

Material seleccionado para estudios MAC-0209

SEMANAS 1 E 2

Chapter 1

Introduction

The importance of computers in physics and the nature of computer simulation is discussed. The nature of object-oriented programming and various computer languages is also considered.

1.1 Importance of computers in physics

Computation is now an integral part of contemporary science and is having a profound effect on the way we do physics, on the nature of the important questions, and on the physical systems we choose to study. Developments in computer technology are leading to new ways of thinking about physical systems. Asking “How can I formulate this problem on a computer?” has led to the understanding that it is practical and natural to formulate physical laws as rules for a computer rather than only in terms of differential equations.

For the purposes of discussion, we will divide the use of computers in physics into the following categories: numerical analysis, symbolic manipulation, visualization, simulation, and the collection and analysis of data. *Numerical analysis* refers to the solution of well-defined mathematical problems to produce numerical (in contrast to symbolic) solutions. For example, we know that the solution of many problems in physics can be reduced to the solution of a set of simultaneous linear equations. Consider the equations

$$\begin{aligned}2x + 3y &= 18 \\ x - y &= 4.\end{aligned}$$

It is easy to find the analytical solution $x = 6$, $y = 2$ using the method of substitution. Suppose we wish to solve a set of four simultaneous equations. We again can find an analytical solution, perhaps using a more sophisticated method. However, if the number of simultaneous equations becomes much larger, we would need to use a computer to find a solution. In this mode the computer is a tool of numerical analysis. Because it is often necessary to compute multidimensional integrals, manipulate large matrices, or solve nonlinear differential equations, this use of the computer is important in physics.

One of the strengths of mathematics is its ability to use the power of abstraction, which allows us to solve many similar problems simultaneously by using symbols. Computers can be used to do much of the *symbolic manipulation*. As an example, suppose we want to know the solution of the quadratic equation, $ax^2 + bx + c = 0$. A symbolic manipulation program can give the solution as $x = [-b \pm \sqrt{b^2 - 4ac}]/2a$. In addition, such a program can give the usual numerical solutions for specific values of a , b , and c . Mathematical operations such as differentiation,

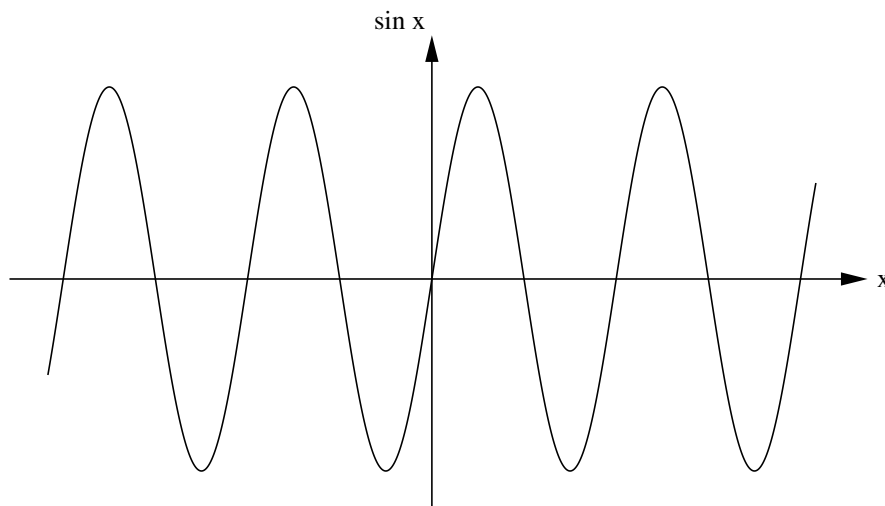


Figure 1.1: What is the meaning of the sine function?

integration, matrix inversion, and power series expansion can be performed using symbolic manipulation programs. The calculation of Feynman diagrams, which represent multidimensional integrals of importance in quantum electrodynamics, has been a major impetus to the development of computer algebra software that can manipulate and simplify symbolic expressions. Maxima, Maple, and Mathematica are examples of software packages that have symbolic manipulation capabilities as well as many tools for numerical analysis. Matlab and Octave are examples of software packages that are convenient for computations involving matrices and related tasks.

As the computer plays an increasing role in our understanding of physical phenomena, the *visual representation* of complex numerical results is becoming even more important. The human eye in conjunction with the visual processing capacity of the brain is a very sophisticated device. Our eyes can determine patterns and trends that might not be evident from tables of data and can observe changes with time that can lead to insight into the important mechanisms underlying a system's behavior. The use of graphics can also increase our understanding of the nature of analytical solutions. For example, what does a sine function mean to you? We suspect that your answer is not the series, $\sin x = x - x^3/3! + x^5/5! + \dots$, but rather a periodic, constant amplitude curve (see Figure 1.1). What is most important is the mental image gained from a visualization of the form of the function.

Traditional modes of presenting data include two- and three-dimensional plots including contour and field line plots. Frequently, more than three variables are needed to understand the behavior of a system, and new methods of using color and texture are being developed to help researchers gain greater insights into their data.

An essential role of science is to develop models of nature. To know whether a model is consistent with observation, we have to understand the behavior of the model and its predictions. One way to do so is to implement the model on a computer. We call such an implementation a *computer simulation* or simulation for short. For example, suppose a teacher gives \$10 to each student in a class of 100. The teacher, who also begins with \$10 in her pocket, chooses a student at random and flips a coin. If the coin is heads, the teacher gives \$1 to the student; otherwise, the student gives \$1 to the teacher. If either the teacher or the student would go into debt by

this transaction, the transaction is not allowed. After many exchanges, what is the probability that a student has s dollars? What is the probability that the teacher has t dollars? Are these two probabilities the same? Although these particular questions can be answered by analytical methods, many problems of this nature cannot be solved in this way (see Problem 1.1).

One way to determine the answers to these questions is to do a classroom experiment. However, such an experiment would be difficult to arrange, and it would be tedious to do a sufficient number of transactions.

A more practical way to proceed is to convert the rules of the model into a computer program, simulate many exchanges, and estimate the quantities of interest. Knowing the results might help us gain more insight into the nature of an analytical solution if one exists. We can also modify the rules and ask “what if?” questions. For example, would the probabilities change if the students could exchange money with one another? What would happen if the teacher was allowed to go into debt?

Simulations frequently use the computational tools of numerical analysis and visualization, and occasionally symbolic manipulation. The difference is one of emphasis. Simulations are usually done with a minimum of analysis. Because simulation emphasizes an exploratory mode of learning, we will stress this approach.

Computers are also involved in all phases of a laboratory experiment, from the design of the apparatus to the *collection and analysis of data*. LabView is an example of a data acquisition program. Some of the roles of the computer in laboratory experiments, such as the varying of parameters and the analysis of data, are similar to those encountered in simulations. However, the tasks involved in *real-time control* and interactive data analysis are qualitatively different and involve the interfacing of computer hardware to various types of instrumentation. We will not discuss this use of the computer.

1.2 The importance of computer simulation

Why is computation becoming so important in physics? One reason is that most of our analytical tools such as differential calculus are best suited to the analysis of *linear* problems. For example, you probably have analyzed the motion of a particle attached to a spring by assuming a linear restoring force and solving Newton’s second law of motion. In this case a small change in the displacement of the particle leads to a small change in the force. However, many natural phenomena are *nonlinear*, and a small change in a variable might produce a large change in another. Because relatively few nonlinear problems can be solved by analytical methods, the computer gives us a new tool to explore nonlinear phenomena.

Another reason for the importance of computation is the growing interest in systems with many variables or with many degrees of freedom. The money exchange model described in Section 1.1 is a simple example of a system with many variables. A similar problem is given at the end of this chapter.

Computer simulations are sometimes referred to as *computer experiments* because they share much in common with laboratory experiments. Some of the analogies are shown in Table 1.1. The starting point of a computer simulation is the development of an idealized model of a physical system of interest. We then need to specify a procedure or *algorithm* for implementing the model on a computer and decide what quantities to measure. The results of a computer simulation can serve as a bridge between laboratory experiments and theoretical calculations. In some cases we can obtain essentially exact results by simulating an idealized model that has no

Laboratory Experiment	Computer Simulation
sample	model
physical apparatus	computer program
calibration	testing of program
measurement	computation
data analysis	data analysis

Table 1.1: Analogies between a computer simulation and a laboratory experiment.

laboratory counterpart. The results of the idealized model can serve as a stimulus to the development of the theory. On the other hand, we sometimes can do simulations of a more realistic model than can be done theoretically, and hence make a more direct comparison with laboratory experiments. Computation has become a third way of doing physics and complements both theory and experiment.

Computer simulations, like laboratory experiments, are not substitutes for thinking, but are tools that we can use to understand natural phenomena. The goal of all our investigations of fundamental phenomena is to seek explanations of natural phenomena that can be stated concisely.

1.3 Programming languages

There is no single best programming language any more than there is a best natural language. Fortran is the oldest of the more popular scientific programming languages and was developed by John Backus and his colleagues at IBM between 1954 and 1957. Fortran is commonly used in scientific applications and continues to evolve. Fortran 90/95/2000 has many modern features that are similar to C/C++.

The Basic programming language was developed in 1965 by John Kemeny and Thomas Kurtz at Dartmouth College as a language for introductory courses in computer science. In 1983 Kemeny and Kurtz extended the language to include platform independent graphics and advanced control structures necessary for structured programming. The programs in the first two editions of our textbook were written in this version of Basic, known as True Basic.

C was developed by Dennis Ritchie at Bell Laboratories around 1972 in parallel with the Unix operating system. C++ is an extension of C designed by Bjarne Stroustrup at Bell laboratories in the mid-eighties. C++ is considerably more complex than C and has object oriented features, as well as and other extensions. In general, programs written in C/C++ have high performance, but can be difficult to debug. C and C++ are popular choices for developing operating systems and software applications because they provide direct access to memory and other system resources.

Python, like Basic, was designed to be easy to learn and use. Python enthusiasts like to say that C and C++ were written to make life easier for the computer, but Python was designed to be easier for the programmer. Guido van Rossum created Python in the late 80's and early 90's. It is an interpreted, object-oriented, general-purpose programming language that is also good for prototyping. Because Python is interpreted, its performance is significantly less than optimized languages like C or Fortran.

Java is an object-oriented language that was created by James Gosling and others at Sun Microsystems. Since Java was introduced in late 1995, it has rapidly evolved and is the language of choice in most introductory computer science courses. Java borrows much of its syntax from

in Problem 1.1b is not known. More information about this model can be found in the article by Patriarca, Chakraborti, and Kaski (see the references at the end of this chapter).

Problem 1.1 illustrates some of the characteristics of simulations that we will consider in the following chapters. Implementing this model on a computer would help you to gain insight into its behavior and might encourage you to explore variations of the model. Note that the model lends itself to asking a relatively simple “what if” question, which in this case leads to qualitatively different behavior. Asking similar questions might require modifying only a few lines of code. However, such a change might convert an analytically tractable problem into one for which the solution is unknown.

Problem 1.2. Questions to consider

- a. You are familiar with the fall of various objects near the earth’s surface. Suppose that a ball is in the earth’s atmosphere long enough for air resistance to be important. How would you simulate the motion of the ball?
- b. Suppose that you wish to model a simple liquid such as liquid Argon. Why is such a liquid simpler to simulate than water? What is the maximum number of atoms that can be simulated in a reasonable amount of time using present computer technology? What is the maximum real time that is possible to simulate? That is, if we run our program for a week of computer time, what would be the equivalent time that the liquid has evolved?
- c. Discuss some examples of systems that would be interesting to you to simulate. Can these systems be analyzed by analytical methods? Can they be investigated experimentally?
- d. An article by Post and Votta (see references) claims that “...(computers) have largely replaced pencil and paper as the theorist’s main tool.” Do you agree with this statement? Ask some of the theoretical physicists that you know for their opinions.

Appendix 1A: Laboratory reports

Laboratory reports should reflect clear writing style and obey proper rules of grammar and correct spelling. Write in a manner that can be understood by another person who has not done the research. In the following, we give a suggested format for your reports.

Introduction. Briefly summarize the nature of the physical system, the basic numerical method or algorithm, and the interesting or relevant questions.

Method. Describe the algorithm and how it is implemented in the program. In some cases this explanation can be given in the program itself. Give a typical listing of your program. Simple modifications of the program can be included in an appendix if necessary. The program should include your name and date and be annotated in a way that is as self-explanatory as possible. Be sure to discuss any important features of your program.

Verification of program. Confirm that your program is not incorrect by considering special cases and by giving at least one comparison to a hand calculation or known result.

Data. Show the results of some typical runs in graphical or tabular form. Additional runs can be included in an appendix. All runs should be labeled, and all tables and figures must be referred to in the body of the text. Each figure and table should have a caption with complete information, for example, the value of the time step.

Analysis. In general, the analysis of your results will include a determination of qualitative and quantitative relationships between variables and an estimation of numerical accuracy.

Interpretation. Summarize your results and explain them in simple physical terms whenever possible. Specific questions that were raised in the assignment should be addressed here. Also give suggestions for future work or possible extensions. It is not necessary to answer every part of each question in the text.

Critique. Summarize the important physical concepts for which you gained a better understanding and discuss the numerical or computer techniques you learned. Make specific comments on the assignment and suggestions for improvements or alternatives.

Log. Keep a log of the time spent on each assignment and include it with your report.

References and suggestions for further reading

Programming

We list some of our favorite Java programming books here. There are many useful online tutorials.

Joshua Bloch, *Effective Java* (Addison–Wesley, 2001). This excellent book is for advanced Java programmers and should be read after you have become familiar with Java.

Rogers Cadenhead and Laura Lemay *Teach Yourself Java in 21 Days* 4th ed. (Sams, 2004). An inexpensive self-study guide that uses a step by step tutorial approach to cover the basics.

Stephen J. Chapman, *Java for Engineers and Scientists*, 2nd ed. (Prentice Hall, 2004).

Wolfgang Christian, *Open Source Physics: A User's Guide with Examples* (Addison–Wesley, 2006). This guide is a useful supplement to our text.

Bruce Eckel, *Thinking in Java*, 3rd ed. (Prentice Hall, 2003). This text discusses the finer points of object-oriented programming and is recommended after you have become familiar with Java. See also <www.mindview.net/Books/>.

David Flanagan, *Java in a Nutshell*, 5th ed. (O'Reilly, 2005) and *Java Examples in a Nutshell*, 3rd ed. (O'Reilly, 2004). A fast-paced Java tutorial for those who already know another programming language.

Brian D. Hahn and Katherine M. Malan, *Essential Java for Scientists and Engineers* (Butterworth-Heinemann, 2002).

Cay S. Horstmann and Gary Cornell, *Core Java 2: Fundamentals* and *Core Java 2: Advanced Features*, both in 7th ed. (Prentice Hall, 2005). A two-volume set that covers all aspects of Java programming.

Patrick Niemeyer and Jonathan Knudsen, *Learning Java*, 2nd ed. (O'Reilly, 2002). A comprehensive introduction to Java that starts with HelloWorld and ends with a discussion of XML. The book contains many examples showing how the core Java API is used. This book is one of our favorites for beginning Java programmers. However, it might be intimidating to someone who does not have some familiarity with computers.

Sherry Shavor, Jim D’Anjou, Pat McCarthy, John Kellerman, and Scott Fairbrothe, *The Java Developer’s Guide to Eclipse* (Addison–Wesley Professional, 2003). A good reference for the open source Eclipse development environment. Check for new versions because Eclipse is evolving rapidly.

General References on Physics and Computers

Richard E. Crandall, *Projects in Scientific Computation* (Springer–Verlag, 1994).

Paul L. DeVries, *A First Course in Computational Physics* (John Wiley & Sons, 1994).

Alejandro L. Garcia, *Numerical Methods for Physics*, 2nd ed. (Prentice Hall, 2000). Matlab, C++, and Fortran are used.

Neil Gershenfeld, *The Nature of Mathematical Modeling* (Cambridge University Press, 1998).

Nicholas J. Giordano and Hisao Nakanishi, *Computational Physics*. 2nd ed. (Prentice Hall, 2005).

Dieter W. Heermann, *Computer Simulation Methods in Theoretical Physics*, 2nd ed. (Springer–Verlag, 1990). A discussion of molecular dynamics and Monte Carlo methods directed toward advanced undergraduate and beginning graduate students.

David Landau and Kurt Binder, *A Guide to Monte Carlo Simulations in Statistical Physics*, 2nd ed. (Cambridge University Press, 2005). The authors emphasize the complementary nature of simulation to theory and experiment.

Rubin H. Landau, *A First Course in Scientific Computing* (Princeton University Press, 2005).

P. Kevin MacKeown, *Stochastic Simulation in Physics* (Springer, 1997).

Tao Pang, *Computational Physics* (Cambridge University Press, 1997).

Franz J. Vesely, *Computational Physics*, 2nd ed. (Plenum Press, 2002).

Michael M. Wolfson and Geoffrey J. Perl, *Introduction to Computer Simulation* (Oxford University Press, 1999).

Other References

Ruth Chabay and Bruce Sherwood, *Matter & Interactions* (John Wiley & Sons, 2002). This two-volume text uses computer models written in VPython to present topics not typically discussed in introductory physics courses.

H. Gould, “Computational physics and the undergraduate curriculum,” *Computer Physics Communications* **127** (1), 6–10 (2000).

Brian Hayes, “g-LOGY,” *Am. Scientist* **92** (3), 212–216 (2004) discusses the g-factor of the electron and the importance of algebraic and numerical calculations.

Problem 1.1 is based on a paper by Marco Patriarca, Anirban Chakraborti, and Kimmo Kaski, “Gibbs versus non-Gibbs distributions in money dynamics,” *Physica A* **340**, 334–339 (2004).

An interesting article on the future of computational science by Douglass E. Post and Lawrence G. Votta, “Computational science demands a new paradigm,” *Physics Today* **58** (1), 35–41 (2005) raises many interesting questions.

Introdução

1.1 PARA QUE SERVE A FÍSICA?

A ciência desempenha um papel muito importante no mundo contemporâneo. Não era assim há poucas gerações: o desenvolvimento científico tem-se acelerado enormemente. Tornou-se lugar comum dizer que vivemos numa sociedade tecnológica e medir o progresso pelo grau de desenvolvimento tecnológico. A tecnologia depende crucialmente da ciência para renovar-se, e também contribui para ela, mas não devem ser confundidas.

Sem dúvida, nossas vidas são profundamente afetadas pela tecnologia, e de forma que, muitas vezes, está longe de ser benéfica. Basta lembrar os problemas da poluição e do aquecimento global. Os cientistas são frequentemente responsabilizados pelos aspectos negativos decorrentes de suas descobertas, embora o uso que delas se faz dependa de fatores políticos e econômicos alheios a sua vontade. Por mais benéfica que seja a intenção original, ela é frequentemente deturpada. Por isso mesmo, os cientistas devem ter consciência de sua responsabilidade.

Vários problemas cruciais de nossa época dependem para sua solução de avanços científicos e tecnológicos, inclusive aqueles que se originam direta ou indiretamente desses avanços. Os problemas da energia e do meio ambiente adquiriram uma importância vital.

A reação anticientífica existiu desde os primórdios da história da física. Basta lembrar o exemplo de Galileu. Goethe atacou Newton por sua teoria das cores, dizendo que a essência das cores se percebe num pôr do sol, e não fazendo experimentos com um prisma. É preciso reconhecer que a visão científica do mundo não exclui nem invalida outras variedades da experiência. Podemos aplicar a acústica, a neurofisiologia e a psicologia ao estudo das sensações provocadas pela audição de uma sonata de Mozart, mas ainda estaríamos omitindo provavelmente o aspecto mais importante.

A consciência das limitações do método científico não nos deve impedir de apreciar sua imensa contribuição ao conhecimento da natureza. A motivação básica da ciência sempre tem sido a de entender o mundo. É a mesma curiosidade que leva um menino a desmontar um relógio para saber como funciona. De que são feitas as coisas?

Como e por que se movem os corpos celestes? Qual é a natureza da eletricidade e do magnetismo? O que é a luz? Qual a origem do universo? Estas são algumas das grandes questões que têm sido abordadas pelos físicos.

A experiência tem demonstrado que o trabalho de pesquisa básica, motivado exclusivamente pela curiosidade, leva com frequência a aplicações inesperadas de grande importância prática. Conta-se que o grande experimentador Michael Faraday, questionado pelo primeiro-ministro da Inglaterra sobre para que serviria sua recente descoberta do fenômeno da indução eletromagnética, teria respondido: “Quem sabe um dia será taxado pelo governo”. Quase todas as aplicações que fazemos hoje em dia da energia elétrica decorrem do efeito descoberto por Faraday. O transistor, o laser, os computadores resultaram de pesquisas básicas em física.

O trabalho de muitas gerações demonstrou a existência de ordem e regularidade nos fenômenos naturais, daquilo que chamamos de leis da natureza. O estudo que ora iniciamos pode ser empreendido pelos mais diversos motivos, mas uma de suas maiores recompensas é uma melhor apreciação da simplicidade, beleza e harmonia dessas leis. É uma espécie de milagre, como disse Einstein: “O que a natureza tem de mais incompreensível é o fato de ser compreensível”.

1.2 RELAÇÕES ENTRE FÍSICA E OUTRAS CIÊNCIAS

A física é, em muitos sentidos, a mais fundamental das ciências naturais, e é também aquela cuja formulação atingiu o maior grau de refinamento.

Com a explicação da estrutura atômica fornecida pela mecânica quântica, a química pode ser considerada até certo ponto como um ramo da física. A física forneceu a explicação da ligação química, e a estrutura e as propriedades das moléculas podem ser calculadas “em princípio” resolvendo problemas de física. Isso não significa que o sejam na prática, exceto em alguns casos extremamente simples. De fato, na imensa maioria dos casos, os sistemas químicos são demasiado complexos para serem tratáveis fisicamente, mesmo com auxílio dos computadores mais poderosos disponíveis, o que significa que os métodos específicos extremamente engenhosos elaborados pelos químicos para tratar esses problemas continuam sendo indispensáveis. Entretanto, não temos razões para duvidar de que as interações básicas responsáveis pelos processos químicos sejam já conhecidas e reduzidas a termos físicos.

A situação com respeito à biologia é até certo ponto análoga, se bem que a compreensão em termos de leis físicas se encontre ainda num estágio menos desenvolvido. Muitas das peculiaridades dos sistemas biológicos resultam de ser fruto de uma evolução histórica – a teoria de Darwin da evolução é fundamental na biologia. Esse fator não é usualmente considerado para sistemas físicos. É certo que na cosmologia a evolução do universo a partir de sua origem é um tema central, mas não no sentido de evolução darwiniana. Os avanços recentes da biologia molecular vêm atuando no sentido de estabelecer uma aproximação cada vez maior entre a biologia e a física.

A física deve grande parte de seu sucesso como modelo de ciência natural ao fato de que sua formulação utiliza uma linguagem que é, ao mesmo tempo, uma ferramenta

muito poderosa: a matemática. Na expressão de Galileu, “A ciência está escrita neste grande livro colocado sempre diante de nossos olhos – o universo – mas não podemos lê-lo sem apreender a linguagem e entender os símbolos em termos dos quais está escrito. Este livro está escrito na linguagem matemática”.

É importante compreender bem as relações entre física e matemática. Bertrand Russell definiu a matemática como: “A ciência onde nunca se sabe de que se está falando nem se o que se está dizendo é verdade” para caracterizar o método axiomático: tudo é deduzido de um conjunto de axiomas, mas a questão da “validade” desses axiomas no mundo real não se coloca. Hilbert, ao axiomatizar a geometria, disse que nada deveria se alterar se as palavras “ponto, reta, plano” fossem substituídas por “mesa, cadeira, copo”. Conforme o conjunto de axiomas adotado, obtém-se a geometria euclidiana ou uma das geometrias não euclidianas, mas não tem sentido perguntar, do ponto de vista da matemática, qual delas é “verdadeira”.

Na física, como ciência natural, essa pergunta faz sentido: qual é a geometria do mundo real? A experiência mostra que, na escala astronômica, aparecem desvios da geometria euclidiana.

A física é muitas vezes classificada como “ciência exata”, para ressaltar seus aspectos quantitativos. Já no século VI a. C., a descoberta pela Escola Pitagórica de algumas das leis das cordas vibrantes, estabelecendo uma relação entre sons musicais harmoniosos e números inteiros (proporção entre comprimentos de cordas que emitem tons musicais), levou à convicção de que: “Todas as coisas são números”.

Embora a formulação em termos quantitativos seja muito importante, a física também lida com muitos problemas interessantes de natureza qualitativa. Isso não significa que não requerem tratamento matemático: algumas das teorias mais difíceis e elaboradas da matemática moderna dizem respeito a métodos qualitativos.

Neste curso, a ênfase não será no tratamento matemático, e sim nos conceitos físicos. Alguns dos conceitos matemáticos básicos que vamos empregar serão introduzidos à medida que se tornarem necessários. Também exemplificaremos algumas aplicações à biologia.

A natureza ignora as distinções que estabelecemos entre diferentes disciplinas. A pesquisa científica de fronteira requer cada vez mais uma abordagem interdisciplinar.

1.3 O MÉTODO CIENTÍFICO

Não se pode codificar um conjunto de regras absolutas para a pesquisa. Cabem apenas algumas observações sobre esse tema.

1. *Observação e experimentação*: são o ponto de partida e, ao mesmo tempo, o teste crucial na formulação das leis naturais. A física, como as demais ciências naturais, é uma ciência experimental. Assim, o bom acordo com a experiência é o juiz supremo da validade de qualquer teoria científica. O diálogo Hegeliano, “Só pode haver sete planetas. Mas isso contradiz os fatos! Tanto pior para os fatos!”, representa o oposto da atitude científica. A única autoridade reconhecida como árbitro decisivo da validade de uma teoria é a verificação experimental de suas consequências. Se não está de acordo com a experiência, tem de ser descartada.

Entretanto, “embora a ciência se construa com dados experimentais, da mesma forma que uma casa se constrói com tijolos, uma coleção de dados experimentais ainda não é ciência, da mesma forma que uma coleção de tijolos não é uma casa” (Poincaré).

2. *Abstração, indução*: Já se disse que a primeira lei da ecologia é: “Tudo depende de tudo”; é por isso que problemas ecológicos são tão complexos. Em certa medida, o mesmo vale para a física ou qualquer outra ciência natural. Quando uma maçã cai da árvore, o movimento da Terra sofre uma (pequeníssima!) perturbação, e ele também é afetado pelo que acontece em galáxias extremamente distantes. Entretanto, seria impossível chegar à formulação de leis naturais se procurássemos levar em conta desde o início, no estudo de cada fenômeno, todos os fatores que possam influenciá-lo, por menor que seja essa influência.

O primeiro passo no estudo de um fenômeno natural consiste em fazer abstração de um grande número de fatores considerados inessenciais, concentrando a atenção apenas nos aspectos mais importantes. O julgamento sobre o que é ou não importante já envolve a formulação de modelos e conceitos teóricos, que representam, segundo Einstein, uma “livre criação da mente humana”.

Um bom exemplo é o conceito de “partícula” na mecânica. Na geografia, em que o globo terrestre é o principal objeto de estudo, é preciso, para muitos fins, levar em conta as irregularidades da crosta terrestre. Ao estudar o movimento de rotação da

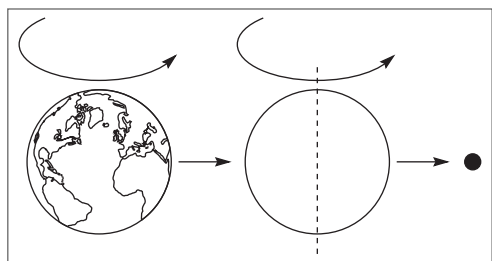


Figura 1.1 Estágios sucessivos de abstração na representação da Terra.

