result in terms of \( q, \dot{\pi}, and \ddot{\pi} \), then \( \Pi_s \) may be computed more easily through the formula
\[ \Pi_s = \frac{\partial R}{\partial \ddot{\pi}_s}, \quad s = 1, \ldots, n. \quad (10.49) \]
Consequently, (10.47) takes on the form
\[ \frac{\partial U}{\partial \ddot{\pi}_s} = \frac{\partial R}{\partial \ddot{\pi}_s}, \quad s = 1, \ldots, n, \]
which motivates the introduction of the Gibbs function
\[ \mathcal{G} = U - R, \quad (10.50) \]
whereby the equations (10.47) are expressed as

\[
\frac{\partial \mathcal{G}}{\partial \dot{\pi}_s} = 0, \quad s = 1, \ldots, n. \tag{10.51}
\]

I haven’t checked the details of the reasoning here, but in a couple of applications which I calculated, (10.47) and (10.51) produce identical results.

10.4 The simple pendulum via Gibbs-Appell

Here we will apply the Gibbs-Appell’s equations (10.20) to derive the familiar equation of motion of a simple pendulum. This is certainly an overkill; deriving the equations of motion of a holonomic system such as a simple pendulum is done much more easily through the Lagrangian approach. The power of the Gibbs-Appell approach manifests itself when applied to nonholonomic systems, as we will see in the next section.

But for now, let’s look at the simple pendulum as illustrated in Figure 1.1 on page 2. We take the angle \( \phi \) as the problem’s generalized coordinate, express the position vector \( r \) of the pendulum’s bob in terms of \( \phi \), then calculate the velocity and the acceleration:

\[
r = \ell (\sin \phi, \cos \phi),
\]

\[
v = \dot{r} = \ell (\dot{\phi} \cos \phi, -\dot{\phi} \sin \phi),
\]

\[
w = \dot{v} = \ell (\ddot{\phi} \cos \phi - \dot{\phi}^2 \sin \phi, -\ddot{\phi} \sin \phi - \dot{\phi}^2 \cos \phi),
\]

whence \( \|w\|^2 = \ell^2 (\dot{\phi}^2 + \dot{\phi}^4) \). Considering that the force vector is \( f = (0, mg) \), we are lead to the Gibbs function

\[
\mathcal{G} = \frac{1}{2} m \|w\|^2 - f \cdot w
\]

\[
= \frac{1}{2} \ell^2 (\dot{\phi}^2 + \dot{\phi}^4) + mg \ell (\dot{\phi} \sin \phi + \dot{\phi}^2 \cos \phi).
\]

Since the Gibbs-Appell equations of motion (10.20) reduce to \( \frac{\partial \mathcal{G}}{\partial \dot{\phi}} = 0 \) in our case, we conclude that

\[
\ell^2 \ddot{\phi} + mg \ell \sin \phi = 0,
\]

which is the familiar equation of motion of a simple pendulum.

10.5 The Čaplygin sleigh

The Čaplygin sleigh consists of two point masses \( m_1 \) and \( m_2 \) connected through a rigid massless rod of length \( \ell \). The two masses slide on a horizontal surface. The mass \( m_1 \) can slide freely on the surface with no resistance at all. Mass \( m_1 \) rides on a sharp blade, as in an ice hockey skate, which allows motion only in the direction of the rod. We wish to describe the dynamics of the sleigh. Figure 10.1 depicts the sleigh. The position vectors \( r_1 \) and \( r_2 \) of the masses \( m_1 \) and \( m_2 \) may be specified through the generalized coordinates \( x(t), y(t), \) and \( \phi(t), \) shown on the figure. We have:

\[
r_1 = (x, y), \quad r_2 = r_1 + (\ell \cos \phi, \ell \sin \phi).
\]

Since the mass \( m_1 \) can slide only along the direction of the rod, the velocity \( \dot{r}_1 = (\dot{x}, \dot{y}) \) is parallel to the vector \( (\cos \phi, \sin \phi) \), therefore \( \dot{x}/\cos \phi = \dot{y}/\sin \phi \), that is,

\[
\dot{x} \sin \phi - \dot{y} \cos \phi = 0. \tag{10.52}
\]
10.5. The Čaplygin sleigh

Figure 10.1: The Čaplygin sleigh’s configuration may be described in terms of the generalized coordinates \( x, y, \) and \( \phi \).

This is the concrete version of the general constraint (10.8).

To calculate the Gibbs function, with begin with

\[
\begin{align*}
v_1 &= r_1 = (x, y), \\
v_2 &= r_2 = (\dot{x} + \ell \dot{\phi} \sin \phi, \dot{y} + \ell \dot{\phi} \cos \phi).
\end{align*}
\]

and then

\[
\begin{align*}
w_1 &= \dot{v}_1 = (\ddot{x}, \ddot{y}), \\
w_2 &= \dot{v}_2 = (\ddot{x} - \ell \ddot{\phi}^2 \cos \phi - \ell \dot{\phi} \sin \phi, \ddot{y} - \ell \ddot{\phi}^2 \sin \phi + \ell \dot{\phi} \cos \phi).
\end{align*}
\]

Thus, we arrive at

\[
G = \frac{1}{2}(m_1 + m_2)(\ddot{x}^2 + \ddot{y}^2) + \frac{1}{2} m_2 \ell^2 \ddot{\phi}^2 - m_2 \ell (\ddot{\phi} \sin \phi + \ddot{\phi}^2 \cos \phi) \ddot{x} + m_2 \ell (\ddot{\phi} \cos \phi - \ddot{\phi}^2 \sin \phi) \ddot{y} + \frac{1}{2} m_2 \ell^2 \ddot{\phi}^4.
\]

Now, following the idea in (10.11) of eliminating redundant generalized velocities, we solve (10.52) for \( \dot{y} \), and eliminate it from the equations. Substituting \( \dot{y} = \dot{x} \tan \phi \) in the above expression for \( G \), we obtain

\[
G = \frac{m_1 + m_2}{2 \cos^2 \phi} \ddot{x}^2 + \frac{1}{\cos^3 \phi} [(m_1 + m_2) \dot{x} \sin \phi - m_2 \ell \dot{\phi} \cos^2 \phi] \ddot{x} \\
+ \frac{1}{2} m_2 \ell \ddot{\phi}^2 + \frac{1}{\cos \phi} m_2 \ell \dot{x} \dot{\phi} \ddot{\phi} + \cdots, \quad (10.53)
\]

where, as before, the ellipsis stands for terms that involve no accelerations.

There are no externally applied forces on the sleigh, therefore the equations of motion (10.20) reduce to

\[
\frac{\partial \mathcal{G}}{\partial \ddot{x}} = 0, \quad \frac{\partial \mathcal{G}}{\partial \ddot{\phi}} = 0,
\]

that is,

\[
\frac{m_1 + m_2}{\cos^2 \phi} \ddot{x} + \frac{1}{\cos^3 \phi} [(m_1 + m_2) \dot{x} \sin \phi - m_2 \ell \dot{\phi} \cos^2 \phi] \ddot{x} = 0, \quad (10.54a)
\]

\[
m_2 \ell \ddot{\phi} + \frac{1}{\cos \phi} m_2 \ell \dot{x} \dot{\phi} = 0. \quad (10.54b)
\]

We may solve this system of two second order differential equations for the two unknowns \( x(t) \) and \( \phi(t) \), subject to initial conditions \( x(0) = x_0, \dot{x}(0) = \dot{x}_0, \phi(0) = \phi_0 \).
and \( \dot{\phi}(0) = \dot{\phi}_0 \). We may compute \( \dot{y} \) retroactively from the constraint equation \( \dot{y}(t) = \dot{x}(t) \tan \phi(t) \), and then integrate it to find \( y(t) \) subject to the initial condition \( y(0) = y_0 \).

In practice, when solving the system on a computer, it is easier to adjoin the constraint equation (10.52) to the two equations, as in

\[
\frac{m_1 + m_2}{\cos^2 \phi} \ddot{x} + \frac{1}{\cos \phi} \left[ (m_1 + m_2) \dot{x} \sin \phi - m_2 \ell \dot{\phi} \cos^2 \phi \right] \dot{\phi} = 0,
\]

\[
m_2 \ell \dot{\phi} + \frac{1}{\cos \phi} m_2 \ell \dot{x} \dot{\phi} = 0,
\]

\[
\dot{x} \sin \phi - \dot{y} \cos \phi = 0,
\]

and solve the whole thing in one fell swoop for the three unknowns \( x(t) \), \( y(t) \), and \( \phi(t) \) by applying the initial conditions

\[
x(0) = x_0, \quad \dot{x}(0) = \dot{x}_0, \quad \phi(0) = \phi_0, \quad \dot{\phi}(0) = \dot{\phi}_0, \quad y(0) = y_0.
\]

Note that there is no initial condition on \( \dot{y}(0) \) since the \( \dot{x}_0 \) and \( \phi_0 \) determine \( \dot{y}(0) \) through the constraint equation (10.52).

### 10.6 The Čaplygin sleigh revisited

In Section 10.5 we obtained the differential equations of motion of the Čaplygin sleigh by applying the Gibbs-Appell equations (10.20). Here we solve the same problem by applying the equations (10.27). Toward that end, let us introduce the generalized velocities \( v \) and \( \omega \) defined through

\[
v = \frac{\dot{x}}{\cos \phi}, \quad \omega = \dot{\phi}.
\]

Thus, \( \omega \) is the rod’s angular velocity. To understand \( v \), note that \( \dot{x} = v \cos \phi \), therefore \( v \) is the speed (not velocity!) of the mass \( m_1 \).

Now, substitute \( \dot{\phi} = \omega \) and \( \dot{x} = v \cos \phi \) in (10.53) and simplify:

\[
\mathcal{G} = \frac{1}{2} (m_1 + m_2) v^2 + \frac{1}{2} m_2 \ell^2 \omega^2 + m_2 \ell v \omega \omega - m_2 \ell \omega^2 \dot{\phi} + \frac{1}{2} ((m_1 + m_2) v^2 + m_2 \ell^2 \omega^2) \omega^2.
\]

Equations (10.27) in this case take the form

\[
\frac{\partial \mathcal{G}}{\partial \dot{v}} = 0, \quad \frac{\partial \mathcal{G}}{\partial \dot{\omega}} = 0,
\]

and thus,

\[
(m_1 + m_2) v - m_2 \ell \omega^2 = 0, \quad (10.55a)
\]

\[
m_2 \ell^2 \omega^2 + m_2 \ell v \omega = 0. \quad (10.55b)
\]

These are much more pleasant-looking equations compared to (10.54) which we had obtained before. In fact, they are quite amenable to solving by hand. To wit, multiply the first equation by \( v \), the second by \( \omega \), and add. We get

\[
(m_1 + m_2) v v + m_2 \ell^2 \omega \omega = 0,
\]
that is
\[ \left( \frac{1}{2} (m_1 + m_2) v^2 + \frac{1}{2} m_2 \ell^2 \omega^2 \right) = 0, \]
whence
\[ E = \frac{1}{2} (m_1 + m_2) v^2 + \frac{1}{2} m_2 \ell^2 \omega^2 \] (10.56)
is a constant of the motion. (In fact, it is the sleigh’s kinetic energy.) The value of \( E \) is determined by the initial conditions \( v(0) \) and \( \omega(0) \).

Now, multiplying (10.55a) by \( \ell/2 \) and adding to (10.56), we get
\[ \frac{1}{2} (m_1 + m_2) \ell \dot{v} + \frac{1}{2} (m_1 + m_2) v^2 = E, \]
that is,
\[ \ell \dot{v} + v^2 = \alpha^2, \quad \text{where} \quad \alpha^2 = \frac{2E}{m_1 + m_2}. \]
We solve this separable equation for \( v \) subject to the initial condition \( v(0) = \nu_0 \), and obtain
\[ v(t) = \alpha \tanh \left[ \frac{\alpha}{\ell} t + \tanh^{-1} \frac{\nu_0}{\alpha} \right]. \] (10.57)
We see that \( \lim_{t \to \infty} v(t) = \alpha \), that is, the sleigh’s speed approaches \( \alpha \) in the long run.

Now that we have \( v(t) \), the angular velocity \( \omega(t) \) may be computed easily. From the differential equation \( \ell \dot{v} + v^2 = \alpha^2 \) we get \( \dot{v} = \frac{1}{\ell} (\alpha^2 - v^2) \). We substitute this expression for \( \dot{v} \) in (10.55a), then solve for \( \omega^2 \):
\[ \omega(t)^2 = \frac{m_1 + m_2}{m_2 \ell^2} (\alpha^2 - v(t)^2), \] (10.58)
where \( v(t) \) is given in (10.57). Note, in particular, that \( v(t) \to \alpha \) implies \( \omega(t) \to 0 \), that is, the sleigh’s spin slows down to zero in the long run.

**Remark 10.1.** Let’s observe that (a) From the definition \( \dot{x} = v \cos \phi \) of the speed \( v \) we see that a positive \( v \) implies motion in which \( m_1 \) trails \( m_2 \); and (b) Since \( \tanh \) is positive when its argument is positive, the solution (10.57) indicates that regardless of the initial conditions, \( v(t) \) will be positive for sufficiently large \( t \). Putting these two observations together, we conclude that regardless of the sleigh’s initial conditions, in the long run it will orient itself so that \( m_1 \) trails \( m_2 \) in its motion.

### 10.7 The problem from page 63 of Gantmacher

A “dumbbell” consists of a rigid weightless rod of length \( \ell \) with point masses of \( m \) each attached to its ends. The rod is free to move in a vertical plane, other than the requirement that the velocity of its center should point along the rod itself. Find the equations of motion.

