1

Introduction

1.1 Differences or Differentials?

Computers have been used to model physical systems from their earliest days. The behaviour of these models is subtly but profoundly affected by the sequential and discrete nature of all digital computers, in which one step in an algorithm follows another, just as a charged particle follows one turn after another around a circular accelerator. Numerical tools – computers – make dynamical problems look like difference equations. Analytic tools – most pre-computer mathematics – make problems look like differential equations. The best choice of representation depends on the nature of the particular system being modelled, and on the questions that are to be answered.

Figure 1.1 illustrates a simple nonlinear system – a gravity pendulum with unit length – that is inherently continuous in time, and which is accurately represented by the differential equation

$$\theta'' = -g \ \sin(\theta) \tag{1.1}$$

where a prime indicates differentiation with respect to time, and g is the acceleration due to gravity. Numerically, this motion is simulated by breaking the motion into two steps that are iterated repeatedly. In the first step, the pendulum 'drifts' with a constant angular velocity θ' for a time Δt. In the second step, gravity is represented by a 'kick' impulse that instantaneously changes the angular velocity. This is written in pseudo-code as

$$\begin{aligned} &\text{until finished } \{ \\ &\qquad \theta = \theta + \theta' \,.\, \Delta t \\ &\qquad \theta' = \theta' - g \,\sin(\theta) \,.\, \Delta t \\ &\} \end{aligned} \tag{1.2}$$

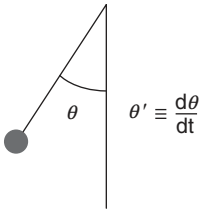

Figure 1.1 Phase space co-ordinates (θ, θ') for a unit length gravity pendulum.

Clearly, the differential and difference representations of a pendulum become identical as Δt approaches zero. Less clear is the potential for the numerical model to introduce artefacts into the simulated motion – behaviour that is not demonstrated by actual pendula.

Circular accelerators are inherently discrete in time, given the turn-by-turn nature of the progress of a test particle. It will be seen, in Chapter 4, that the *standard map* motion of Equation 1.2 (with g set equal to 1) also naturally describes the turn-by-turn longitudinal oscillations of a test particle, relative to an ideal particle at the centre of a bunch. Further, the standard map that naturally models many apparently independent nonlinear accelerator scenarios is a *standard* way to generate chaos, a phenomenon not demonstrated by gravity pendula.

Sometimes analytic solutions to differential equations are known. This is the case for the pendulum, a non-trivial nonlinear system (when large angles are considered) that was completely solved well before computers were invented. Analytic solutions to nonlinear *difference* systems are much rarer. Most traditional mathematical methods are implicitly continuous, and break down more or less completely when truly differential systems, like accelerators, are modelled.

Although computers and accelerators proliferated together, in the second half of the twentieth century, it took time to recognise the limitations of traditional continuous methods. For example, the difference map of Equation 1.2 can be forced into a differential representation by writing

$$\theta'' = -\sum_{n=1}^{\infty} \delta(t - n\Delta t)\; g\; \sin(\theta)\; .\; \Delta t \tag{1.3}$$

where the delta function $\delta()$ is non-zero every Δt units of time. The perceived advantage of this kind of representation of accelerators was that it enabled traditional Hamiltonian, Fourier and perturbation analysis tools to be used. It eventually became clear that traditional methods often break down more or less completely,

and that numerical techniques are not just more convenient, but are also more fundamentally correct. Often, particles must be numerically tracked around a circular accelerator, turn-by-turn.

It is often fast and easy to model a physical system by writing a short program containing an approximate numerical model, even if the system is inherently continuous. Two deceptively simple real-world systems – a tape drive and a betatron – are discussed in Exercises 2.8 and 7.5.

1.2 Phase Space Co-ordinates

Time advances discretely in a circular accelerator, labelled by the integer turn number n.

Figure 1.2 illustrates the observation of the horizontal displacement x and the tangent $x' = dx/ds$ of a test particle as it passes a reference point at $s = s_0$ on the circumference of an accelerator. Here $0 \le s < C$ measures the location of objects – magnets and beam diagnostics, as well as test particles – projected onto the 'design orbit' that is followed by an idealised test particle as it circulates. Usually, but not always, the design orbit goes down the centre of a perfectly aligned beampipe. In general the angle of the test particle is small

$$x' = \tan(\psi) \;\approx\; \psi \;\sim\; 10^{-4} \tag{1.4}$$

and there is little practical distinction between x' and the horizontal angle ψ with respect to the beampipe centre line.