Terra em torno de seu eixo, podemos considerá-la, em primeira aproximação, como uma esfera rígida uniforme. Já quando estudamos o movimento de translação da Terra em torno do Sol, considerando que o diâmetro da Terra é menor que um décimo-milésimo de sua distância ao Sol, podemos desprezar suas dimensões, tratando-a como uma partícula ou “ponto material”. Temos assim estágios sucessivos de abstração (Figura 1.1) na representação de nosso planeta.

A arte do teórico está em julgar o que e como abstrair, o que é essencial e o que é acessório. O experimentador enfrenta problemas análogos: eliminar “efeitos espúrios” e medir apenas o efeito desejado é extremamente difícil. Só recentemente se descobriu que o universo inteiro é atravessado por radiação eletromagnética, proveniente da Grande Explosão da qual se teria originado, e que pode produzir efeitos importantes na escala quântica.

Uma vez atingido certo estágio no desenvolvimento de conceitos e modelos, pode-se procurar, por meio de um processo indutivo, formular leis fenomenológicas, ou seja, obtidas diretamente a partir dos fenômenos observados, como forma sintética e mais econômica de descrevê-los. Convém frisar que esse é apenas um de muitos processos possíveis que têm sido empregados na formulação de leis físicas.

3. *Leis e teorias físicas*: Um exemplo clássico desse processo, que será discutido no Capítulo 10, foi a formulação das leis de Kepler do movimento planetário

a partir das observações feitas por Tycho Brahe. Neste caso, a etapa ulterior, que culminou na obra de Newton, foi a formulação das leis gerais do movimento e da lei da gravitação universal. O resultado foi a elaboração de uma nova teoria física, a teoria da gravitação, situada dentro de uma teoria mais ampla, a mecânica clássica.

Esse exemplo ilustra algumas das características importantes de uma boa teoria:

a) Deve ser capaz de reduzir grande número de fenômenos diversos a um pequeno número de leis simples, mostrando que podem ser deduzidos matematicamente a partir dessas leis básicas; b) Deve ter poder preditivo: a partir das leis básicas, deve ser possível prever fenômenos novos que possam ser comparados com a experiência. Uma teoria deve sempre ser explorada em todas as direções possíveis, no sentido de verificação de suas previsões. Um dos maiores triunfos da teoria da gravitação universal foi a previsão da existência de Netuno, feita por Adams e Leverrier em 1846.

4. *Domínio de validade*: Todas as teorias físicas conhecidas sempre têm representado aproximações aplicáveis num certo domínio da experiência. Assim, por exemplo, as leis da mecânica clássica são aplicáveis aos movimentos usuais de objetos macroscópicos, mas deixam de valer: (i) para velocidades comparáveis com a velocidade da luz, quando aparecem efeitos relativísticos; (ii) para objetos na escala atômica, quando temos de empregar a mecânica quântica.

Entretanto, uma “revolução científica” raramente inutiliza por completo as teorias precedentes. A validade aproximada dessas teorias no domínio em que já haviam sido testadas experimentalmente garante, em geral, sua sobrevivência nesse domínio. Assim, a mecânica clássica continua sendo aplicável a uma grande variedade de movimentos macroscópicos.

Uma nova teoria representa em regra uma generalização da antiga, estendendo-a a um domínio mais amplo, mas contendo-a muitas vezes como caso particular ou caso-limite, válido aproximadamente no domínio anterior. Isso não impede que os conceitos básicos da nova teoria possam diferir radicalmente dos anteriores.

O processo de “seleção natural” pelo qual passam as teorias científicas exige que sejam sempre submetidas a uma ampla crítica pela comunidade científica internacional e ao maior número possível de testes experimentais. Por isso, o segredo e o dogma são inimigos da ciência e a liberdade de comunicação e de pesquisa são vitais para o seu florescimento.

Poderia parecer conveniente iniciar desde logo o estudo da física pelas leis mais exatas conhecidas, uma vez que contêm as formulações anteriores como caso-limite ou caso particular. Entretanto, isso não seria recomendável, e nem mesmo possível, por muitas razões. Do ponto de vista pedagógico, é importante começarmos pelo domínio de fenômenos que nos são mais familiares. A física clássica, que compreende a maior parte do nosso curso, tem um extenso domínio de aplicabilidade, na escala de nossa experiência cotidiana, e uma boa compreensão da mesma tem importância fundamental para a própria formulação da mecânica quântica. Entretanto, convém não perder de vista os limites de aplicabilidade das teorias que vamos estudar. Sempre que possível, chamaremos a atenção sobre esses limites.

1.4 ORDENS DE GRANDEZA. ALGARISMOS SIGNIFICATIVOS

Conta-se que o astrônomo inglês Arthur Eddington iniciou uma de suas aulas, em certa ocasião dizendo: “Acredito que o número total de elétrons no universo (igual ao número de prótons) é dado por 15. 747. 724. 136. 275. 002. 577. 605. 653. 961. 181. 555. 468. 044. 717. 914. 527. 116. 709. 366. 231. 425. 076. 185. 631. 031. 296”. Na opinião dele, esse número representaria uma constante fundamental da natureza, dedutível teoricamente.

Embora as ideias numerológicas de Eddington não tenham encontrado receptividade, esse exemplo serve pelo menos para ilustrar o fato de que na física é frequente termos de lidar com números muito grandes ou muito pequenos, uma vez que ela abrange o estudo de fenômenos que vão desde a escala subatômica até a escala do universo. Torna-se necessário assim o uso de uma notação conveniente.

O número de Eddington é igual a $2 \times 136 \times 2^{256}$, o que ilustra a vantagem da notação exponencial. Convém lembrar algumas regras simples da potenciação:

$$\begin{aligned} a^p \cdot a^q &= a^{p+q} \\ a^{-p} &= 1 / a^p \\ (a^p)^q &= a^{pq} \end{aligned}$$

Usualmente trabalhamos com potências de 10. A tabela abaixo dá as abreviações usadas junto aos nomes das unidades para potências decrescentes e crescentes de 10.

10^1	deca	da	10^{-1}	deci	d
10^2	hecto	h	10^{-2}	centi	c
10^3	kilo	k	10^{-3}	mili	m
10^6	mega	M	10^{-6}	micro	μ
10^9	giga	G	10^{-9}	nano	n
10^{12}	tera	T	10^{-12}	pico	p
10^{15}	peta	P	10^{-15}	femto	f
10^{18}	exa	E	10^{-18}	atto	a
10^{21}	zetta	Z	10^{-21}	zepto	z
10^{24}	yotta	Y	10^{-24}	yocto	y

Exemplo: A velocidade da luz no vácuo é aproximadamente

$$c \approx 300.000 \text{ km / s} = 3 \times 10^5 \text{ km / s}$$

onde “ \approx ” significa: “aproximadamente igual a”.

$$1 \text{ km} = 10^3 \text{ m} = 10^5 \text{ cm} \Rightarrow c \approx 3 \times 10^{10} \text{ cm / s}$$

O número de Eddington, nesta notação, é $\approx 1,6 \times 10^{79}$. Embora não levemos a sério seus argumentos numerológicos, a ordem de grandeza concorda bastante bem com as estimativas atuais sobre o número total de átomos (dominado pelos de hidrogênio) no universo.

SEMANAS 3 E 4

2

Movimento unidimensional

2.1 VELOCIDADE MÉDIA

A análise do movimento é um problema fundamental em física, e a forma mais simples de abordá-la é considerar primeiro os conceitos que intervêm na *descrição* do movimento (*cinemática*), sem considerar ainda o problema de como determinar o movimento que se produz numa dada situação física (*dinâmica*). No presente capítulo, para simplificar ainda mais a discussão, vamo-nos limitar ao movimento em uma só dimensão – por exemplo, o movimento de um automóvel em linha reta ao longo de uma estrada. Como muitos aspectos da cinemática são discutidos no curso secundário, vamos restringir o tratamento a apenas alguns tópicos centrais.

Para descrever o movimento, precisamos em primeiro lugar de um *referencial*, que, no caso unidimensional, é simplesmente uma reta orientada em que se escolhe a origem O ; a posição de uma partícula em movimento no instante t é descrita pela abscissa correspondente $x(t)$.

Concretamente, podemos pensar no seguinte exemplo: $x(t)$ é a posição na estrada, no instante t , ocupada pelo para-choque dianteiro de um carro em movimento ao longo da estrada (em linha reta). Poderíamos determinar $x(t)$, por exemplo, filmando o movimento do carro e depois analisando uma a uma as imagens do filme. Sabendo quantas imagens por segundo são tiradas pelo filmador, saberíamos o intervalo de tempo Δt (uma fração de segundo) entre duas imagens consecutivas do filme, e poderíamos assim obter o valor de x nos instantes: $0, \Delta t, 2 \Delta t, \dots$, bastante próximos entre si (poderíamos também estar filmando simultaneamente um cronômetro fixo em primeiro plano para definir o instante correspondente a cada imagem).

Outro método de “congelar” a posição instantânea de um objeto em movimento é tirar uma fotografia de exposição múltipla em que o objeto é iluminado a intervalos de

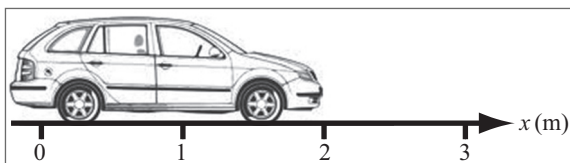


Figura 2.1 Exemplo de movimento unidimensional.

tempo Δt regulares por um “flash” ultrarrápido (estroboscopia). O aspecto de uma fotografia deste tipo para uma bolinha em queda livre ao longo de uma régua graduada está representado na Figura 2.2.

Por qualquer um destes métodos, podemos construir uma “tabela horária” do movimento, do tipo

Tabela 2.1

$t(s)$	0	1	2	3	4	5	6
$x(m)$	0	2	3,5	4	3,5	2	0

ou um gráfico, do tipo

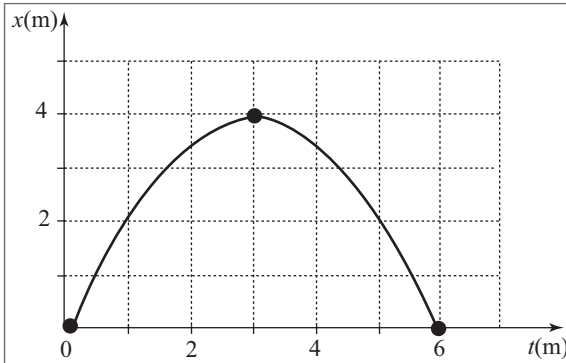


Figura 2.3 Gráfico de movimento.

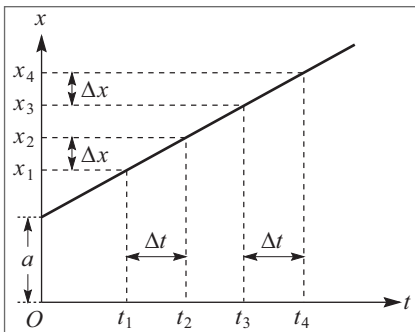


Figura 2.4 Gráfico do movimento retilíneo uniforme.

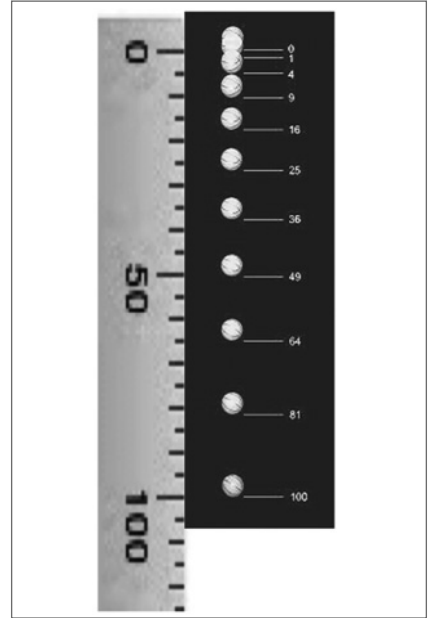


Figura 2.2 Queda livre (observação estroboscópica).

O movimento mais simples é o *movimento uniforme*, em que este gráfico é uma reta:

$$x(t) = at + b \quad (2.1.1)$$

Este movimento se caracteriza pelo fato de que percursos iguais $\Delta x = x_4 - x_3 = x_2 - x_1$ (Figura 2.4) são descritos em intervalos de tempo iguais $\Delta t = t_4 - t_3 = t_2 - t_1$

A *velocidade* v do movimento é definida por

$$v = \frac{\Delta x}{\Delta t} = \frac{x(t_2) - x(t_1)}{t_2 - t_1} \quad (2.1.2)$$

ou seja, é a razão do deslocamento ao intervalo de tempo que ele leva para se produzir. Graficamente, v representa o *coeficiente angular* da reta no gráfico $x \times t$ ($v = a$ para a (2.1.1)).

A velocidade se mede em m/s ($= m \cdot s^{-1}$), ou cm/s, ou km/h, ..., conforme as unidades adotadas. Note que v pode tomar tanto valores positivos como negativos; pela (2.1.2), $v < 0$ quando $\Delta x < 0$ para $\Delta t > 0$, ou seja, quando o movimento se dá no sentido dos x decrescentes (marcha à ré, no exemplo do carro!). Poderíamos chamar de “rapidez” o valor absoluto da velocidade, $|v|$.

Se aplicamos a (2.1.2) tomando para t_2 um instante t qualquer e para t_1 o instante inicial t_0 , com

$$x(t_0) = x_0 \quad (\text{posição inicial}) \quad (2.1.3)$$

obtemos a “lei horária” do movimento retilíneo uniforme:

$$x(t) = x_0 + v(t - t_0) \quad (2.1.4)$$

Qualquer movimento retilíneo não uniforme chama-se “acelerado”. Podemos estender a (2.1.2) a um movimento acelerado definindo $\bar{v}_{t_1 \rightarrow t_2}$, a *velocidade média entre os instantes t_1 e t_2* , com $x(t_1) = x_1$, $x(t_2) = x_2$, $\Delta x = x_2 - x_1$, $\Delta t = t_2 - t_1$, por

$$\bar{v}_{t_1 \rightarrow t_2} = \frac{x(t_2) - x(t_1)}{t_2 - t_1} = \frac{\Delta x}{\Delta t} \quad (2.1.5)$$

que representa geometricamente, conforme vemos na Figura 2.5, o coeficiente angular (= $\tan \theta$) da corda que liga os extremos 1 e 2 do arco de curva correspondente no gráfico $x \times t$. A velocidade média entre t_1 e t_2 corresponde portanto à velocidade de um movimento uniforme que, partindo de $x(t_1)$ em t_1 , chegasse a $x(t_2)$ em t_2 .

Assim, para um carro que percorresse a estrada São Paulo-Rio (suposta retilínea) em dez horas, a velocidade média entre partida e chegada seria de $400 \text{ km}/10 \text{ h} = 40 \text{ km/h}$. Mas isto informa muito pouco sobre o movimento durante o percurso. O carro poderia ter parado durante algumas horas em algum ponto intermediário, e poderia ter desenvolvido velocidades médias bem maiores em algumas etapas do percurso. Seria bem mais informativo dar o valor de \bar{v} em diferentes etapas do percurso, e isto descreveria tanto melhor o movimento quanto mais curtas as etapas, pois o erro cometido ao aproximar trechos curtos do percurso por movimentos uniformes vai diminuindo à medida que encurtamos esses trechos.

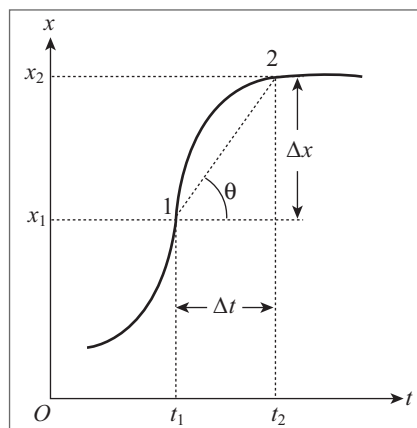


Figura 2.5 Velocidade média.

2.2 VELOCIDADE INSTANTÂNEA

Que significa “velocidade num dado instante t ”?

Para ilustrar este conceito, vamos parafrasear uma anedota utilizada por Feynman em seu curso (veja a Bibliografia no final do livro). Ela tem a forma de um diálogo entre um estudante (E.) que estava dirigindo seu carro de forma a não chegar atrasado na aula de física e o guarda (G.) que o fez parar, acusando-o de excesso de velocidade:

G.: O seu carro estava a 120 km/h , quando o limite de velocidade aqui é de 60 km/h !

- E.:** Como é que eu podia estar a 120 km por hora se só estava dirigindo aqui há cerca de 1 minuto, e não durante uma hora?
- G.:** O que quero dizer é que, se continuasse em frente do jeito que estava, teria percorrido 120 km em uma hora.
- E.:** Se tivesse continuado sempre em frente, eu teria ido bater no prédio da Física!
- G.:** Bem, isso seria verdade se tivesse seguido em frente por uma hora. Mas, se tivesse continuado em frente por 1 minuto, teria percorrido $120 \text{ km}/60 = 2 \text{ km}$, e em 1 s teria percorrido $2 \text{ km}/60 = 33,3 \text{ m}$, e em 0,1 s teria percorrido 3,33 m, e teria dado perfeitamente para prosseguir durante 0,1 s.
- E.:** Mas o limite de velocidade é de 60 km/h, e não de 1,66 m em 0,1s!
- G.:** É a mesma coisa: o que conta é a velocidade instantânea.

[Fizemos apelo a um grau considerável de licença poética nos dotes de **G.** em matéria de conhecimentos de física e de paciência, mas é preciso reconhecer que **E.** também tem um pouco de razão: é permitido exceder o limite de velocidade em intervalos de tempo extremamente curtos, como nas ultrapassagens].

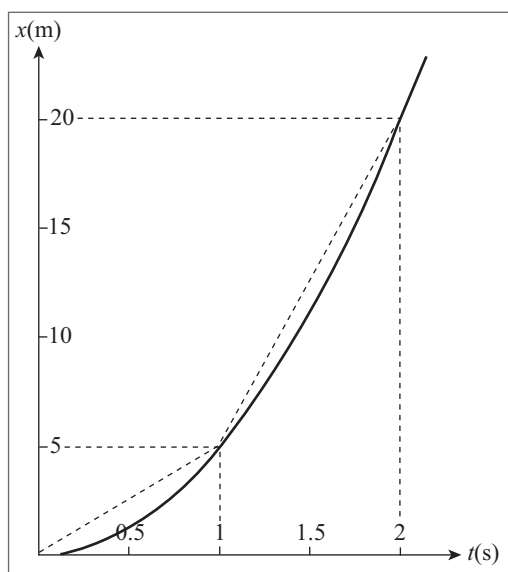


Figura 2.6 Velocidades médias na queda livre.

A velocidade de um carro usualmente não sofre nenhuma alteração apreciável em intervalos de tempo $< 0,1 \text{ s}$, de modo que não é preciso, neste exemplo, tomar intervalos menores. Se necessário, para calcular a velocidade instantânea com precisão cada vez maior, poderíamos considerar o espaço percorrido em 10^{-2} s , 10^{-3} s , ... Quanto menor Δt (e em consequência também o Δx correspondente), mais o valor de $\Delta x/\Delta t$ se aproxima da velocidade instantânea.

Exemplo: Na experiência de queda livre da bolinha (Figura 2.2), o gráfico $x \times t$ tem a forma de uma parábola (Figura 2.6), $x = \alpha t^2$, onde, para x em m e t em s, o valor de α seria $\approx 5 \text{ m/s}^2$; tomemos

$$x(t) = 5t^2 \quad (2.2.1)$$

Qual é a velocidade instantânea para $t = 1 \text{ s}$? Com centro no instante $t = 1 \text{ s}$, calculemos a velocidade média (2.1.5) a partir de instantes anteriores e para instantes posteriores, tomando $\Delta t = 1 \text{ s}$, $0,1 \text{ s}$, $0,01 \text{ s}$, ...

$$\left. \begin{aligned} \bar{v}_{0 \rightarrow 1} &= \frac{x(1) - x(0)}{1 - 0} = \frac{5 - 0}{1 - 0} = 5 \text{ m/s} \\ \bar{v}_{1 \rightarrow 2} &= \frac{x(2) - x(1)}{2 - 1} = \frac{20 - 5}{2 - 1} = 15 \text{ m/s} \end{aligned} \right\} \Delta t = 1 \text{ s}$$

$$\left. \begin{aligned} \bar{v}_{0,9 \rightarrow 1} &= \frac{x(1) - x(0,9)}{1 - 0,9} = \frac{5,00 - 4,05}{1 - 0,9} = 9,5 \text{ m/s} \\ \bar{v}_{1 \rightarrow 1,1} &= \frac{x(1,1) - x(1)}{1,1 - 1} = \frac{6,05 - 5,00}{1,1 - 1} = 10,5 \text{ m/s} \end{aligned} \right\} \Delta t = 0,1 \text{ s}$$

$$\left. \begin{aligned} \bar{v}_{0,99 \rightarrow 1} &= \frac{x(1) - x(0,99)}{1 - 0,99} = \frac{5,0000 - 4,9005}{1,00 - 0,99} = 9,95 \text{ m/s} \\ \bar{v}_{1 \rightarrow 1,01} &= \frac{x(1,01) - x(1)}{1,01 - 1,00} = \frac{5,1005 - 5,0000}{1,01 - 1,00} = 10,05 \text{ m/s} \end{aligned} \right\} \Delta t = 0,01 \text{ s}$$

Como a parábola é uma curva côncava para cima, o coeficiente angular da corda que liga dois pontos da curva vai aumentando à medida que subimos na curva, de forma que a sequência acima deve representar aproximações alternadamente por falta e por excesso da velocidade instantânea v para $t = 1$ s: $5 \text{ m/s} < v < 15 \text{ m/s}$; $9,5 \text{ m/s} < v < 10,5 \text{ m/s}$; $9,95 \text{ m/s} < v < 10,05 \text{ m/s}$, ... o que sugere qual deve ser o valor de v :

$$v = 10 \text{ m/s para } t = 1 \text{ s}$$

Este valor deveria ser obtido como caso limite da sequência quando $\Delta t \rightarrow 0$. Com efeito,

$$\Delta x = x(1 + \Delta t) - x(1) = 5(1 + \Delta t)^2 - 5 = 5[1 + 2\Delta t + (\Delta t)^2] - 5 = 10\Delta t + 5(\Delta t)^2,$$

$$\bar{v}_{1 \rightarrow 1} = \frac{x(1 + \Delta t) - x(1)}{1 + \Delta t - 1} = \frac{\Delta x}{\Delta t} = \frac{10\Delta t + 5(\Delta t)^2}{\Delta t} = 10 + 5\Delta t \rightarrow 10 \text{ quando } \Delta t \rightarrow 0$$

Note que, quando $\Delta t \rightarrow 0$, também $\Delta x \rightarrow 0$, mas o quociente $\Delta x/\Delta t$ tende a um valor finito, $= 10 \text{ m/s}$ no exemplo acima.

Para uma função $x(t)$, o limite

$$\lim_{\Delta t \rightarrow 0} \left[\frac{x(t_0 + \Delta t) - x(t_0)}{\Delta t} \right] = \lim_{\Delta t \rightarrow 0} \left(\frac{\Delta x}{\Delta t} \right) \Big|_{t=t_0} = \left(\frac{dx}{dt} \right)_{t=t_0} \quad (2.2.2)$$

chama-se *derivada de x em relação a t no ponto t_0* . Note que dx e dt são notações.

No exemplo acima da função (2.2.1), obtivemos:

$$\left(\frac{dx}{dt} \right) \Big|_{t=1} = 10.$$

A notação “lim”(limite) para $\Delta t \rightarrow 0$ significa que podemos nos aproximar tanto quanto quisermos do resultado exato tomando Δt suficientemente pequeno, como fizemos nos cálculos numéricos acima. O limite nem sempre existe para qualquer função de t ;

quando existe, a função chama-se *diferenciável* no ponto t_0 . Geralmente, estaremos lidando com funções diferenciáveis.

Exemplo: Calcular a derivada de

$$x(t) = at^2 + bt + c \quad (2.2.3)$$

onde a , b e c são constantes, num ponto t qualquer.

Temos

$$x(t + \Delta t) = a(t + \Delta t)^2 + b(t + \Delta t) + c = a[t^2 + 2t\Delta t + (\Delta t)^2] + b(t + \Delta t) + c$$

$$\Delta x = x(t + \Delta t) - x(t) = 2at\Delta t + a(\Delta t)^2 + b\Delta t$$

$$\frac{\Delta x}{\Delta t} = 2at + a\Delta t + b; \quad \lim_{\Delta t \rightarrow 0} \left(\frac{\Delta x}{\Delta t} \right) = 2at + b$$

ou seja

$$\frac{dx}{dt} = 2at + b \quad (2.2.4)$$

Este exemplo também ilustra os seguintes resultados imediatos: a derivada de uma constante é nula; a derivada de uma soma é a soma das derivadas;

$$\frac{d}{dt}[ax(t)] = a \frac{dx}{dt} \quad (a = \text{constante}) \quad (2.2.5)$$

O resultado anterior para a (2.2.1) é um caso particular, com $b = c = 0$; $a = 5$; $t = 1$.

A velocidade instantânea $v(t)$ num instante t qualquer num movimento descrito por $x = x(t)$ é dada por

$$v(t) = \frac{dx}{dt} \quad (2.2.6)$$

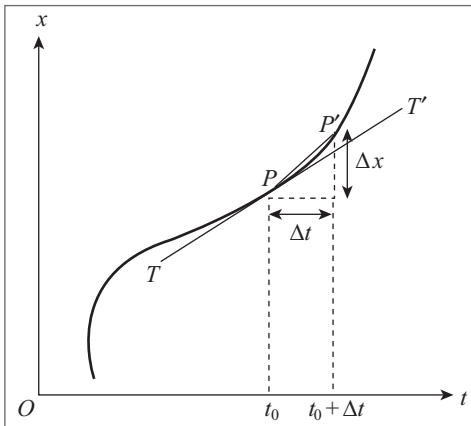


Figura 2.7 Interpretação geométrica da derivada.

Da mesma forma que a velocidade média, a velocidade instantânea também tem uma interpretação geométrica simples no gráfico $x \times t$. Vimos que $\bar{v}_{t_0 \rightarrow t_0 + \Delta t}$ é o coeficiente angular da corda PP' que liga os pontos P e P' do gráfico associados aos instantes t_0 e $t_0 + \Delta t$ (Figura 2.7). À medida que $\Delta t \rightarrow 0$, P' se aproxima de P e $\Delta x / \Delta t$ tende ao coeficiente angular da *tangente* TT' à curva no ponto P . Logo, a *velocidade instantânea* $v(t_0)$ representa o coeficiente angular da tangente ao gráfico $x \times t$ no ponto t_0 ; é o que se chama de “declive” da curva neste ponto. Esta é também, de forma mais geral, a interpretação geométrica da derivada dx/dt ; ela mede a “taxa de variação” de x com t .

A interpretação geométrica da derivada mostra imediatamente (Figura 2.8) que $dx/dt > 0$ num ponto onde x está crescendo com t , $dx/dt < 0$ num ponto onde x está decrescendo com t (marcha à ré, no exemplo do carro), e $dx/dt = 0$ quando a curva tem tangente horizontal no ponto considerado (pode ser um máximo ou um mínimo ou um ponto de inflexão \sphericalangle)

Quanto mais rapidamente x está crescendo com t , mais abrupta é a curva, e maior é portanto a velocidade instantânea, o que concorda com a ideia intuitiva.