We introduce a Cartesian coordinate system \( xy \), where \( x \) is horizontal and \( y \) points up. Let \( x(t) \) and \( y(t) \) be the coordinates of the rod’s center, and let \( \phi(t) \) be the rod’s angle relative to the \( x \) axis. We use \( x, y, \) and \( \phi \) as the problem’s generalized coordinates. Then \( r_c = (x, y) \) is the position vector of the rod’s midpoint. The position vectors of the two masses are
\[ r_1 = r_c - \frac{\ell}{2} (\cos \phi, \sin \phi), \quad r_2 = r_c + \frac{\ell}{2} (\cos \phi, \sin \phi). \]
Chapter 10. The Gibbs-Appell formulation of dynamics

Now we calculate the velocities of the masses:
\[
r_1 = \dot{r}_c - \frac{\ell}{2} \dot{\phi} (-\sin \phi, \cos \phi), \quad r_2 = \dot{r}_c + \frac{\ell}{2} \dot{\phi} (-\sin \phi, \cos \phi),
\]
and their accelerations:
\[
\ddot{r}_1 = \ddot{r}_c - \frac{\ell}{2} \ddot{\phi} (-\sin \phi, \cos \phi) - \frac{\ell}{2} \dot{\phi}^2 (-\cos \phi, -\sin \phi),
\]
\[
\ddot{r}_2 = \ddot{r}_1 + \frac{\ell}{2} \ddot{\phi} (-\sin \phi, \cos \phi) + \frac{\ell}{2} \dot{\phi}^2 (-\cos \phi, -\sin \phi).
\]

It follows that
\[
S = \frac{1}{2} m ||\dot{r}_1||^2 + \frac{1}{2} m ||\dot{r}_2||^2 = m||\dot{r}_c||^2 + \frac{m\ell}{2} (\dot{\phi}^2 + \dot{\phi}_1^2) = m(\dot{x}^2 + \dot{y}^2) + \frac{m\ell}{2} (\dot{\phi}^2 + \dot{\phi}_1^2)
\]
(10.60)

The formulation above indicates that we have made the tacit choice of using \(x\), \(y\) and \(\phi\) as generalized coordinates for this problem. Although these coordinates are independent of each other algebraically, they are related to each other through their derivatives, since according to the problem’s statement, the velocity of the rod’s midpoint is constrained to lie along the rod’s direction. This says that the velocity vector \(\dot{r}_c = (\dot{x}, \dot{y})\) makes an angle of \(\phi\) with the \(x\) axis, as the rod does, therefore \(\dot{y}/\dot{x} = \tan \phi\), or equivalently,
\[
\dot{x} \sin \phi = \dot{y} \cos \phi.
\]
(10.61)

This is the problem’s nonholonomic constraint.

To relate this to the general theorem developed in the previous sections, the problem has \(m = 3\) generalized coordinates \(x\), \(y\) and \(\phi\); and it has \(g = 1\) nonholonomic constraint (10.61). Therefore its degrees of freedom are \(n = m - g = 2\). Again, according to the preceding general theory, we should express the generalized velocities \(\dot{x}, \dot{y}\) and \(\dot{\phi}\), and the constraint (10.61), in terms of \(n = 2\) independent and unconstrained pseudovelocities \(\dot{\pi}_1\) and \(\dot{\pi}_2\).

There is no unique choice of the pseudovelocities \(\dot{\pi}_1\) and \(\dot{\pi}_2\). All is needed is the requirement that the pseudovelocities (10.33) together with the constraint equations (10.32) be solvable for the generalized velocities (10.34). Table 10.1 shows a few possible choices for the problem at hand. In Choice (1) we take pseudovelocities \(\dot{\pi}_1\) and \(\dot{\pi}_2\) to be the actual generalized velocities \(\dot{x}\) and \(\dot{\phi}\). The generalized velocity \(\dot{y}\) then is determined through the constraint (10.61).

Choice (2) defines the pseudovelocity \(\dot{\pi}_1\) to be the left-hand side of the constraint equation (10.61) which is a reasonable choice to make.

Choice (3) sets \(\dot{\pi}_1\) to the seemingly odd expression \(\dot{x}/\cos \phi\). The resulting expressions for \(\dot{x}\) and \(\dot{y}\) are quite simple, in the sense that they are easy to differentiate; we are going to need \(\dot{x}\) and \(\dot{y}\) to calculate the energy of acceleration. The derivatives of \(\dot{x}\) and \(\dot{y}\) will be messier expressions with the Choices (1) and (2).

On a closer scrutiny, the setting of \(\dot{\pi}_1 = \dot{x}/\cos \phi\) is not that odd after all since it has a pleasant physical interpretation. Let \(v\) be the (scalar) speed of the rod’s midpoint. Since the velocity points along the rod’s direction, then it should be clear that \(\dot{x} = v \cos \phi\) and \(\dot{y} = v \sin \phi\). Comparing the \(\dot{x}\) and \(\dot{y}\) values in Choice (3), we see that \(\dot{\pi}_1\) is the speed \(v\).

We adopt the pseudovelocities of Choice (3) for the rest of this section. To ease the notational burden, however, we replace the generic symbols \(\dot{\pi}_1\) and \(\dot{\pi}_2\) with the more meaningful notation \(v = \dot{\pi}_1\) and \(\omega = \dot{\pi}_2\). In particular, according to Choice (3) we have
\[
\dot{x} = v \cos \phi, \quad \dot{y} = v \sin \phi.
\]
(10.62)
10.7. The problem from page 63 of Gantmacher

In the last step we have substituted for \(\ddot{\ell}\) where the ellipses indicates terms which are free of the acceleration terms \(\dot{\omega}\).

To express the energy of acceleration \(S\) in (10.60) in terms of the quasivelocities, we differentiate (10.62):

\[
\begin{align*}
\dot{x} &= \dot{\ell} \cos \varphi - v \dot{\varphi} \sin \varphi = \dot{\ell} \cos \varphi - v \omega \sin \varphi, \\
\dot{y} &= \dot{\ell} \sin \varphi + v \dot{\varphi} \cos \varphi = \dot{\ell} \sin \varphi + v \omega \cos \varphi,
\end{align*}
\]

(10.63a) (10.63b)

whence \(\dot{x}^2 + \dot{y}^2 = \dot{\ell}^2 + v^2 \omega^2\), and (10.60) changes to

\[
S = m(\dot{\varphi}^2 + v^2 \omega^2) + \frac{m r^2}{2}(\omega^2 + \dot{\omega}^2) = m \dot{\varphi}^2 + \frac{m r^2}{2} \dot{\omega}^2 + \cdots,
\]

where the ellipses indicates terms which are free of the acceleration terms \(\dot{\varphi}\) and \(\dot{\omega}\).

A force of \(\mathbf{F} = (0, -mg)\) acts on each of the two masses. Therefore

\[
\mathbf{F} \cdot \mathbf{\ddot{r}_1} + \mathbf{F} \cdot \mathbf{\ddot{r}_2} = \mathbf{F} \cdot (\mathbf{\ddot{r}_1} + \mathbf{\ddot{r}_2}) = 2 \mathbf{F} \cdot \mathbf{\ddot{r}_c}
\]

\[
= 2(0, -mg) \cdot (\dot{x}, \dot{y}) = -2mg \dot{y} = -2mg(\dot{\varphi} \sin \varphi + v \omega \cos \varphi).
\]

In the last step we have substituted for \(\dot{y}\) from (10.63b).

We conclude that the Gibbs function is

\[
G = m \dot{\varphi}^2 + \frac{m r^2}{2} \omega^2 + 2mg \dot{\varphi} \sin \varphi + \cdots,
\]

where, as usual, the ellipses indicates terms which are free of the accelerations.

Then the Appell equations \(\partial G/\partial \dot{\varphi} = 0\) and \(\partial G/\partial \dot{\omega} = 0\) lead to

\[
\dot{\varphi} + g \sin \varphi = 0, \quad \dot{\omega} = 0.
\]

To complete the system, we append the constraint equations (10.62). Therefore the complete set of differential equations of motion are

\[
\begin{align*}
\dot{x} &= v \cos \varphi, \\
\dot{y} &= v \sin \varphi, \\
\dot{\varphi} &= \omega, \\
\dot{\ell} &= \ddot{\ell} \sin \varphi, \\
\dot{\ell} &= \ddot{\ell} \cos \varphi.
\end{align*}
\]

Table 10.1: There is no unique choice of pseudovelocities for a given problem. The choices given here are three out of infinite such possibilities for the dumbbell problem.
These form a set of five first order differential equations in the five unknowns \( x, y, \varphi, \omega, \) and \( v. \)

This system is solvable in terms of elementary functions, as demonstrated in Gantmacher. Plugging the raw system into Maple does not produce a good result immediately, but we can maneuver it toward the right solution as follows.

The equations above imply that \( \omega \) is a constant, and therefore \( \varphi = \omega t + \varphi_0. \) Therefore the system shrinks to

\[
\begin{align*}
\dot{x} &= v \cos \varphi, \\
\dot{y} &= v \sin \varphi, \\
\dot{v} + g \sin \varphi &= 0,
\end{align*}
\]

where \( \varphi = \omega t + \varphi_0. \) In soling this reduced system, Maple distinguishes between the cases where \( \omega \) is zero and nonzero. If \( \omega \) is nonzero, we get something like Gantmacher's. In particular, \( y \) consists of a linear term in \( t \) plus trigonometric function. If, however, \( \omega \) is zero and \( \varphi_0 = \pi/2, \) we get a \(-\frac{1}{2}g t^2\) term, as expected.

\textbf{Note:} In the above, I calculated the Gibbs function a la Desloge. We may calculate it a la Gantmacher if we wish. Here is how.

What I have written above as the quasivelocity \( v, \) Gantmacher calls it \( \dot{\pi}. \) The equation \( \dot{y} = v \sin \varphi \) then takes the form \( \dot{\gamma} = \dot{\pi} \sin \varphi, \) whereby we introduce the virtual displacement of the quasicordinate \( \pi \) through \( \delta y = (\sin \varphi) \delta \pi. \)

In accordance with (10.59), the virtual displacements of the masses are

\[
\begin{align*}
\delta r_1 &= \delta r_c - \frac{\ell}{2}(-\sin \varphi, \cos \varphi) \delta \varphi, \\
\delta r_2 &= \delta r_c + \frac{\ell}{2}(-\sin \varphi, \cos \varphi) \delta \varphi,
\end{align*}
\]

A force of \( F = \langle 0, -mg \rangle \) acts on each of the two masses therefore their virtual work is

\[
\begin{align*}
\delta W &= F \cdot \delta r_1 + F \cdot \delta r_2 = F \cdot (\delta r_1 + \delta r_2) = 2 F \cdot \delta r_c \\
&= 2 \langle 0, -mg \rangle \cdot \langle \delta x, \delta y \rangle = -2mg \delta y = -2m g \sin \varphi \delta \pi. \quad (10.64)
\end{align*}
\]

However \( \delta W = \Pi \delta \pi + \Phi \delta \varphi, \) where \( \Pi \) and \( \Phi \) are the generalized forces corresponding to the quasicordinates \( \pi \) and \( \varphi. \) We conclude that \( \Pi = -2mg \sin \varphi \) and \( \Phi = 0. \) The equations of motion are

\[
\begin{align*}
\frac{\partial S}{\partial \dot{\pi}} &= \Pi, \\
\frac{\partial S}{\partial \dot{\varphi}} &= \Phi,
\end{align*}
\]

which agree with what we obtained earlier.

\section*{10.8 Rigid body dynamics}

The previous sections have focused on the dynamics of point masses. Extending the results to rigid bodies is a matter of replacing the summations with integrals, which is easy in principle, but provides some challenges in practice. It is possible, with some work, to systematize the treatment and arrive at a general formula for the Gibbs function which applies to all rigid bodies. Lurie \cite{9} and Desloge \cite{2}, for instance, have such formulas. Lurie's formula is packaged quite nicely. Desloge's result is a little more general but it is not packaged as nicely. Furthermore, Desloge's formula is actually incorrect(!) due to

In the following sections I will offer a blending of Lurie’s and Desloge’s approaches which combines the nice packaging of the former with the generality of the latter.

10.8.1 Three frames of reference

It turns out that to derive the equation of motion of a rigid body, it is quite convenient to use not one, not two, but three(!) special purpose frames of reference simultaneously. In most applications a frame of reference takes the form of a right-handed orthonormal triad and a point called the frame’s origin.

The stationary frame is defined by a triad \( \{i, j, k\} \) and the associated origin \( O \). As the name implies, the stationary frame is fixed (not moving). Typically the \( i \) and \( j \) vectors lie in a horizontal plane and the vector \( k \) points upward, but that’s not a requirement.

With the stationary frame we associate a Cartesian coordinate system whose \( x, y, \) and \( z \) axes are aligned with the \( i, j \) and \( k \) vectors, respectively.

Ultimately, the purpose of rigid body dynamics is to express a body’s motion relative to the stationary frame.

The body frame of reference is defined by a triad \( \{b_1, b_2, b_3\} \) and an associated origin \( o \). The triad is firmly attached to the body, and therefore moves with it. There are no restrictions on the choice of the origin, nor on the triad’s orientation. Whenever possible, however, we set the origin at the body’s center of mass, and orient the triad along the body’s principal axes of inertia, since that results in significant simplifications.

Since the triad \( \{b_1, b_2, b_3\} \) is affixed to the body, the triad’s angular velocity is exactly the body’s angular velocity \( \omega \).

The intermediate frame of reference is defined by a triad \( \{e_1, e_2, e_3\} \) and an associated origin \( O' \). The choice of this frame is entirely up to you. You choose it to fit the specific application at hand. The triad may rotate and the origin may move as you wish.

Remark 10.2. What’s the purpose of the intermediate frame of reference? In formulating the equations of motion, some quantities are more easily expressed in one coordinate system than the other. For instance, the position of the body’s center of mass is best expressed in terms of the stationary frame, while its spin is best express in terms of the body reference frame. Some other quantities may be difficult to express in either of those, but may be easy in terms of a special purpose intermediate frame of reference. All three frames are related through orthogonal transformations, therefore we may readily translate the information from one frame to another, as needed.

10.8.2 The energy of acceleration for a rigid body

Consider a rigid body \( \mathcal{B} \) of mass \( m \) equipped with a body reference frame with an origin \( o \) and a orthonormal triad \( \{b_1, b_2, b_3\} \), as in Section 10.8.1. Let \( r_o \) be the position vector of the point \( o \) relative to the stationary frame, that is \( r_o = \overrightarrow{Oo} \), and let \( \mathcal{J} \) be body’s moment...
of inertia relative to \( o \). Furthermore, let \( c \) be the body’s center of mass, and let \( \rho_c = \overrightarrow{oc} \). Then we have\(^8\)

\[
\int_{B_{\text{cal}}} \frac{1}{2} ||\ddot{r}||^2 \, dm = \frac{1}{2} m ||\ddot{r}_o||^2 + m(\dot{r}_o \times \omega) \cdot \rho_c + m(\dot{r}_o \times \omega) \cdot (\omega \times \rho_c) + \frac{1}{2} \omega \cdot \mathcal{J}_o \omega + (\omega \times \omega) \cdot \mathcal{J}_o \omega + \cdots ,
\]

where the ellipsis indicates terms that do not involve accelerations, and are, therefore, immaterial to the Gibbs-Appell equations of motion.