Figure 1.3 illustrates the first four observations of the horizontal phase space motion at the reference point $s = s_0$, in the infinite series $(x, x')_n$; $n = 0, 1, \ldots, \infty$.

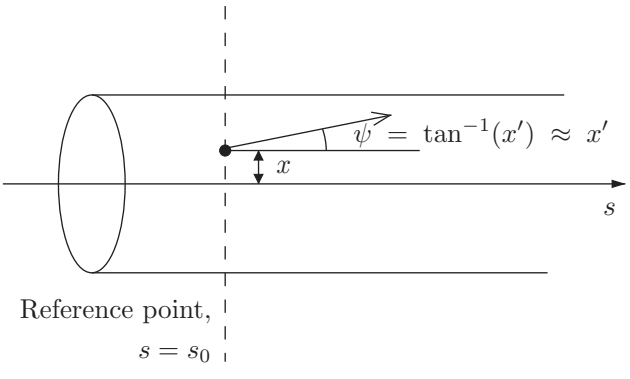

Figure 1.2 Horizontal phase space co-ordinates (x, x') for a test particle moving through a beampipe, and repeatedly passing a reference point at $s = s_0$, turn-by-turn.

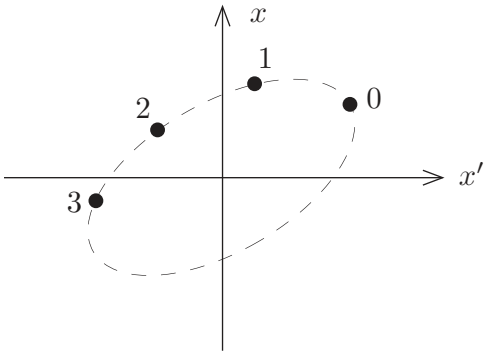

Figure 1.3 Motion of a test particle in horizontal phase space $(x, x')_n$ as it advances for 3 turns (with $n = 0$, 1, 2 and 3) around an accelerator. This Poincaré surface of section is observed at a reference point at $s = s_0$.

In this idealised case the co-ordinate pairs $(x, x')_n$ all lie on an ellipse in phase space, corresponding (as will be seen) to an ideal situation that is completely linear and stable. In practice, the phase space motion observed in this *Poincaré surface of section* is at least a little nonlinear, and the ellipse is at least a little distorted. If conditions are bad enough, the test particle collides with the beampipe wall or some other obstacle when $|x| > x_{max}$, at a time n that falls far short of infinity.

This raises a very practical question: under what circumstances do test particles remain within the beampipe for a sufficiently long time? A particle moving close to the speed of light c in an accelerator with a circumference C of 1 kilometre has a revolution frequency $f_{rev} \approx c/C$ of about 0.3 MHz. If this beam is to be stored for about a day, then we are interested in stability over times as long as $n \approx 3 \times 10^{10}$ turns, a timescale comparable to the age of the solar system, measured in periods (years) of the earth's rotation. In fact, the nonlinear dynamics of solar system and test particle stability are quite closely related.

Phase space descriptions of motion are also convenient when time is continuous. Figure 1.4 illustrates the motion of a gravity pendulum in (θ, θ') phase space, for the unit length gravity pendulum shown in Figure 1.1, and modelled by the standard map of Equation 1.2. The pendulum angle is constrained to lie in the range $-\pi \leq \theta < \pi$ by subtracting or adding 2π if it winds forwards (or backwards) beyond $\theta = \pm\pi$. This graphic was generated by reducing the time step sufficiently far in the limit $\Delta t \to 0$ that neighbouring dots appear to overlap. The more interesting and relevant case when Δt has finite values that correspond to accelerator experience is discussed in Chapter 4.