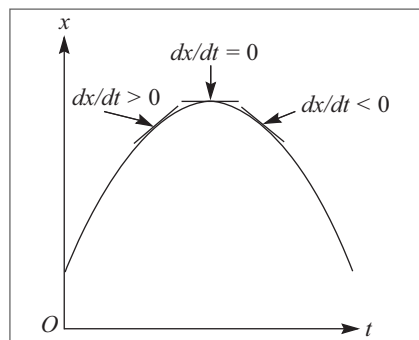


Figura 2.8 Sinal da derivada.

2.3 O PROBLEMA INVERSO

Vimos como, conhecendo a “lei horária” de um movimento, ou seja, a função $x = x(t)$, é possível calcular a velocidade instantânea $v(t)$ no decurso do movimento: basta tomar dx/dt . Assim, por exemplo, para a lei horária (2.2.3), a velocidade é dada pela (2.2.4).

Frequentemente temos de resolver o problema inverso: conhecendo a velocidade instantânea $v(t)$ entre um dado instante inicial t_1 , e um instante final t_2 , calcular o espaço percorrido entre estes dois instantes, ou seja, $x(t_2) - x(t_1)$. Poderíamos pensar num filme do painel de instrumentos de um automóvel que mostrasse simultaneamente o velocímetro e um relógio, permitindo traçar o gráfico de $v \times t$ entre t_1 e t_2 (tomamos sempre $t_2 > t_1$).

Se o movimento for uniforme, como na (2.1.4), velocidade instantânea e velocidade média se confundem, $v = \bar{v} = \text{constante}$, e o gráfico é uma reta paralela ao eixo das abscissas (Figura 2.9). Pela definição de velocidade média, o espaço percorrido entre t_1 e t_2 é:

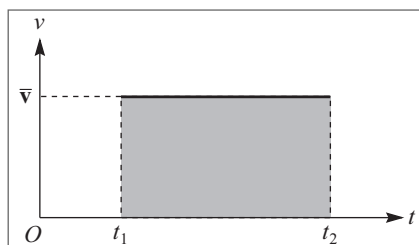


Figura 2.9 Interpretação geométrica do espaço percorrido.

$$(\Delta x)_{t_1 \rightarrow t_2} \equiv x(t_2) - x(t_1) = \bar{v}_{t_1 \rightarrow t_2} \Delta t = \bar{v} \Delta t = v(t_2 - t_1) \quad (2.3.1)$$

que, conforme mostra a Figura 2.9, tem uma interpretação geométrica simples: é a *área* da porção do gráfico $v \times t$ situada entre o gráfico e o eixo das abscissas e limitada pelas ordenadas em t_1 e t_2 .

Note-se que a “área” assim definida pode ser positiva ou negativa, conforme seja $\bar{v} > 0$ ou $\bar{v} < 0$ (Figura 2.10), ou seja, *uma área situada abaixo do eixo Ot tem de ser tomada como negativa* (o que significa simplesmente $x(t_2) < x(t_1)$, ou seja, movimento para trás).

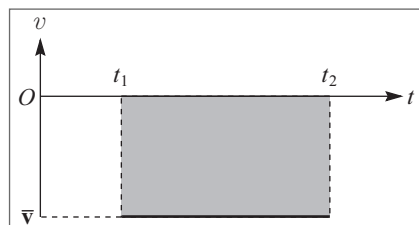


Figura 2.10 Área negativa.

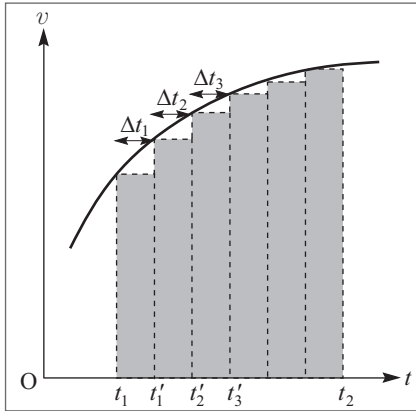


Figura 2.11 Espaço percorrido num movimento qualquer.

Consideremos agora um movimento não uniforme, em que v é uma função qualquer de t (Figura 2.11). Imaginemos o intervalo $[t_1, t_2]$ subdividido em um grande número de pequenos intervalos de larguras $\Delta t_1, \Delta t_2, \Delta t_3, \dots$ por pontos de subdivisão t'_1, t'_2, t'_3, \dots (Figura), onde $t'_1 = t_1 + \Delta t_1$; $t'_2 = t'_1 + \Delta t_2$; $t'_3 = t'_2 + \Delta t_3$; ... Se os intervalos Δt_i ($i = 1, 2, 3, \dots$) forem suficientemente pequenos, a velocidade variará muito pouco em cada um desses intervalos, e podemos calcular a distância percorrida em cada um aproximando a velocidade média nele pela velocidade num de seus pontos, por exemplo, o extremo esquerdo de cada intervalo:

$$\begin{aligned}\Delta x_{t_1 \rightarrow t'_1} &= x(t'_1) - x(t_1) = \bar{v}_{t_1 \rightarrow t'_1} \Delta t_1 \approx v(t_1) \Delta t_1 \\ \Delta x_{t'_1 \rightarrow t'_2} &= x(t'_2) - x(t'_1) = \bar{v}_{t'_1 \rightarrow t'_2} \Delta t_2 \approx v(t'_1) \Delta t_2 \\ \Delta x_{t'_2 \rightarrow t'_3} &= x(t'_3) - x(t'_2) = \bar{v}_{t'_2 \rightarrow t'_3} \Delta t_3 \approx v(t'_2) \Delta t_3\end{aligned}$$

Somando membro a membro estas 3 relações, obtemos o deslocamento total entre t_1 e t'_3 :

$$x(t'_3) - x(t_1) \approx v(t_1) \Delta t_1 + v(t'_1) \Delta t_2 + v(t'_2) \Delta t_3$$

e é claro que, se prosseguirmos até t_2 , obteremos a soma das contribuições de todos os subintervalos em que $[t_1, t_2]$ foi dividido:

$$x(t_2) - x(t_1) \approx \sum_i v(t'_i) \Delta t'_i \quad (2.3.2)$$

Graficamente, conforme mostra a Figura 2.11, cada termo da soma é a área de um retângulo, e a soma (2.3.2) é a área compreendida entre o eixo Ot e uma linha poligonal “em escada” inscrita na curva $v \times t$ entre t_1 e t_2 .

A soma (2.3.2) se aproxima tanto mais do resultado exato quanto menores forem as subdivisões Δt_i . Logo, no limite em que os $\Delta t'_i$ tendem a zero, devemos obter:

$$x(t_2) - x(t_1) = \lim_{\Delta t'_i \rightarrow 0} \sum_i v(t'_i) \Delta t'_i = \text{Área entre a curva } v \times t \text{ e o eixo } Ot, \text{ de } t_1 \text{ a } t_2 \quad (2.3.3)$$

O limite (2.3.3) é chamado de *integral definida* de $v(t)$ entre os extremos t_1 e t_2 , e representado pela notação

$$\lim_{\Delta t'_i \rightarrow 0} \sum_i v(t'_i) \Delta t'_i \equiv \int_{t_1}^{t_2} v(t) dt \quad (2.3.4)$$

O símbolo \int de “integral” é uma deformação do S de “Soma”; t_1 e t_2 são, respectivamente, o *extremo inferior* e o *extremo superior* da integral. A função $v(t)$ sob o sinal de \int chama-se o *integrando*. Note que t tem no integrando um papel análogo ao do índice i na soma (2.3.4): é a *variável de integração*, e pode ser representada por qualquer outra letra (t' , u , v), da mesma forma que podemos chamar o índice de soma de j , k , ..., em lugar de i .

Métodos de cálculo de integrais serão vistos no curso de Cálculo Diferencial e Integral; do ponto de vista do cálculo aproximado, a (2.3.3) mostra que o problema se reduz a calcular a área compreendida entre uma curva e o eixo das abscissas (levando em conta que áreas situadas abaixo do eixo devem ser contadas como negativas), o que pode ser feito aproximadamente traçando a curva em papel quadriculado e contando quadrículas.

Como aplicação, consideremos um movimento cuja velocidade $v(t)$ é dada pela (2.2.4):

$$v(t) = 2at + b \quad (2.3.5)$$

A área a calcular neste caso é a do trapézio sombreado na Figura 2.12.

Sejam

$$v(t_1) = 2at_1 + b \equiv v_1$$

$$v(t_2) = 2at_2 + b \equiv v_2$$

Temos então, pela (2.3.3), $x(t_2) - x(t_1) = \text{Área do trapézio} = (\text{Semissoma das bases}) \times \text{Altura} = \frac{1}{2}(v_1 + v_2)(t_2 - t_1)$ o que, comparando com a (2.1.5), implica que, neste movimento,

$$\bar{v}_{t_1 \rightarrow t_2} = \frac{1}{2}[v(t_1) + v(t_2)] \quad (2.3.6)$$

ou seja, que a velocidade média num intervalo é a média aritmética das velocidades nos extremos do intervalo. Substituindo na (2.3.6) os valores de $v(t_1)$ e $v(t_2)$, vem

$$\bar{v}_{t_1 \rightarrow t_2} = a(t_2 + t_1) + b$$

o que dá, pela (2.3.1),

$$x(t_2) - x(t_1) = a(t_2^2 - t_1^2) + b(t_2 - t_1) \quad (2.3.7)$$

que coincide com o resultado obtido a partir da lei horária (2.2.3) (e dá o valor da integral definida (2.3.4) quando $v(t)$ é dado pela (2.3.5)).

A (2.3.7) pode ser aplicada, em particular, tomando para t_1 o instante inicial $t_1 = 0$, e para t_2 um instante genérico t . Chamando $x(0) = c$ (valor inicial de x), a (2.3.7) dá então

$$x(t) = x(0) + at^2 + bt = at^2 + bt + c \quad (2.3.8)$$

ou seja, este processo de “integração” nos permitiu recuperar a lei horária (2.2.3) a partir da expressão (2.2.4) da velocidade e do valor inicial de x .

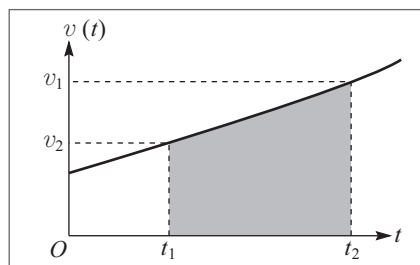


Figura 2.12 Área do trapézio limitado pela reta (2.3.5).

Matematicamente, a (2.2.4) se chama uma *equação diferencial* para a função incógnita $x(t)$ (porque a derivada da função incógnita aparece na equação). Passamos da (2.2.4) à (2.3.8) *integrando a equação diferencial* com a *condição inicial* $x(0) = c$.

2.4 ACELERAÇÃO

Temos todos uma noção intuitiva do conceito de “aceleração” (por exemplo, o efeito do acelerador num automóvel), como medida da rapidez de variação da velocidade com o tempo. Assim, dizemos que um carro tem “boa aceleração” se é capaz de acelerar de 0 a 120 km/h em 10 s. Conforme vemos neste exemplo, a aceleração mede a “velocidade de variação da velocidade”. Por analogia com a (2.1.5), podemos definir primeiro a *aceleração média* no intervalo $[t_1, t_2]$ por

$$\bar{a}_{t_1 \rightarrow t_2} \equiv \frac{v(t_2) - v(t_1)}{t_2 - t_1} = \frac{\Delta v}{\Delta t} \quad (2.4.1)$$

Assim, no exemplo acima do carro, a aceleração média no intervalo de 0 s a 10 s seria de

$$\frac{33,3 \text{ m/s}}{10 \text{ s}} = 3,33 \text{ m/s}^2$$

ilustrando o fato de que, se tomarmos o metro como unidade de comprimento e o segundo como unidade de tempo, a unidade de aceleração é $1 \text{ m} \cdot \text{s}^{-2}$.

Tomando sempre $t_2 > t_1$, vemos que a aceleração média é *positiva* quando v cresce de t_1 para t_2 , e *negativa* quando decresce; se $v > 0$, v cresce ou decresce conforme $|v|$ cresça ou decresça, mas se $v < 0$ é o contrário: v cresce quando $|v|$ decresce. Assim, no exemplo do carro, em marcha à frente, a aceleração é negativa quando o carro está freiando, mas em marcha à ré é o contrário: freiar em marcha à ré corresponde a uma aceleração positiva.

A aceleração média pode geralmente ser variável durante o movimento, e considerações análogas às da Seção 2.2 levam-nos a definir a *aceleração instantânea* $a(t)$ num instante t por (cf. (2.2.2) e (2.2.5))

$$a(t) = \lim_{\Delta t \rightarrow 0} \left[\frac{v(t + \Delta t) - v(t)}{\Delta t} \right] = \lim_{\Delta t \rightarrow 0} \frac{\Delta v}{\Delta t} = \frac{dv}{dt} \quad (2.4.2)$$

ou seja, a *aceleração instantânea* é a *derivada em relação ao tempo da velocidade instantânea*.

Substituindo $v(t)$ na (2.4.2) pela (2.2.5), obtemos

$$a(t) = \frac{d}{dt} \left(\frac{dx}{dt} \right) = \frac{d^2 x}{dt^2} \quad (2.4.3)$$

onde introduzimos a definição de *derivada segunda de x em relação a t* , indicada pela notação $d^2 x/dt^2$.

A interpretação geométrica da derivada (Seção 2.2) se aplica à (2.4.2): num gráfico $v \times t$, $a(t)$ é o coeficiente angular da tangente à curva no ponto correspondente ao instante t .

Consideremos por exemplo o seguinte gráfico $x \times t$, que poderia representar o movimento do automóvel no exemplo da Seção 2.1:

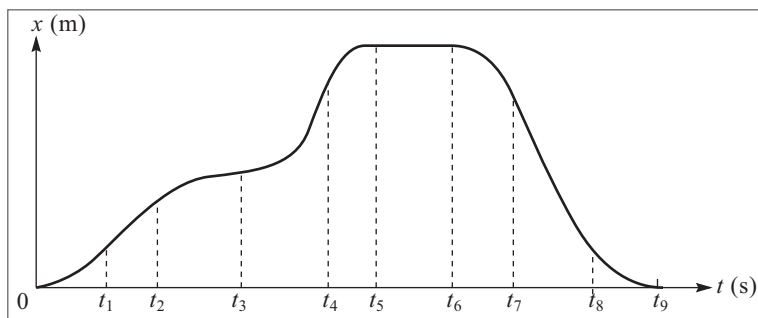


Figura 2.13 Gráfico da posição.

Note que o carro parte da origem em $t = 0$ e acaba regressando à origem em $t = t_9$.

Com o auxílio da interpretação geométrica da derivada como coeficiente angular da tangente à curva, podemos esboçar pelo menos qualitativamente o gráfico da velocidade instantânea associada a esse movimento:

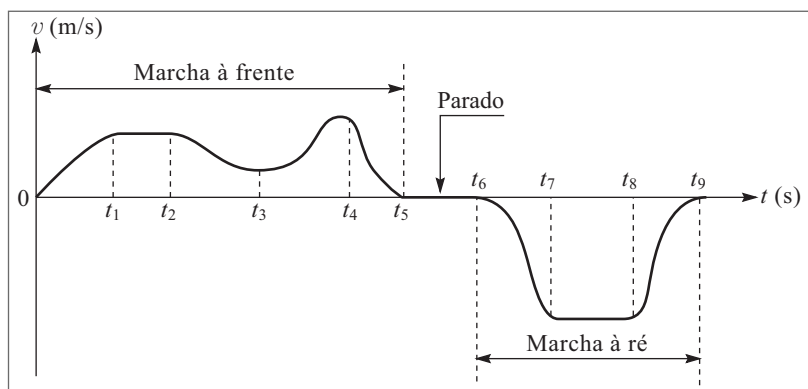


Figura 2.14 Gráfico da velocidade.

Como $x(t_9) - x(0) = 0$, a (2.3.3) implica que a área total entre a curva e o eixo Ot é $= 0$, ou seja, que a área positiva (acima do eixo) é exatamente cancelada pela área negativa (abaixo do eixo).

O gráfico da aceleração instantânea se obtém de forma análoga do gráfico $v \times t$:

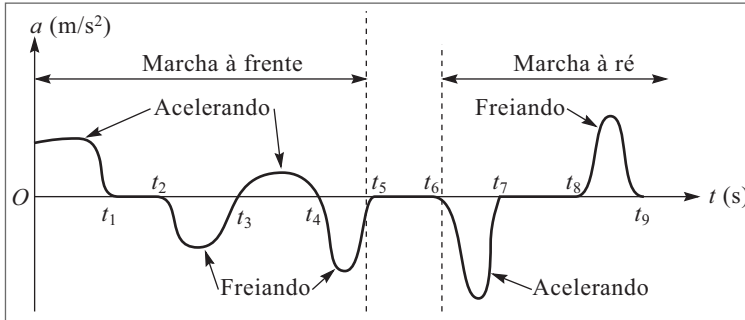


Figura 2.15 Gráfico da aceleração.

Note a correlação entre o sinal de $a(t)$ e a interpretação em termos de acelerar ou freiar o carro, que é diferente para marcha à frente e marcha à ré, conforme a discussão acima.

Aqui também podemos considerar o problema inverso, de determinar a variação de velocidade entre dois instantes, conhecendo $a(t)$. A solução se obtém imediatamente das (2.3.3) – (2.3.4), bastando trocar $x \rightarrow v$, $v \rightarrow a$:

$$v(t_2) - v(t_1) = \int_{t_1}^{t_2} a(t) dt \quad (2.4.4)$$

que também se interpreta graficamente em termos da área entre a curva de $v(t)$ e o eixo Ot . No exemplo acima, como $v(0) = v(t_5) = 0$, a área negativa abaixo de Ot entre $t = 0$ e t_5 deve cancelar exatamente a área acima de Ot (no gráfico $a \times t$), e o mesmo vale para o intervalo $[t_6, t_9]$.

2.5 MOVIMENTO RETILÍNEO UNIFORMEMENTE ACELERADO

Um movimento retilíneo chama-se *uniformemente acelerado* quando a aceleração instantânea é constante (independentemente do tempo):

$$\boxed{\frac{dv}{dt} = \frac{d^2x}{dt^2} = a = \text{constante}} \quad (2.5.1)$$

Podemos usar as técnicas de solução do “problema inverso” (Seção 2.3) para determinar a lei horária de um movimento uniformemente acelerado.

Para isto, consideremos o movimento durante um intervalo de tempo $[t_0, t]$, onde t_0 é o “instante inicial” (frequentemente se toma $t_0 = 0$).

A (2.4.4) dá:

$$v(t) - v(t_0) = \int_{t_0}^t a dt = a(t - t_0) \quad (2.5.2)$$

que é a área do retângulo sombreado na Figura 2.16 (compare com a (2.3.1)).

O valor

$$v(t_0) \equiv v_0 \quad (2.5.3)$$

da velocidade no instante inicial chama-se *velocidade inicial*. A (2.5.2) dá então

$$v(t) = v_0 + a(t - t_0) \quad (2.5.4)$$

mostrando que a velocidade é uma função linear do tempo no movimento uniformemente acelerado. Esta é precisamente a situação já analisada no caso da (2.3.5), de forma que o resultado (2.3.6) vale para qualquer movimento uniformemente acelerado: a velocidade média num intervalo é a média aritmética das velocidades nos extremos do intervalo.

Poderíamos obter a lei horária simplesmente adaptando a (2.3.7) à notação da (2.5.4) (em particular, $2a$ na (2.3.5) corresponde a a na (2.5.4), mas é instrutivo recalculer o resultado de forma um pouco diferente. Pelas (2.3.3) e (2.3.4),

$$x(t) - x(t_0) = \int_{t_0}^t v(t') dt' \quad (2.5.5)$$

onde chamamos de t' a variável de integração (veja a discussão após a (2.3.4)), para evitar confusão com t , o extremo superior da integral. A área do trapézio, conforme mostra a Figura 2.17, pode também ser calculada como a soma da área do retângulo sombreado, que é $v_0(t - t_0)$, com a área do triângulo sombreado, que é

$$\frac{1}{2} a(t - t_0) \cdot (t - t_0)$$

ou seja

$$x(t) - x(t_0) = v_0(t - t_0) + \frac{1}{2} a(t - t_0)^2 \quad (2.5.6)$$

Analogamente à (2.5.3), definimos

$$x(t_0) \equiv x_0 \quad (2.5.7)$$

como a *posição inicial*. A (2.5.6) dá então finalmente a *lei horária do movimento retilíneo uniformemente acelerado*,

$$x(t) = x_0 + v_0(t - t_0) + \frac{1}{2} a(t - t_0)^2 \quad (2.5.8)$$

em função dos *valores iniciais* x_0 e v_0 da posição e da velocidade no instante inicial t_0 .

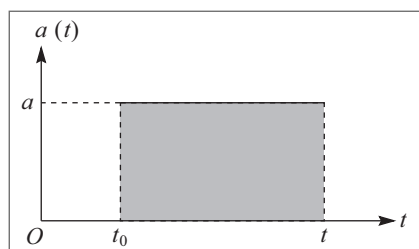


Figura 2.16 Variação da velocidade como área.

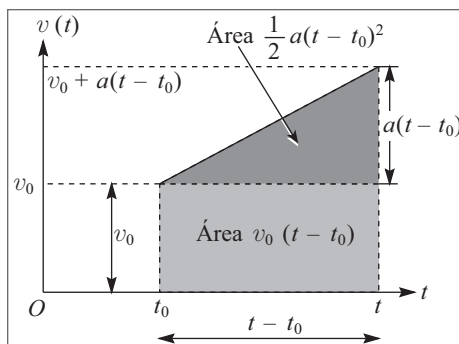


Figura 2.17 Integração gráfica.

Do ponto de vista matemático, a passagem da (2.5.1) à (2.5.8) corresponde à “integração” da equação diferencial de 2ª ordem (2.5.1) para a função incógnita $x(t)$ (de segunda ordem porque entra a derivada segunda d^2x/dt^2), com as condições iniciais (2.5.3) e (2.5.7). O gráfico $x \times t$ de um movimento uniformemente acelerado é uma parábola.

Frequentemente interessa também exprimir a velocidade no movimento uniformemente acelerado em função da posição x (em lugar do tempo t). Para obter esta expressão, basta substituir a (2.5.4) na (2.5.8), eliminando $t - t_0$:

$$t - t_0 = \frac{v - v_0}{a} \Rightarrow x - x_0 = v_0 \left(\frac{v - v_0}{a} \right) + \frac{a}{2} \left(\frac{v - v_0}{a} \right)^2 = \frac{(v - v_0)(v + v_0)}{2a} = \frac{v^2 - v_0^2}{2a}$$

ou seja,

$$\boxed{v^2 = v_0^2 + 2a(x - x_0)} \quad (2.5.9)$$

que é a expressão procurada.

Exemplo: Um motorista freia seu carro uniformemente, de tal maneira que a velocidade cai de 60 km/h a 30 km/h em 5 s. Que distância o carro ainda percorrerá depois disso até parar, e quanto tempo levará para percorrer essa distância adicional?

Como o freiamento é uniforme, a aceleração instantânea e média se confundem; podemos calcular o seu valor pela (2.5.4)

$$\left\{ \begin{array}{l} t - t_0 = 5 \text{ s} \\ v_0 = 60 \text{ km/h} = 16,66 \text{ m/s} \\ v = 30 \text{ km/h} = 8,33 \text{ m/s} \end{array} \right\} a = \frac{v - v_0}{t - t_0} = -\frac{8,33}{5} \text{ m/s}^2 = -1,66 \text{ m/s}^2,$$

onde o sinal negativo está associado ao freiamento.

Para calcular a distância que o carro percorrerá até parar depois de atingir 30 km/h, podemos aplicar a (2.5.9) com $v_0 = 30 \text{ km/h} = 8,33 \text{ m/s}$ e $v = 0$ (velocidade final), o que dá

$$x - x_0 = \frac{v^2 - v_0^2}{2a} = \frac{0 - (8,33)^2}{-3,33} \text{ m} = 20,83 \text{ m}.$$

O tempo que o carro levará para percorrer esta distância adicional se calcula de novo pela (2.5.4), tomando $v_0 = 30 \text{ km/h}$ e $v = 0$:

$$t - t_0 = \frac{v - v_0}{a} = \frac{0 - 8,33}{-1,66} \text{ s} = 5 \text{ s}.$$

Logo, o carro leva mais 5 s para parar depois de freiado.

2.6 GALILEU E A QUEDA DOS CORPOS

O exemplo mais familiar de movimento retilíneo uniformemente acelerado é a queda livre de um corpo solto em repouso. Este foi um dos problemas analisados por Galileu em seus trabalhos, que deram origem à era moderna da física.

Chapter 2

Tools for Doing Simulations

We introduce some of the core syntax of Java in the context of simulating the motion of falling particles near the Earth's surface. A simple algorithm for solving first-order differential equations numerically is also discussed.

2.1 Introduction

If you were to take a laboratory-based course in physics, you would soon be introduced to the oscilloscope. You would learn the function of many of the knobs, how to read the display, and how to connect various devices so that you could measure various quantities. If you did not know already, you would learn about voltage, current, impedance, and AC and DC signals. Your goal would be to learn how to use the oscilloscope. In contrast, you would learn only a little about the inner workings of the oscilloscope.

The same approach can be easily adopted with an object-oriented language such as Java. If you are new to programming, you will learn how to make Java do what you want, but you will not learn everything about Java. In this chapter, we will present some of the essential syntax of Java and introduce the Open Source Physics library, which will facilitate writing programs with a graphical user interface and visual output such as plots and animations.

One of the ways that science progresses is by making models. If the model is sufficiently detailed, we can determine its behavior and then compare the behavior with experiment. This comparison might lead to verification of the model, changes in the model, and further simulations and experiments. In the context of computer simulation, we usually begin with a set of initial conditions, determine the dynamical behavior of the model numerically, and generate data in the form of tables of numbers, plots, and animations. We begin with a simple example to see how this process works.

Imagine a particle such as a ball near the surface of the Earth subject to a single force, the force of gravity. We assume that air friction is negligible, and the gravitational force is given by

$$F_g = -mg \quad (2.1)$$

where m is the mass of the ball and $g = 9.8 \text{ N/kg}$ is the gravitational field (force per unit mass) near the Earth's surface. To make our example as simple as possible, we first assume that there is only vertical motion. We use Newton's second law to find the motion of the ball,

$$m \frac{d^2 y}{dt^2} = F \quad (2.2)$$

where y is the vertical coordinate defined so that up is positive, t is the time, F is the total force on the ball, and m is the inertial mass [which is the same as the gravitational mass in (2.1)]. If we set $F = F_g$, (2.1) and (2.2) lead to

$$\frac{d^2y}{dt^2} = -g. \quad (2.3)$$

Equation (2.3) is a statement of a model for the motion of the ball. In this case the model is in the form of a second-order differential equation.

You are probably familiar with the model summarized in (2.3) and know the analytic solution:

$$y(t) = y(0) + v(0)t - \frac{1}{2}gt^2 \quad (2.4a)$$

$$v(t) = v(0) - gt. \quad (2.4b)$$

Nevertheless, we will determine the motion of a freely falling particle numerically in order to introduce the tools that we will need in a familiar context.

We begin by expressing (2.3) as two first-order differential equations:

$$\frac{dy}{dt} = v \quad (2.5a)$$

$$\frac{dv}{dt} = -g \quad (2.5b)$$

where v is the vertical velocity of the ball. We next *approximate* the derivatives by small (finite) differences:

$$\frac{y(t + \Delta t) - y(t)}{\Delta t} = v(t) \quad (2.6a)$$

$$\frac{v(t + \Delta t) - v(t)}{\Delta t} = -g. \quad (2.6b)$$

Note that in the limit $\Delta t \rightarrow 0$, (2.6) reduces to (2.5). We can rewrite (2.6) as

$$y(t + \Delta t) = y(t) + v(t)\Delta t \quad (2.7a)$$

$$v(t + \Delta t) = v(t) - g\Delta t. \quad (2.7b)$$

The finite difference approximation we used to obtain (2.7) is an example of the *Euler* algorithm. Equation (2.7) is an example of a *finite difference* equation, and Δt is the time step.