If the origin \( o \) of the body frame is chosen to coincide with the body’s center of mass, then \( \rho_c = 0 \) and (10.65) reduces to

\[
\int_{B_{\text{cal}}} \frac{1}{2} ||\ddot{r}||^2 \, dm = \frac{1}{2} m ||\ddot{r}_c||^2 + \frac{1}{2} \dot{\omega} \cdot \mathcal{J}_c \dot{\omega} + (\omega \times \omega) \cdot \mathcal{J}_c \omega + \cdots ,
\]

where \( \ddot{r}_c \) is the position vector, relative to the stationary frame, of the body’s center of mass, and \( \mathcal{J}_c \) is the moment of inertia tensor relative to the center of mass.

10.9 ■ The rolling coin

Here we analyze the somewhat challenging dynamics of a coin, modeled as a thin homogeneous disk of radius \( a \), rolling without slipping on a horizontal floor; see Figure 10.2.

10.9.1 ■ The three frames

Figure 10.2 depicts the coin along with the three frames of reference discussed in Section 10.8.1. Specifically, the \{i, j, k\} triad of the stationary frame has its origin on the floor and the \( k \) vector points upward; the \{b_1, b_2, b_3\} triad of the body frame has its origin at the disk’s center and the vector \( b_3 \) is perpendicular to the disk.

We choose an intermediate frame with its origin at the disk’s center, and an associated orthonormal triad \( \{e_1, e_2, e_3\} \) oriented as follows. The vector \( e_3 \) is perpendicular to the disk, and therefore coincides with \( b_3 \). The vector \( e_1 \) is horizontal. Then we set \( e_2 = e_3 \times e_1 \).

The body frame is fixed to the coin, by definition, and rotates with it. The position of the body frame relative to the intermediate frame is described by a single parameter, the angle \( \psi \), between the vectors \( b_1 \) and \( e_1 \) as indicated on the figure.

The coin’s position relative to the stationary frame is specified through five generalized coordinates \( x, y, \theta, \phi, \) and \( \psi \), where \( \theta \) is the angle between the coin’s plane and the floor; \( \phi \) is the angle between the vectors \( i \) and \( e_1 \) (it measures how the coin’s plane is rotated relative to the \( k \) vector); and \( (x, y, a \sin \theta) \) are the coordinates of the coin’s center.

10.9.2 ■ The angular velocity

From the geometry of Figure 10.2 it is evident that \( e_1 = \cos \phi \, i + \sin \phi \, j \). The unit vector \( e_2 \) makes an angle \( \theta \) with the floor and its horizontal projection is perpendicular to \( e_1 \), therefore its horizontal and vertical projections are \( (-\sin \phi \, i + \cos \phi \, j) \cos \theta \) and \( \sin \theta \, k \).

We conclude that \( e_3 = (-\sin \phi \, i + \cos \phi \, j) \cos \theta + \sin \theta \, k \). The vector \( e_3 \) may be found

\(^8\)This is a combination of equation (4.11.8) in Lurie [9] and (22) in Desloge [11]. I haven’t personally verified them yet.
Figure 10.2: The rolling coin’s configuration is specified through five generalized coordinates $x$, $y$, $\theta$, $\phi$, and $\psi$, as shown. The no-slip conditions at the contact point with the floor imposes two nonholonomic constraints.

with a similar geometric reasoning, or just by computing $e_3 = e_1 \times e_2$. Here is a summary:

\begin{align*}
  e_1 &= \cos \phi \hat{i} + \sin \phi \hat{j}, \quad (10.67a) \\
  e_2 &= -\sin \phi \cos \theta \hat{i} + \cos \phi \cos \theta \hat{j} + \sin \theta \hat{k}, \quad (10.67b) \\
  e_3 &= \sin \phi \sin \theta \hat{i} - \cos \phi \sin \theta \hat{j} + \cos \theta \hat{k}. \quad (10.67c)
\end{align*}

We will need the time derivatives of these vectors shortly, so let’s compute them right now. We have

\[ \dot{e}_1 = -\dot{\phi} \sin \phi \hat{i} + \dot{\phi} \cos \phi \hat{j}. \]

It follows that

\[ \dot{e}_1 \cdot e_1 = 0, \quad \dot{e}_1 \cdot e_2 = \dot{\phi} \cos \theta, \quad \dot{e}_1 \cdot e_3 = -\dot{\phi} \sin \theta, \]

therefore, according to (8.7), we have $\dot{e}_1 = \dot{\phi} \cos \theta e_2 - \dot{\phi} \sin \theta e_3$. In a similar manner we compute $\dot{e}_2$ and $\dot{e}_3$ and express them in terms of the basis $\{e_1, e_2, e_3\}$. Here is what we get:

\begin{align*}
  \dot{e}_1 &= \dot{\phi} \cos \theta e_2 - \dot{\phi} \sin \theta e_3, \quad (10.68a) \\
  \dot{e}_2 &= -\dot{\phi} \cos \theta e_1 + \dot{\theta} e_3, \quad (10.68b) \\
  \dot{e}_3 &= \dot{\phi} \sin \theta e_1 - \dot{\theta} e_2. \quad (10.68c)
\end{align*}
Chapter 10. The Gibbs-Appell formulation of dynamics

Now let us examine the body frame. Referring to Figure 10.2 we have:

\[ b_1 = \cos \psi e_1 + \sin \psi e_2, \]  
\[ b_2 = -\sin \psi e_1 + \cos \psi e_2, \]  
\[ b_3 = e_3. \]  

(10.69a)
(10.69b)
(10.69c)

We differentiate these with respect to \( t \), and substitute from (10.68) for the derivatives of \( e_j \), and obtain

\[ \dot{b}_1 = -(\cos \theta + \dot{\psi}) \sin \psi e_1 + (\cos \theta + \dot{\psi}) \cos \psi e_2 + (-\dot{\psi} \sin \theta \cos \psi + \dot{\theta} \sin \psi) e_3, \]  
\[ \dot{b}_2 = -(\cos \theta + \dot{\psi}) \cos \psi e_1 - (\cos \theta + \dot{\psi}) \sin \psi e_2 + (\dot{\psi} \sin \theta \sin \psi + \dot{\theta} \cos \psi) e_3, \]  
\[ \dot{b}_3 = \dot{\psi} \sin \theta e_1 - \dot{\theta} e_2. \]  

(10.70a)
(10.70b)
(10.70c)

Remark 10.3. The expressions for \( \dot{b}_3 \) in (10.70c) and \( \dot{e}_3 \) in (10.68c) agree since \( b_3 = e_3 \).

Equations (10.70) may be used in conjunction with (7.5) to calculate the body’s angular velocity vector. Toward that end we compute

\[ \omega^b_1 = b_2 \cdot \dot{b}_3 = -(\cos \theta + \dot{\psi}) \cos \psi (e_1 \cdot b_3) - (\cos \theta + \dot{\psi}) \sin \psi (e_2 \cdot b_3) \]
\[ + (\dot{\psi} \sin \theta \sin \psi + \dot{\theta} \cos \psi) (e_3 \cdot b_3), \]

where the superscript \( b \) is to remind us that \( \omega^b \) is the component of the vector \( \omega \) in the the body frame. On the right-hand side we substitute for \( b_3 \) from (10.69), simplify the result, and arrive at \( \omega^b_1 = \dot{\psi} \sin \theta \sin \psi + \dot{\theta} \cos \psi \). Computing the \( \omega^b_2 \) and \( \omega^b_3 \) components in the same way, we arrive at:

\[ \omega^b_2 = b_3 \cdot \dot{b}_1 = \dot{\psi} \sin \theta \cos \psi - \dot{\theta} \sin \psi, \]
\[ \omega^b_3 = b_1 \cdot \dot{b}_2 = \dot{\psi} \cos \theta + \dot{\psi}. \]

Then from the definition (7.3) of the angular velocity vector we conclude that

\[ \omega = (\dot{\psi} \sin \theta \sin \psi + \dot{\theta} \cos \psi) b_1 + (\dot{\psi} \sin \theta \cos \psi - \dot{\theta} \sin \psi) b_2 + (\dot{\psi} \cos \theta + \dot{\psi}) b_3, \]  

(10.71)

This is the coin’s angular velocity vector expressed in the \( \{b_1, b_2, b_3\} \) basis. By substituting for \( \{b_1, b_2, b_3\} \) from (10.69), we obtain and expression for \( \omega \) in the \( \{e_1, e_2, e_3\} \) basis:

\[ \omega = \dot{\theta} e_1 + \dot{\psi} \sin \theta e_2 + (\dot{\psi} \cos \theta + \dot{\psi}) e_3. \]  

(10.72)

We see that the components of \( \omega \) along the intermediate basis are particularly simple. Therein lies the significance of the choice of our intermediate basis. Those components play a central role in what comes next, therefore we name them simply \( \omega^e_i \),

\[ \omega_1 = \dot{\theta}, \quad \omega_2 = \dot{\psi} \sin \theta, \quad \omega_3 = \dot{\psi} \cos \theta + \dot{\psi}, \]  

(10.73)

To be consistent with the \( b \) superscript introduced earlier for the components of \( \omega \) in the \( \{b_1, b_2, b_3\} \) basis, we should have named its \( \{e_1, e_2, e_3\} \) components with superscripts \( e \). We refrain from doing that, however, to reduce clutter.
Note that each $\omega_i$ is a linear combination of the generalized velocities $\dot{\theta}$, $\dot{\phi}$, and $\dot{\psi}$, and therefore each $\omega_i$ is a quasi-velocity as defined in Section 10.1.9. Well, actually $\omega_1$ is not a true quasi-velocity since it is the derivative of the generalized velocity $\dot{\theta}$, but the other two are honest quasi-velocities since they are not derivatives of anything.

Solving (10.73) as a linear system of three equations for the three unknowns $\dot{\theta}$, $\dot{\phi}$, and $\dot{\psi}$, we get:

$$
\dot{\theta} = \omega_1, \quad \dot{\phi} = \frac{1}{\sin \theta} \omega_2, \quad \dot{\psi} = \omega_3 - \omega_2 \cot \theta.
$$

Equations (10.73) and (10.74) establish a one-to-one correspondence between the generalized velocities $\dot{\theta}$, $\dot{\phi}$, and $\dot{\psi}$, and the quasi-velocities $\omega_1$, $\omega_2$, $\omega_3$. We may formulate the rest of the analysis in terms of one or the other set of variables. We will do it in terms of the $\omega$'s since the derivations are easier that way.

For future reference, let us make a note of the following formula whose derivation is left as an exercise:

$$
\dot{\omega} = [\dot{\omega}_1 + \omega_2 \omega_3 - \omega_2^2 \cot \theta] e_i + [\dot{\omega}_2 - \omega_3 \omega_1 + \omega_1 \omega_2 \cot \theta] e_j + \omega_3 e_k.
$$

10.9.3 The no-slip constraint

The position vector of the coin’s center in the stationary frame is

$$
r_c = x \hat{i} + y \hat{j} + a \sin \theta \hat{k}.
$$

The vector $\rho = -ae_2$ extends from the coin’s center to its contact point with the floor. Therefore, according to (7.7), the velocity, relative to the coin’s center, of a point on the coin’s rim at the contact point is given by $\omega \times \rho$. Consequently, the velocity of that point relative to the stationary frame is $\dot{r}_c + \omega \times \rho$. Since the velocity of the ground at the contact point is zero, then to prevent slippage, we need

$$
\dot{r}_c + \omega \times \rho = 0.
$$

Upon substituting for $r_c$ from (10.76), for $\omega$ from (10.72), for the $e_i$ vectors from (10.67), and simplifying the result, we get

$$
[\dot{x} - a \dot{\theta} \sin \theta \sin \phi + (\dot{\phi} \cos \theta + \dot{\psi}) \cos \phi] \hat{i} + [\dot{y} + a \dot{\theta} \cos \theta \sin \phi + (\dot{\phi} \cos \theta + \dot{\psi}) \cos \phi] \hat{j} = 0.
$$

Then we apply (10.74) to change over to quasi-velocities and arrive at

$$
[\dot{x} + a \omega_1 \cos \phi - a \omega_2 \sin \phi \sin \theta] \hat{i} + [\dot{y} + a \omega_3 \sin \phi + a \omega_1 \cos \phi \sin \theta] \hat{j} = 0.
$$

This leads us to a pair of scalar equations

$$
\dot{x} = -a(\omega_2 \cos \phi - \omega_1 \sin \phi \sin \theta), \quad \dot{y} = -a(\omega_3 \sin \phi + \omega_1 \cos \phi \sin \theta)
$$

which express the rolling coin’s nonholonomic constraints. We use these to eliminate $\dot{x}$ and $\dot{y}$ from the rest of the computations.

10.9.4 The acceleration of the coin’s center

From (10.76) we have $\dot{r}_c = \dot{x} \hat{i} + \dot{y} \hat{j} + a \dot{\theta} \cos \theta \hat{k}$. We substitute for $\dot{x}$ and $\dot{y}$ from (10.78), then we apply (10.74) and arrive at

$$
\dot{r}_c = -a(\omega_2 \cos \phi - \omega_1 \sin \phi \sin \theta) i - a(\omega_3 \sin \phi + \omega_1 \cos \phi \sin \theta) i + a \omega_1 \cos \theta k.
$$
Then we differentiate this once again to find the acceleration:

\[ \ddot{r}_c = a\left[\dot{\omega}_1 \sin \varphi \sin \theta - \dot{\omega}_2 \cos \varphi + \dot{\omega}_1^2 \sin \varphi \cos \theta + \omega_1 \omega_2 \cos \varphi + \omega_2 \omega_3 \sin \varphi / \sin \theta \right] i \\
+ a\left[-\dot{\omega}_1 \cos \varphi \sin \theta - \dot{\omega}_3 \sin \varphi - \dot{\omega}_1^2 \cos \varphi \cos \theta + \omega_1 \omega_2 \sin \varphi - \omega_2 \omega_3 \cos \varphi / \sin \theta \right] j \\
+ a\left[\dot{\omega}_1 \cos \theta - \dot{\omega}_1^2 \sin \theta \right] k. \]  

(10.79)

Here we have applied (10.74) to eliminate the generalized velocities \( \dot{\Theta}, \dot{\varphi}, \dot{\psi} \) in favor of the quasi-velocities \( \omega_1, \omega_2, \omega_3 \). Finally, we compute \( ||\ddot{r}_c||^2 \) which forms a part of the problem’s Gibbs function:

\[ ||\ddot{r}_c||^2 = a^2[\dot{\omega}_1^2 + \dot{\omega}_2^2 + 2\omega_1 \omega_2 \omega_3 - 2\omega_1 \omega_2 \omega_3 + \cdots], \]

(10.80)

where the ellipsis indicate terms that involve no acceleration terms \( \dot{\omega}_1, \dot{\omega}_2, \dot{\omega}_3 \).