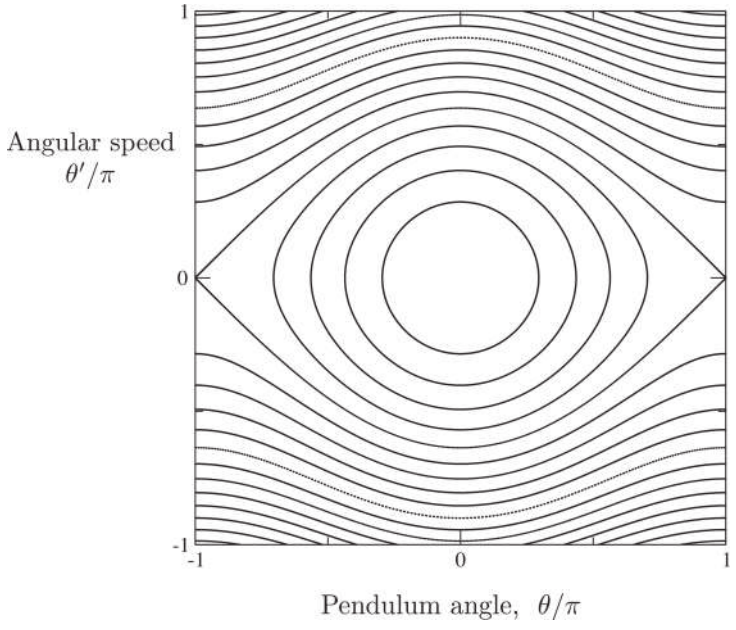

Figure 1.4 Motion of a unit length gravity pendulum (with $g = 1$) in (θ, θ') phase space. The pendulum angle θ is reduced to always lie in the range from $-\pi$ to $+\pi$, while the angular speed θ' is unconstrained: the pendulum can wind forwards or backwards.

1.3 Iterations, Ancient and Modern

Although computers are relatively recent, iterative algorithms have been around for a long time. One well-known example is the Newton search method for solving the equation $f(x) = 0$

$$x = \text{guess} \tag{1.5}$$

until converged {

$$x = x \; - \; \frac{f(x)}{f'(x)}$$

}

where $f'(x) = df/dx$ is the differential of the general function f. Algorithms that are specific examples of the general method precede Newton himself. For example, for centuries before the arrival of tables of logarithms and calculators, the square root of y was found as follows

$$x = 1 \tag{1.6}$$

until converged {

$$x = \left(x + \frac{y}{x}\right) / 2$$

}

This algorithm converges remarkably rapidly – the square root of 10 is correct to 10 decimal places after only six iterations – and it continues to be used in computerised mathematical library functions. Good algorithms die hard.

The initial guess in a Newton search must be close enough to the right answer for the solution to converge. It is remarkable that the initial guess in the following algorithm for inverting the matrix Y guarantees convergence [2, 37, 51].

$$X = \tilde{Y} / \left(\max_i \sum_j |A_{ij}| \; . \; \max_j \sum_i |A_{ij}| \right)$$

until converged {

$$X = X + X(I - XY) \tag{1.7}$$

}

Here $\tilde{Y}$ is the transpose of Y, and I is the identity matrix. The most difficult requirement of this algorithm is that the user be able to multiply matrices. A program using this algorithm to invert matrices is not the fastest, but it is surely one of the easiest to understand and construct. These are two essential elements of good software engineering.

In solving a problem numerically there is often a choice between a more or less direct transcription of an analytic method (for example, for matrix inversion), and a less traditional iterative method. An iterative approach is often easier to understand, and is more concise and flexible. This is true for solar system and accelerator dynamics, as well as for computational algorithms.

1.4 Accelerator History: The Two Golden Ages

The first golden age of accelerator construction and innovation began after the second world war, almost exclusively in support of particle physics experimentation. In those days an accelerator – typically a synchrotron – accelerated electrons, protons, or charged ions to relativistic speeds before diverting the beam onto a fixed target that may have been composed mainly of protons (in a liquid hydrogen target), or perhaps of heavier nuclei. Accelerator technology advanced rapidly in pursuit of ever higher centre-of-mass collisions, in order to create new, heavier, particles.

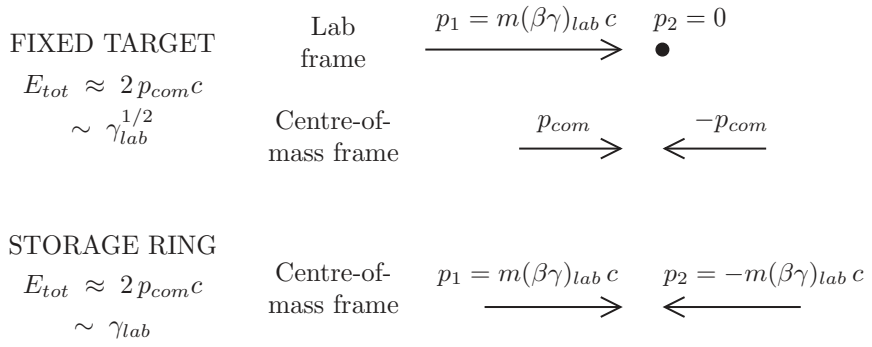

Figure 1.5 Storage rings have a significant advantage over fixed target accelerators, because the total useful centre-of-mass energy E_{tot} scales like γ with the beam energy, in comparison to only $\gamma^{1/2}$ in fixed target accelerators, for large values of γ. The lab frame and the centre-of-mass frame are the same for a (symmetric) storage ring collider.