Now we are ready to follow $y(t)$ and $v(t)$ in time. We begin with an initial value for y and v and then *iterate* (2.7). If Δt is sufficiently small, we will obtain a numerical answer that is close to the solution of the original differential equations in (2.6). In this case we know the answer, and we can test our numerical results directly.

Exercise 2.1. A simple example

Consider the first-order differential equation

$$\frac{dy}{dx} = f(x) \quad (2.8)$$

where $f(x)$ is a function of x . The approximate solution as given by the Euler algorithm is

$$y_{n+1} = y_n + f(x_n)\Delta x. \quad (2.9)$$

Note that the rate of change of y has been approximated by its value at the *beginning* of the interval, $f(x_n)$.

- (a) Suppose that $f(x) = 2x$ and $y(x = 0) = 0$. The analytic solution is $y(x) = x^2$, which we can confirm by taking the derivative of $y(x)$. Convert (2.8) into a finite difference equation using the Euler algorithm. For simplicity, choose $\Delta x = 0.1$. It would be a good idea to first use a calculator or pencil and paper to determine y_n for the first several time steps.
- (b) Sketch the difference between the exact solution and the approximate solution given by the Euler algorithm. What condition would the rate of change $f(x)$ have to satisfy for the Euler algorithm to give the exact answer? \square

Problem 2.2. Invent your own numerical algorithm

As we have mentioned, the Euler algorithm evaluates the rate of change of y by its value at the beginning of the interval, $f(x_n)$. The choice of where to approximate the rate of change of y during the interval from x to $x + \Delta x$ is arbitrary, although we will learn that some choices are better than others. All that is required is that the finite difference equation must reduce to the original differential equation in the limit $\Delta x \rightarrow 0$. Think of several other algorithms that are consistent with this condition. \square

2.2 Simulating Free Fall

The source code for the *class* `FirstFallingBallApp` shown in Listing 2.1 is defined in a file named `FirstFallingBallApp.java`. The code consists of a sequence of *statements* that create *variables* and define *methods*. Each statement ends with a semicolon. Each source code file is compiled into *byte code* that can then be executed. The compiler places the byte code in a file with the same name as the Java source code file with the extension `class`. For example, the compiler converts `FirstFallingBallApp.java` into byte code and produces the `FirstFallingBallApp.class` file. One of the features of Java is that this byte code can be used by any computer that can run Java programs.

A Java application is a class that contains a *main* method. The following application is an implementation of the Euler algorithm given in (2.7). The program also compares the numerical and analytic results. We will next describe the syntax used in each line of the program.

Listing 2.1: First version of a simulation of a falling particle.

```

1 // example of a single line comment statement (ignored by compiler)
2 package org.opensourcephysics.sip.ch02;           // location of file
3 // beginning of class definition
4 public class FirstFallingBallApp {
5     // beginning of method definition
6     public static void main(String[] args) {
7         // braces { } used to group statements.
8         // indent statements within a block so that
9         // they can be easily identified
10        // following statements form the body of main method
11        // example of declaration and assignment statement
12        double y0 = 10;
13        double v0 = 0;    // initial velocity
14        double t = 0;    // time
15        double dt = 0.01; // time step
16        double y = y0;
17        double v = v0;
18        double g = 9.8;  // gravitational field

```

4

Os princípios da dinâmica

4.1 FORÇAS EM EQUILÍBRIO

Até aqui discutimos somente a *descrição* de movimentos, sem nos preocuparmos com a determinação do tipo de movimento que ocorrerá em dadas circunstâncias físicas. Essa determinação constitui o problema fundamental da dinâmica.

Os princípios básicos da dinâmica foram formulados por Galileu e por Newton. Procuraremos chegar a eles, baseando-nos o máximo possível em noções intuitivas. Sabemos todos por experiência que o movimento é afetado pela ação do que costumamos chamar de “forças”. Nossa ideia intuitiva de forças está relacionada com o esforço muscular, e sabemos que, exercendo “forças” desse tipo, somos capazes de colocar objetos em movimento ou, mais geralmente, alterar seu estado de movimento.

Historicamente, as forças e seus efeitos foram analisadas primeiro em situações “estáticas”, ou seja, de equilíbrio, e vamos inicialmente considerar situações desse tipo afim de formular um método (que será ainda provisório) de medir o efeito de uma força. Também vamos nos limitar, por enquanto, a forças aplicadas a uma partícula, ou seja, um corpo de dimensões desprezíveis.

Dizemos que uma partícula que permanece em repouso em relação a um dado referencial está em *equilíbrio* nesse referencial. Por enquanto, podemos pensar no referencial como estando ligado ao laboratório; mais tarde, voltaremos a discutir o problema da escolha do referencial.

Podemos medir o efeito de uma força aplicada a uma partícula P, conforme mostra a Figura 4.1, pela distensão que ela produz puxando uma mola, presa rigidamente pela outra extremidade a um suporte fixo. A posição indicada por um ponteiro ligado à mola permite graduar uma escala, cuja indicação “0” corresponde à posição do ponteiro antes da aplicação da força. Podemos provisoriamente definir a unidade de força nesta escala de forma

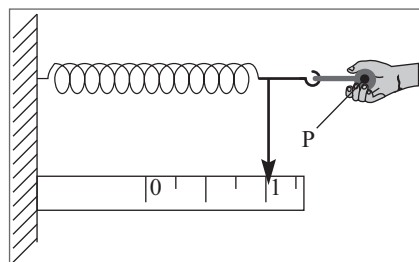


Figura 4.1 Distensão de uma mola.

bastante arbitrária, associando-lhe a graduação “1” na escala (é preciso que a força não seja tão grande que produza uma deformação permanente da mola, ou seja, que ela volte ao “0” quando a soltarmos, removendo a força). Diremos então, por exemplo, que duas pessoas diferentes produzem “a mesma força” sobre a partícula quando levam o ponteiro à mesma posição de equilíbrio sobre a escala.

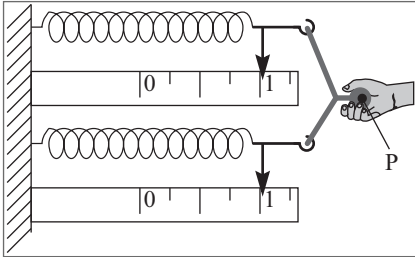


Figura 4.2 Força dupla.

A Figura 4.2 mostra como podemos definir uma força de *duas* unidades na escala acima, utilizando duas molas idênticas, cada uma das quais sofre uma distensão correspondente a uma unidade de força. Podemos definir analogamente outros múltiplos da unidade de força.

Uma força produz efeitos diferentes conforme a direção e sentido em que é aplicada, o que sugere uma representação de tipo vetorial. Na Figura 4.3

(b), \mathbf{F}_1 , \mathbf{F}_2 e \mathbf{F}_3 representam as forças aplicadas à partícula P em (a), em magnitude (medida pela distensão das molas), direção e sentido. É um fato experimental que a partícula P permanece em equilíbrio sob a ação simultânea dessas três forças quando a resultante (vetorial!) das três forças se anula (polígono fechado na Figura 4.3 (b)).

$$\mathbf{F}_1 + \mathbf{F}_2 + \mathbf{F}_3 = 0 \quad (4.1.1)$$

A experiência mostra portanto que as forças se combinam como vetores, e a condição de equilíbrio (resultante nula) permanece válida para um número qualquer de forças aplicadas a uma partícula.

Em particular, na Figura 4.1, podemos dizer que, como a partícula P está em equilíbrio, a mola aplica sobre ela uma força igual e contrária à força aplicada pela pessoa que está puxando.

Consideremos agora a situação da Figura 4.4, em que a partícula está sus-

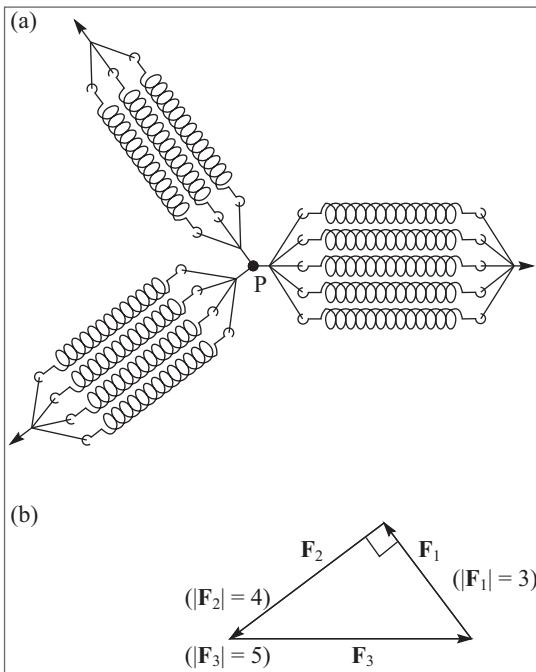


Figura 4.3 Equilíbrio de forças.

pensa verticalmente da mola. É a base da “balança de mola”, instrumento tipicamente usado por feirantes para pesar mercadorias. O ponteiro acusa uma distensão da mola, na situação de equilíbrio. Temos portanto duas forças iguais e contrárias, \mathbf{F} e $-\mathbf{F}$ na figura, agindo sobre a partícula. Uma delas, $-\mathbf{F}$, é devida à mola, como no exemplo acima. E a outra? A força \mathbf{F} não é devida ao puxão de uma pessoa; sabemos que se deve à atração gravitacional da Terra e representa a *força-peso*. O peso de um corpo é a magnitude

desta força, ou seja, é a *magnitude da força* (vertical, dirigida para cima) *que é preciso aplicar ao corpo para mantê-lo em equilíbrio quando suspenso livremente, sob a ação da gravidade*. No exemplo acima, esta força está sendo aplicada pela mola.

A força-peso é um exemplo de uma força que atua sobre uma partícula sem que precise haver *contato direto* com o agente responsável pela força (no caso a Terra). Forças elétricas e magnéticas sobre partículas carregadas são exemplos análogos.

Se a partícula P que estava suspensa da mola é agora colocada sobre uma mesa, onde também permanece em equilíbrio, inferimos que a força $-\mathbf{F}$ que equilibra a força-peso, e que estava sendo aplicada pela mola, está sendo aplicada agora pela mesa sobre a partícula. Esta força $-\mathbf{F}$ é um exemplo de uma *reação de contato*, normal à superfície da mesa, e que tem origem na deformação elástica da mesa devida a seu contato com o objeto colocado sobre ela.

Voltaremos adiante a analisar, de forma mais detalhada, os diferentes tipos de forças discutidas acima.

4.2 A LEI DA INÉRCIA

Segundo Aristóteles, tanto para colocar um corpo em movimento como para mantê-lo em movimento é necessária a ação de uma força. Isto parece concordar com nossa experiência imediata de que um objeto deslizando sobre o solo, por exemplo, tende a parar se pararmos de empurrá-lo. Entretanto, um projétil como uma pedra ou uma flecha continua em movimento depois de lançado. Aristóteles explicava isto afirmando que é o ar “empurrado para os lados” pelo projétil que se desloca para trás dele e produz a força que o impulsiona. Logo, segundo Aristóteles, se a força que atua sobre um corpo é nula, o corpo permanecerá sempre em repouso.

Vejamos agora o que diz Galileu nos “Diálogos Sobre os Dois Principais Sistemas do Mundo”:

SALVIATI: ...Diga-me agora: Suponhamos que se tenha uma superfície plana lisa como um espelho e feita de um material duro como o aço. Ela não está horizontal, mas inclinada, e sobre ela foi colocada uma bola perfeitamente esférica, de algum material duro e pesado, como o bronze. A seu ver, o que acontecerá quando a soltarmos?

SIMPLÍCIO: Não acredito que permanecerá em repouso; pelo contrário, estou certo de que rolaria espontaneamente para baixo...

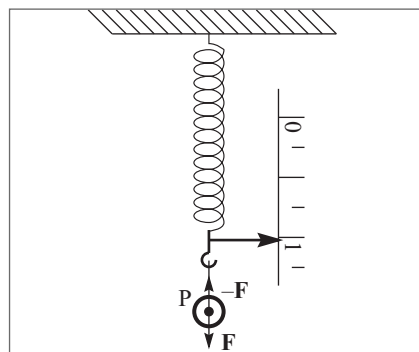


Figura 4.4 Força-peso.

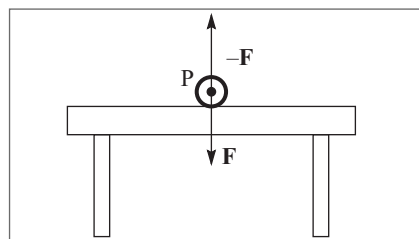


Figura 4.5 Reação de contato.

SALVIATI: ...E por quanto tempo a bola continuaria a rolar, e quão rapidamente? Lembre-se de que eu falei de uma bola perfeitamente redonda e de uma superfície altamente polida, afim de remover todos os impedimentos externos e acidentais. Analogamente, não leve em consideração qualquer impedimento do ar causado por sua resistência à penetração, nem qualquer outro obstáculo acidental, se houver.

SIMPLÍCIO: Compreendo perfeitamente, e em resposta a sua pergunta digo que a bola continuaria a mover-se de modo indefinido, enquanto permanecesse sobre a superfície inclinada, e com um movimento continuamente acelerado...

SALVIATI: Mas se quiséssemos que a bola se movesse para cima sobre a mesma superfície, acha que ela subiria?

SIMPLÍCIO: Não espontaneamente; mas ela o faria se fosse puxada ou lançada para cima.

SALVIATI: E se fosse lançada com um certo impulso, qual seria seu movimento, e de que amplitude?

SIMPLÍCIO: O movimento seria constantemente freado e retardado, sendo contrário à tendência natural, e duraria mais ou menos tempo conforme o impulso e a inclinação do plano fossem maiores ou menores.

SALVIATI: Muito bem, até aqui você me explicou o movimento sobre dois planos diferentes. Num plano inclinado para baixo, o corpo móvel desce espontaneamente e continua acelerando, e é preciso empregar uma força para mantê-lo em repouso. Num plano inclinado para cima, é preciso uma força para lançar o corpo ou mesmo mantê-lo parado, e o movimento impresso no corpo diminui de modo contínuo até cessar de todo. Você diria ainda que, nos dois casos, surgem diferenças conforme a inclinação do plano seja maior ou menor, de forma que um declive mais acentuado implica maior velocidade, ao passo que, num aclave, um corpo lançado com uma dada força se move tanto mais longe quanto menor o aclave.

Diga-me agora o que aconteceria ao mesmo corpo móvel colocado sobre uma superfície sem nenhum aclave nem declive.

SIMPLÍCIO: Aqui preciso pensar um instante sobre a resposta. Não havendo declive, não pode haver tendência natural ao movimento; e, não havendo aclave, não pode haver resistência ao movimento. Parece-me portanto que o corpo deveria naturalmente permanecer em repouso. Mas eu me esqueci; faz pouco tempo que Sagredo me deu a entender que isto é o que aconteceria.

SALVIATI: Acredito que aconteceria se colocássemos a bola firmemente num lugar. Mas que sucederia se lhe déssemos um impulso em alguma direção?

SIMPLÍCIO: Ela teria que se mover nessa direção.

SALVIATI: Mas com que tipo de movimento? Seria continuamente acelerado, como no declive, ou continuamente retardado, como no aclave?

SIMPLÍCIO: Não posso ver nenhuma causa de aceleração, uma vez que não há aclave nem declive.

SALVIATI: Exatamente. Mas se não há razão para que o movimento da bola se retarde, ainda menos há razão para que ele pare; por conseguinte, por quanto tempo você acha que a bola continuaria se movendo?

SIMPLÍCIO: Tão longe quanto a superfície se estendesse sem subir nem descer.

SALVIATI: Então, se este espaço fosse ilimitado, o movimento sobre ele seria também ilimitado? Ou seja, perpétuo?

SIMPLÍCIO: Parece-me que sim, desde que o corpo móvel fosse feito de material durável”.

Temos aqui formulada pela primeira vez a *lei da inércia*, na situação ideal contemplada por Galileu, com uma esfera lançada sobre um plano horizontal perfeitamente polido (sem atrito), desprezando a resistência do ar. O movimento não seria nem acelerado nem desacelerado: não havendo forças na direção horizontal, teríamos um *movimento retilíneo uniforme*. Ao contrário do que dizia Aristóteles, não há necessidade de forças para manter um movimento retilíneo uniforme: pelo contrário, uma aceleração nula ($\mathbf{v} = \text{constante}$) está necessariamente associada à ausência de força resultante sobre a partícula ($\mathbf{F} = 0$).

A situação imaginada por Galileu é muito difícil de realizar na prática, na escala do laboratório. Podemos pensar nela como um caso limite. Em circunstâncias em que procuramos minimizar o atrito, como na patinação no gelo, um impulso adquirido tende a persistir durante muito tempo. Em demonstrações de laboratório, costumam-se empregar discos de base bem polida, deslizando sobre uma camada de ar ou de gás carbônico (proveniente da evaporação de gelo seco) que escapa através de orifícios, produzindo um “colchão de gás” sobre o qual o disco flutua, como um aerobarco sobrevoando a água, tornando muito pequeno o efeito do atrito. Nessas condições, podemos verificar aproximadamente a lei da inércia.

Em seu monumental tratado “Os Princípios Matemáticos da Filosofia Natural”, publicado em 1687, Newton formulou três “Axiomas ou Leis do Movimento”.

A 1ª Lei é a Lei da Inércia:

“Todo corpo persiste em seu estado de repouso, ou de movimento retilíneo uniforme, a menos que seja compelido a modificar esse estado pela ação de forças impressas sobre ele”.

O que significa realmente esta lei? Como podemos saber que não existem “forças impressas sobre o corpo”? Pelo fato de que permanece em repouso ou movimento retilíneo e uniforme? se assim fosse, Eddington teria tido razão quando criticou o enunciado da 1ª lei, dizendo ser equivalente a “... persiste... exceto quando não persiste” (o que corresponderia à bem conhecida previsão meteorológica: “Tempo bom, salvo se chover”). Esta crítica é injusta. Se todas as forças fossem devidas ao contato com outros corpos, bastaria a ausência de contato para estabelecer a ausência de forças. O exemplo da força-peso, e das forças elétricas e magnéticas, mostra, porém, a existência de forças que atuam sem que haja contato direto com o corpo responsável pela força.

Entretanto, estas forças tendem a diminuir à medida que os corpos em interação se afastam um do outro. Em média, a distância típica entre uma estrela e sua vizinha mais próxima é $\sim 10^{18}$ cm (veja a Tabela 1.1, Seç. 1.5), o que deveria ser suficiente para que forças entre elas possam ser desprezadas. A observação das estrelas confirma que elas

obedecem com muito boa aproximação à lei da inércia. Em relação a que referencial? Não é em relação à Terra, pois um observador terrestre vê as estrelas girarem no céu noturno.

Isso indica outro ponto importante na compreensão da 1ª lei: ela não pode ser válida em qualquer referencial. Os referenciais em que é válida chamam-se *referenciais inerciais*. A Terra não é um referencial inercial. Entretanto, o movimento de rotação da Terra em torno do eixo afeta muito pouco os movimentos usuais, na escala de laboratório, e na prática, nessa escala, empregar o laboratório como referencial inercial é uma boa aproximação (o movimento de rotação da Terra pode ser evidenciado, conforme veremos no Capítulo 13, pela experiência do pêndulo de Foucault). Por outro lado, um referencial ligado às estrelas fixas é, com excelente aproximação, um referencial inercial, e é a este tipo de referencial que nos referiremos, em princípio, daqui por diante.

Decorre imediatamente da (3.9.2) que um *referencial em movimento retilíneo uniforme em relação a um referencial inercial é também inercial* (porque um corpo em repouso ou em movimento retilíneo uniforme em relação a um deles também estará em repouso ou em movimento retilíneo uniforme em relação ao outro). Logo, dispondo de um referencial inercial (ligado às estrelas fixas), dispomos em consequência de uma infinidade deles.

A expressão “movimento retilíneo” refere-se à geometria euclidiana que, conforme já foi mencionado (Seç. 1.2) não é um conceito válido *a priori*, mas está sujeito à verificação experimental. Em escala cosmológica, observam-se desvios, mas na escala em que estaremos aplicando as leis da mecânica clássica tais desvios são desprezíveis.

4.3 A 2ª LEI DE NEWTON

Uma das implicações da 1ª lei é que qualquer variação da velocidade \mathbf{v} de um corpo (em módulo ou em direção!) em relação a um referencial inercial, ou seja, qualquer *aceleração*, deve estar associada à ação de *forças*. Isso sugere procurar uma relação mais precisa entre força e aceleração.

Consideremos o exemplo da queda livre de um corpo. Já vimos que, neste caso, a aceleração é constante [cf. (3.6.1)]: $\mathbf{a} = \mathbf{g}$, onde \mathbf{g} é vertical e dirigido para baixo. Qual é a força que atua sobre o corpo? Vimos na Seç. 4.1 que esta força (atração gravitacional), é também vertical, dirigida para baixo, e constante para um dado corpo (ou seja, a mesma em qualquer altura, na vizinhança dum dado ponto da superfície da Terra: a distensão da mola que equilibra esta força é a mesma em qualquer altura). Isso sugere que a aceleração devida a uma força seja proporcional à força (vetorialmente!), ou seja, $\mathbf{a} = k\mathbf{F}$. Que podemos dizer sobre o coeficiente de proporcionalidade k ?

Sabemos que a mesma força (medida em termos da distensão de uma mola), quando aplicada a corpos diferentes, produz em geral acelerações diferentes. Logo, o coeficiente k mede uma *propriedade do corpo*, que caracteriza sua *resposta* à força aplicada.

Acelerar ou frear um carro requer uma força bem maior do que para uma bicicleta, para a mesma variação de velocidade (as consequências de uma colisão com um ou com outro, à mesma velocidade, também são bem diferentes!). Dizemos usualmente que um

carro tem *inércia* muito maior que uma bicicleta, resistindo portanto bem mais a variações de velocidade. O coeficiente k deve medir então uma propriedade inversamente proporcional à “inércia” do corpo.

As Figuras 4.6(a), (b) e (c) mostram uma série de experiências idealizadas que poderiam ser feitas com discos deslizantes sobre uma camada de gás, para minimizar o atrito (Seç. 4.2). Em (a), a força \mathbf{F} , medida pela distensão de uma mola, como foi discutido na Seç. 4.2, é aplicada ao disco D, que desliza com movimento retilíneo uniformemente acelerado de aceleração \mathbf{a} na direção de \mathbf{F} . Em (b) o disco D é o mesmo, mas a força aplicada é $2\mathbf{F}$, e verifica-se que a aceleração de D é $2\mathbf{a}$. Logo, temos de fato proporcionalidade entre aceleração e força para um mesmo corpo D:

$$\mathbf{a} = \mathbf{F} / m \quad (4.3.1)$$

onde o coeficiente de proporcionalidade ($1/m$) é característico do disco D.

Em (c), a força voltou a ser \mathbf{F} , mas empilhamos dois discos idênticos D e D', e a aceleração caiu à metade. Comparando então as experiências (a) e (c), vemos que, na (4.3.1), é preciso atribuir ao sistema de dois discos idênticos D e D' o coeficiente de proporcionalidade $1/(2m)$, ou seja, que a “inércia” de dois objetos idênticos agregados num objeto único é o dobro da inércia de um deles. O “coeficiente de inércia” m mede portanto, nesse sentido, a “quantidade de matéria” do objeto.

Repetindo experiências como (a) e (c) com objetos diferentes sujeitos à mesma força \mathbf{F} , obteríamos de forma mais geral

$$\mathbf{F} = m_1 \mathbf{a}_1 = m_2 \mathbf{a}_2 \quad (4.3.2)$$

ou seja, $|\mathbf{a}_2|/|\mathbf{a}_1| = m_1/m_2$: as acelerações adquiridas por objetos diferentes submetidos à mesma força são inversamente proporcionais aos respectivos “coeficientes de inércia”.

Experiências deste tipo nos permitem inferir assim a 2ª Lei de Newton

$$\mathbf{F} = m\mathbf{a} \quad (4.3.3)$$

onde o “coeficiente de inércia” m associado à partícula sobre a qual age a força \mathbf{F} chama-se *massa inercial* dessa partícula.

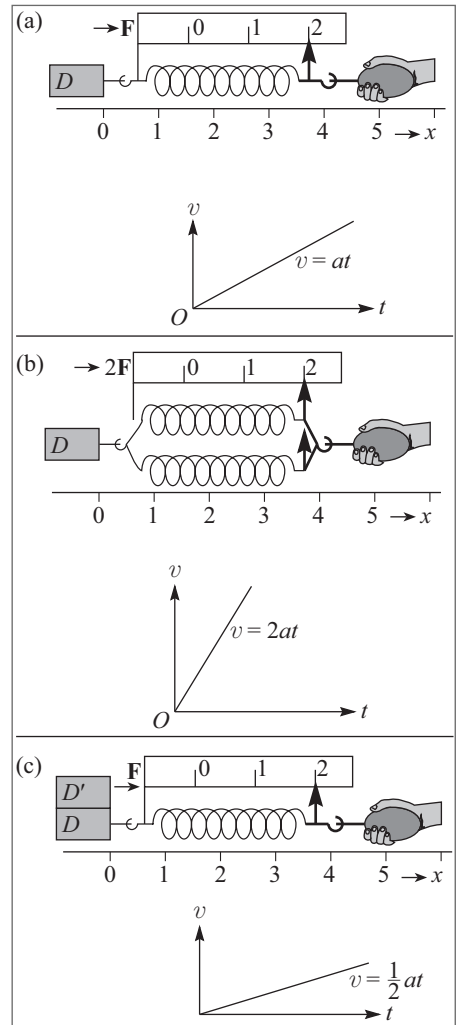


Figura 4.6 Coeficiente de inércia.

Utilizando uma mesma força padrão \mathbf{F} , como na (4.3.2), podemos estabelecer uma *escala relativa de massas inerciais*. Em lugar de escolher arbitrariamente uma “unidade de força”, como fizemos na Seq. 4.1, é mais conveniente escolher arbitrariamente uma unidade de massa inercial. Em geral, omitiremos a palavra “inercial”, falando simplesmente de “massa”. Veremos depois que se pode definir também a chamada “massa gravitacional”.

A *unidade de massa* é definida em termos de um protótipo, um padrão de platina iridiada, depositado no Ofício Internacional de Pesos e Medidas em Paris, que representa o *quilograma* (kg), e foi construído originalmente para corresponder à massa de 1 l de água à pressão atmosférica e à temperatura de 4°C. Por definição, 1 kg é a massa desse protótipo. Poderíamos pensar em adotar também unidades atômicas para a massa, mas isto seria atualmente desvantajoso do ponto de vista de precisão nas aplicações práticas, uma vez que não podemos contar diretamente o número de átomos contido num corpo macroscópico, e o número de Avogadro (n° de moléculas por mol) é conhecido com precisão muito inferior à precisão com a qual podemos medir massas em termos de quilograma padrão.

Usualmente adotaremos o Sistema Internacional (SI) de unidades, em que a unidade de comprimento é o metro, a de massa é o quilograma, e a de tempo é o segundo. Todas as demais unidades mecânicas poderão ser expressas em termos destas três.

Assim, a (4.3.3) permite agora substituir nossa definição provisória de unidade de força da Seq. 4.1 pela definição de *newton* (N), unidade de força do sistema SI. Por definição, 1 N é a força que, quando aplicada a um corpo de massa de 1 kg, imprime-lhe uma aceleração de 1 m/s^2 . Para ter uma ideia concreta da ordem de grandeza do newton, lembremos que, pela (2.6.1), 1 N é a ordem de grandeza da força-peso exercida pela gravidade sobre um objeto de massa $\approx 100 \text{ g}$ (uma maçã, p. ex.!).