**10.9.5 The rotational acceleration**

The body frame \( \{b_1, b_2, b_3\} \) is lined up with the coin’s principal moment of inertia axes, therefore the moment of inertia tensor relative to the coin’s center is

\[ \mathcal{I} = a b_1 \otimes b_1 + a b_2 \otimes b_2 + \beta b_3 \otimes b_3, \]

where

\[ a = \frac{1}{4} m a^2, \quad \beta = \frac{1}{2} m a^2. \]

Therefore we have:

\[ \mathcal{I} \omega = a(b_1 \cdot \omega)b_1 + a(b_2 \cdot \omega)b_2 + \beta(b_3 \cdot \omega)b_3, \]

\[ \mathcal{I} \ddot{\omega} = a(b_1 \cdot \ddot{\omega})b_1 + a(b_2 \cdot \ddot{\omega})b_2 + \beta(b_3 \cdot \ddot{\omega})b_3. \]

Inserting for \( \ddot{\omega} \) from (10.75) we get:

\[ \ddot{\omega} \cdot \mathcal{I} \ddot{\omega} = a(\dot{\omega}_1 + \omega_2 \omega_3 - \omega_3 \omega_2) \omega_1 + a(\omega_2 - \omega_3 \omega_1 + \omega_1 \omega_2 \cot \theta)^2 + \omega_1^2 \\
= a \omega_1^2 + a \omega_2^2 + \beta \omega_3^2 - 2a(\omega_1 \omega_2 - \omega_1 \omega_2)(\omega_2 \cot \theta - \omega_3) \\
+ a(\omega_1^2 + \omega_2^2)(\omega_2 \cot \theta - \omega_3)^2. \]

(10.82)

The final term is free of accelerations, therefore it may be dropped when forming the Gibbs function.

In a similar manner we also calculate

\[ (\dddot{\omega} \times \omega) \cdot \mathcal{I} \ddot{\omega} = (\beta - a)(\omega_1 \omega_2 - \omega_1 \omega_2) \omega_3 \\
- (\beta - a)(\omega_1^2 + \omega_2^2)(\omega_2 \cot \theta - \omega_3) \omega_3. \]

(10.83)

Again, the final term is free of accelerations, therefore it may be dropped when forming the Gibbs function.

According to (10.66) and (10.7), The Gibbs function is

\[ \mathfrak{G} = \int_{\mathcal{M}} \frac{1}{2} ||\ddot{r}_c||^2 \, dm - \int_{\mathcal{M}} \mathcal{F} \cdot \ddot{r} \, dm \\
= \frac{1}{2} m ||\ddot{r}_c||^2 + \frac{1}{2} \dot{\omega} \cdot \mathcal{I} \ddot{\omega} + (\dddot{\omega} \times \omega) \cdot \mathcal{I} \ddot{\omega} - (\mathbf{mgk}) \cdot \ddot{r}_c, \]

where \( \mathcal{F} \) is the force due to friction and the like.
which may be evaluated by putting together (10.79), (10.80), (10.82), and (10.83):

\[ \mathbf{G} = \frac{1}{2} ma^2 [x\dot{\omega}_1^2 + \frac{5}{8} \dot{\omega}_1^2 + \frac{1}{8} \dot{\omega}_2^2 + \frac{3}{4} \dot{\omega}_1 \dot{\omega}_2 \cos \theta - 6\omega_1^2] + \frac{\alpha}{2} \omega_1^2 + \frac{\beta}{2} \dot{\omega}_2^2 \]

\[ - (\omega_1 \omega_2 - \omega_1 \dot{\omega}_2)(\alpha \omega_2 \cot \theta - \beta \dot{\omega}_1) + m g a \dot{\omega}_1 \cos \theta + \cdots, \]

where, as always, we have dropped terms which do not depend on the accelerations. Finally, substituting for \( \alpha \) and \( \beta \) from (10.81), we arrive at

\[ \mathbf{G} = ma^2 \left[ \frac{5}{8} \dot{\omega}_1^2 + \frac{1}{8} \dot{\omega}_2^2 + \frac{3}{4} \dot{\omega}_1 \dot{\omega}_2 \cos \theta - 6\omega_1^2 \right] + \frac{1}{4} \omega_1 \dot{\omega}_2 (\omega_2 \cot \theta - 2 \omega_3) - \omega_1 \omega_2 \dot{\omega}_1 + m g a \dot{\omega}_1 \cos \theta. \]

Then from the equations \( \partial \mathbf{G} / \partial \dot{\omega}_1 = 0, \partial \mathbf{G} / \partial \dot{\omega}_2 = 0, \partial \mathbf{G} / \partial \dot{\omega}_3 = 0 \) we get

\[ \frac{5}{4} \dot{\omega}_1 - \frac{1}{4} \dot{\omega}_2 (\omega_2 \cot \theta - 6 \omega_1) + \frac{8}{d} \cos \theta = 0, \]
\[ \frac{1}{4} \dot{\omega}_2 + \frac{1}{4} \dot{\omega}_1 (\omega_2 \cot \theta - 2 \omega_3) = 0, \]
\[ \frac{3}{2} \dot{\omega}_3 - \omega_1 \omega_2 = 0. \]

Clearly this set of differential equations is under-determined, since it depend on \( \theta \) which is also an unknown. However, from (10.74) we see that \( \dot{\theta} = \omega_1 \). Adjoining this to the above gives us a system of four first order differential equations in the four unknowns \( \omega_1, \omega_2, \omega_3, \) and \( \dot{\theta} \).

In practice, we extend the system by adjoining all three of the equations from (10.74), and the two equations from (10.78). Thus, we obtain a system of eight first order differential equations in the eight unknowns \( \omega_1, \omega_2, \omega_3, \theta, \phi, \psi, x, \) and \( y \), whose solution completely determines the coin's motion. Here is the system in its full glory:

\[ \begin{align*}
\dot{\omega}_1 &= -\frac{1}{5} \omega_2 (\omega_2 \cot \theta - 6 \omega_1) + \frac{4g}{5a} \cos \theta = 0, \\
\dot{\omega}_2 + \omega_1 (\omega_2 \cot \theta - 2 \omega_3) &= 0, \\
\dot{\omega}_3 &= \frac{2}{3} \omega_1 \omega_2 = 0, \\
\dot{\theta} &= \omega_1, \\
\dot{\phi} &= \frac{1}{\sin \theta} \omega_2, \\
\dot{\psi} &= \omega_3 - \omega_2 \cot \theta, \\
\dot{x} &= -a (\omega_3 \cos \phi - \omega_1 \sin \phi \sin \theta), \\
\dot{y} &= -a (\omega_3 \sin \phi + \omega_1 \cos \phi \sin \theta).
\end{align*} \]

**Remark 10.4.** We need eight initial conditions to go with these equations. The initial conditions on \( x, y, \dot{x}, \dot{y}, \phi, \psi \), \( \dot{\theta} \), \( \dot{\phi} \), and \( \dot{\psi} \) determine the coin's initial position relative to the stationary axes, so they may be specified in the obvious way. The initial conditions on \( \omega_1, \omega_2, \) and \( \omega_3 \) require a more careful consideration. You may recall from (10.73) that the \( \omega \)'s were defined as the components of the coin's angular velocity along the intermediate
frame \{e_1, e_2, e_3\}. Since the intermediate frame has no immediate physical manifestation, it is not easy to make up meaningful initial values for the \(\omega\)'s. We do note, however, that (10.73) defines the \(\omega\)'s in terms of the generalized coordinates and their velocities, which are easier to grasp. Therefore in practice you will make up initial values for \(\dot{\theta}, \dot{\phi}, \text{ and } \dot{\psi}\) as desired, then use (10.73) to determine the initial values for \(\omega_1, \omega_2, \text{ and } \omega_3\).

### Exercises

10.1. Derive (10.75).

10.2. Give the details of the computation which leads from (10.77) to (10.78).

10.3. Derive (10.83).

10.4. Equip a rigid body \(B\) with a body reference frame \(\{b_1, b_2, b_3\}\) whose origin coincides with the body’s center of mass. Additionally, suppose that the \(\{b_1, b_2, b_3\}\) is aligned with the body’s moment of inertia axes, so that the moment of inertia tensor takes the form

\[
\mathcal{I} = I_1 b_1 \otimes b_1 + I_2 b_2 \otimes b_2 + I_3 b_3 \otimes b_3.
\]

Let \(\omega_1, \omega_2, \omega_3\) be the components of the body’s angular velocity along the \(\{b_1, b_2, b_3\}\) vectors, that is

\[
\omega = \omega_1 b_1 + \omega_2 b_2 + \omega_3 b_3.
\]

Show that:

1. \(\dot{\omega} \cdot \mathcal{I} \dot{\omega} = I_1 \dot{\omega}_1^2 + I_2 \dot{\omega}_2^2 + I_3 \dot{\omega}_3^2\).

2. \((\dot{\omega} \times \omega) \cdot \mathcal{I} \omega = I_1 (\dot{\omega}_2 \omega_3 - \dot{\omega}_3 \omega_2) \omega_1 + I_2 (\dot{\omega}_3 \omega_1 - \dot{\omega}_1 \omega_3) \omega_2 + I_3 (\dot{\omega}_1 \omega_2 - \dot{\omega}_2 \omega_1) \omega_3\).

Conclude that the equations of motion of a freely spinning rigid body are

\[
\begin{align*}
I_1 \ddot{\omega}_1 + (I_3 - I_2) \omega_2 \omega_3 &= 0, \\
I_2 \ddot{\omega}_2 + (I_1 - I_3) \omega_3 \omega_1 &= 0, \\
I_3 \ddot{\omega}_3 + (I_2 - I_1) \omega_1 \omega_2 &= 0.
\end{align*}
\]

10.5. Solve the system of differential equations (10.84) and produce an animation of the coin’s motion.
Chapter 11
Quaternions

11.1 The quaternion algebra

A quaternion, as a mathematical object, is a pair \([a, u]\), where \(a \in \mathbb{R}\) and \(u \in E_3\). Multiplication by a scalar \(c\), the sum \(p + q\), and the product \(p \circ q\) of the quaternions \(p = [a, u]\) and \(q = [b, v]\) are defined according to

\[
cp = c[a, u] = [ca, cu], \quad (c \in \mathbb{R})
\]

\[
p + q = [a, u] + [b, v] = [a + b, u + v],
\]

\[
p \circ q = [a, u] \circ [b, v] = [ab - u \cdot v, av + bu + u \times v].
\]

Note that addition is commutative but multiplication, due to the presence of the \(u \times v\) term, is not.

As a special case of (11.3), we have

\[
[c, 0] \circ [a, u] = c[a, u] \quad \text{and} \quad [a, u] \circ [c, 0] = c[a, u],
\]

therefore the quaternion \([c, 0]\) acts exactly like the scalar \(c\) under quaternion multiplication. In particular, the quaternion \([1, 0]\) is the multiplicative identity in the quaternion algebra.

In view of the observation above, we adopt the convention of writing \(c\) instead of \([c, 0]\) when there is no chance of confusion. This is analogous to writing \(a\), instead of \(a + 0i\) when dealing with complex numbers. By the same token, we write \(u\) instead of \([0, u]\). The quaternion \([c, 0]\) is called a scalar quaternion or just a scalar if the meaning is clear from the context. Similarly, the quaternion \([0, u]\) is called a vectorial quaternion, or just a vector, for short.\footnote{The terminology is not quite standardized. What we have called a scalar quaternion is also called a real quaternion, and what we have called a vectorial quaternion is also called a pure quaternion or imaginary quaternion.}

As a consequence of the convention adopted above, the notation \(u \circ v\), where \(u\) and \(v\) are vectors, makes sense, and evaluates to

\[
u \circ v = [0, u] \circ [0, v] = [-u \cdot v, u \times v].
\]

You may want to think of \(u \circ v\) as the “quaternion product” of the vectors \(u\) and \(v\), but note that the product is a quaternion, not a vector, unless \(u \cdot v = 0\), in which case \(u \circ v = u \times v\).
The conjugate $p^*$ of the quaternion $p = [a, u]$ is defined as $p^* = [a, -u]$. Let us note that
\[ p^* \circ p = [a^2 + \|u\|^2, 0] = (a^2 + \|u\|^2)[1, 0] = a^2 + \|u\|^2 = |p|^2, \tag{11.5} \]
where we have defined $|p| = (a^2 + \|u\|^2)^{1/2}$.

The quantity $|p|$ is called the norm (or modulus) of the quaternion $p$. A quaternion $p$ such that $|p| = 1$ is called a unit quaternion.

If $|p|$ is nonzero, then we may rearrange (11.5) into
\[ \left( \frac{1}{|p|^2} p^* \right) \circ p = 1, \]
which shows that the parenthesized expression is the left multiplicative inverse of $p$. Repeating the calculation, beginning with $p \circ p^*$, we see that the parenthesized expression is also the right multiplicative inverse of $p$. We conclude that if $|p|$ is nonzero, then $p$ has a multiplicative inverse, $p^{-1}$, given by
\[ p^{-1} = \frac{1}{|p|^2} p^*, \quad (|p| \neq 0). \]

The special case unit criterion arises quite frequently:
\[ p^{-1} = p^*, \quad (|p| = 1). \]

**Remark 11.1.** This section has touched on just a few basic concepts of the quaternion algebra which will be needed in what follows. For an in-depth study of quaternions see Altman [12]. For a leisurely introduction to quaternions, with applications to computer graphics, see Hanson [13].

### 11.2 The geometry of the quaternions

The goal of this section is to elucidate the geometric interpretations of a few features related to quaternion which play significant roles in the dynamics of rigid bodies.