Storage ring colliders deliver centre-of-mass energies that scale like the relativistic factor γ of the beam, rather than the $\gamma^{1/2}$ scaling inherent to fixed target accelerators. This is illustrated in Figure 1.5. The enhanced efficiency of storage rings comes at the cost of the added complexity of ensuring that two counter-rotating beams pass through each other at (only) a few collision points. Electron storage rings – with intrinsically larger values of γ – began operation in the 1960s, followed by proton, ion and *ep* colliders in the 1980s and beyond, as summarised by Table 1.1.

Electron–positron collisions are experimentally more desirable than electron–electron collisions, because of the cancellation of quantum numbers in the collision. Practical e^+e^- colliders began to operate in the 1970s, while copious supplies of anti-protons only became available (thanks to the invention of stochastic cooling) for proton–antiproton $p\bar{p}$ colliders in the 1980s. Particle–antiparticle collisions are fortunately possible in a single storage ring, thanks to an important feature of the Lorentz equation

$$\vec{F} = q\,(\vec{E} + \vec{v} \times \vec{B}) \tag{1.8}$$

where q is the charge of a particle with a velocity $\vec{v}$ passing through an electric field $\vec{E}$ and magnetic field $\vec{B}$. If there are no electric fields ($\vec{E} = 0$), then the same force $\vec{F}$ is experienced at the same location by counter-rotating particles and anti-particles, under the transformation

$$\begin{aligned} q &\to -q \\ \vec{v} &\to -\vec{v} \end{aligned} \tag{1.9}$$

Table 1.1 *A partial list of storage ring colliders, showing the historical progression from a proliferation of low-energy electron rings to a small number of very high-energy hadron rings (Wikipedia).*

Name	Location	Operating years	Energy [GeV]	Comment
Electron–positron rings				
AdA	Frascati	1961-64	0.25	e^-e^-
P-S	Stanford	1962-67	0.30	e^-e^-
VEP-1	Novosibirsk	1964-68	0.13	e^-e^-
SPEAR	Stanford	1972-90	3	
DORIS	Hamburg	1974-93	5	
CESR, -c	Ithaca	1979-08	6	
LEP	Geneva	1989-00	104	
BEPC	Beijing	1989-04	2.2	
PEP-II	Stanford	1998-08	9, 3.5	Asymmetric
KEKB	Tsukuba	1999-09	8, 3.5	Asymmetric
Hadron rings				
ISR	Geneva	1971-84	32	pp
Sp$\bar{p}$S	Geneva	1981-84	270	$p\bar{p}$
HERA	Hamburg	1992-07	28, 920	ep
Tevatron	Batavia	1992-11	980	$p\bar{p}$
RHIC	Upton	2000-	250, 100	$\vec{p}$, ions
LHC	Geneva	2008-	4000	pp, ions

Thus, particle and anti-particle follow the same trajectory in opposite directions, and bunches of counter-rotating particles are guaranteed to collide (pass through each other) head-on. Further, perturbative electrostatic fields can be used to manipulate a beam of many bunches to collide at desirable locations in the middle of particle detectors, but not at other locations around the arcs of the storage ring.

Luminosity L is the second key measure of storage ring performance, in addition to centre-of-mass energy. It measures the ability of particle collisions to generate events with a cross section σ, through the equation

$$R = L\sigma \tag{1.10}$$

where R is the average rate of event generation. In the simple case of constant cross section Gaussian beams, the luminosity is

$$L = f_{rev} M \frac{N^2}{4\pi\sigma_H^*\sigma_V^*} \tag{1.11}$$

where f_{rev} is the revolution frequency of M identical bunches, each with N particles per bunch and with horizontal and vertical RMS transverse sizes of σ_H^* and σ_V^* at

the collision point. A third quantity of interest is the total current in each beam,

$$I = Ze f_{rev} NM \tag{1.12}$$

where Z is the particle charge (1 for protons, electrons and positrons, and 79 for gold ions). Finally, the energy stored in each beam is

$$U = NM A\, m_0 c^2 \gamma \tag{1.13}$$

where for ions $m_0 = 1.661 \times 10^{-27}$ kg is the atomic mass unit, and A is the atomic weight (1.008 for hydrogen and 196.967 for gold). See Exercise 1.2 for a discussion of typical parameters of the Relativistic Heavy Ion Collider, RHIC [18].