No *sistema CGS* de unidades, em que as unidades básicas são cm, g e s, a unidade de força é 1 dina, a força que comunica uma aceleração de 1 cm/s^2 a uma massa de 1 g. Como $1 \text{ g} = 10^{-3} \text{ kg}$, $1 \text{ cm} = 10^{-2} \text{ m}$, temos

$$1 \text{ dina} = 10^{-5} \text{ N}$$

4.4 DISCUSSÃO DA 2ª LEI

A 2ª lei de Newton é o *princípio fundamental da dinâmica*; conforme veremos, é a lei básica que permite determinar a evolução de um sistema na mecânica clássica. A 1ª lei pode ser considerada como um caso particular da 2ª: se a força resultante \mathbf{F} que atua sobre uma partícula é nula, a (4.3.3) mostra que $\mathbf{a} = 0$, e já demonstramos [cf. (3.5.9)] que isto acarreta para a partícula a permanência em repouso ou em movimento retilíneo uniforme. Note-se que a 2ª lei, como a 1ª, só é válida num referencial inercial.

Muitas vezes se diz que a 2ª lei não passa de uma *definição de força*. Se assim fosse, ela seria desprovida de conteúdo físico, e não poderíamos questionar sua validade: no máximo, poderíamos argumentar sobre se é ou não uma definição conveniente.

Se \mathbf{F} fosse dado apenas pela (4.3.3), ela seria realmente uma definição de força. Entretanto, isto não é verdade: as forças que atuam sobre uma partícula resultam de sua interação com outras partículas, e veremos que são dadas por *leis de forças*, que definem \mathbf{F} em termos da situação em que a partícula se encontra. Exemplos disso são a lei da gravitação universal e as leis que dão as forças elétricas e magnéticas que atuam sobre uma partícula carregada. A (4.3.3) é uma espécie de molde, que permanece vazio enquanto não substituirmos \mathbf{F} pela sua expressão em termos de leis de forças, mas que adquire todo o seu significado uma vez que isto é feito. De fato, a 2ª lei define uma espécie de *programa* para a física clássica: encontrar as leis de forças correspondentes a todas as interações possíveis.

Vimos que a 2ª lei permite estabelecer uma escala de massas inerciais, e neste sentido ela pode ser considerada como permitindo definir o conceito de massa inercial, mas não é tão pouco apenas uma definição deste conceito. De fato, a ideia implícita na 2ª lei é que a massa inercial m é uma característica da partícula; uma vez determinada quando atua sobre a partícula uma força conhecida, devemos empregar o mesmo valor de m para descrever o movimento da partícula sob a ação de quaisquer outras forças. Admite-se também tacitamente que m (ou seja, o efeito de uma força em produzir aceleração) é independente da posição e velocidade da partícula, pelo menos enquanto se mantém a sua identidade (isto não se aplicaria a uma gota de chuva que aumenta de volume enquanto cai, ou a um foguete que ejeta combustível à medida que sobe; discutiremos mais tarde sistemas de massa variável como estes).

Na relatividade restrita, conforme veremos, verifica-se que m de fato depende da velocidade da partícula. Entretanto, este efeito é desprezível enquanto a partícula não atinge velocidades comparáveis à velocidade da luz no vácuo. Temos de excluir este domínio relativístico de velocidades elevadas do campo de aplicabilidade da mecânica newtoniana, limitando-nos ao *domínio não relativístico*. Usualmente teremos de nos limitar também ao domínio macroscópico, excluindo objetos pertencentes à escala atômica, aos quais se aplicam as leis da mecânica quântica (o conceito de força não tem muita utilidade do ponto de vista quântico). Em alguns casos, porém, ainda poderemos aplicar os resultados obtidos pela mecânica clássica a objetos atômicos, por exemplo, ao movimento de feixes de partículas como os que foram considerados no fim da Seq. 3.6.

A 2ª lei tem ainda diversas outras implicações. Uma delas é que só intervêm na dinâmica deslocamentos, velocidades e acelerações das partículas; não é preciso considerar, por exemplo, derivadas temporais da aceleração, tais como $d\mathbf{a}/dt$ ou $d^2\mathbf{a}/dt^2$.

Outra implicação importante está relacionada com o caráter vetorial da (4.3.3). Como \mathbf{a} é um vetor e m um escalar, segue-se que \mathbf{F} é um vetor. Assim, se $\mathbf{F}_1, \mathbf{F}_2, \dots, \mathbf{F}_n$ são forças de diferentes origens que atuam sobre a mesma partícula, \mathbf{F} na (4.3.3) é a força resultante que atua sobre a partícula, ou seja,

$$\mathbf{F} = \mathbf{F}_1 + \mathbf{F}_2 + \dots + \mathbf{F}_n \quad (4.4.1)$$

onde a soma é *vetorial* (para $n = 2$, obedece à regra do paralelogramo). Este é um *resultado experimental*, conhecido como *princípio de superposição, de forças*, que já foi mencionado no caso particular do equilíbrio [cf. (4.1.1)].

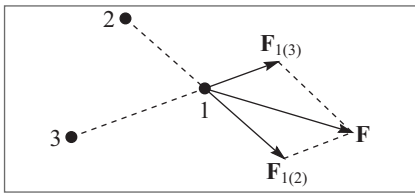


Figura 4.7 Partícula em interação com duas outras.

Consideremos por exemplo uma partícula 1 interagindo com duas outras 2 e 3 (Figura 4.7), e seja $\mathbf{F}_{1(2)}$ a força sobre 1 devida à partícula 2, e $\mathbf{F}_{1(3)}$ a força sobre 1 devida à partícula 3. A força resultante sobre a partícula 1 será então $\mathbf{F} = \mathbf{F}_{1(2)} + \mathbf{F}_{1(3)}$. Em muitos casos, $\mathbf{F}_{1(2)}$ é também a força que agiria sobre 1 se somente 2 estivesse presente (e $\mathbf{F}_{1(3)}$ a força que agiria sobre 1 em presença apenas de 3), mas isto não é necessariamente verdade, ou seja, $\mathbf{F}_{1(2)}$ pode ser modificada pela presença de 3 e $\mathbf{F}_{1(3)}$ pela presença de 2. O princípio de superposição continua valendo, mas é preciso calcular cada força levando em conta a presença de todas as partículas. Geralmente, porém, nos problemas de que vamos tratar, esta complicação não aparecerá, e as forças devidas a uma partícula poderão ser calculadas como se as outras não estivessem presentes.

A (4.3.3) não corresponde à formulação original de Newton da 2ª lei. Newton começou definindo o que chamou de “quantidade de movimento”, também conhecido como *momento linear*, ou simplesmente *momento*. A definição de Newton foi:

“A quantidade de movimento é a medida do mesmo, que se origina conjuntamente da velocidade e da massa”.

Ou seja: o *momento (linear)* de uma partícula é o produto de sua massa por sua velocidade:

$$\boxed{\mathbf{p} = m\mathbf{v}} \quad (4.4.2)$$

Decorre imediatamente desta definição que \mathbf{p} é um vetor.

Se m não varia com o tempo, ou seja, se excluirmos sistemas de massa variável, obtemos, derivando em relação ao tempo ambos os membros da (4.4.2) (cf. (2.2.5)),

$$\frac{d\mathbf{p}}{dt} = m \frac{d\mathbf{v}}{dt} = m\mathbf{a} \quad (4.4.3)$$

e, comparando com a (4.3.3),

$$\boxed{\frac{d\mathbf{p}}{dt} = \mathbf{F}} \quad (4.4.4)$$

o que corresponde à **formulação de Newton da 2ª lei**:

“A variação do momento é proporcional à força impressa, e tem a direção da força”.

Ou seja: a *força é a taxa de variação temporal do momento*. Embora essa formulação da 2ª lei pareça inteiramente equivalente à (4.3.3), veremos que ela tem vantagens. Uma delas, que revela a importância do conceito de momento, aparecerá na próxima Seção. Outra, que será vista posteriormente, é que a (4.4.4), ao contrário da (4.3.3), permanece válida na mecânica relativística.

Vejam agora alguns *exemplos simples da aplicação da 2ª lei*:

Exemplo 1 – Força-peso: Substituindo a (3.6.1) na (4.3.3), vemos que a força \mathbf{P} que atua sobre um corpo na vizinhança da superfície da Terra devido à atração gravitacional por ela exercida sobre o corpo é

$$\mathbf{P} = m\mathbf{g} \quad (4.4.5)$$

onde m é a massa inercial do corpo e \mathbf{g} a aceleração da gravidade, vertical, dirigida para baixo e de magnitude g . A (4.4.5) chama-se *força-peso*; pode ser medida em equilíbrio pela balança de mola (Seç. 4.1). Para uma partícula em queda livre, a 2ª lei de Newton leva à (3.6.1),

$$\mathbf{a} = \mathbf{g} \quad (4.4.6)$$

A proporcionalidade da força-peso à massa inercial é uma peculiaridade notável dessa força, que voltaremos a discutir no capítulo sobre gravitação. É graças a ela que a aceleração da gravidade é a mesma para qualquer partícula [cf. (4.4.6) e (2.6.1)]. É também graças a ela que podemos medir a massa inercial pelo peso, por exemplo, por pesagem com uma balança de mola. É importante, porém, evitar confusão entre os conceitos de massa e peso, que são totalmente diferentes. Num ponto muito distante da superfície da Terra (na superfície da Lua, por exemplo), o peso de uma partícula, indicado pela distensão da balança de mola, seria muito diferente, embora sua massa não se tenha alterado. Aliás, o peso sofre pequenas variações mesmo de ponto a ponto da superfície da Terra, devido às variações locais de g .

Em engenharia, é comum utilizar como unidade de força o quilograma-força (kgf), definido como a força-peso sobre uma massa de 1 kg ao nível do mar e na latitude de 45° N (onde $g \approx 9,81 \text{ m/s}^2$). Na prática, podemos tomar: $1 \text{ kgf} \approx 9,8 \text{ N}$.

Exemplo 2 – Plano inclinado: Consideremos uma partícula de massa m colocada sobre um plano inclinado de ângulo de inclinação θ (Figura 4.8.). Além da força-peso $\mathbf{P} = m\mathbf{g}$, atua sobre a partícula a reação de contato \mathbf{N} devida a seu contato com o plano. Já vimos um exemplo de uma tal reação no caso de uma partícula em equilíbrio sobre uma mesa (Seç. 4.1).

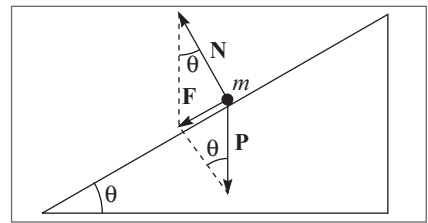


Figura 4.8 Plano inclinado.

Em geral, a reação de contato pode ter componentes tanto na direção normal ao plano como na direção tangencial. A componente tangencial está associada às forças de atrito, que serão discutidas no capítulo 5. Para simplificar, vamos tomar o caso limite ideal em que a superfície do plano é perfeitamente polida, “sem atrito”, o que elimina a componente tangencial: a força de reação \mathbf{N} é normal ao plano.

A Figura 4.9 mostra então que a magnitude da resultante \mathbf{F} é

$$F = P \sin \theta = mg \sin \theta \quad (4.4.7)$$

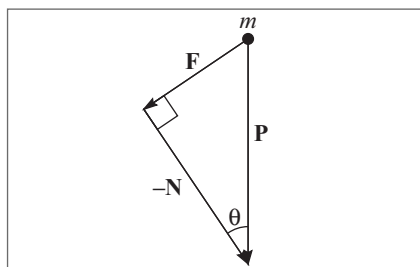


Figura 4.9 Cálculo da resultante.

e que \mathbf{F} é dirigida tangencialmente ao plano, para baixo. A aceleração \mathbf{a} do movimento da partícula ao longo do plano inclinado tem a direção de \mathbf{F} ; pelas (4.4.7) e (4.3.3), temos

$$a = g \sin \theta \quad (4.4.8)$$

Logo, o efeito do plano inclinado é reduzir a aceleração da queda livre por um fator igual ao seno do ângulo de inclinação. Este resultado, que

já havia sido obtido por Galileu, foi empregado por ele, como vimos (Seç. 2.6), no estudo experimental do movimento uniformemente acelerado.

Exemplo 3 – Funda: Voltemos agora a considerar o exemplo de uma partícula em movimento circular uniforme (Seç. 3.7). Vimos que este é um movimento acelerado, de forma que só pode ser mantido pela ação de uma força. Para uma partícula de massa m , a força \mathbf{F} necessária para mantê-la em movimento circular uniforme de velocidade v num círculo de raio r é dada pelas (4.3.3) e (3.7.13);

$$\mathbf{F} = -\frac{mv^2}{r} \hat{\mathbf{r}} \quad (4.4.9)$$

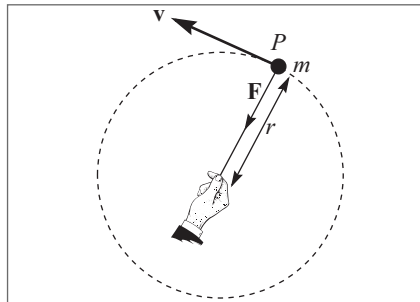


Figura 4.10 Funda.

Esta é a chamada *força centrípeta*. A Figura 4.10 mostra um exemplo familiar da atuação desta força: fazemos girar em torno de nossa mão uma pedra amarrada num fio, em movimento circular uniforme. Neste caso, a força centrípeta \mathbf{F} é aplicada pela nossa mão e transmitida à pedra através do fio.

Se soltarmos o fio quando a pedra se encontra num determinado ponto P de sua órbita, e se desprezarmos o efeito da força-peso (gravidade), \mathbf{F} subitamente se anula, e a lei da inércia implica en-

tão que a pedra se move, a partir do ponto P , com movimento retilíneo uniforme de velocidade \mathbf{v} igual à velocidade do movimento circular no ponto P da órbita, ou seja (cf. (3.7.3)), tangente ao círculo em P . A pedra “sai pela tangente”.

4.5 CONSERVAÇÃO DO MOMENTO E 3ª LEI DE NEWTON

Até aqui, consideramos apenas as forças exercidas sobre uma única partícula; sabemos que são devidas à ação de outras partículas, mas não consideramos ainda o que acontece com estas partículas. A situação mais simples imaginável é aquela em que há apenas duas partículas em interação, que podemos designar por 1 e 2; as únicas forças

SEMANAS 5 E 6

3

Movimento bidimensional

3.1 DESCRIÇÃO EM TERMOS DE COORDENADAS

Neste capítulo, vamos passar do movimento retilíneo à descrição do movimento num plano, que inclui muitos casos importantes, como o movimento dos projéteis e o movimento da Terra em torno do Sol.

Conforme já foi mencionado na Seção 1.6, podemos especificar a posição de um ponto num plano através de 2 parâmetros, que são suas *coordenadas* em relação a um dado referencial. Se adotarmos coordenadas cartesianas, por exemplo, a posição de uma partícula em movimento no plano será descrita pelo par de funções

$$(x(t), y(t)) \quad (3.1.1)$$

onde $x(t)$ é a abscissa e $y(t)$ a ordenada da partícula no instante t . Podemos dizer que, à medida que o ponto P se move, descrevendo a trajetória da partícula no plano, suas projeções sobre os eixos Ox e Oy se movem correspondentemente, descrevendo movimentos unidimensionais. Reduzimos assim a descrição de um movimento bidimensional à de dois movimentos unidimensionais simultâneos, cuja composição leva ao movimento no plano.

Em muitos casos, os movimentos ao longo de dois eixos ortogonais são independentes um do outro (embora isto nem sempre aconteça: veja a Seção 3.6). Este fato foi reconhecido por Galileu e permitiu-lhe descrever corretamente, pela primeira vez, o movimento dos projéteis. Já em seu “Diálogo sobre os Dois Principais Sistemas do Mundo” Galileu havia empregado a independência dos movimentos para refutar um dos principais argumentos usados pelos partidários de Ptolomeu para provar a imobilidade da Terra. Eis o argumento, conforme expresso por Salviati:

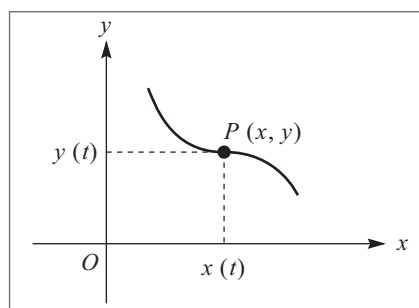


Figura 3.1 Movimento num plano.

“...Se a Terra tivesse um movimento diurno de rotação, uma torre do alto da qual se deixasse cair uma pedra, sendo transportada pela Terra em sua rotação, já se teria deslocado de muitas centenas de jardas para Leste durante o tempo de queda da pedra, e a pedra deveria atingir o solo a essa distância da base da torre. Também é mencionada a experiência em que se deixa cair uma bola de chumbo do topo do mastro de um navio parado, notando que ela cai ao pé do mastro, mas, se se deixa cair a mesma bola do mesmo ponto com o navio em movimento, ela cairá a uma distância do pé do mastro igual à distância de que o navio se tiver deslocado durante a queda...” (a comparação entre o exemplo da Terra e o do navio levaria a concluir pela imobilidade da Terra).

Ao refutar o argumento, Salviati começa por perguntar a Simplicio:

“SALVIATI: Muito bem. Você jamais fez esta experiência do navio?

SIMPLÍCIO: Nunca fiz, mas certamente acredito que as autoridades que formularam o argumento tinham feito uma observação cuidadosa...

SALVIATI: ...você o toma como certo sem tê-lo feito... e eles fizeram o mesmo tendo fé em seus antecessores, e assim por diante, sem jamais chegar a alguém que o tenha feito. Pois quem quer que faça a experiência verá que ela mostra exatamente o contrário do que foi escrito, ou seja, que a pedra sempre cai no mesmo ponto do navio, quer ele esteja parado, quer esteja se movendo em qualquer velocidade que se queira. O mesmo vale para a Terra: nada pode ser inferido sobre o movimento ou imobilidade da Terra pelo fato de que a pedra sempre cai ao pé da torre.”

E Salviati explica a Simplicio que a pedra que se deixa cair do topo do mastro de um navio em movimento já compartilha, desde o início, do movimento (horizontal) do navio, ao passo que o movimento de queda livre na direção vertical é *independente* deste, de forma que não há diferença no ponto de queda em relação ao navio.

Galileu dá uma série de outros exemplos: um cavaleiro que lança uma bola para cima, enquanto seu cavalo galopa, pode recapturá-la mais adiante, como faria num lançamento vertical com seu cavalo parado (em relação a um observador em repouso no solo, a trajetória da bola, com o cavalo em movimento, seria uma parábola). E Sagredo dá outro exemplo:

“Analogamente, se um canhão horizontal numa torre atira paralelamente ao horizonte, não importa se a carga de pólvora é grande ou pequena, de forma que a bala caia a mil jardas de distância, ou quatro mil, ou seis mil; todos estes tiros levam o mesmo tempo (para atingir o chão), e este tempo é igual ao que a bala levaria da boca do canhão até o solo se caísse diretamente para baixo sem qualquer impulso”.

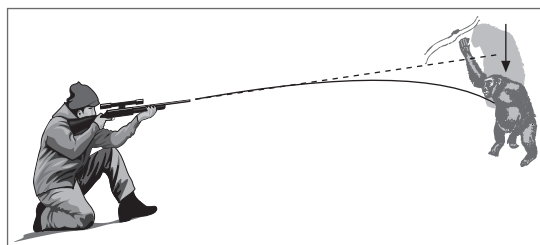


Figura 3.2 O caçador e o macaco.

canhão até o solo se caísse diretamente para baixo sem qualquer impulso”.

Este efeito é também ilustrado pela tradicional história do caçador e do macaco (Figura 3.2). O caçador aponta para o macaco dependurado num galho; ao ver o clarão do disparo, o macaco se assusta e cai, mas isto não o salva, porque

a aceleração da gravidade atua da mesma maneira sobre ele e a bala, no movimento de queda livre, e o desvio vertical da bala e do macaco em relação à linha de mira original, é o mesmo.

3.2 VETORES

O sistema de coordenadas escolhido para descrever o movimento na Seção 3.1 tem um caráter acessório; o mesmo movimento pode ser descrito com eixos de orientação diferente, ou em coordenadas polares, por exemplo. Vamos ver agora que é possível dar uma descrição *intrínseca* do movimento, independente da escolha do sistema de coordenadas, com o auxílio do conceito de *vetores*.

Para dar uma caracterização intrínseca do deslocamento de uma partícula em sua trajetória em relação a uma origem dada, não basta conhecermos a *magnitude* do deslocamento (distância à origem): é preciso também especificarmos a *direção* e o *sentido* do deslocamento. Por exemplo, não basta para determinar a posição de um carro dizer que ele se deslocou de 100 Km em relação ao ponto de partida. Definiríamos completamente o deslocamento, por outro lado, dizendo que ele se deu segundo a direção Norte-Sul, e no sentido Sul para o Norte.

Uma representação geométrica do deslocamento pode ser obtida por uma *seta*, que dá diretamente a direção e sentido, e cujo, comprimento mede a magnitude do deslocamento (Figura 3.3); usamos a notação \mathbf{r} para nos referimos ao deslocamento assim representado. Representaremos vetores sempre por símbolos em **negrito**.

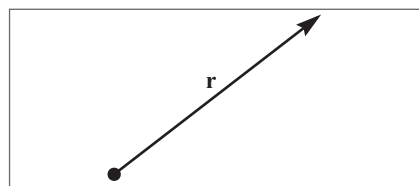


Figura 3.3 Deslocamento como vetor.

Uma propriedade fundamental dos deslocamentos é ilustrada pelo exemplo do navio dado por Galileu (Figura 3.4). O deslocamento total \mathbf{r} da pedra que se deixa cair do topo do mastro pode ser considerado como *resultante* do deslocamento \mathbf{r}_x na direção horizontal (que é o deslocamento do navio) com o deslocamento \mathbf{r}_y devido à queda livre da pedra na direção vertical.

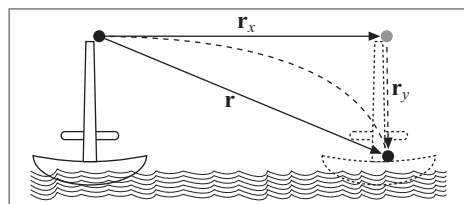


Figura 3.4 Composição de deslocamentos.

Isto independe de se tratar de deslocamentos em direções perpendiculares: o deslocamento \mathbf{r} resultante de dois deslocamentos \mathbf{r}_1 e \mathbf{r}_2 em direções diferentes se obtém unindo a “origem” do primeiro à “extremidade” do segundo (Figura 3.5), ou, o que é equivalente, pela “regra do paralelogramo”, tomando a diagonal do paralelogramo construído sobre \mathbf{r}_1 e \mathbf{r}_2 (Figura 3.6).

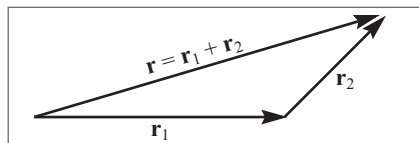


Figura 3.5 Deslocamento resultante.

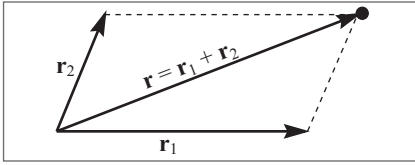


Figura 3.6 Regra do paralelograma.

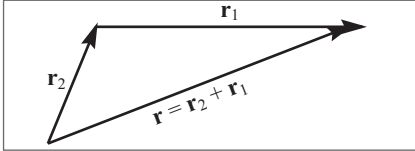


Figura 3.7 Comutatividade da soma.

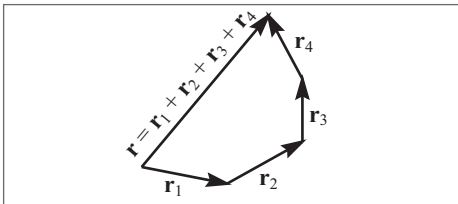


Figura 3.8 Soma de vários deslocamentos.

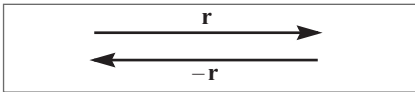


Figura 3.9 Deslocamento oposto.

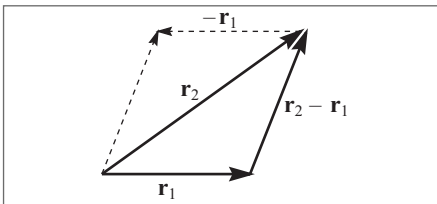


Figura 3.10 Diferença de deslocamento.

Vamos chamar o deslocamento resultante de “soma” dos deslocamentos \mathbf{r}_1 e \mathbf{r}_2 :

$$\mathbf{r} = \mathbf{r}_1 + \mathbf{r}_2 \quad (3.2.1)$$

Note que, com esta definição, a soma é comutativa (Figura 3.7)

$$\mathbf{r}_1 + \mathbf{r}_2 = \mathbf{r}_2 + \mathbf{r}_1 \quad (3.2.2)$$

e associativa (verifique!)

$$\mathbf{r}_1 + (\mathbf{r}_2 + \mathbf{r}_3) = (\mathbf{r}_1 + \mathbf{r}_2) + \mathbf{r}_3 \quad (3.2.3)$$

A soma de um número qualquer de deslocamentos obtém-se da forma indicada na Figura 3.8, unindo a origem do primeiro à extremidade do último.

Designamos por $\mathbf{0}$ um deslocamento nulo, $\mathbf{r} + \mathbf{0} = \mathbf{r}$. Para cada deslocamento \mathbf{r} , existe um deslocamento oposto, que designaremos por $-\mathbf{r}$, que leva de volta ao ponto de partida, e que difere de \mathbf{r} apenas pelo sentido (Figura 3.9): $\mathbf{r} + (-\mathbf{r}) = \mathbf{0}$.

Isto nos permite definir a *diferença* de dois deslocamentos por (Figura 3.10)

$$\mathbf{r}_2 - \mathbf{r}_1 = \mathbf{r}_2 + (-\mathbf{r}_1) \quad (3.2.4)$$

A Figura 3.10 mostra que $\mathbf{r}_2 - \mathbf{r}_1$ se obtém unindo a extremidade de \mathbf{r}_1 à extremidade de \mathbf{r}_2 (o que corresponde à outra diagonal na “regra do paralelogramo”).

Para $\lambda > 0$, o deslocamento $\lambda \mathbf{r}$ é um deslocamento de mesma direção e sentido que \mathbf{r} , mas de magnitude λ vezes maior; se $\lambda < 0$, o sentido muda. Com esta definição, é imediato que $\lambda (\mathbf{r}_1 + \mathbf{r}_2) = \lambda \mathbf{r}_1 + \lambda \mathbf{r}_2$, e que $(\lambda + \mu) \mathbf{r} = \lambda \mathbf{r} + \mu \mathbf{r}$.

As propriedades acima dos deslocamentos caracterizam o que chamamos de *vetores*. Grandezas físicas representadas apenas por um número, como o tempo ou a distância, chamam-se *grandezas escalares*; as que são representadas por vetores, como o deslocamento, chamam-se *grandezas vetoriais*. Por conseguinte, *uma grandeza física*

é um vetor quando é caracterizada por magnitude, direção e sentido e se comporta como um deslocamento, ou seja, obedece a leis de composição do mesmo tipo, que correspondem à soma de vetores e ao produto de um vetor por um escalar. As definições e propriedades destas operações para vetores quaisquer são idênticas às que foram vistas acima, bastando substituir a palavra “deslocamento” por “vetor”. Do ponto de vista matemático, um vetor é um elemento de um “espaço vetorial”, que se caracteriza precisamente pelas propriedades acima, da adição e produto por um escalar.