#### 11.2.1 The reflection operator

As we noted in connection with (11.4), the quaternion product of two vectors is not a vector in general. Thus, it may come as a surprise that the triple product $v \circ u \circ v$ is always a vector:
\[ v \circ u \circ v = \|v\|^2 u - 2(v \cdot u) v, \quad \text{for all } u, v \in E_3. \tag{11.6} \]

The proof of this identity is left as an exercise.

In particular, if $v$ is a unit vector, let’s call it $n$, then we have
\[ n \circ u \circ n = u - 2(n \cdot u) n, \quad \text{if } \|n\| = 1. \tag{11.7} \]

The right-hand side has a well-known geometric interpretation. Suppose the tails of the vectors $n$ and $u$ are attached to a common point $o$, and let $P$ be a plane through $o$ and perpendicular to $n$, as illustrated in Figure 11.1, and let $Q$ be the plane of the vectors $n$. 

11.2. The geometry of the quaternions

and \( u \). Then it is simple exercise to show that \( u - 2(n \cdot u)n \) is the mirror reflection of the vector \( u \) through the plane \( P \). Thus, we have established the following:

**Theorem 11.1.** The reflection operator \( R_n \) into a plane with a unit normal \( n \) is given by

\[
R_n u = n \circ u \circ n, \quad u \in E_3.
\]

(11.8)

11.2.2 **The rotation operator**

Consider two planes, \( P_1 \) and \( P_2 \), which have unit normals \( n_1 \) and \( n_2 \) and make a dihedral angle \( \phi/2 \) with each other. In the sketch shown in Figure 11.2a, the planes are shown edgewise, therefore they appear as lines. The line of intersection of the two planes appears as a point in that figure since we are looking at it head on. The unit vector \( n = (n_1 \times n_2)/\|n_1 \times n_2\| \) lies in the direction of that line of intersection.

Let \( R_{n_1} \) and \( R_{n_2} \) be the reflection operators, as in the previous subsection, into the planes \( P_1 \) and \( P_2 \). Pick an arbitrary vector \( u \), apply \( R_{n_1} \) to it, and then apply \( R_{n_2} \) to the result. Figure 11.2b shows the effect.

To gain insight into the structure of this geometric construction, it helps to introduce the plane \( P'_1 \) which is the reflection of \( P_1 \) into \( P_2 \). The plane \( P'_1 \), seen edgewise, is shown as a dashed line in Figure 11.2c.

In Figure 11.2d we have added the vector \( R_{n_2} u \). It should be evident from the symmetries of the diagram that \( R_{n_2} u \) coincides with the reflection of \( R_{n_1} R_{n_2} u \) into \( P'_1 \).

As a rotation by the angle \( \phi \) about the vector \( n \) takes the plane \( P_1 \) to the plane \( P'_1 \), it is evident that the same rotation takes the vector \( u \) to the vector \( R_{n_1} R_{n_2} u \). Thus, we have established the following:

**Lemma 11.2.** Suppose the planes \( P_1 \) and \( P_2 \), with unit normals \( n_1 \) and \( n_2 \), make a dihedral angle of \( \phi/2 \) with each other. Let \( R_{n_1} \) and \( R_{n_2} \) be the reflection operators into the planes \( P_1 \) and \( P_2 \), and let \( R_{n_1 \phi} \) be the rotation operator by angle \( \phi \) about the vector \( n = (n_1 \times n_2)/\|n_1 \times n_2\| \). Then \( R_{n_1 \phi} = R_{n_2} R_{n_1} \).

In the previous section we characterized a reflection operator as a triple quaternion product. That, along with the lemma above enables us to express a rotation operator in terms of quaternions. This is the subject of the following:
104 Chapter 11. Quaternions

(a) Planes \( P_1 \) and \( P_2 \), shown edge-wise, have unit normals \( n_1 \) and \( n_2 \), and make a dihedral angle of \( \phi/2 \). The unit vector 
\[
\mathbf{n} = (n_1 \times n_2) / \|n_1 \times n_2\| \nolinebreak 
\]
lies along their intersection, and points out of the picture, toward you, therefore it is not visible.

(b) Reflecting the arbitrary vector \( \mathbf{u} \) into the plane \( P_1 \) produces \( R_{n_1} \mathbf{u} \). Reflecting the latter into the plane \( P_2 \) results in \( R_{n_2} R_{n_1} \mathbf{u} \).

(c) We introduce the auxiliary plane \( P_1' \) (shown as a dashed line) which is the reflection of the plane \( P_1 \) into the plane \( P_2 \), therefore it makes a dihedral angle of \( \phi/2 \) with \( P_2 \).

(d) The vector \( R_{n_2} \mathbf{u} \), produced by reflecting \( \mathbf{u} \) into \( P_2 \) coincides with the reflection of \( R_{n_1} R_{n_2} \mathbf{u} \) into \( P_1' \). From the symmetries of the diagram it is evident that a rotation by an angle \( \phi \) about the vector \( \mathbf{n} \) takes the plane \( P_1 \) to the plane \( P_1' \). It follows that the same rotation takes the vector \( \mathbf{u} \) to the vector \( R_{n_2} R_{n_1} \mathbf{u} \), thus proving that \( R_{n_2} R_{n_1} = R_{n_1} R_{n_2} \).

Figure 11.2: The sequence of diagrams is in effect a “proof without words” of Lemma 11.2. It shows clearly that the composition of the two reflections \( R_{n_1} \) and \( R_{n_2} \) into the planes \( P_1 \) and \( P_2 \) is equivalent to a rotation by an angle \( \phi \) about the vector \( \mathbf{n} \) which lies in the intersection of those planes.
11.2. The geometry of the quaternions

Theorem 11.3. The operator $R_{n, \varphi}$ of rotation by angle $\varphi$ about a unit vector $n$ has an associated unit quaternion

$$q = [\cos \frac{\varphi}{2}, n \sin \frac{\varphi}{2}]$$

so that

$$R_{n, \varphi} u = q \circ u \circ q^*, \text{ for all } u \in \mathbb{R}^3.$$ (11.10)

Proof. Pick any two planes $P_1$ and $P_2$ which intersect along a line parallel to $n$, and whose dihedral angle is $\varphi/2$. Then according to Theorem 11.1 we have

$$R_{n_1} u = n_1 \circ u \circ n_1, \quad R_{n_2} u = n_2 \circ u \circ n_2,$$

and according to Lemma 11.2 we have $R_{n, \varphi} = R_{n_2} R_{n_1}$. Thus, we compute

$$R_{n, \varphi} u = R_{n_2} R_{n_1} u$$

$$= R_{n_2} (n_1 \circ u \circ n_1)$$

$$= n_2 \circ (n_1 \circ u \circ n_1) \circ n_2$$

$$= (n_2 \circ n_1) \circ u \circ (n_1 \circ n_2).$$

Let $q = n_2 \circ n_1$. We leave it as an exercise for you to show that $q^* = n_1 \circ n_2$. We conclude that $R_{n, \varphi} = q \circ u \circ q^*$, which proves (11.10). To show that $q$ has the form given in (11.9), observe that the dihedral angle between the planes $P_1$ and $P_2$ is $\varphi/2$, therefore the vectors $n_1$ and $n_2$, which are perpendicular to those planes, also form an angle of $\varphi/2$. Then applying (11.4) we see that

$$q = n_2 \circ n_1 = [-n_2 \cdot n_1, n_2 \times n_1] = [-\cos \frac{\varphi}{2}, -n \sin \frac{\varphi}{2}] = [-\cos \frac{\varphi}{2}, n \sin \frac{\varphi}{2}].$$

The negative sign in the final expression is immaterial since $q$ appears twice in the rotation formula (11.10). Finally, the assertion that $q$ is a unit quaternion follows immediately from the form of (11.9).

Remark 11.2. The planes $P_1$ and $P_2$ and their associated unit normals $n_1$ and $n_2$ which enter the proof above are conveniences for the proof. They do not appear in the theorem’s final result.

Remark 11.3. If we split the quaternion $q$ in (11.10) into components, as in $q = [q_0, q]$, and expand the triple quaternion product, we get

$$R_{n, \varphi} u = (q_0^2 - \|q\|^2)u + 2(q \cdot u)q + 2q_0(q \times u).$$ (11.11)

We may be more explicit by setting $q_0 = \cos \varphi/2$ and $q = n \sin \varphi/2$ and simplifying (11.10) to get

$$R_{n, \varphi} u = (n \cdot u)n + (u - (n \cdot u)n) \cos \varphi + (n \times u) \sin \varphi$$

which has an obvious geometric interpretation if you look at it closely.
11.3 Angular velocity

Theorem 11.3 shows that a rotation operator in $E_3$ may be expressed in terms of an associated unit quaternion $q$ defined in (11.9). A rotation that varies with the time $t$ then may be expressed in terms of a time-dependent unit quaternion $q(t)$ of that form. An arbitrarily fixed vector $r_0 \in E_3$ then rotates to the position $r(t)$ according

$$ r(t) = q(t) \circ r_0 \circ q^*(t). \quad (11.12) $$

Suppose that the instantaneous angular velocity vector under this rotation at time $t$ is $\omega(t)$. Then, according to (7.7), we have:

$$ \dot{r}(t) = \omega(t) \times r(t). \quad (11.13) $$

Comparing (11.12) and (11.13) we expect a relationship between $q(t)$ and $\omega(t)$. This is stated in the following:

**Theorem 11.4.** Let $q(t)$ be any time-varying unit quaternion, and let $\omega(t)$ the angular velocity of the rotation (11.12) engendered by $q(t)$. Then we have

$$ \omega = 2\dot{q} \circ q^*. \quad (11.14) $$

**Proof.** Since $q(t)$ is a unit quaternion, we have $|q(t)|^2 = q(t) \circ q^*(t) = 1$, and therefore $\dot{q} \circ q^* + q \circ \dot{q}^* = 0$. But since $q \circ q^* = (\dot{q} \circ q^*)^*$, it follows that $\dot{q} \circ q^* + (\dot{q} \circ q^*)^* = 0$, which indicates that the scalar part of the quaternion $\dot{q} \circ q^*$ is zero, that is, $\dot{q} \circ q^*$ is a vectorial quaternion. This justifies defining the vector $\omega$ through the quaternion product (11.14).

It remains to show that $\omega$ thus defined is indeed the motion’s angular velocity.

Toward that end, let us compute $\dot{r}$ by differentiating (11.12):

$$ \dot{r} = \dot{q} \circ r_0 \circ q^* + q \circ r_0 \circ \dot{q}^*. $$

We note $q^{-1} = q^*$ by virtue of $q$ being a unit quaternion, therefore $r = q \circ r_0 \circ q^*$ implies that

$$ r_0 \circ q^* = q^* \circ r \quad \text{and} \quad q \circ r_0 = r \circ q. $$

Therefore, the expression for $\dot{r}$ takes the form

$$ \dot{r} = \dot{q} \circ q^* \circ r + r \circ \dot{q} \circ q^* $$

$$ = \dot{q} \circ q^* \circ r + r \circ (\dot{q} \circ q^*)^* $$

$$ = \left(\frac{1}{2}\omega\right) \circ r + r \circ \left(\frac{1}{2}\omega^*\right) = \frac{1}{2}(\omega \circ r + r \circ \omega^*). $$

We leave it as an easy exercise to show that evaluating the right-hand side with the help of the identity (11.4) leads to

$$ \dot{r} = \omega \times r \quad (11.15) $$

which shows that $\omega$ defined in (11.14) is indeed the sought-after angular velocity. \qed
11.4 A differential equation for the quaternion rotation

Recall the definitions of the three frames of reference in Section 10.8.1 in connection with the motion of a rigid body. The stationary frame is equipped with a non-moving orthonormal triad \( \{i, j, k\} \). The body frame is equipped with an orthonormal triad \( \{b_1, b_2, b_3\} \) which is fixed to the body and moves with it.

At any time \( t \), the body triad \( \{b_1, b_2, b_3\} \) may be viewed as the rotated version of the stationary triad \( \{i, j, k\} \). Let \( q(t) \) be the unit quaternion associated with that rotation. Thus,

\[
b_1(t) = q(t) \circ i \circ q^*(t), \quad b_2(t) = q(t) \circ j \circ q^*(t), \quad b_3(t) = q(t) \circ k \circ q^*(t),
\]

which may be expanded with the help of (11.11), if desired.

Let \( \omega \) be the body’s angular velocity vector, and \( \omega_1, \omega_2, \) and \( \omega_3 \) be \( \omega \)'s components in the body frame:

\[
\omega = \omega_1 b_1 + \omega_2 b_2 + \omega_3 b_3.
\]

Substituting from (11.16) we get

\[
\omega = \omega_1 q \circ i \circ q^* + \omega_2 q \circ j \circ q^* + \omega_3 q \circ k \circ q^* = q \circ (\omega_1 i + \omega_2 j + \omega_3 k) \circ q^*.
\]

Then from (11.14) it follows that

\[
2\dot{q} \circ q^* = q \circ (\omega_1 i + \omega_2 j + \omega_3 k) \circ q^*,
\]

whence

\[
\dot{q} = \frac{1}{2} q \circ (\omega_1 i + \omega_2 j + \omega_3 k).
\]

This system of first order differential equations expresses the evolution of \( q \) in time. Together with the equations in (11.16), these determine the body’s orientation in space. We will adjoin them to the differential equations of dynamics to obtain a complete system of differential equations that describes the body’s motion.

The differential equation’s initial condition, \( q(0) \), which specifies the body’s initial orientation, is a unit quaternion, as any quaternion associated with a rotation ought to be. A question arises whether the solution \( q(t) \) of the differential equation is a unit quaternion for all \( t \). The answer is yes, as it follows from Theorem 11.5 below.

**Remark 11.4.** Equation (11.17) defines the components \( \omega_1, \omega_2, \omega_3 \) of the angular velocity vector \( \omega \) along the body triad. Beware that the sum \( \omega_1 i + \omega_2 j + \omega_3 k \) that appears in (11.18) does not add up to \( \omega \).

**Theorem 11.5.** The differential equation

\[
\dot{q}(t) = q(t) \circ a(t), \quad q(0) = q_0
\]

for the quaternion \( q(t) \), where \( a(t) \) is an arbitrary time dependent vector, preserves the norm, that is \( |q(t)| = |q_0| \) for all \( t \). In particular, if \( q_0 \) is a unit quaternion, then \( q(t) \) is a unit quaternion for all \( t \).