The trend to ever-higher luminosities is illustrated for gold ion beam collisions in the RHIC two-ring collider in Figure 1.6, for the period 2007 to 2014. Equation 1.11 shows that there are three main paths to higher luminosity: more bunches M, more particles N and smaller beam sizes σ^*. The number of bunches is increased by separating the beams into two rings, so that counter-circulating bunches do not see and feel the electromagnetic interactions between each other that are inevitable, even if the beams are transversely separated, in a single ring.

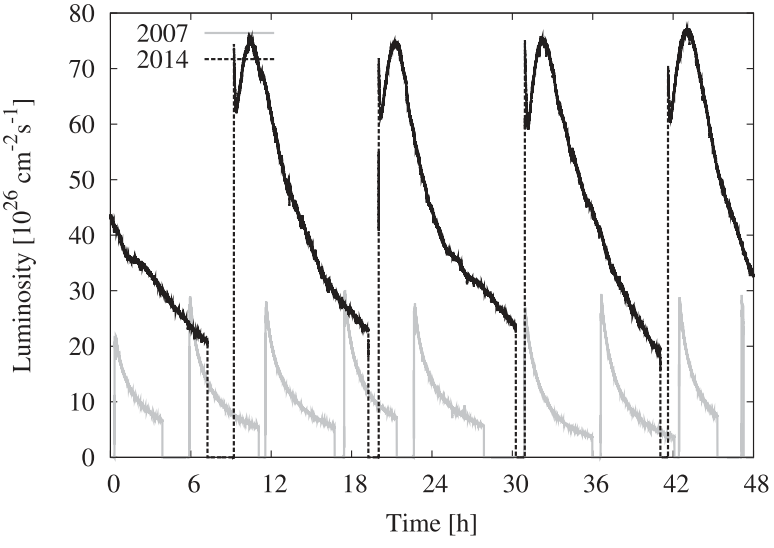

Figure 1.6 Luminosity versus time in the Relativistic Heavy Ion Collider, and its enhancement from 2007 to 2014. The grey curve shows two days of good running with gold-on-gold collisions in 2007, in which shorter stores show a classic exponential luminosity decay. The black curve shows longer stores in 2014, with approximately 90% of the gold ions productively 'burnt-off' in luminosity production. The luminosity grows initially, thanks to increasing bunch density due to transverse and longitudinal cooling. (Courtesy of W. Fischer [15].)

Increasing N and decreasing σ^* both increase the transverse density of a bunch, enhancing the transverse electromagnetic fields that accompany it, and exacerbating the violent impulse visited on counter-rotating test particles in a second bunch that passes through the first bunch. The nonlinear dynamics of this beam–beam effect (discussed in Chapter 15) limit the maximum achievable beam density, and therefore also the luminosity. Before discussing such nonlinear dynamics, however, it is first necessary to establish the linear language of beam optics that (for example) describes and enables the controlled manipulation of beam sizes like σ^*.

The second golden age of accelerator development began with the practical application of the fact that an electron that is forced to bend its trajectory radiates photons, typically X-rays, while moving through an accelerator. The strength of this synchrotron radiation increases inversely with a high power of the particle mass m, which is almost 2,000 times smaller for an electron than for a proton. For most practical purposes, GeV-range electrons radiate copiously, while protons and ions do not radiate 'at all'.

Synchrotron radiation was a curiosity in the early days of electron particle physics experimentation – even a nuisance, because the power emitted from the beam must be replaced, and paid for. Later, synchrotron radiation was used parasitically, to perform surface science, biological and other experiments, with little disturbance to the particle physics experiments. Today, electron accelerators – both circular storage rings and linear free electron lasers – are designed, built and commissioned as dedicated facilities solely for synchrotron light uses.

Today, accelerators continue their rapid diversification far beyond their particle physics roots, in the on-going second golden age. In addition to a new generation of synchrotron light sources, cutting edge and overlapping accelerator research and development is active in pursuit of room-size medical accelerators for proton and ion beam cancer therapy, medium-size accelerators for industrial applications and large-scale high-power MW-class superconducting proton linear accelerators (linacs) for neutron science. In the future, perhaps multi-MW proton sources in accelerator driven reactors will help to generate electrical power, at the same time as burning nuclear waste generated by conventional uranium power reactors. The future is hard to predict, but the language of single particle dynamics helps to describe and design it.

Exercises

1.1 The *logistic map* advances a population x from generation n to generation $n + 1$ through the equation

$$x_{n+1} = \alpha\, x_n(1 - x_n) \tag{1.14}$$