É importante notar que não basta que uma grandeza física seja caracterizada por sua magnitude, direção e sentido para que ela tenha caráter vetorial. É preciso ainda que ela obedeça às leis de composição consideradas acima, com todas as suas propriedades. Uma boa ilustração deste ponto é fornecida pelas rotações finitas em torno de eixos diferentes.

Com efeito, consideremos uma rotação por um ângulo θ em torno de um eixo. Poderíamos tentar associar-lhe um “vetor” “ θ ” que caracterizaria completamente a rotação, tomando “ θ ” na direção do eixo e de magnitude dada pelo ângulo de rotação θ ; o sentido de “ θ ” poderia ser associado ao sentido de rotação, convencionando-se que a rotação, vista a partir da “extremidade da seta” de “ θ ”, é no sentido anti-horário (Figura 3.11). Entretanto, embora “ θ ” tenha magnitude, direção e sentido, *não* é um vetor.

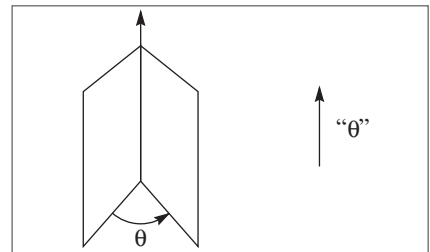


Figura 3.11 Representação de rotação finita.

Para ver isto notemos que a operação de composição de duas rotações finitas, representadas por “ θ_1 ” e “ θ_2 ” (em torno de eixos quaisquer), deveria corresponder à soma dos “vetores” correspondentes, “ $\theta_1 + \theta_2$ ”, da mesma forma que o deslocamento resultante de dois deslocamentos é a soma dos vetores correspondentes. Vamos mostrar agora que esta operação de “soma” deixaria de satisfazer à propriedade comutativa (3.2.2), ou seja, que em geral

$$“\theta_1 + \theta_2” \neq “\theta_2 + \theta_1” \quad (3.2.5)$$

No exemplo a seguir, “ θ_1 ” é uma rotação de $+90^\circ$ em torno do eixo Ox , e “ θ_2 ” é uma rotação de $+90^\circ$ em torno do eixo Oz . As Figuras 3.12 (a), (b) e (c) mostram respectivamente, a posição inicial de um objeto (livro) e os efeitos de aplicar primeiro “ θ_1 ”, depois “ θ_2 ”, levando em (c) ao resultado “ $\theta_1 + \theta_2$ ”; as Figuras (a’), (b’) e (c’) mostram os efeitos de tomar a ordem inversa; comparando (c) e (c’), vemos que vale a (3.2.5). Logo, as rotações finitas não são vetores.

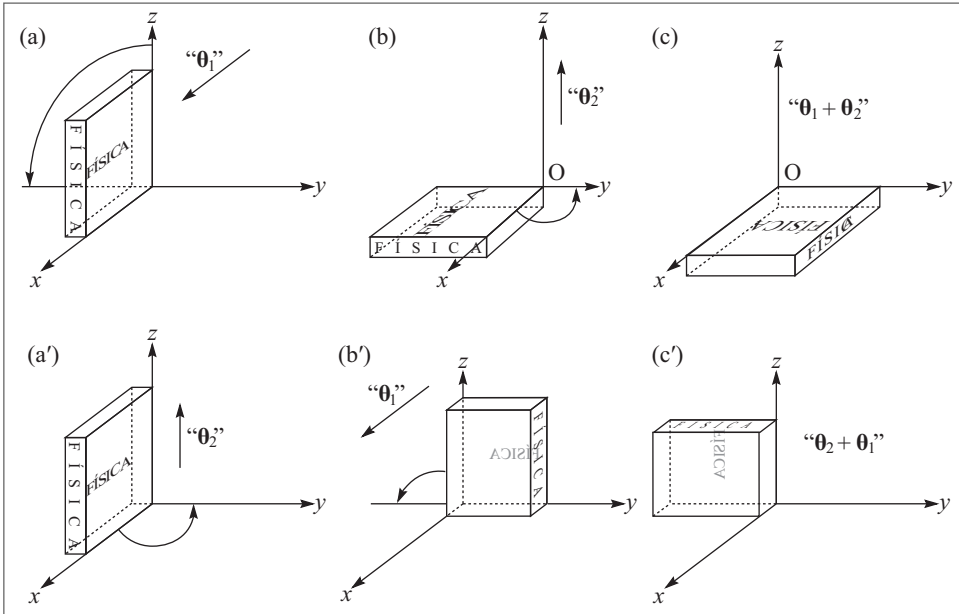


Figura 3.12 Não comutatividade da resultante de rotações finitas.

3.3 COMPONENTES DE UM VETOR

Podemos agora relacionar a descrição “intrínseca” de um deslocamento por um vetor (Seç. 3.2) com sua descrição em termos de coordenadas (Seç. 3.1), introduzindo as *componentes* de um vetor em relação a um sistema de coordenadas. Vamo-nos limitar por enquanto a vetores num plano, onde tomamos um sistema de coordenadas cartesianas.

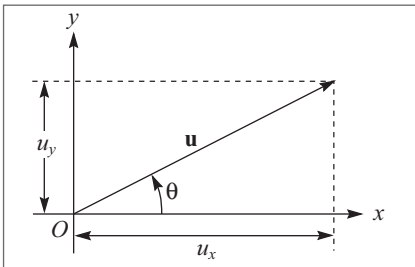


Figura 3.13 Componentes de um vetor.

Seja \mathbf{u} um vetor qualquer (na Figura 3.13, tomamos a origem de \mathbf{u} no ponto O, origem das coordenadas, o que não tem nada de restritivo, porque um vetor não está associado a uma origem determinada: um vetor obtido de \mathbf{u} por qualquer translação é igual a \mathbf{u}).

Chamam-se *componentes de \mathbf{u}* segundo os eixos Ox e Oy as projeções u_x e u_y de \mathbf{u} sobre esses eixos (Figura). A magnitude de \mathbf{u} (ou *módulo de \mathbf{u}*) é dada por

$$|\mathbf{u}| = \sqrt{u_x^2 + u_y^2} \quad (3.3.1)$$

Chama-se *vetor unitário* um vetor de módulo = 1. Costuma-se designar um vetor unitário na direção de \mathbf{u} por $\hat{\mathbf{u}}$, de forma que

$$\hat{\mathbf{u}} = \mathbf{u} / |\mathbf{u}| \quad (3.3.2)$$

Os vetores unitários nas direções de Ox e Oy são designados por \mathbf{i} e \mathbf{j} , respectivamente (ou então por $\hat{\mathbf{x}}$ e $\hat{\mathbf{y}}$). É imediato que (Figura 3.14)

$$\mathbf{u} = u_x \mathbf{i} + u_y \mathbf{j} = u_x \hat{\mathbf{x}} + u_y \hat{\mathbf{y}} \quad (3.3.3)$$

Se θ é o ângulo entre \mathbf{u} e Ox (Figura 3.13), temos

$$\begin{cases} u_x = |\mathbf{u}| \cos \theta \\ u_y = |\mathbf{u}| \sin \theta \end{cases} \quad (3.3.4)$$

o que permite obter θ em termos de u_x e u_y , com o auxílio da (3.3.1):

$$\begin{aligned} \cos \theta &= \frac{u_x}{\sqrt{u_x^2 + u_y^2}} \\ \sin \theta &= \frac{u_y}{\sqrt{u_x^2 + u_y^2}} \quad \text{tg } \theta = \frac{u_y}{u_x} \end{aligned} \quad (3.3.5)$$

A (3.3.3) mostra que

$$\mathbf{u} + \mathbf{v} = (u_x + v_x) \mathbf{i} + (u_y + v_y) \mathbf{j} \quad (3.3.6)$$

ou seja, que as componentes da soma de dois vetores são as somas das componentes correspondentes, o que também é óbvio pela Figura 3.15. Vemos também que

$$\lambda \mathbf{v} = \lambda v_x \mathbf{i} + \lambda v_y \mathbf{j} \quad (3.3.7)$$

ou seja, as componentes de $\lambda \mathbf{v}$ são $(\lambda v_x, \lambda v_y)$.

Num dado sistema de coordenadas, vemos assim que um vetor está associado a um par ordenado $\mathbf{u} \rightarrow (u_x, u_y)$, com $\mathbf{u} + \mathbf{v} \rightarrow (u_x + v_x, u_y + v_y)$, $\lambda \mathbf{u} \rightarrow (\lambda u_x, \lambda u_y)$. Entretanto, a recíproca não é verdadeira: nem todo par ordenado (u_x, u_y) define um vetor. Como vemos pela (3.3.3), o vetor \mathbf{u} só fica definido quando são dados também os vetores unitários (\mathbf{i}, \mathbf{j}) que definem as direções dos eixos do sistema de coordenadas, o que permite construir o vetor como entidade *intrínseca*, representável geometricamente de forma independente do sistema de coordenadas.

Se passarmos de um sistema de coordenadas Oxy a outro $Ox'y'$ de orientação diferente (o que equivale a uma *rotação dos eixos* em torno da origem), o significado intrínseco do vetor \mathbf{u} nos deve permitir calcular suas componentes (u'_x, u'_y) no novo sistema a partir de (u_x, u_y) e do ângulo de rotação ϕ . A Figura 3.16 mostra como isto se faz por projeção sobre os novos eixos, uma vez cons-

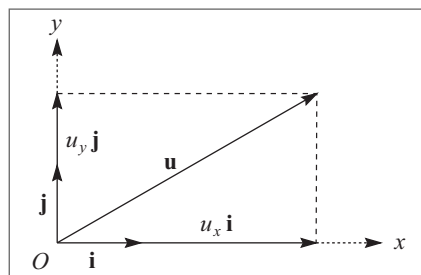


Figura 3.14 Decomposição de vetor.

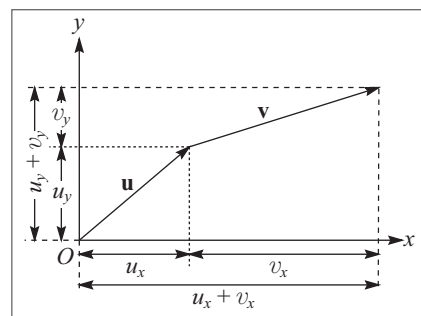


Figura 3.15 Componentes da soma.

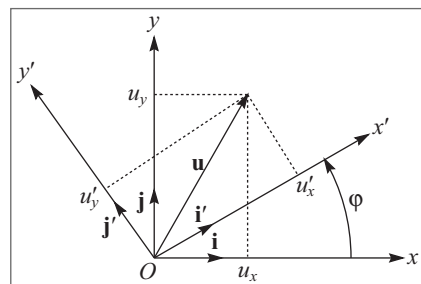


Figura 3.16 Rotação de eixos.

truído \mathbf{u} pela (3.3.3). Veremos depois a forma analítica da relação entre (u'_x, u'_y) e (u_x, u_y) , que dá a *lei de transformação* das componentes de um vetor numa rotação de eixos. O importante é perceber que essa lei de transformação é bem definida e característica de um vetor, refletindo o seu caráter intrínseco. Assim, por exemplo, se tivermos

$$u_x + v_x = w_x, \quad u_y + v_y = w_y \quad (3.3.8)$$

correspondendo à relação intrínseca $\mathbf{u} + \mathbf{v} = \mathbf{w}$, a lei de transformação tem de ser tal que, em relação ao novo sistema de coordenadas, tenha-se também

$$u'_x + v'_x = w'_x, \quad u'_y + v'_y = w'_y \quad (3.3.9)$$

Exemplo: Deslocamento relativo

Sejam $\mathbf{r}_1 = \mathbf{OP}_1$ e $\mathbf{r}_2 = \mathbf{OP}_2$, os deslocamentos de dois pontos P_1 e P_2 em relação à origem O (Figura 3.17). Chama-se *deslocamento relativo* de P_2 em relação a P_1 o vetor $\mathbf{P}_1\mathbf{P}_2 = \mathbf{r}_{12}$ definido por

$$\mathbf{r}_{12} = \mathbf{r}_2 - \mathbf{r}_1 \quad (3.3.10)$$

A Figura 3.17 mostra a construção gráfica que dá \mathbf{r}_{12} . Uma caracterização intrínseca de \mathbf{r}_{12} pode ser obtida dando seu módulo e o ângulo β entre \mathbf{r}_{12} e \mathbf{r}_1 (por exemplo).

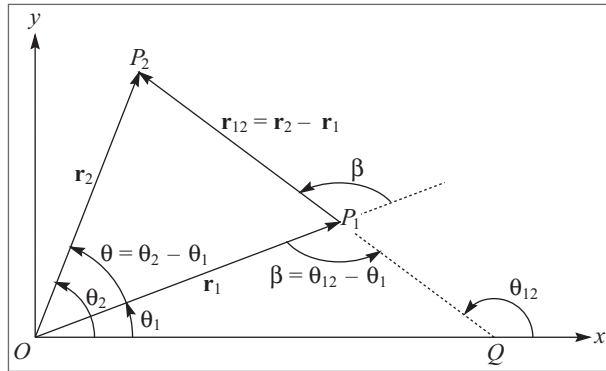


Figura 3.17 Deslocamento relativo.

Se θ é o ângulo entre \mathbf{r}_1 e \mathbf{r}_2 , a lei dos cossenos aplicada ao triângulo OP_1P_2 dá

$$|\mathbf{r}_{12}|^2 = r_1^2 + r_2^2 - 2r_1r_2 \cos \theta \quad (3.3.11)$$

o que determina $|\mathbf{r}_{12}|$. A lei dos senos dá

$$\frac{|\mathbf{r}_{12}|}{\sin \theta} = \frac{r_2}{\sin(\pi - \beta)} = \frac{r_2}{\sin \beta} \left\{ \begin{array}{l} \sin \beta = \frac{r_2 \sin \theta}{|\mathbf{r}_{12}|} \end{array} \right. \quad (3.3.12)$$

Por outro lado, tomando um sistema de coordenadas Oxy (Figura 3.17), as (3.3.10) e (3.3.1) dão

$$|\mathbf{r}_{12}|^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 = (x_2^2 + y_2^2) + (x_1^2 + y_1^2) - 2(x_2x_1 + y_2y_1) \quad (3.3.13)$$

onde, pelas (3.3.4), $x_i = r_i \cos \theta_i$, $y_i = r_i \sin \theta_i$ ($i = 1, 2$), de modo que a (3.3.13) se escreve

$$|\mathbf{r}_{12}|^2 = r_2^2 + r_1^2 - 2r_1r_2 \underbrace{(\cos \theta_2 \cos \theta_1 + \sin \theta_2 \sin \theta_1)}_{\cos(\theta_2 - \theta_1)}$$

o que coincide com a (3.3.11), pois $\theta = \theta_2 - \theta_1$ (Figura 3.17).

Analogamente, se θ_{12} é o ângulo entre \mathbf{r}_{12} e Ox , as (3.3.5) dão

$$\begin{aligned} \cos \theta_{12} &= \frac{x_2 - x_1}{|\mathbf{r}_{12}|} = \frac{r_2 \cos \theta_2 - r_1 \cos \theta_1}{|\mathbf{r}_{12}|} \\ \sin \theta_{12} &= \frac{y_2 - y_1}{|\mathbf{r}_{12}|} = \frac{r_2 \sin \theta_2 - r_1 \sin \theta_1}{|\mathbf{r}_{12}|} \end{aligned}$$

o que dá

$$|\mathbf{r}_{12}| \underbrace{(\sin \theta_{12} \cos \theta_1 - \cos \theta_{12} \sin \theta_1)}_{\sin(\theta_{12} - \theta_1)} = r_2 \overbrace{(\sin \theta_2 \cos \theta_1 - \sin \theta_1 \cos \theta_2)}^{\sin(\theta_2 - \theta_1)}$$

o que coincide com a (3.3.12), pois $\beta = \theta_2 - \theta_1$, conforme vemos pelo triângulo OP_1Q . Vemos assim que os resultados (3.3.11) e (3.3.12) também podem ser obtidos a partir das componentes dos vetores num dado sistema de coordenadas.

3.4 VELOCIDADE E ACELERAÇÃO VETORIAIS

Consideremos uma partícula, em movimento num plano, que descreve uma trajetória APB, em relação a um sistema de referência Oxy . Seja $\mathbf{r}(t) = \mathbf{OP}$ o deslocamento da partícula em relação à origem O no instante t , onde P é a posição ocupada pela partícula no instante t ; seja $\mathbf{r}(t + \Delta t) = \mathbf{OP}'$ o deslocamento no instante $t + \Delta t$. Pela (3.3.10), o deslocamento relativo da partícula entre os instantes t e $t + \Delta t$ é o vetor (Figura 3.18)

$$\mathbf{PP}' = \Delta \mathbf{r} = \mathbf{r}(t + \Delta t) - \mathbf{r}(t) \quad (3.4.1)$$

Por analogia com a (2.1.5), é natural definirmos a *velocidade média* entre os instantes t e $t + \Delta t$ por

$$\mathbf{v}_{t \rightarrow t + \Delta t} = \frac{\mathbf{r}(t + \Delta t) - \mathbf{r}(t)}{\Delta t} = \frac{\Delta \mathbf{r}}{\Delta t} \quad (3.4.2)$$

Como a diferença entre dois vetores e o produto de um vetor por um escalar são vetores, a (3.4.2) mostra que a velocidade média é um vetor, cuja direção e sentido são os da corda \mathbf{PP}' que liga as posições nos instantes t e $t + \Delta t$ sobre a trajetória.

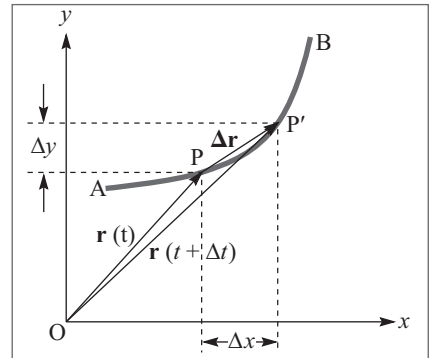


Figura 3.18 Trajetória plana.

As componentes da velocidade média são

$$\left. \begin{aligned} v_{x(t \rightarrow t + \Delta t)} &= \frac{\Delta x}{\Delta t} \\ v_{y(t \rightarrow t + \Delta t)} &= \frac{\Delta y}{\Delta t} \end{aligned} \right\} \quad (3.4.3)$$

ou seja, são exatamente as velocidades médias dos movimentos unidimensionais descritos pelas projeções $x(t)$, $y(t)$ do deslocamento instantâneo $\mathbf{r}(t)$ sobre os eixos.

Quando $\Delta t \rightarrow 0$, sabemos pelas (2.2.2) e (2.2.5) que as (3.4.3) levam a

$$\left. \begin{aligned} v_x(t) &= \lim_{\Delta t \rightarrow 0} \left(\frac{\Delta x}{\Delta t} \right) = \frac{dx}{dt} \\ v_y(t) &= \lim_{\Delta t \rightarrow 0} \left(\frac{\Delta y}{\Delta t} \right) = \frac{dy}{dt} \end{aligned} \right\} \quad (3.4.4)$$

que representam as velocidades instantâneas dos movimento unidimensionais descritos pelas projeções. Isto sugere definir a *velocidade instantânea no instante t* por

$$\mathbf{v}(t) = \lim_{\Delta t \rightarrow 0} \left(\frac{\Delta \mathbf{r}}{\Delta t} \right) = \frac{d\mathbf{r}}{dt} = \frac{dx}{dt} \mathbf{i} + \frac{dy}{dt} \mathbf{j} = v_x(t) \mathbf{i} + v_y(t) \mathbf{j} \quad (3.4.5)$$

o que define ao mesmo tempo o conceito de *derivada de um vetor* dependente de um parâmetro (t) em relação a este parâmetro.

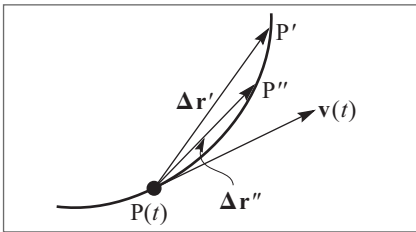


Figura 3.19 Velocidade vetorial.

Observando o comportamento de $\Delta \mathbf{r}$ à medida que $\Delta t \rightarrow 0$ (Figura 3.19), vemos também que a direção da velocidade instantânea $\mathbf{v}(t)$ é a da *tangente à trajetória* em $P(t)$, e o sentido é o sentido de percurso da trajetória para t crescente. Obtemos assim a direção e o sentido de \mathbf{v} , mas como sabemos que é um vetor? Como vimos após a (2.2.4),

$$\frac{d}{dt} [x_1(t) + x_2(t)] = \frac{dx_1}{dt} + \frac{dx_2}{dt}, \quad \frac{d}{dt} [\lambda x(t)] = \lambda \frac{dx}{dt}$$

de modo que (cf. (3.3.6), (3.3.7)) a definição (3.4.5) satisfaz a todas as leis de composição que caracterizam um vetor. Podemos concluir, de forma mais geral, que a *derivada de um vetor é um vetor*.

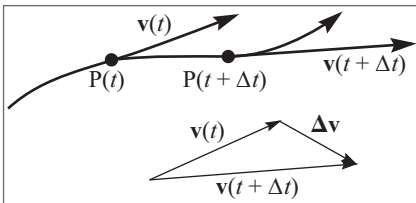


Figura 3.20 Aceleração vetorial.

Para definir a *aceleração média* de forma análoga, consideremos um intervalo $[t, t + \Delta t]$ e sejam $\mathbf{v}(t)$, $\mathbf{v}(t + \Delta t)$ os vetores velocidade instantâneos nos extremos do intervalo, que são tangentes à trajetória nos pontos correspondentes $P(t)$ e $P(t + \Delta t)$ (Figura 3.20). Por definição (cf. (3.4.2)),

$$\bar{\mathbf{a}}_{t \rightarrow t+\Delta t} = \frac{\mathbf{v}(t+\Delta t) - \mathbf{v}(t)}{\Delta t} = \frac{\Delta \mathbf{v}}{\Delta t} \quad (3.4.6)$$

é o vetor *aceleração média* no intervalo $t \rightarrow t + \Delta t$.

A *aceleração instantânea* no instante t é o vetor

$$\mathbf{a}(t) = \lim_{\Delta t \rightarrow 0} \left[\frac{\mathbf{v}(t+\Delta t) - \mathbf{v}(t)}{\Delta t} \right] = \lim_{\Delta t \rightarrow 0} \left(\frac{\Delta \mathbf{v}}{\Delta t} \right) = \frac{d\mathbf{v}}{dt} \quad (3.4.7)$$

ou seja, é a derivada do vetor velocidade instantânea em relação ao tempo. Pela (3.4.5), também podemos escrever

$$\mathbf{a}(t) = \frac{d^2 \mathbf{r}}{dt^2} = \frac{d^2 x}{dt^2} \mathbf{i} + \frac{d^2 y}{dt^2} \mathbf{j} \quad (3.4.8)$$

introduzindo assim ao mesmo tempo a *derivada segunda* de um vetor.

Para ter uma interpretação geométrica do vetor aceleração instantânea, basta aplicar a interpretação geométrica da derivada de um vetor, discutida após a (3.4.5). Se, a partir de uma origem comum O , representarmos os vetores velocidade associados aos diferentes pontos da trajetória (Figura 3.21), a extremidade do vetor $\mathbf{v}(t)$ descreverá uma curva (em linha interrompida na Figura) que se chama *hodógrafo* do movimento. O vetor $\mathbf{a}(t)$ é tangente ao hodógrafo no ponto correspondente $\mathbf{v}(t)$. Vemos que em geral $\mathbf{a}(t)$ não será tangente à trajetória. Geralmente omitiremos a palavra “instantânea”. Quando nos referirmos à *velocidade* e à *aceleração* no instante t , estas expressões designarão $\mathbf{v}(t)$ e $\mathbf{a}(t)$, respectivamente. Um resultado fundamental da discussão acima é que a aceleração não está associada apenas a uma variação do *módulo* da velocidade: conforme ilustrado na Figura 3.22, *uma variação de direção da velocidade também representa uma aceleração*. Assim, se um carro percorre uma pista circular, ele tem aceleração, mesmo quando o ponteiro do velocímetro indica sempre o mesmo valor! Este resultado decorre do caráter vetorial da velocidade e da aceleração. O motorista percebe que tem de pisar no acelerador para descrever a curva.

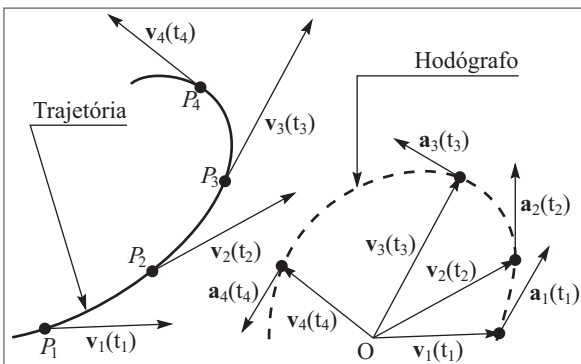


Figura 3.21 Hodógrafo.

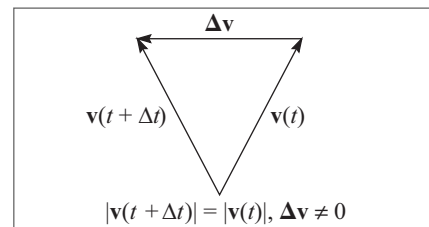


Figura 3.22 Variação da direção da velocidade.

3.5 MOVIMENTO UNIFORMEMENTE ACELERADO

Um movimento qualquer chama-se *uniformemente acelerado* quando a aceleração é constante (independentemente do tempo):

$$\boxed{\mathbf{a}(t) = \mathbf{a} = \text{constante}} \quad (3.5.1)$$

onde “constante”, para um vetor, significa *constante em módulo, direção e sentido*.

Analogamente à discussão da Seq. 2.5, para determinar o movimento é preciso ainda dar as *condições iniciais*:

$$\boxed{\begin{aligned} \mathbf{v}(t_0) &= \mathbf{v}_0 \\ \mathbf{r}(t_0) &= \mathbf{r}_0 \end{aligned}} \quad (3.5.2)$$

No instante $t_0 + \Delta t$, \mathbf{v} e \mathbf{r} terão variado respectivamente de $\Delta \mathbf{v}$ e $\Delta \mathbf{r}$, onde, para Δt pequeno o suficiente (de modo que possamos confundir aceleração e velocidade médias e instantâneas), teremos

$$\left. \begin{aligned} \Delta \mathbf{v} &= \mathbf{a} \Delta t \\ \Delta \mathbf{r} &= \mathbf{v}_0 \Delta t \end{aligned} \right\} \quad (3.5.3)$$

Se \mathbf{v}_0 é paralelo a \mathbf{a} , as (3.5.3) mostram que o movimento será retilíneo, segundo a direção paralela a \mathbf{v}_0 e \mathbf{a} que passa por \mathbf{r}_0 (com efeito, podemos repetir o raciocínio a partir de $t_0 + \Delta t$, porque $\mathbf{v}_0 + \Delta \mathbf{v}$ continua neste caso sendo paralelo a \mathbf{a}). Recaímos então no caso do movimento retilíneo uniformemente acelerado, já estudado na Seq. 2.5.

Vamos supor então que \mathbf{v}_0 não é paralelo a \mathbf{a} , de forma que as direções de \mathbf{v}_0 e \mathbf{a} definem um plano, ou melhor, uma família de planos paralelos. As (3.5.3) mostram então que o movimento estará contido no plano dessa família que passa pela posição inicial \mathbf{r}_0 , ou seja, *o movimento é bidimensional*. Podemos assim, sem restrição de generalidade, tomar a origem neste plano.