**Proof.** We calculate the derivative of \( |q(t)|^2 \) and substitute for \( \dot{q} \) from the differential
equation:

\[
\frac{d}{dt}|q|^2 = \frac{d}{dt}(q \circ q^*) = \dot{q} \circ q^* + q \circ \dot{q}^* \\
= (q \circ a) \circ q^* + q \circ (q \circ a)^* = (q \circ a) \circ q^* + q \circ (a^* \circ q^*).
\]

In the last step we have used the quaternion conjugation property \((p \circ q)^* = q^* \circ p^*\); see Exercise 5. Now note that for any vector \(a\) we have \(a^* = [0, a]^* = [0, -a] = -a\). Then it follows that \(\frac{d}{dt}|q|^2 = 0\), therefore the norm of \(q\) remains constant. \(\Box\)

For computing purposes, we express \(q\) in components, as in

\[q = [q_0, q_1i + q_2j + q_3k],\]

whereby (11.18) expands to the system of differential equations

\[
\frac{d}{dt}\begin{pmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 0 & -\omega_1 & -\omega_2 & -\omega_3 \\ \omega_1 & 0 & -\omega_3 & -\omega_2 \\ -\omega_2 & -\omega_3 & 0 & -\omega_1 \\ \omega_3 & \omega_2 & \omega_1 & 0 \end{pmatrix} \begin{pmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{pmatrix}.
\] (11.19)

### 11.5 Unbalanced ball rolling on a horizontal plane

Consider a solid ball of radius \(R\) and uniformly distributed total mass of \(M\). We know that the ball’s moment of inertia tensor relative to its center is \(2/5MR^2\) times the identity tensor.

We embed a point of mass \(m\), let’s call it a “slug”, within that ball at a distance \(a \leq R\) away from its center. We wish to study the motion of the composite object when it rolls without slipping on a horizontal plane.

We set up the usual \([i, j, k]\) orthonormal triad as a stationary frame, with the origin on the horizontal plane and \(k\) pointing upward. We set up the body frame with its origin at the ball’s center, and the body triad \([b_1, b_2, b_3]\) oriented so that the slug is at \(ab_3\) relative to the ball’s center, and we let \(q(t)\) be the unit quaternion associated with the rotation that takes \([i, j, k]\) to \([b_1, b_2, b_3]\).

Let \(r_r = x i + y j + R k\) be the position vector of the ball’s center relative to the stationary frame. The body’s position at any time is given by the vector \(r_r\) and the quaternion \(q\).

We express the ball’s angular velocity \(\omega\) in terms of its components along the body triad as in (11.17), and note that the derivative \(\dot{\omega}\) is given by (7.4) on page 58.

#### 11.5.1 The no-slip condition

The velocity of the point on the ball which is at contact with the horizontal plane is \(\dot{r}_r + \omega \times (-Rk)\), which should be zero if there is to be no slippage at the contact point. Therefore we have \(\dot{r}_r = R\omega \times k\), that is, \(\dot{x} i + \dot{y} j = R\omega \times k\). It follows that \(\dot{x} = R\omega \times k \cdot i = Rk \cdot i \cdot \omega = Rj \cdot \omega\). Similarly, \(\dot{y} = -Ri \cdot \omega\). The conditions

\[
\dot{x} = Rj \cdot \omega \quad \text{and} \quad \dot{y} = -Ri \cdot \omega
\] (11.20)

are this problem’s nonholonomic constraints. We will use these to eliminate \(\dot{x}\) and \(\dot{y}\) in favor of \(\omega\).
Additionally, we use these to eliminate the second derivatives, $\ddot{x}$ and $\ddot{y}$ from the Gibbs function. Let us make a note here:

$$\ddot{x} = R \dot{y} \cdot \dot{\omega}, \quad \ddot{y} = -R \dot{x} \cdot \dot{\omega}. \quad (11.21)$$

### 11.5.2 The Gibbs function and the equations of motion

The Gibbs function of the system is the sum of the Gibbs functions corresponding to the homogeneous ball and to the slug. We calculate them separately, and then add them up.

The Gibbs function of the homogeneous ball is

$$\mathfrak{G}_1 = \frac{1}{2} \left( M (\dddot{x}^2 + \dddot{y}^2) + \frac{2}{5} MR^2 (\dot{\omega}_1^2 + \dot{\omega}_2^2 + \dot{\omega}_3^2) \right).$$

Let us note that due to (11.21) we have

$$||\ddot{r}_c||^2 = \dddot{x}^2 + \dddot{y}^2 = R^2 \left( (\dot{y} \cdot \dot{\omega})^2 + (\dot{x} \cdot \dot{\omega})^2 \right) = R^2 \left( \dot{\omega}_1^2 + \dot{\omega}_2^2 + \dot{\omega}_3^2 - (k \cdot \dot{\omega})^2 \right), \quad (11.22)$$

therefore

$$\mathfrak{G}_1 = \frac{1}{2} MR^2 \left( \frac{7}{5} (\dot{\omega}_1^2 + \dot{\omega}_2^2 + \dot{\omega}_3^2) - (k \cdot \dot{\omega})^2 \right).$$

To compute the slug’s Gibbs function, let us write $r$ for its position vector. We have:

$$r = r_c + a b_3.$$ 

Then we calculate the slug’s velocity

$$\dot{r} = \dot{r}_c + a \dot{b}_3 = \dot{r}_c + a \dot{\omega} \times b_3,$$

and its acceleration

$$\ddot{r} = \ddot{r}_c + a (\dot{\omega} \times b_3 + \dot{\omega} \times \dot{b}_3)
\quad = \ddot{r}_c + a (\dot{\omega} \times b_3 + \dot{\omega} \times (\dot{\omega} \times b_3))
\quad = \ddot{r}_c + a (\dot{\omega} \times b_3 + (\omega \times b_3) \dot{\omega} - (\omega \times \omega) b_3).$$

Noting that $\omega \cdot b_3 = \omega_3$ and

$$\dot{\omega} \times b_3 = (\dot{\omega}_1 b_1 + \dot{\omega}_2 b_2 + \dot{\omega}_3 b_3) \times b_3 = -\omega_1 b_2 + \omega_2 b_1,$$

the acceleration simplifies to

$$\ddot{r} = \ddot{r}_c + a \left( -\omega_1 b_2 + \omega_2 b_1 + \omega_3 \omega - ||\omega||^2 b_3 \right)
\quad = \ddot{r}_c + a \left( -\omega_1 b_2 + \omega_2 b_1 + \omega_3 (\omega_1 b_1 + \omega_2 b_2 + \omega_3 b_3) - (\omega_1^2 + \omega_2^2 + \omega_3^2) b_3 \right),$$

which we group as

$$\ddot{r} = \ddot{r}_c + a \left( (\omega_2 + \omega_3 \omega_1) b_1 - (\omega_1 - \omega_2 \omega_3) b_2 - (\omega_1^2 + \omega_2^2) b_3 \right), \quad (11.23)$$

then for $\ddot{r}_c$ substitute from (11.22), and finally we compute

$$\mathfrak{G}_2 = \frac{1}{2} m ||\ddot{r}||^2 - (mg \cdot k) \cdot \ddot{r}.$$
The fully expanded expression of $\mathbf{G}_2$ is horrendously large. For all practical purposes it is impossible to compute it with bare hands. Doing it with a symbolic computational software, such as MAPLE or MATHEMATICA, however, is not a problem.

The composite ball’s Gibbs function os $\mathbf{G} = \mathbf{G}_1 + \mathbf{G}_2$. The equations of motion are

$$\frac{\partial \mathbf{G}}{\partial \omega_1} = 0, \quad \frac{\partial \mathbf{G}}{\partial \omega_2} = 0, \quad \frac{\partial \mathbf{G}}{\partial \omega_3} = 0,$$

These along with the two equations (11.20) and the four equations (11.19) form a system of nine first order differential equations in the nine unknowns $x$, $y$, $\omega_1$, $\omega_2$, $\omega_3$, $q_0$, $q_1$, $q_2$, $q_3$.

The initial conditions for the first five have immediate physical meanings and their specifications are intuitive and easy. The $x$ and $y$ place the ball’s center at a desired location, and $\omega_1$, $\omega_2$, $\omega_3$ impart it an initial velocity. (Remember that are the components of the angular velocity in the body triad. The ball’s initial orientation is specified through $q_0$, $q_1$, $q_2$, $q_3$ as follows.

Begin with the ball oriented so that the body triad $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ is lined up with the stationary $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ triad. Then apply a rotation $A_0$ about a unit vector $\mathbf{n}$ and rotation angle $p$ to orient the ball as desired. Any orientation in space may be achieved through the appropriate choices of $\mathbf{n}$ and $p$. The quaternion associated with that rotation is given in (11.9). Therefore

$$q_0(0) = \cos \frac{p}{2}, \quad q_1(0) = n_1 \sin \frac{p}{2}, \quad q_2(0) = n_2 \sin \frac{p}{2}, \quad q_3(0) = n_3 \sin \frac{p}{2}. \quad (11.24)$$

Don’t forget that $\mathbf{n} = n_1 \mathbf{i} + n_2 \mathbf{j} + n_3 \mathbf{k}$ must be a unit vector.

Remark 11.5. To make an animation of the rolling ball, it is suffices to design the ball graphics in a reference position, say where the body axes is aligned with the stationary axes. Then, after having solved the differential equations and computed $x(t)$, $y(t)$, $q_0(t)$, $q_1(t)$, $q_2(t)$, $q_3(t)$, rotate the reference ball about the vector $q_1(t) \mathbf{i} + q_2(t) \mathbf{j} + q_3(t) \mathbf{k}$ by the angle $p(t)$ obtained from $q_0(t) = \cos \frac{p(t)}{2}$, that is, $p(t) = 2 \cos^{-1} q_0(t)$. Then translate the rotated ball’s center to the location $x(t) \mathbf{i} + y(t) \mathbf{j} + R \mathbf{k}$. MAPLE’s plottools package provides the commands rotate() and translate() for rotating and translating graphics objects. Figure 11.3 shows the ball in its reference position (on the left), and a frame from an animation sequence (on the right).

---

**Exercises**

11.1. Prove the identity (11.6).

11.2. In Figure 11.1, $\mathbf{u}$ is a unit vector to the plane $P$, and $\mathbf{u}$ is an arbitrary vector. Why is $\mathbf{u} - 2(\mathbf{n} \cdot \mathbf{u}) \mathbf{n}$ the mirror reflection of $\mathbf{u}$ through the plane $P$?

11.3. Let $\mathbf{q} = n_2 \circ n_1$, where $n_1$ and $n_2$ are any two vectors. Show that $\mathbf{q}^* = n_1 \circ n_2$.

11.4. Verify that the quaternion multiplication is associative, that is

$$\mathbf{p} \circ (\mathbf{q} \circ r) = (\mathbf{p} \circ \mathbf{q}) \circ r.$$

11.5. Show that $(\mathbf{p} \circ \mathbf{q})^* = \mathbf{q}^* \circ \mathbf{p}^*$ for all quaternions $\mathbf{p}$ and $\mathbf{q}$.

11.6. Show that $|\mathbf{p} \circ \mathbf{q}| = |\mathbf{p}| |\mathbf{q}|$ for all quaternions $\mathbf{p}$ and $\mathbf{q}$.
11.7. Show that $\mathbf{n} \circ \mathbf{n} = -1$ for all unit vectors $\mathbf{n}$. Note the parallel with the imaginary numbers: $i^2 = -1$.

11.8. Supply the details that lead to (11.15).

11.9. Produce an animation of the unbalanced ball of Section 11.5.
Appendix A

Maple basics

This appendix provides a very terse summary of the basic features of MAPLE that you will need to solve this book’s problems. For details you should consult MAPLE’s User Manual (about 350 pages) and Programming Guide (about 650 pages) which you may download from

http://www.maplesoft.com/documentation_center/

That page provides documentation for the latest version of MAPLE. You will find the documentation of earlier versions in


I have broken the long URL into two lines so that it fits between the margins of this book. You will enter those two lines as a single line in your browser.

A.1 Configuring Maple

Maple’s user interface is quite customizable. Unfortunately its default settings are less than desirable. In fact, in my opinion they are so awful as to make Maple almost unusable. Fortunately you may adjust those settings by following the simple instructions in:


(Again, I have broken the long URL into two lines to fit). The screenshots there were produced a few years ago, using an early version of MAPLE, therefore they may not correspond exactly to what you will see on your version. Nevertheless, you should have no trouble in interpreting what needs to be done.

A.2 The execution group

The MAPLE prompt which looks like > on the screen, starts what is called an execution group. You may enter one or more MAPLE commands at the execution group, ending each command with a semicolon, then hit Enter to execute those commands at once. Here is an example of a single command in an execution group:
and here is an example of two commands in an execution group:

\[ > \text{expand((x+1)^2)}; \quad \text{int}(\% , x=0..1); \]
\[ x^2 + 2x + 1 \]
\[ \frac{7}{3} \]

The % character seen above, called a \textit{ditto operator}, picks up the result of the previous command.

As we see in the examples shown above, Maple attaches a \textit{label} to the final computed result in execution groups. You may refer to that result through its label. See the description of the Ctrl–L key in the next section.

\section*{A.3 \ Maple key bindings}

\textbf{Ctrl–L}. pops up a dialog window in which you enter the \textit{label} produced in a prior execution group. The label entered here stands for the result of that execution group.

\textbf{Ctrl–K} and \textbf{Ctrl–J} insert a fresh execution group above or below, respectively, of the current execution group.

\textbf{Ctrl–Delete} deletes the current execution group.

\textbf{Shift–Enter} breaks a line, enabling you to format your input, as in

\[ > \text{expand((x^2+1)^2);} \]
\[ 1/\%; \]
\[ \text{int}(\%, x=-1..1); \]
\[ x^4 + 2x^2 + 1 \]
\[ \frac{1}{x^4 + 2x^2 + 1} \]
\[ \frac{1}{2} + \frac{\pi}{4} \]

\section*{A.4 \ Expression sequences, lists, and sets}

An \textit{expression sequence} is a comma-separated collection of MAPLE objects. For instance, each of the following commands produces an expression sequence containing three objects:

\[ > 1, \ 2, \ 3; \]
\[ 1,2,3 \]

The following expression sequence consists of two equations and one fraction:

\[ > x = a + b, \ y = \cos(t), \ 3/4; \]
\[ x = a + b, \ y = \cos(t), \frac{3}{4} \]
A list is an expression sequence enclosed in square brackets. For instance, \([a, b, c]\) is the list associated with the expression sequence \(a, b, c\). To see the distinction between the two, consider a function \(f\) of three variables, and a function \(g\) of one variable. These may be invoked as \(f(a, b, c)\) and \(g(p)\). Note, however, that \(g(a, b, c)\) is illegal, since \(g\) cannot be invoked with more than one argument. On the other hand, \(g([a, b, c])\) is legal, since the brackets encapsulate the expression sequence \(a, b, c\) into a single object, that is, the list \([a, b, c]\).

The order of the objects in a list is essential. The lists \([a, b]\) and \([b, a]\) are not the same. To wit:

\[\text{is}([a, b] = [b, a]); \quad \text{false} \]

A set is an expression sequence enclosed in curly braces. The order of the entries in a set is immaterial. Moreover, in accordance with the mathematical definition of a set, \({b, a, a, b}\) is the same as \({a, b}\):

\[\text{is}({b, a, a, b} = {a, b}); \quad \text{true} \]

Sets may be combined through the union and intersection operations, as in

\[\{ a, b, c \} \cup \{ c, d, e \}; \quad \{a, b, c, d, e\} \]
\[\{ a, b, c \} \cap \{ c, d, e \}; \quad \{c\} \]

To pick element \(i\) of a list or a set \(A\), we use the selection operation \(A[i]\). For instance:

\[A := [ a, b, c ];\]
\[B := \{ a, b, c \};\]
\[A[2]; \quad b \]
\[B[2]; \quad b \]

Specifying an empty bracket in a selection operation produces the expression sequence that makes up the list or the set. Thus, with \(A\) and \(B\) as above, we get:

\[A[]; \quad [a, b, c]\]
\[B[]; \quad [a, b, c]\]

### A.5 Selecting and removing subsets

Consider the set

\[S := \{ a, b, x = a^2 + b^2, y = a^2 - b^2 \}; \quad \{a, b, x = a^2 + b^2, y = a^2 - b^2\} \]

Two elements of \(S\) are equations, the other two are not. Suppose we wish to pick the subset \(S\) consisting of equations. We do

\[\text{select(type, S, ‘equation’)}; \quad \{x = a^2 + b^2, y = a^2 - b^2\} \]
To select the subset of $S$ consisting of everything other than equations, we do

$$> \text{remove(type, S, 'equation')} ;\{a, b\}$$  \hfill (18)

### A.6 Solving equations symbolically

The `solve()` function solves algebraic equations or systems of equations. It attempts to obtain a solution in symbolic form and gives up if no symbolic solution can be found. In its most basic form, `solve()` takes two arguments. The first argument is an equation, or a set, or a list of equations. The second argument is the unknown, or a set, or a list of unknowns. For example:

$$> \text{solve}(a*x^2 + b*x + c = 0, x);$$

$$\frac{1}{2} - \frac{b + \sqrt{-4ac + b^2}}{a}, \quad \frac{1}{2} - \frac{b + \sqrt{-4ac + b^2}}{a}$$  \hfill (19)


**Remark A.1.** If the right-hand side of an equation is zero, then the " = 0" is optional. Thus, the first of two commands entered above may have been entered as

$$> \text{solve}(a*x^2 + b*x + c, x);$$

$$\frac{1}{2} - \frac{b + \sqrt{-4ac + b^2}}{a}, \quad \frac{1}{2} - \frac{b + \sqrt{-4ac + b^2}}{a}$$  \hfill (21)

If the equation is entered by hand, then there is no great savings in omitting the zero. The value of this convention becomes apparent when solving machine-generated equations, or systems of equations, where only the left-hand sides are present, and the zeros of the right-hand sides are implicitly assumed.

### A.7 Solving equations numerically

We are often faced with equations which do not have explicit symbolic solutions. In those cases we turn to numerical approximation of their solutions. For instance, the transcendental equation

$$3(x + 1)e^{-x} - 1 = 0$$

has two solutions, as is evident in its graph:

$$> \text{plot}(3*(x+1)*\exp(-x)-1, x=-1..4, \text{view}=[-1.5..4, -1..2.5]);$$
A.8. The eval() function

The function fsolve() is MAPLE’s numeric equation solver. It is used just like solve():

\[ \text{fsolve}(3(x+1)\times \exp(-x)-1, x); \]
\[ -0.8587727590 \] (22)

We see that it picked one of the two solutions in this case. To pick the other one, we specify a range in which to search for it:

\[ \text{fsolve}(3(x+1)\times \exp(-x)-1, x=2..3); \]
\[ 2.289281415 \] (23)

Remark A.2. In the two invocations of fsolve() above, I have omitted the “= 0 from the equation’s right-hand side, as this is permissible according to Remark A.1 in the previous section.

A.8. The eval() function

The eval() function plays several diverse roles in MAPLE. The one feature of eval() that we need for our purposes is to evaluate an expression with given values. For instance

\[ \text{eval}(x^2 + x + 1, [x = 1]); \]
\[ 3 \] (24)

This is equivalent to the usual mathematical notation \( (x^2 + x + 1)_{x=1} \).

Multiple substitutions are allowed. E.g.:

\[ \text{eval}(ax^2 + bx + c, [a=1, b=2, c=1]); \]
\[ x^2 + 2x + 1 \] (25)

A very common use of eval() is the following. Consider the solution of the linear system given under MAPLE output (20) in Section A.6. To extract the solution’s y value, we do:

\[ Y := \text{eval}(y, (20)); \]
\[ Y := \frac{a_1c_2 - a_2c_1}{a_1b_2 - a_2b_1} \] (26)

Remark A.3. The second argument of eval() may be given as a list \([\ldots]\) or as a set \(\{\ldots\}\). They have similar effects, but in the case of a list, the arguments are substituted in order, from left to right, while in the case of a set, the order of substitution may vary.

A.9. Expressions and functions

In MAPLE, as in customary mathematics, we distinguish between a function and the result of the application of that function. In calculus we write, for instance, \( f : \mathbb{R} \to \mathbb{R} \) indicates a function \( f \), while \( f(x) \) indicates the result of applying \( f \) to \( x \). In casual mathematical discussion we mix the two concepts, by referring, for instance, to “the function \( x^2 \)” which is a short way of saying “the function whose application to \( x \) produces \( x^2 \).” A computer algebra system such as MAPLE does not permit such a loose talk. As far as MAPLE is concerned, if \( f : \mathbb{R} \to \mathbb{R} \) defined through \( f(x) = x^2 \), then \( f \) is a function while \( x^2 \) is an expression, and the two shall not to be confused.

In MAPLE a function is defined like this:
> f := x -> x^2;
\[ f := x \rightarrow x^2 \] (27)

Then it may be used as

> f(5);
25 (28)

> f(a+b);
\[(a + b)^2\] (29)

### A.10 Vectors and Matrices

A row vector with entries \(a, b, c\) is produced by

> < a | b | c >;
\[
\begin{bmatrix}
a \\
b \\
c
\end{bmatrix}
\] (30)

while a column vector is produced by

> < a, b, c >;
\[
\begin{bmatrix}
a \\
b \\
c
\end{bmatrix}
\] (31)

A matrix may be constructed as column vector whose entries are row vectors, as in

> < < a[1,1] | a[1,2] | a[1,3] >, 
\[
\begin{bmatrix}
a_{11} & a_{12} & a_{13} \\
a_{21} & a_{22} & a_{23}
\end{bmatrix}
\] (32)

Equivalently, that matrix may be constructed as row vector whose entries are column vectors, as in

> < < a[1,1], a[2,1] > | < a[1,2], a[2,2] > |
| < a[1,3], a[2,3] > >;
\[
\begin{bmatrix}
a_{11} & a_{12} & a_{13} \\
a_{21} & a_{22} & a_{23}
\end{bmatrix}
\] (33)

The transpose of a vector or a matrix is constructed through the \(\%T\) superscript, as in

> u := < a | b >;
\[
\begin{bmatrix}
a \\
b
\end{bmatrix}
\] (34)

> u^\%T;
\[
\begin{bmatrix}
a \\
b
\end{bmatrix}
\] (35)

Multiplication between two vectors, or two matrices, or a vector and a matrix, is done through the dot operator. For instance,

> A := < < a[1,1] | a[1,2] >, < a[2,1] | a[2,2] > >;
\[
\begin{bmatrix}
a_{11} & a_{12} \\
a_{21} & a_{22}
\end{bmatrix}
\] (36)

\[
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix}
\] (37)
Example A.1. The position vector of a particle is a vector that extends from a point of reference to the particle. The motion of the particle is tracked by a time-dependent position vector. For instance, the position of a simple pendulum’s bob may be given as

\[
r := \phi \rightarrow \langle \sin(\phi), \cos(\phi) \rangle;
\]

(40)

where \( \phi \) is the angle that the pendulum makes with the downward pointing vertical. The bob’s position at any time is given by

\[
r(\phi(t)) = \begin{bmatrix}
\sin(\phi(t)) \\
\cos(\phi(t))
\end{bmatrix}
\]

(41)

A.11 Differentiation

Differentiation, ordinary or partial, is performed through the \texttt{diff()} function:

\[
> \text{diff}(a*x^2 + b*x + c, x);
\]

\[
2ax + b
\]

(42)

\[
> \text{diff}(x^2 + y^2, x);
\]

\[
2x
\]

(43)

Higher order derivatives are produced by extending \texttt{diff()}’s arguments, as seen below:

\[
> \text{diff}(x^7, x, x);
\]

\[
42x^5
\]

(44)

\[
> \text{diff}(x^7, x, x, x);
\]

\[
210x^4
\]

(45)

\[
> \text{diff}(g(x,y), x, x);
\]

\[
\frac{\partial^2}{\partial x^2} g(x, y)
\]

(46)

\[
> \text{diff}(g(x,y), x, y);
\]

\[
\frac{\partial^2}{\partial y \partial x} g(x, y)
\]

(47)

The \( D \) operator offers an alternative syntax for computing derivatives. For instance, if \( f(x) = x^2 \), then

\[
> D(f)(x);
\]

\[
2x
\]

(48)

\[
> D(f)(5);
\]

\[
10
\]

(49)

A frequent use of the \( D \) operator is in specifying initial conditions for differential equations of second order and higher. For instance, the initial condition \( \dot{x}(0) = a \) is entered as \( D(x)(0) = a \). See subsection A.12.1 for an example.
Remark A.4. The \texttt{diff()} function is designed to differentiate scalar quantities. To differentiate a vector, we use \texttt{diff~()}. For instance, the velocity vector of Example A.1’s position vector \( r(\phi(t)) \) is obtained by

\begin{verbatim}
> v := diff~(r(phi(t)), t);
\end{verbatim}

\[
\begin{bmatrix}
\cos(\phi(t)) \left( \frac{d}{dt} \phi(t) \right) \\
-\sin(\phi(t)) \left( \frac{d}{dt} \phi(t) \right)
\end{bmatrix}
\] (50)

Remark A.5. If the particle’s mass in the previous remark is \( m \), then its kinetic energy is \( \frac{1}{2} m \|v\|^2 \):

\begin{verbatim}
> 1/2 * m * v^%T . v;
\end{verbatim}

\[
\frac{1}{2} m \cos(\phi(t))^2 \left( \frac{d}{dt} \phi(t) \right)^2 + \frac{1}{2} m \sin(\phi(t))^2 \left( \frac{d}{dt} \phi(t) \right)^2
\] (51)

The result may be simplified through the application of MAPLE’s \texttt{simplify()} command:

\begin{verbatim}
> simplify((51));
\end{verbatim}

\[
\frac{1}{2} m \left( \frac{d}{dt} \phi(t) \right)^2
\] (52)

\section*{A.12 Solving differential equations}

The \texttt{dsolve()} function solves differential equations or systems of differential equations. The equations may be linear or nonlinear. By default, \texttt{dsolve()} attempts a symbolic solution. Numeric solutions may be obtained by supplying extra arguments.

\subsection*{A.12.1 Solving differential equations symbolically}

Here are a couple of example of \texttt{dsolve()} in symbolic mode:

\begin{verbatim}
> de1 := diff(y(t),t) + a*y(t) = 0;
\end{verbatim}

\[
de1 := \frac{d}{dt} y(t) + ay(t) = 0
\] (53)

\begin{verbatim}
> dsolve(de1, y(t));
\end{verbatim}

\[
y(t) = _C1 e^{-at}
\] (54)

\begin{verbatim}
> de2 := diff(y(t),t,t) + omega^2*y(t) = 0;
\end{verbatim}

\[
de2 := \frac{d^2}{dt^2} y(t) + \omega^2 y(t) = 0
\] (55)

\begin{verbatim}
> dsolve(de2, y(t));
\end{verbatim}

\[
y(t) = _C1 \sin(\omega t) + _C2 \cos(\omega t)
\] (56)

\begin{verbatim}
> de Sys := diff(x(t),t) + x(t) = y(t),
\end{verbatim}

\[
diff(y(t),t) + y(t) = x(t);
\]

\[
de_{sys} := \frac{d}{dt} x(t) + x(t) = y(t), \quad \frac{d}{dt} y(t) + y(t) = x(t)
\] (57)

\begin{verbatim}
> dsolve({de Sys}, {x(t),y(t)});
\end{verbatim}

\[
x(t) = _C1 + _C2 e^{-2t}, \quad y(t) = _C1 - _C2 e^{-2t}
\] (58)
Arbitrary constants that enter general solutions of differential equations appear as \( C_1, C_2, \ldots \). Underscores guard against a potential conflict between user-defined versus MAPLE-supplied symbols.