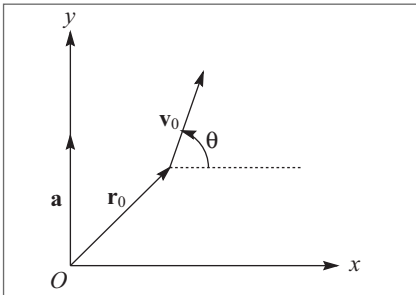


Figura 3.23 Condições iniciais.

Vamos adotar um sistema de coordenadas cartesianas com eixo Oy segundo a direção de \mathbf{a} (Figura 3.23). Temos então

$$\left. \begin{aligned} \mathbf{a} &= a \mathbf{j} \\ \mathbf{v}_0 &= v_{0x} \mathbf{i} + v_{0y} \mathbf{j} \\ \mathbf{r}_0 &= x_0 \mathbf{i} + y_0 \mathbf{j} \end{aligned} \right\} \quad (3.5.4)$$

As projeções do movimento sobre os eixos x e y obedecerão então a

$$\left. \begin{aligned} a_y &= a = \text{constante}; & v_y(t_0) &= v_{0y}; & y(t_0) &= y_0 \\ a_x &= 0; & v_x(t_0) &= v_{0x}; & x(t_0) &= x_0 \end{aligned} \right\} \quad (3.5.5)$$

que correspondem a movimentos unidimensionais do tipo já considerado na Seq. 2.5. Podemos então aplicar imediatamente as (2.5.4) e (2.5.8), obtendo.

$$\left. \begin{aligned} v_y(t) &= v_{0y} + a(t - t_0) \\ v_x(t) &= v_{0x} \end{aligned} \right\} \quad (3.5.6)$$

$$\left. \begin{aligned} y(t) &= y_0 + v_{0y}(t - t_0) + \frac{1}{2}a(t - t_0)^2 \\ x(t) &= x_0 + v_{0x}(t - t_0) \end{aligned} \right\} \quad (3.5.7)$$

Em forma vetorial, estes resultados se tornam:

$$\mathbf{v}(t) = \mathbf{v}_0 + \mathbf{a}(t - t_0) \quad (3.5.8)$$

$$\mathbf{r}(t) = \mathbf{r}_0 + \mathbf{v}_0(t - t_0) + \frac{1}{2}\mathbf{a}(t - t_0)^2 \quad (3.5.9)$$

que dão a solução do problema de valores iniciais posto pelas (3.5.1) e (3.5.2).

No caso particular em que $\mathbf{a} = 0$, recaímos no movimento retilíneo uniforme (Figura 3.24). Com efeito, neste caso, as (3.5.7) dão

$$\left. \begin{aligned} x &= x_0 + v_{0x}(t - t_0) \\ y &= y_0 + v_{0y}(t - t_0) \end{aligned} \right\} \frac{y - y_0}{v_{0y}} = \frac{x - x_0}{v_{0x}} \quad (3.5.10)$$

que é a equação de uma reta (trajetória).

Para obter a forma da trajetória no caso geral do movimento uniformemente acelerado, basta eliminar $t - t_0$ entre as (3.5.7). A condição de que \mathbf{v}_0 não é paralelo a \mathbf{a} dá

$$v_{0x} \neq 0 \quad (3.5.11)$$

permitindo obter $t - t_0$ da segunda (3.5.7):

$$t - t_0 = \frac{x - x_0}{v_{0x}} \quad (3.5.12)$$

Substituindo na primeira (3.5.7), obtemos

$$y - y_0 = \left(\frac{v_{0y}}{v_{0x}} \right) (x - x_0) + \frac{1}{2} \frac{a}{v_{0x}^2} (x - x_0)^2 \quad (3.5.13)$$

que é a equação de uma *parábola* de eixo vertical, que passa por (x_0, y_0) , e cuja tangente neste ponto tem a direção de \mathbf{v}_0 (por construção).

As (3.5.7) mostram que o movimento ao longo da parábola (3.5.13) pode ser considerado como resultante da composição de um movimento uniforme na direção horizontal com um movimento uniformemente acelerado na direção vertical.

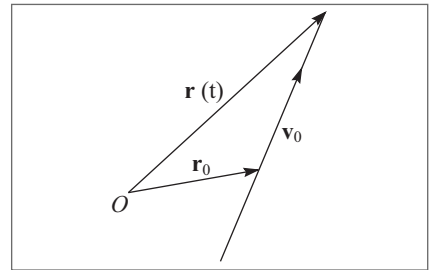


Figura 3.24 Movimento retilíneo uniforme.

3.6 MOVIMENTO DOS PROJÉTEIS

Uma aplicação importante dos resultados da Seção 3.5, é o movimento dos projéteis na vizinhança da superfície da Terra. Na balística usual, podemos considerar a Terra como plana e a aceleração da gravidade como constante (cf. (2.6.1)) (isto não seria verdade para foguetes balísticos intercontinentais!). Desprezaremos, também, o efeito da resistência do ar.

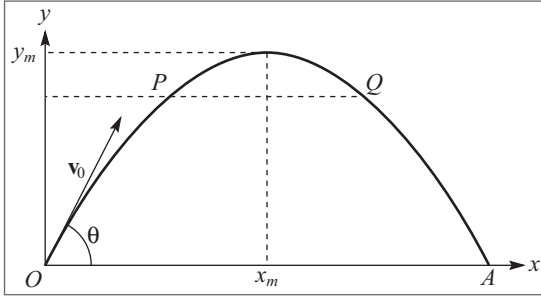


Figura 3.25 Trajetória parabólica.

Pela convenção da Seção 3.5, temos de tomar o eixo Oy segundo a vertical. Vamos orientá-lo apontando para cima, de modo que, na (3.5.4), $a = -g$:

$$\mathbf{a} = -g \mathbf{j} \quad (3.6.1)$$

Vamo-nos limitar também ao caso em que $x_0 = y_0 = 0$, tomando a posição inicial na origem, e vamos tornar $t_0 = 0$. Seja θ o ângulo entre \mathbf{v}_0 e Ox (Figura 3.25), de modo que

$$v_{0x} = v_0 \cos \theta \quad v_{0y} = v_0 \sin \theta \quad (3.6.2)$$

As (3.5.6) e (3.5.7) ficam

$$v_y = v_0 \sin \theta - gt \quad v_x = v_0 \cos \theta \quad (3.6.3)$$

$$\left. \begin{aligned} y &= v_0 \sin \theta t - \frac{1}{2} g t^2 \\ x &= v_0 \cos \theta t \end{aligned} \right\} \quad (3.6.4)$$

e a equação da trajetória (3.5.13) fica

$$y = \tan \theta \cdot x - \frac{g x^2}{2 v_0^2 \cos^2 \theta} \quad (3.6.5)$$

Conforme mostra a Figura 3.25, a altura máxima y_m atingida pelo projétil corresponde ao instante t_m em que v_y se anula, ou seja, pela (3.6.3),

$$t_m = \frac{v_0 \sin \theta}{g} \quad (3.6.6)$$

e o valor correspondente de y é dado pela (3.6.4):

$$y_m = v_0 \sin \theta \cdot \frac{v_0 \sin \theta}{g} - \frac{1}{2} g \cdot \frac{v_0^2 \sin^2 \theta}{g^2}$$

ou seja,

$$y_m = \frac{v_0^2 \sin^2 \theta}{2g} \quad (3.6.7)$$

Quanto tempo o projétil leva para atingir o solo no ponto $x = A$ (Figura 3.25)? Fazendo $y = 0$ na primeira das (3.6.4), obtemos uma equação do 2º grau em t , em que uma das raízes é $t = 0$, correspondendo ao ponto de lançamento, e a outra é

$$t = t_A = \frac{2v_0 \sin \theta}{g} = 2t_m \quad (3.6.8)$$

ou seja, é o dobro do tempo que leva para atingir a altura máxima, o que poderíamos ter inferido pela simetria da trajetória com respeito a $x = x_m$.

Com que velocidade o projétil atinge o solo? Basta fazer $t = t_A$ na (3.6.3):

$$\left. \begin{aligned} v_y(t_A) &= v_0 \sin \theta - g t_A = -v_0 \sin \theta \\ v_z(t_A) &= v_0 \cos \theta \end{aligned} \right\} |\mathbf{v}(t_A)| = |\mathbf{v}_0| \quad (3.6.9)$$

Logo, ao atingir o solo, a velocidade do projétil só difere da velocidade inicial \mathbf{v}_0 pela inversão da componente vertical ($v_y \rightarrow -v_y$), e tem o mesmo módulo. Como $y = 0$ é um plano arbitrário, o mesmo vale em qualquer plano horizontal ($y = \text{constante}$), ou seja, também se aplica às velocidades nos dois pontos P e Q em que a parábola corta um dado plano horizontal (Figura 3.25).

Podemos exprimir as componentes da velocidade diretamente em função da altura y com o auxílio da (2.5.9):

$$v_y = \pm \sqrt{v_0^2 \sin^2 \theta - 2gy}, \quad v_x = v_0 \cos \theta \quad (3.6.10)$$

onde o sinal é + ou – conforme o projétil esteja subindo ou descendo.

A distância $x = A$ entre o ponto de lançamento O e o ponto em que o projétil volta a passar pelo plano $y = 0$ chama-se *alcance* do projétil, e se obtém substituindo a (3.6.8) na segunda (3.6.4):

$$A = v_0 \cos \theta \cdot \frac{2v_0 \sin \theta}{g} = \frac{v_0^2}{g} \sin(2\theta) \quad (3.6.11)$$

onde usamos a bem conhecida relação trigonométrica: $\sin(2\theta) = 2 \sin \theta \cos \theta$. Uma consequência imediata da (3.6.11) é que o alcance é máximo quando o “ângulo de elevação” θ vale 45°.

Na última jornada dos “Diálogos”, Galileu discute o movimento dos projéteis. Precursores de Galileu acreditavam que uma bala de canhão se move em linha reta até esgotar seu impulso, e depois cai verticalmente (um deles propôs juntar esses dois segmentos de reta por um arco de círculo tangente a ambos para descrever a trajetória). Galileu foi o primeiro a demonstrar que a trajetória é uma parábola. Além disso, obteve vários dos resultados discutidos acima, inclusive que o alcance é máximo para $\theta = 45^\circ$, enunciando ainda o seguinte resultado:

“As amplitudes das parábolas descritas por projéteis disparados com a mesma velocidade, mas em ângulos de elevação acima e abaixo de 45° e equidistantes de 45°, são iguais entre si”.

Isto significa que, para o mesmo valor de $|\mathbf{v}_0|$, os alcances correspondentes a $\theta = 45^\circ + \delta$ e $\theta = 45^\circ - \delta$ são iguais. Verifique este resultado!

Galileu também observou o fato de que todos estes resultados sobre o movimento de projéteis são bastante idealizados, uma vez que não foi levado em conta o efeito da *resistência do ar*, que tende a diminuir o alcance e alterar o caráter do movimento. Este efeito é bastante complicado, porque a resistência do ar depende da forma do projétil e do *módulo* da velocidade instantânea, $|\mathbf{v}| = \sqrt{v_x^2 + v_y^2}$, de modo que *acopla* os movimentos horizontal e vertical, que não podem mais ser considerados como independentes. Entretanto, para projéteis de forma aerodinâmica (como as balas de armas de fogo) e lançados com velocidades iniciais elevadas, os resultados acima constituem geralmente uma boa aproximação.

É interessante observar que nos podemos aproximar melhor das condições ideais em que a resistência do ar está ausente utilizando feixes de partículas atômicas ou subatômicas (como elétrons) lançados numa região de alto vácuo.

No caso de elétrons, o efeito foi observado nas primeiras experiências que levaram à descoberta do elétron, feitas por J. J. Thomson em 1897.

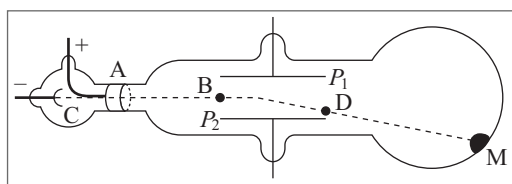


Figura 3.26 Tubo de raios catódicos.

retilínea ao passar entre as placas P_1 e P_2 (Figura), entre as quais se estabelece uma diferença de potencial. Veremos depois que isto equivale a superpor ao movimento retilíneo uniforme dos elétrons no feixe, segundo a horizontal, um movimento uniformemente acelerado na vertical, onde a aceleração se deve ao campo elétrico entre as placas. A porção de trajetória BD na região entre as placas (Figura) é então um arco de parábola, e o feixe assim defletido é detectado pela mancha luminosa M que produz ao incidir sobre um depósito fluorescente na parede interna do tubo. Thomson também investigou o efeito de um campo magnético sobre o feixe (deflexão magnética), e essas experiências lhe permitiram medir a razão da carga elétrica para a massa do elétron, conforme será visto posteriormente (Seç. 5.4).

No caso de feixes atômicos, foi detectado o efeito de *queda livre dos átomos*, em experiências realizadas por Estermann, Simpson e Stern* em 1947. Numa delas, foi utilizado um feixe de átomos de cézio, proveniente de um “forno” à temperatura de 450 K. Átomos de cézio a essa temperatura têm velocidades médias da ordem de 300 m/s; um feixe colimado, propagando-se na direção horizontal, era extraído do forno, penetrando num tubo onde se fazia alto vácuo, e nele viajando uma distância de ~2 m. O tempo de

O aparelho utilizado por Thomson (Figura 3.26) era uma versão primitiva do moderno tubo de osciloscópio ou de televisão, conhecido como “tubo de raios catódicos”. Um feixe de elétrons (“raios catódicos”) produzido numa descarga elétrica num gás rarefeito entre os eletrodos C (catodo) e A (anodo), é defletido de sua trajetória

* 1. Estermann, O.C. Sinipson e O. Stern, *Phys. Rev.* 71, 238 (1947).

percurso correspondente é $(2/300)$ s. De quanto os átomos caem nesse tempo sob a ação da gravidade?

Caem de $1/2 g t^2 = 1/2 \times 9,8 \times (1/150)^2 \text{ m} \approx 2 \times 10^{-4} \text{ m}$, ou seja, de $\sim 0,2 \text{ mm}$. Embora se trate de um deslocamento pequeno, ele pode ser detectado com relativa facilidade, porque o diâmetro do feixe e do fio utilizado no detector são ~ 10 vezes menores (é um “detector de fio quente”, em que os átomos são ionizados e depois coletados por um eletrodo, medindo-se a corrente elétrica resultante). Devido ao alto vácuo, os resultados obtidos nesta seção sobre trajetórias parabólicas se aplicam com grande precisão. O objetivo das experiências não era estudar a queda livre, mas sim testar resultados da teoria cinética dos gases, sobre a distribuição de velocidades dos átomos. Também têm sido feitas recentemente tentativas de observar a queda livre de elétrons, mas a experiência, realizada dentro de um cilindro evacuado, se torna muito mais difícil neste caso, devido a forças de origem elétrica, oriundas das paredes dos cilindro, que atuam sobre os elétrons e têm efeitos dominantes, de modo que não há ainda evidência clara de que o efeito tenha sido observado.

Em 1999, foi produzido um “chafariz de átomos” em que eles foram lançados no vácuo, descrevendo trajetórias parabólicas. O tempo de atraso entre subida e descida foi empregado na construção do relógio atômico NIST-F1 (Seç. 1.7).

3.7 MOVIMENTO CIRCULAR UNIFORME

Um tipo de movimento plano de grande importância na física é o movimento circular uniforme, em que a trajetória é um círculo e o *módulo* da velocidade instantânea é constante, de modo que a partícula descreve arcos de círculo iguais em tempos iguais. Temos assim um movimento periódico, em que o período corresponde ao tempo levado para descrever uma volta completa, o que define um “relógio”.

De fato, o movimento da extremidade dos ponteiros de um relógio é deste tipo. O movimento da Lua em torno da Terra também pode ser aproximado por um movimento circular uniforme. Outro exemplo são as órbitas de partículas carregadas em aceleradores de tipo circular.

Seja r o raio da trajetória circular. A posição instantânea P da partícula fica definida pelo ângulo θ entre o vetor deslocamento $\mathbf{r} = \mathbf{OP}$ correspondente e o eixo Ox de um sistema cartesiano com origem no centro do círculo (Figura 3.27), onde θ é positivo no sentido anti-horário. O arco s correspondente ao ângulo θ sobre o círculo é dado por

$$s = r\theta$$

(3.7.1)

onde θ é medido em radianos ($2\pi \text{ rad} = 360^\circ$). Vamos introduzir $\hat{\mathbf{r}}$, o vetor unitário na direção de \mathbf{r} , que aponta radialmente para fora, e $\hat{\boldsymbol{\theta}}$, o vetor unitário tangente ao círculo

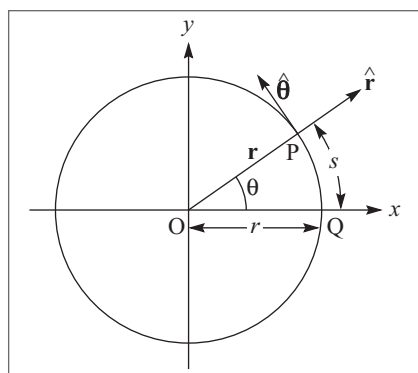


Figura 3.27 Movimento circular.

```

if (v*vold < 0) {
    control.println("maximum height = " + y);
}

```

where $v = v_{n+1}$ and $vold = v_n$. Why is this criterion preferable to other criteria that you might imagine using? \square

3.8 Two-Dimensional Trajectories

You are probably familiar with two-dimensional trajectory problems in the absence of air resistance. For example, if a ball is thrown in the air with an initial velocity v_0 at an angle θ_0 with respect to the ground, how far will the ball travel in the horizontal direction, and what is its maximum height and time of flight? Suppose that a ball is released at a nonzero height h above the ground. What is the launch angle for the maximum range? Are your answers still applicable if air resistance is taken into account? We consider these and similar questions in the following.

Consider an object of mass m whose initial velocity \mathbf{v}_0 is directed at an angle θ_0 above the horizontal [see Figure 3.3(a)]. The particle is subjected to gravitational and drag forces of magnitude mg and F_d ; the direction of the drag force is opposite to \mathbf{v} (see Figure 3.33.3(b)). Newton's equations of motion for the x and y components of the motion can be written as

$$m \frac{dv_x}{dt} = -F_d \cos \theta \quad (3.17a)$$

$$m \frac{dv_y}{dt} = -mg - F_d \sin \theta. \quad (3.17b)$$

For example, let us maximize the range of a round steel ball of radius 4 cm. A reasonable assumption for a steel ball of this size and typical speed is that $F_d = C_2 v^2$. Because $v_x = v \cos \theta$ and $v_y = v \sin \theta$, we can rewrite (3.17) as

$$m \frac{dv_x}{dt} = -C_2 v v_x \quad (3.18a)$$

$$m \frac{dv_y}{dt} = -mg - C_2 v v_y. \quad (3.18b)$$

Note that $-C_2 v v_x$ and $-C_2 v v_y$ are the x and y components of the drag force $-C_2 v^2$. Because (3.18a) and (3.18b) for the change in v_x and v_y involve the square of the velocity, $v^2 = v_x^2 + v_y^2$, we cannot calculate the vertical motion of a falling body without reference to the horizontal component, that is, the motion in the x and y direction is *coupled*.

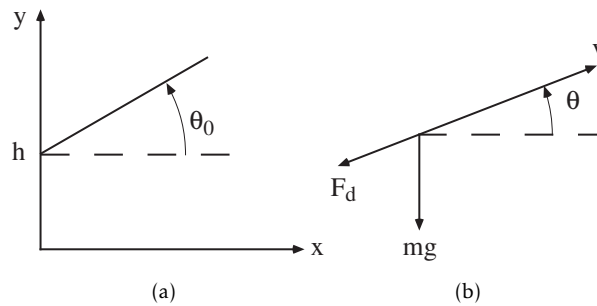


Figure 3.3: (a) A ball is thrown from a height h at a launch angle θ_0 measured with respect to the horizontal. The initial velocity is \mathbf{v}_0 . (b) The gravitational and drag forces on a particle.

Problem 3.10. Trajectory of a steel ball

- Use `Projectile` and `ProjectileApp` to compute the two-dimensional trajectory of a ball moving in air without air friction and plot y as a function of x . Compare your computed results with the exact results. For example, assume that a ball is thrown from ground level at an angle θ_0 above the horizontal with an initial velocity of $v_0 = 15$ m/s. Vary θ_0 and show that the maximum range occurs at $\theta_0 = \theta_{\max} = 45^\circ$. What is R_{\max} , the maximum range, at this angle? Compare your numerical result to the analytic result $R_{\max} = v_0^2/g$.
- Suppose that a steel ball is thrown from a height h at an angle θ_0 above the horizontal with the same initial speed as in part (a). If you neglect air resistance, do you expect θ_{\max} to be larger or smaller than 45° ? What is θ_{\max} for $h = 2$ m? By what percent is the range R changed if θ is varied by 2% from θ_{\max} ?
- Consider the effects of air resistance on the range and optimum angle of a steel ball. For a ball of mass 7 kg and cross-sectional area 0.01 m^2 , the parameter $C_2 \approx 0.1$. What are the units of C_2 ? It is convenient to exaggerate the effects of air resistance so that you can more easily determine the qualitative nature of the effects. Hence, compute the optimum angle for $h = 2$ m, $v_0 = 30$ m/s, and $C_2/m = 0.1$ and compare your answer to the value found in part (b). Is R more or less sensitive to changes in θ_0 from θ_{\max} than in part (b)? Determine the optimum launch angle and the corresponding range for the more realistic value of $C_2 = 0.1$. A detailed discussion of the maximum range of the ball has been given by Lichtenberg and Wills. \square

Problem 3.11. Comparing the motion of two objects

Consider the motion of two identical objects that both start from a height h . One object is dropped vertically from rest and the other is thrown with a horizontal velocity v_0 . Which object reaches the ground first?

- Give reasons for your answer assuming that air resistance can be neglected.
- Assume that air resistance cannot be neglected and that the drag force is proportional to v^2 . Give reasons for your anticipated answer for this case. Then perform numerical simulations using, for example, $C_2/m = 0.1$, $h = 10$ m, and $v_0 = 30$ m/s. Are your qualitative results consistent with your anticipated answer? If they are not, the source of the discrepancy

SEMANA 7

Chapter 4

Oscillations

We explore the behavior of oscillatory systems, including the simple harmonic oscillator, a simple pendulum, and electrical circuits, and introduce the concept of phase space.

4.1 Simple Harmonic Motion

There are many physical systems that undergo regular, repeating motion. Motion that repeats itself at definite intervals, for example, the motion of the earth about the sun, is said to be *periodic*. If an object undergoes periodic motion between two limits over the same path, we call the motion *oscillatory*. Examples of oscillatory motion that are familiar to us from our everyday experience include a plucked guitar string and the pendulum in a grandfather clock. Less obvious examples are microscopic phenomena such as the oscillations of the atoms in crystalline solids.

To illustrate the important concepts associated with oscillatory phenomena, consider a block of mass m connected to the free end of a spring. The block slides on a frictionless, horizontal surface (see Figure 4.1). We specify the position of the block by x and take $x = 0$ to be the equilibrium position of the block, that is, the position when the spring is relaxed. If the block is moved from $x = 0$ and then released, the block oscillates along a horizontal line. If the spring is not compressed or stretched too far from $x = 0$, the force on the block at position x is proportional to x :

$$F = -kx. \quad (4.1)$$

The force constant k is a measure of the stiffness of the spring. The negative sign in (4.1) implies that the force acts to restore the block to its equilibrium position. Newton's equation of motion for the block can be written as

$$\frac{d^2x}{dt^2} = -\omega_0^2 x \quad (4.2)$$

where the angular frequency ω_0 is defined by

$$\omega_0^2 = \frac{k}{m}. \quad (4.3)$$

The dynamical behavior described by (4.2) is called *simple harmonic motion* and can be solved analytically in terms of sine and cosine functions. Because the form of the solution will

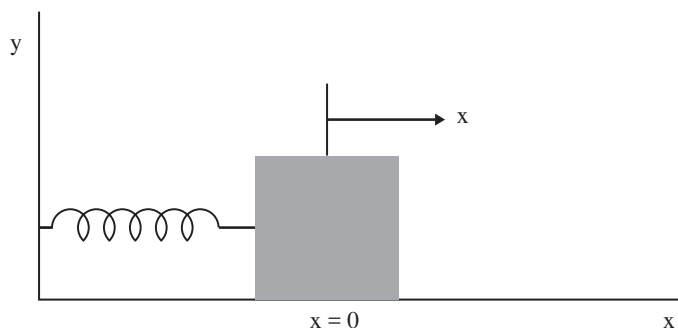


Figure 4.1: A one-dimensional harmonic oscillator. The block slides horizontally on the frictionless surface.

help us introduce some of the terminology needed to discuss oscillatory motion, we include the solution here. One form of the solution is

$$x(t) = A \cos(\omega_0 t + \delta) \quad (4.4)$$

where A and δ are constants and the argument of the cosine is in radians. It is straightforward to check by substitution that (4.4) is a solution of (4.2). The constants A and δ are called the amplitude and the phase, respectively, and are determined by the initial conditions for x and the velocity $v = dx/dt$.

Because the cosine is a periodic function with period 2π , we know that $x(t)$ in (4.4) is also periodic. We define the period T as the smallest time for which the motion repeats itself, that is,

$$x(t + T) = x(t). \quad (4.5)$$

Because $\omega_0 T$ corresponds to one cycle, we have

$$T = \frac{2\pi}{\omega_0} = \frac{2\pi}{\sqrt{k/m}}. \quad (4.6)$$

The frequency ν of the motion is the number of cycles per second and is given by $\nu = 1/T$. Note that T depends on the ratio k/m and not on A and δ . Hence, the period of simple harmonic motion is independent of the amplitude of the motion.

Although the position and velocity of the oscillator are continuously changing, the total energy E remains constant and is given by

$$E = \frac{1}{2}mv^2 + \frac{1}{2}kx^2 = \frac{1}{2}kA^2. \quad (4.7)$$

The two terms in (4.7) are the kinetic and potential energies, respectively.

Problem 4.1. Energy conservation

- (a) Use the Euler ODEsolver to solve the dynamical equations for a simple harmonic oscillator by extending `AbstractSimulation` and implementing the `doStep` method. (See Section 4.2 for an example of such a program for the pendulum.) Have your program plot $\Delta E_n = E_n - E_0$, where E_0 is the initial energy and E_n is the total energy at time $t_n = t_0 + n\Delta t$. (It is necessary

percurso correspondente é $(2/300)$ s. De quanto os átomos caem nesse tempo sob a ação da gravidade?

Caem de $1/2 g t^2 = 1/2 \times 9,8 \times (1/150)^2 \text{ m} \approx 2 \times 10^{-4} \text{ m}$, ou seja, de $\sim 0,2 \text{ mm}$. Embora se trate de um deslocamento pequeno, ele pode ser detectado com relativa facilidade, porque o diâmetro do feixe e do fio utilizado no detector são ~ 10 vezes menores (é um “detector de fio quente”, em que os átomos são ionizados e depois coletados por um eletrodo, medindo-se a corrente elétrica resultante). Devido ao alto vácuo, os resultados obtidos nesta seção sobre trajetórias parabólicas se aplicam com grande precisão. O objetivo das experiências não era estudar a queda livre, mas sim testar resultados da teoria cinética dos gases, sobre a distribuição de velocidades dos átomos. Também têm sido feitas recentemente tentativas de observar a queda livre de elétrons, mas a experiência, realizada dentro de um cilindro evacuado, se torna muito mais difícil neste caso, devido a forças de origem elétrica, oriundas das paredes dos cilindro, que atuam sobre os elétrons e têm efeitos dominantes, de modo que não há ainda evidência clara de que o efeito tenha sido observado.