To solve initial value problems, include the initial values in \texttt{dsolve()}’s first argument:

\[
\begin{align*}
> & \texttt{dsolve(\{de1, y(0)=b\}, y(t));} \\
& y(t) = b e^{-at} \quad (59) \\
> & \texttt{dsolve(\{de2, y(0)=a, D(y)(0)=b\}, y(t));} \\
& y(t) = \frac{b \sin(\omega t)}{\omega} + a \cos(\omega t) \quad (60) \\
> & \texttt{dsolve(\{de\_sys, x(0)=5, y(0)=7\}, \{x(t), y(t)\});} \\
& x(t) = 6 - e^{-2t}, \quad y(t) = 6 + e^{-2t} \quad (61)
\end{align*}
\]

### A.12.2 Solving differential equations numerically

MAPLE offers a plethora of methods for calculating and displaying numerical solution of differential equations. Here I present the one method which is most useful for our purposes, through solving the differential equation, and plotting the solution of the Van der Pol oscillator. Note the use of \texttt{eval()} from Section A.8 to extract the parts \( x(t) \) and \( \frac{dx}{dt} \) from the list returned by \texttt{dsolve()}.

\[
\begin{align*}
> & \texttt{vdp := diff(x(t),t,t) - (1-x(t)^2)*diff(x(t),t) + x(t) = 0;} \\
& \frac{d^2}{dt^2} x(t) - (1-x(t)^2) \frac{d}{dt} x(t) + x(t) = 0 \quad (62) \\
> & \texttt{tmax := 10;} \\
& tmax := 10 \quad (63) \\
> & \texttt{dsol := dsolve(\{vdp, x(0)=2, D(x)(0)=-2,} \\
& \texttt{numeric, output=listprocedure, range=0..10);} \\
& \texttt{dsol := [t = proc(t) \ldots endproc, x(t) = proc(t) \ldots endproc,} \\
& \texttt{\frac{d}{dt} x(t) = proc(t) \ldots endproc]} \quad (64) \\
> & \texttt{my\_x := eval(x(t), \texttt{dsol};} \\
& my\_x := \texttt{proc(t) \ldots endproc} \quad (65) \\
> & \texttt{my\_dx := eval(diff(x(t),t), \texttt{dsol};} \\
& my\_dx := \texttt{proc(t) \ldots endproc} \quad (66) \\
> & \texttt{plot(\{my\_x(t), t=0..tmax, labels=[t,x(t)]\);} \\
> & \texttt{plot(\{my\_dx(t), t=0..tmax, labels=[t,dx(t)/dt]\);} \\
& \texttt{plot(\{\{\{my\_x(t), my\_dx(t), t=0..tmax, labels=[x, dx(t)/dt]\};}
\end{align*}
\]

The \texttt{plot(...)} commands in the final three lines produce the graphs of \( x \) versus \( t \), \( dx/dt \) versus \( t \), and \( dx/dt \) versus \( x \), shown side by side in Figure A.1. The details of MAPLE’s plotting functions are given in Section A.13.
Figure A.1: The graphs of $x$ versus $t$, $dx/dt$ versus $t$, and $dx/dt$ versus $x$, of a solution of the Van der Pol oscillator.

A.13.2 • Plotting multiple function together

If `plot()`’s first argument is a list of expressions, then their graphs are plotted all together in one coordinate system, and in varying colors:

```maple
> plot([x^2, 1-x^2, 0.4*exp(x)], x=-1..1, labels=[x,y]);
```

A.13.3 • Parametric plot

A parametric plot, such as \((x(t), y(t))\), \(a \leq t \leq b\) is specified as \([x(t), y(t), t=a..b]\). The following example plots a cycloid:

```maple
> plot([t - sin(t), 1 - cos(t), t=0..2*Pi],
     scaling=constrained, labels=[x,y], color=red,
     tickmarks=[piticks,default], view=[0..7, 0..2.5]);
```

The `scaling=constrained` option seen above, requests a graph with a 1:1 aspect ratio, that is, the unit lengths are equal along the horizontal and vertical axes. Addi-
tionally, the command above shows the effects of the (self-explanatory) tickmarks and view options.

A.13.4 • Plotting points and more

More specialized plotting constructs are delegated to the plots package which you will load with \texttt{with(plots)} in order to access them. The \texttt{pointplot()} function in that package is designed to plot points on the plane, and optionally, connect them with straight lines. Here is an example of its usage, along with a few options which modify its behavior:

\begin{verbatim}
> with(plots):
> pointplot([[0,0], [1,3], [2,1], [3,2], [4,2]],
  symbol=solidcircle, symbolsize=30, color=blue,
  labels=[x,y]);
\end{verbatim}

\begin{verbatim}
> pointplot([[0,0], [1,3], [2,1], [3,2], [4,2]],
  symbol=solidcircle, symbolsize=30, color=red,
  labels=[x,y], connect=true);
\end{verbatim}

The last execution group consists of two commands. The first produces the plot. The second assigns the name \texttt{p1} to that plot. We may access the plot later on in the worksheet by referring to its name. If you don’t expect to be referring to the plot, then there is no need to name it.

The \texttt{connect=true} option connects the consecutive points with straight line segments, and thus produces a piecewise linear curve. The point symbols themselves are not plotted in that case.

\begin{verbatim}
> pointplot([[0,0], [1,3], [2,1], [3,2], [4,2]],
  symbol=solidcircle, symbolsize=30, color=red,
  labels=[x,y], connect=true);
\end{verbatim}

\begin{verbatim}
> pointplot([[0,0], [1,3], [2,1], [3,2], [4,2]],
  symbol=solidcircle, symbolsize=30, color=blue,
  labels=[x,y]);
\end{verbatim}

A.13.5 • Overlaying multiple plots

To superpose the two or more graphs, use the \texttt{display()} function, which is defined in the \texttt{plots} package. For instance, the plots \texttt{p1} and \texttt{p2} produced above may be superposed.
A.13.6 Reflecting a plot

The plottools package provides several additional plotting tools. Among these are functions to plot ellipses, polygons, as well as three-dimensional solids. It also provides functions for transforming plots, such as scaling, translating, rotating, and reflecting.

To illustrate some of the plottools functionality, let us make an ellipse:

```maple
> myellipse := ellipse([0,2],1,2, filled=true, color=pink):
```

This produces a plot structure of an ellipse; it does not plot the ellipse by itself. To see the ellipse, apply display() to the plot structure:

```maple
> display(myellipse, scaling=constrained);
```

Now, suppose that we wish to reflect the ellipse about the line that connects the points (0,5) and (3,0). The following sequence of commands plots the points, the line that connects them, the ellipse, the reflected ellipse, and then applies display() to display all four objects together in one diagram:

```maple
> display([
    pointplot([[0,5], [3,0]], symbol=solidcircle, symbolsize=20, color=blue),
    pointplot([[0,5], [3,0]], connect=true, color=gray),
    myellipse,
    reflect(myellipse, [[0,5], [3,0]])
], scaling=constrained);
```
A.14. The Euler–Lagrange equations

MAPLE’s EulerLagrange() function, which is defined in the VariationalCalculus package, produces the Euler–Lagrange equations the functional

\[ J(x) = \int_a^b L(t, x(t), x'(t)) \, dt \]

subject to \( x(a) = A \) and \( x(b) = B \). Here \( x \) may be a scalar- or vector-valued function. Here is how it is done. First, we load the package:

\[ > \text{with(VariationalCalculus)}: \]

If \( x \) is scalar-valued, the corresponding Euler–Lagrange equation is obtained through

\[ > \text{EulerLagrange}(L, t, x(t)); \]

If \( x \) is vector-valued, with components \( x_1(t), x_2(t), \ldots, x_n(t) \), we do

\[ > \text{EulerLagrange}(L, t, [x[1](t), x[2](t), \ldots, x[n](t)]); \]

In either case, EulerLagrange() returns a set consisting of

- \( n \) second order differential equations in the \( n \) unknowns \( x_1(t), x_2(t), \ldots, x_n(t) \).
- zero or more first order differential equations of the type

\[ F_i(x_1(t), x_2(t), \ldots, x_n(t), \dot{x}_1(t), \dot{x}_2(t), \ldots, \dot{x}_n(t)), \quad i = 1, \ldots, \]

each of which represents a first integral of the problem, and where \( K_i \) are arbitrary constants. A first integral is the result of integrating, if possible, a combination of the previously noted second order differential equations.

In our application we won’t be interested in the first integrals, therefore we remove them from the set by applying the remove() function described in Section A.5, and thus leaving only the second order differential equations.

A.15. The animation of a simple pendulum

In this section we will present a complete and self-contained MAPLE worksheet in which we derive the equation of motion of a simple pendulum through the Lagrangian formulation, solve the resulting differential equation, and produce an animation that depicts the pendulum’s motion. All MAPLE functions used in this worksheet have been described earlier in this chapter.
The worksheet begins in Listing A.1 and continues into Listing A.2. The purpose of the restart command which appears at the top of the worksheet is to unassign all assigned variables and free the memory that MAPLE has allocated for its work. The command has no effect in a freshly started worksheet but its inclusion at the top of a worksheet is a good practice. As you experiment with a worksheet’s contents, you may find it useful to execute restart to reset MAPLE to its initial state and then start a new set of experiments. Without the restart you will have to exit and restart MAPLE to achieve the same effect.

Next we load three packages which supply some of the special-purpose functions used in the rest of the worksheet. Specifically, the VariationalCalculus package supplies the EulerLagrange() function; the plots package supplies the pointplot() and display() functions; and the plottools package supplies the reflect() functions.

Now I will continue with commenting on the rest of the worksheet. A “Line numbers” in this context refers to the numerical labels produced by MAPLE which appear along the right edge of the listing.

Line (1). We define the position vector \( \mathbf{r}(\varphi) \) of the pendulum’s bob as a function of the deflection angle \( \varphi \). The components are in accordance with the Cartesian coordinate system shown in Figure 1.1. In the worksheet I have used “\( a \)” for the length of the pendulum’s rod instead of “\( \ell \)” shown in the figure because the symbol “\( \ell \)” is not available in a MAPLE worksheet.

Lines (2)–(5). We compute the velocity vector \( \mathbf{v} \), the kinetic energy \( T \), the potential energy \( V \), and finally, the Lagrangian \( L \). Note the “[2]” in computing \( V \); it extracts the second, that is the \( y \), component of the position vector \( \mathbf{r} \).

Lines (6)–(8). We apply the EulerLagrange() function to produce the differential equations of motion. The result is a set consisting of a differential equation, and a first integral. (See section A.14 for explanation.) We apply remove() (see Section A.5) to remove the first integral, and thus are left with a set consisting of a single equation. Finally, we apply the empty selector [ ] to the set (see Section A.4) to produce its contents, and assign it to the variable de.

Lines (9) and (10). We select a set of numerical values for the pendulum’s parameters, substitute them in the differential equation, and name the result myde.

Lines (11) and (12). We define the motion’s initial conditions \( \varphi(0) \) and \( \dot{\varphi}(0) \), and a variable \( t_{\text{max}} \) which sets the upper bound on the time range of interest.

Lines (13) and (14). We apply dsolve() to solve (numerically) the differential equation myde along with the initial conditions ic. Then we extract the \( \varphi(t) \) component of the solution and name it myphi.

Continuing on to listing A.2, we plot the function \( \text{myphi}(t) \) which was computed in the preceding listing. As expected, the graph periodic oscillations of the angle.

Next, on Line 15, we evaluate the vector \( \mathbf{r}(\varphi) \) with \( \varphi \) set to \( \text{myphi}(t) \). Applying the eval() is necessary; otherwise the result will contain the undefined symbols \( g \) and \( a \) in it.

The next longish block of code beginning with frames := ... produces a sequence of frames which when played quickly in succession, will produce and animation of the pendulums motion. The definition of frames may look somewhat complex in
Listing A.1: A complete MAPLE session, demonstrating the derivation of the equation of a simple pendulum, and its numerical solution. The final command produces a plot (not shown) of the pendulum’s angle $\phi(t)$ versus the time $t$.

```maple
> restart;
> with(VariationalCalculus): with(plots): with(plottools):
> r := phi -> < a*sin(phi), a*cos(phi) >;
> v := diff~(r(phi(t)), t);
> T := simplify( 1/2 * m * v^%T . v );
> V := -m*g*r(phi(t))[2];
> L := T - V;
> EulerLagrange(L, t, phi(t));
> remove(type, (6), 'equation');
> de := (7)[];
> params := [ m = 1, a = 1, g = 1 ];
> myde := eval(de, params);
> ic := phi(0)=Pi/6, D(phi)(0)=0;
> tmax := 30;
> dsol := dsolve({myde, ic}, numeric, output=listprocedure, range=0..tmax);
> myphi := eval(phi(t), dsol);
```

...continued in Listing A.2...
Listing A.2: Continued from Listing A.1.

```maple
> plot(myphi(t), t=0..tmax, tickmarks=[default,piticks],
   view=[0..tmax,-Pi/4..Pi/4], labels=[t,phi(t)]);

> myr := eval(r(myphi(t)), params);

> frames := 
   seq(
      display(
         [ 
            pointplot([<0,0>, myr], connect=true, thickness=2),
            pointplot(<0,0>, symbol=soliddiamond, symbolsize=30, 
              color=black),
            pointplot(myr, symbol=solidcircle, symbolsize=50, 
              color=red)
         ]),
      t=0..tmax, tmax/250):

> reflect(
   display([frames], insequence=true, scaling=constrained,
    tickmarks=[0,0]),
   [[0,0],[1,0]]);
```

First encounter, but fortunately it is not as complex as it looks. What it really says is that `frames` consists of a sequence of 251 plots:

```maple
frames := seq( p(t), t=0..tmax, tmax/250):
```

where each `p(t)` represents a drawing of the pendulum at some time `t`. Each plot `p(t)` is a composite, consisting of three subplots overlaid with the help of the `display()` command; see subsection A.13.5. The three subplots consist of (a) the pendulum’s “rod” drawn with `pointplot()` `connect=true`; (b) a black diamond placed at the origin
A.15. The animation of a simple pendulum

as an indicator of the pendulum’s support; and (c) a red solid circle that represents the pendulum’s bob.

The final command, beginning with reflect(... plays the animation. To understand what it does, ignore the reflect part for the moment and look at its argument, which, in principle, says

\[
\text{display([frames], insequence=true, other options);}
\]

This displays the 251 individual frames stored in frames, one at a time in quick succession, and thus “plays” a movie.

The purpose of applying the reflect() function to frames is to flip the plots relative to the horizontal axis. This is because our pendulum model is derived in a Cartesian coordinate system where the vertical axis points downward; see Figure 1.1. MAPLE, however, is not aware of this, so without the reflect() it will show the vertical axis in the standard way, that is pointing upward, and consequently the animation will come out inverted. See subsection A.13.6 for a description of MAPLE’s reflect() function.

Remark A.6. Upon executing the reflect(display[...] command, MAPLE will display the first frame of the animation. To set the animation into motion, click on the displayed frame. You will observe that animation control buttons, similar to

![Animation control buttons](image)

appears in the worksheet’s toolbar area. Experiment with the various parts to find out what they do.