Em 1999, foi produzido um “chafariz de átomos” em que eles foram lançados no vácuo, descrevendo trajetórias parabólicas. O tempo de atraso entre subida e descida foi empregado na construção do relógio atômico NIST-F1 (Seç. 1.7).

3.7 MOVIMENTO CIRCULAR UNIFORME

Um tipo de movimento plano de grande importância na física é o movimento circular uniforme, em que a trajetória é um círculo e o *módulo* da velocidade instantânea é constante, de modo que a partícula descreve arcos de círculo iguais em tempos iguais. Temos assim um movimento periódico, em que o período corresponde ao tempo levado para descrever uma volta completa, o que define um “relógio”.

De fato, o movimento da extremidade dos ponteiros de um relógio é deste tipo. O movimento da Lua em torno da Terra também pode ser aproximado por um movimento circular uniforme. Outro exemplo são as órbitas de partículas carregadas em aceleradores de tipo circular.

Seja r o raio da trajetória circular. A posição instantânea P da partícula fica definida pelo ângulo θ entre o vetor deslocamento $\mathbf{r} = \mathbf{OP}$ correspondente e o eixo Ox de um sistema cartesiano com origem no centro do círculo (Figura 3.27), onde θ é positivo no sentido anti-horário. O arco s correspondente ao ângulo θ sobre o círculo é dado por

$$s = r\theta$$

(3.7.1)

onde θ é medido em radianos ($2\pi \text{ rad} = 360^\circ$). Vamos introduzir $\hat{\mathbf{r}}$, o vetor unitário na direção de \mathbf{r} , que aponta radialmente para fora, e $\hat{\boldsymbol{\theta}}$, o vetor unitário tangente ao círculo

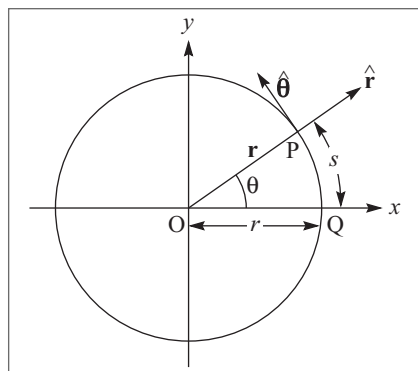


Figura 3.27 Movimento circular.

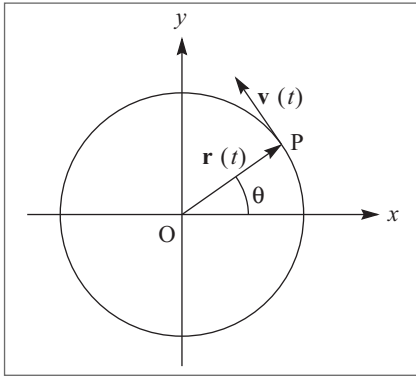


Figura 3.28 Velocidade instantânea.

(portanto perpendicular a $\hat{\mathbf{r}}$) em P, orientado no sentido de θ crescente (anti-horário). Note que ao contrário de $\hat{\mathbf{i}}$ e $\hat{\mathbf{j}}$, que são vetores fixos nas direções dos eixos, as direções de $\hat{\mathbf{r}}$ e $\hat{\boldsymbol{\theta}}$ variam com a posição P ocupada pela partícula ao longo do círculo.

Pela definição de movimento circular uniforme, a lei horária é

$$s = s_0 + v(t - t_0) \quad (3.7.2)$$

onde s_0 é o valor do arco no instante inicial t_0 e v é a “*velocidade linear*” com que o arco s é descrito.

Lembrando a definição (3.4.5) da velocidade instantânea e o fato de que $|\Delta \mathbf{r}|$ se confunde com Δs (corda e arco se confundem) quando $\Delta t \rightarrow 0$, vemos que $v = |\mathbf{v}|$ dá o módulo da velocidade instantânea $\mathbf{v}(t)$, que é tangente ao círculo em P. A velocidade instantânea $\mathbf{v}(t)$ é dada por

$$\mathbf{v} = v\hat{\boldsymbol{\theta}} \quad (3.7.3)$$

(note que isto continua valendo quando o círculo é descrito no sentido horário e $v < 0$). Temos ainda

$$v = ds / dt \quad (3.7.4)$$

como consequência imediata da (3.7.2).

O *período* T do movimento é o tempo para dar uma volta completa, ou seja,

$$T = 2\pi r / |v| \quad (3.7.5)$$

Chama-se *frequência* ν o inverso do período:

$$\nu = 1 / T \quad (3.7.6)$$

A frequência dá portanto o *número de rotações por unidade de tempo*. Assim, um disco LP tem 33 1/3 rpm (rotações por minuto), o que corresponde a $\nu \approx 0,5 \text{ s}^{-1}$ e $T \approx 2 \text{ s}$.

Podemos empregar a (3.7.1) para exprimir a lei horária (3.7.2) em termos do ângulo θ descrito em função do tempo:

$$\theta = \theta_0 + \omega(t - t_0) \quad (3.7.7)$$

onde

$$\omega = v / r \quad (3.7.8)$$

chama-se *velocidade angular*. Temos, analogamente à (3.7.4),

$$\omega = \frac{d\theta}{dt} \quad (3.7.9)$$

e as (3.7.5) e (3.7.8) mostram que

$$|\omega| = \frac{2\pi}{T} = 2\pi\nu \quad (3.7.10)$$

A velocidade angular se mede em rad/s, ou simplesmente em s^{-1} . Assim, por exemplo, a velocidade angular do ponteiro dos segundos de um relógio, para o qual $T = 1 \text{ min}$, é

$$\omega = (2\pi / 60) s^{-1} \approx 0,1 s^{-1} \quad (0,1 \text{ rad} / s)$$

A (3.7.8), escrita sob a forma $v = \omega r$, nos mostra ainda que, num disco em rotação uniforme (por exemplo, um disco LP num toca-discos), a velocidade linear cresce linearmente com a distância ao centro, sendo nula no centro e máxima na periferia.

As (3.7.3) e (3.7.8) dão

$$\mathbf{v} = \omega r \hat{\boldsymbol{\theta}} \quad (3.7.11)$$

Embora o movimento circular uniforme tenha uma velocidade de *módulo* constante, a *direção* da velocidade \mathbf{v} varia de ponto a ponto da trajetória. Logo, conforme foi mencionado no fim da Seç. 3.4, ele é um movimento *acelerado*, ou seja, a aceleração é $\neq 0$. Vamos agora ver como se obtém a aceleração \mathbf{a} .

Uma forma possível de determinar \mathbf{a} é pelo processo geométrico do hodógrafo, descrito na Seç. 3.4. O hodógrafo de um movimento circular uniforme também é um movimento circular uniforme, sobre um círculo de raio v (em linha interrompida na Figura 3.29). Pelo que vimos na Seç. 3.4, a velocidade do movimento sobre o hodógrafo (“velocidade de variação da velocidade”) é a aceleração \mathbf{a} . Como a velocidade angular com que é descrito o hodógrafo é a mesma do movimento circular uniforme, mas o raio do hodógrafo é v , em lugar de r , obtemos da (3.7.8), aplicada ao hodógrafo, o módulo da aceleração:

$$|\mathbf{a}| = \omega v = \omega^2 r = v^2 / r \quad (3.7.12)$$

Por outro lado, o exame da Figura 3.29 mostra que o vetor \mathbf{a} , tangente ao hodógrafo, está dirigido radialmente para dentro no círculo original (trajetória). Logo

$$\mathbf{a} = -|\mathbf{a}| \hat{\mathbf{r}} = -\omega^2 r \hat{\mathbf{r}} = -\frac{v^2}{r} \hat{\mathbf{r}} \quad (3.7.13)$$

Esta é a chamada *aceleração centrípeta* (porque aponta para o centro do círculo).

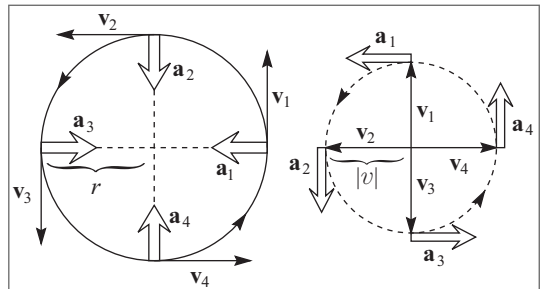


Figura 3.29 Hodógrafo do movimento circular uniforme.

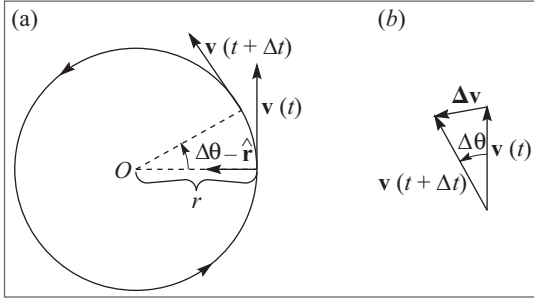


Figura 3.30 Incremento de velocidade.

que $\Delta t \rightarrow 0$, podemos confundir o comprimento de $\Delta \mathbf{v}$ (corda) com o do arco de círculo de raio $|\mathbf{v}|$ que subentende o ângulo $|\Delta \theta|$:

$$|\Delta \mathbf{v}| \approx |\mathbf{v}| |\Delta \theta| \left\{ \frac{|\Delta v|}{\Delta t} \approx |\mathbf{v}| \frac{|\Delta \theta|}{\Delta t} \right. \quad (3.7.14)$$

o que se torna exato no limite em que $\Delta t \rightarrow 0$, levando novamente à (3.7.12) (cf. (3.7.9)).

A título de ilustração, vamos calcular a *aceleração centrípeta da Lua* em sua órbita em redor da Terra, supondo a órbita circular, o que é uma boa aproximação. O raio da órbita é $r \approx 380.000 \text{ km} = 3,8 \times 10^8 \text{ m}$ (distância Terra-Lua). O período de rotação da Lua em redor da Terra (mês lunar) é $T \approx 27,3 \text{ dias} \approx 27,3 \times 8,64 \times 10^4 \text{ s} \approx 2,4 \times 10^6 \text{ s}$. Logo, pelas (3.7.12) e (3.7.10),

$$|\mathbf{a}| = \omega^2 r = \frac{4\pi^2}{T^2} r \approx 2,7 \times 10^{-3} \frac{\text{m}}{\text{s}^2} \quad (3.7.15)$$

o que podemos comparar com a aceleração da gravidade (2.6.1):

$$\frac{|\mathbf{a}|}{g} \approx \frac{2,7 \times 10^{-3}}{9,8} \sim \frac{1}{3.600} \quad (3.7.16)$$

3.8 ACELERAÇÕES TANGENCIAL E NORMAL

Consideremos agora um movimento não uniforme sobre um círculo. Embora a velocidade instantânea continue naturalmente sendo tangente ao círculo, a não uniformidade do movimento circular significa que o *módulo* da velocidade, além da sua direção, também variará com o tempo. É natural então considerar separadamente os dois fatores que contribuem para $\Delta \mathbf{v}$.

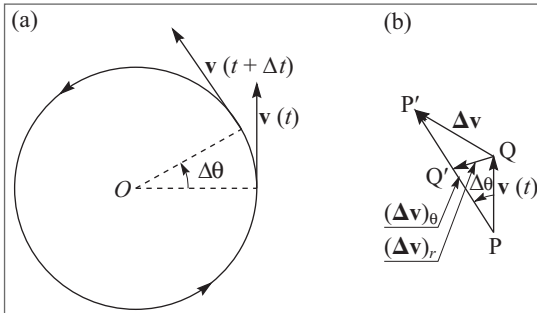


Figura 3.31 Movimento circular acelerado.

Na figura 3.31 (a), consideramos a situação em que $|\mathbf{v}(t + \Delta t)| \neq |\mathbf{v}(t)|$. Na (b), \mathbf{PQ} representa $\mathbf{v}(t)$ e $\mathbf{PP'}$ corresponde a $\mathbf{v}(t + \Delta t)$. O ponto Q' é tomado de tal forma que $|\mathbf{PQ}'| = |\mathbf{PQ}| = |\mathbf{v}(t)|$. Vemos que

$$\Delta \mathbf{v} = \mathbf{QP}' = \mathbf{QQ}' + \mathbf{Q}'\mathbf{P}' \quad (3.8.1)$$

Já calculamos \mathbf{QQ}' na (3.7.14):

$$|\mathbf{QQ}'| = |(\Delta \mathbf{v})_r| \approx |\mathbf{v}| |\Delta \theta| \quad (3.8.2)$$

onde o índice r se refere ao fato de que, para $\Delta t \rightarrow 0$, $(\Delta \mathbf{v})_r$ dará a componente radial de \mathbf{a} (aceleração centrípeta), já calculada na Seq. 3.7:

$$\lim_{\Delta t \rightarrow 0} \left[\frac{(\Delta \mathbf{v})}{\Delta t} \right] = -\omega^2 r \hat{\mathbf{r}} = -r \left(\frac{d\theta}{dt} \right)^2 \hat{\mathbf{r}} \quad (3.8.3)$$

A componente nova que temos de calcular é $\mathbf{Q}'\mathbf{P}'$. No caso ilustrado na Figura 3.31 (a), temos $v > 0$ e a (b) mostra que, no limite em que $\Delta t \rightarrow 0$, a direção e sentido de $\mathbf{Q}'\mathbf{P}'$ tendem a coincidir com os de $\mathbf{v}(t)$, ou seja (cf. (3.7.3)) com os de $\hat{\boldsymbol{\theta}}$, e

$$\mathbf{Q}'\mathbf{P}' = (\Delta \mathbf{v})_{\theta} \approx [v(t + \Delta t) - v(t)] \hat{\boldsymbol{\theta}} = \Delta v \cdot \hat{\boldsymbol{\theta}} \quad (3.8.4)$$

de modo que

$$\lim_{\Delta t \rightarrow 0} \left[\frac{(\Delta \mathbf{v})_{\theta}}{\Delta t} \right] = \lim_{\Delta t \rightarrow 0} \left(\frac{\Delta v}{\Delta t} \right) \hat{\boldsymbol{\theta}} = \frac{dv}{dt} \hat{\boldsymbol{\theta}} \quad (3.8.5)$$

Esta relação permanece válida quer seja $v > 0$ ou $v < 0$, e quer $|v|$ esteja crescendo ou decrescendo, como se vê por uma discussão análoga à da Seq. 2.4. Pelas (3.7.8) e (3.7.9), temos ainda a relação

$$\boxed{\frac{dv}{dt} = r \frac{d\omega}{dt} = r \frac{d^2\theta}{dt^2} = r\alpha} \quad (3.8.6)$$

onde α se chama *aceleração angular*.

Combinando as expressões acima, obtemos finalmente a expressão da *aceleração num movimento circular qualquer*:

$$\boxed{\mathbf{a} = a_r \hat{\mathbf{r}} + a_{\theta} \hat{\boldsymbol{\theta}}} \quad (3.8.7)$$

onde

$$\boxed{a_r = -\omega^2 r = -r \left(\frac{d\theta}{dt} \right)^2 = -\frac{v^2}{r}} \quad (3.8.8)$$

e

$$\boxed{a_{\theta} = \alpha r = r \frac{d^2\theta}{dt^2} = \frac{dv}{dt}} \quad (3.8.9)$$

O termo $a_r \hat{\mathbf{r}}$ na (3.8.7) continua sendo chamado de *aceleração centrípeta*; o outro termo, $a_{\theta} \hat{\boldsymbol{\theta}}$ é a componente da aceleração tangente ao círculo, e se chama por isso de

aceleração *tangencial*; a_r e a_θ na (3.8.7) definem também as *componentes do vetor \mathbf{a} em coordenadas polares* (Seç. 1.6).

Exemplo – Movimento circular uniformemente acelerado: É, por definição, aquele em que a *aceleração angular α é constante*:

$$\alpha = \frac{d^2\theta}{dt^2} = \text{constante} \quad (3.8.10)$$

Sejam

$$\omega_0 = \left(\frac{d\theta}{dt} \right)_{t=t_0}, \quad \theta_0 = \theta(t_0) \quad (3.8.11)$$

os valores iniciais da velocidade angular e do ângulo θ . Pela analogia entre as (3.8.10) – (3.8.11) e as (2.5.1), (2.5.3) e (2.5.7), vemos imediatamente que a lei horária do movimento uniformemente acelerado é (cf. (2.5.8))

$$\theta(t) = \theta_0 + \omega_0(t - t_0) + \frac{1}{2}\alpha(t - t_0)^2 \quad (3.8.12)$$

que a velocidade angular instantânea é (cf. (2.5.4))

$$\omega(t) = \omega_0 + \alpha(t - t_0) \quad (3.8.13)$$

e que (cf. (2.5.9))

$$\omega^2 = \omega_0^2 + 2\alpha(\theta - \theta_0) \quad (3.8.14)$$

A expressão de \mathbf{a} se obtém substituindo as (3.8.10) e (3.8.13) nas (3.8.7) a (3.8.9).

Consideremos agora um movimento plano sobre uma trajetória curva qualquer AB (Figura 3.32), e sejam P e P' as posições nos instantes t e $t + \Delta t$, respectivamente. As

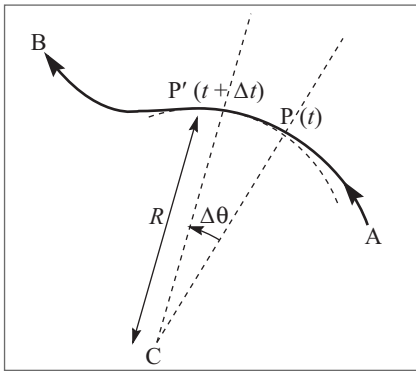


Figura 3.32 Movimento plano geral.

normais à curva em P e P' (perpendiculares às tangentes nesses pontos) se encontram geralmente num ponto C (Figura), que, para Δt muito pequeno, é equidistante de P e P', ou seja, é o centro de um círculo de raio $R = \overline{CP}$ tal que o arco de círculo PP' e arco de curva PP' tendem a ter as mesmas tangentes em P e P' quando $\Delta t \rightarrow 0$ (é o arco de círculo que melhor se aproxima do arco de curva PP' nesse limite). O ponto C chama-se *centro de curvatura* da curva no ponto P e R é o *raio de curvatura* correspondente (para um segmento de reta, R é infinito). O círculo de centro C e raio R é o *círculo de curvatura* da curva no ponto P; em geral, C e R variam de ponto a ponto.

SEMANAS 8 E 9

Chapter 6

The Chaotic Motion of Dynamical Systems

We study simple nonlinear deterministic models that exhibit chaotic behavior. We will find that the use of the computer to do numerical experiments will help us gain insight into the nature of chaos.

6.1 Introduction

Most natural phenomena are intrinsically nonlinear. Weather patterns and the turbulent motion of fluids are everyday examples. Although we have explored some of the properties of nonlinear systems in Chapter 4, it is easier to introduce some of the important concepts in the context of ecology. Our first goal will be to motivate and analyze the one-dimensional difference equation

$$x_{n+1} = 4rx_n(1 - x_n) \tag{6.1}$$

where x_n is the ratio of the population in the n th generation to a reference population. We shall see that the dynamical properties of (6.1) are surprisingly intricate and have important implications for the development of a more general description of nonlinear phenomena. The significance of the behavior of (6.1) is indicated by the following quote from the ecologist Robert May:

"Its study does not involve as much conceptual sophistication as does elementary calculus. Such study would greatly enrich the student's intuition about nonlinear systems. Not only in research but also in the everyday world of politics and economics we would all be better off if more people realized that simple nonlinear systems do not necessarily possess simple dynamical properties."

The study of chaos is of much current interest, but the phenomena is not new and has been of interest, particularly to astronomers and mathematicians, for over one hundred years. Much of the current interest is due to the use of the computer as a tool for making empirical observations. We will use the computer in this spirit.

6.2 A Simple One-Dimensional Map

Imagine an island with an insect population that breeds in the summer and leaves eggs that hatch the following spring. Because the population growth occurs at discrete times, it is appropriate to model the population growth by a difference equation rather than by a differential equation. A simple model of population growth that relates the population in generation $n + 1$ to the population in generation n is given by

$$P_{n+1} = aP_n \quad (6.2)$$

where P_n is the population in generation n and a is a constant. In the following, we will assume that the time interval between generations is unity and will refer to n as the time.

If $a < 1$, the population decreases at each generation, and eventually the population becomes extinct. If $a > 1$, each generation will be a times larger than the previous one. In this case (6.2) leads to geometrical growth and an unbounded population. Although the unbounded nature of geometrical growth is clear, it is remarkable that most of us do not integrate our understanding of geometrical growth into our everyday lives. Can a bank pay 4% interest each year indefinitely? Can the world's human population grow at a constant rate forever?

It is natural to formulate a more realistic model in which the population is bounded by the finite carrying capacity of its environment. A simple model of density-dependent growth is

$$P_{n+1} = P_n(a - bP_n). \quad (6.3)$$

Equation (6.3) is nonlinear due to the presence of the quadratic term in P_n . The linear term represents the natural growth of the population; the quadratic term represents a reduction of this natural growth caused, for example, by overcrowding or by the spread of disease.

It is convenient to rescale the population by letting $P_n = (a/b)x_n$ and rewriting (6.3) as

$$x_{n+1} = ax_n(1 - x_n). \quad (6.4)$$

The replacement of P_n by x_n changes the units used to define the various parameters. To write (6.4) in the standard form (6.1), we define the parameter $r = a/4$ and obtain

$$x_{n+1} = f(x_n) = 4rx_n(1 - x_n). \quad (6.5)$$

The rescaled form (6.5) has the desirable feature that its dynamics are determined by a single control parameter r instead of the two parameters a and b . Note that if $x_n > 1$, x_{n+1} will be negative. To avoid this nonphysical feature, we impose the conditions that x is restricted to the interval $0 \leq x \leq 1$ and $0 < r \leq 1$, respectively. Because the function $f(x)$ defined in (6.5) transforms any point on the one-dimensional interval $[0, 1]$ into another point in the same interval, the function f is called a *one-dimensional map*.

The form of $f(x)$ in (6.5) is known as the *logistic map*. The logistic map is a simple example of a *dynamical system*; that is, the map is a deterministic, mathematical prescription for finding the future state of a system given its present state.

The sequence of values x_0, x_1, x_2, \dots is called the *trajectory*. To check your understanding, suppose that the initial value of x_0 or *seed* is $x_0 = 0.5$ and $r = 0.2$. Do a calculation to show that the trajectory is $x_1 = 0.2, x_2 = 0.128, x_3 = 0.089293, \dots$. The first thirty iterations of (6.5) are shown for two values of r in Figure 6.1.

The class `IterateMapApp` computes the trajectory of the logistic map in (6.5). Note that we have extended the `AbstractCalculation` class, which is appropriate because many of the results of Sections 6.1–6.4 were discovered using a programmable calculator.

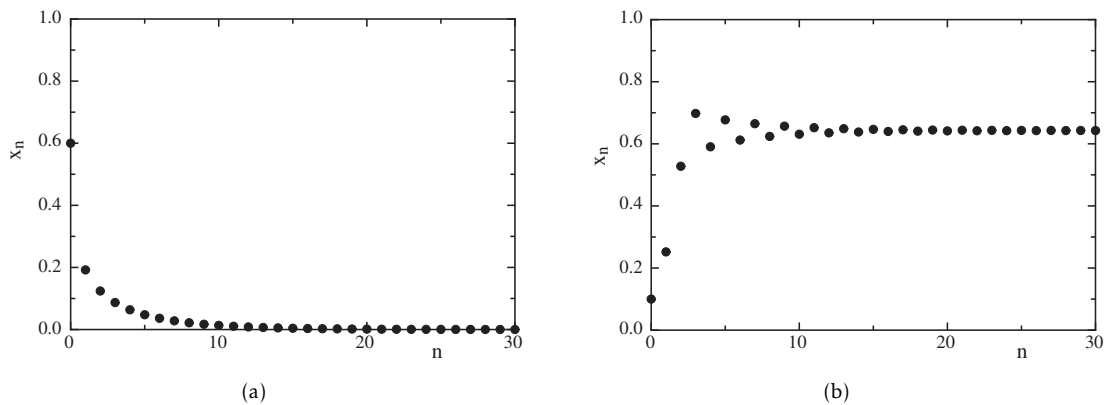


Figure 6.1: (a) The trajectory of x for $r = 0.2$ and $x_0 = 0.6$. The stable fixed point is at $x = 0$. (b) The trajectory for $r = 0.7$ and $x_0 = 0.1$. Note the initial transient behavior.

Listing 6.1: The `IterateMapApp` class iterates the logistic map and plots the resulting trajectory

```
package org.opensourcephysics.sip.ch06;
import org.opensourcephysics.frames.*;
import org.opensourcephysics.controls.*;

public class IterateMapApp extends AbstractCalculation {
    int datasetIndex = 0;
    PlotFrame plotFrame = new PlotFrame("iterations", "x",
                                         "trajectory");

    public IterateMapApp() {
        // keep data between calls to calculate
        plotFrame.setAutoClear(false);
    }

    public void reset() {
        control.setValue("r", 0.2);
        control.setValue("x", 0.6);
        control.setValue("iterations", 50);
        datasetIndex = 0;
    }

    public void calculate() {
        double r = control.getDouble("r");
        double x = control.getDouble("x");
        int iterations = control.getInt("iterations");
        for(int i = 0; i <= iterations; i++) {
            plotFrame.append(datasetIndex, i, x);
            x = map(r, x);
        }
        plotFrame.setMarkerSize(datasetIndex, 1);
        plotFrame.setXYColumnNames(datasetIndex, "iteration",
                                   "calc #" + datasetIndex);
        datasetIndex++;
    }
}
```

```

    }

    double map(double r, double x) {
        return 4*r*x*(1-x); // iterate map
    }

    public static void main(String[] args) {
        CalculationControl.createApp(new IterateMapApp());
    }
}

```

Problem 6.1. The trajectory of the logistic map

- (a) Explore the dynamical behavior of the logistic map in (6.5) with $r = 0.24$ for different values of x_0 . Show numerically that $x = 0$ is a *stable fixed point* for this value of r . That is, the iterated values of x converge to $x = 0$ independently of the value of x_0 . If x represents the population of insects, describe the qualitative behavior of the population.
- (b) Explore the dynamical behavior of (6.5) for $r = 0.26, 0.5, 0.74$, and 0.748 . A fixed point is *unstable* if for almost all values of x_0 near the fixed point, the trajectories diverge from it. Verify that $x = 0$ is an unstable fixed point for $r > 0.25$. Show that for the suggested values of r , the iterated values of x do not change after an initial *transient*; that is, the long time dynamical behavior is *period 1*. In Appendix 6A we show that for $r < 3/4$ and for x_0 in the interval $0 < x_0 < 1$, the trajectories approach the *stable attractor* at $x = 1 - 1/4r$. The set of initial points that iterate to the attractor is called the *basin* of the attractor. For the logistic map, the interval $0 < x < 1$ is the basin of attraction of the attractor $x = 1 - 1/4r$.
- (c) Explore the dynamical properties of (6.5) for $r = 0.752, 0.76, 0.8$, and 0.862 . For $r = 0.752$ and 0.862 , approximately 1000 iterations are necessary to obtain convergent results. Show that if r is greater than 0.75 , x oscillates between two values after an initial transient behavior. That is, instead of a stable cycle of period 1 corresponding to one fixed point, the system has a stable cycle of period 2. The value of r at which the single fixed point x^* splits or *bifurcates* into two values x_1^* and x_2^* is $r = b_1 = 3/4$. The pair of x values, x_1^* and x_2^* , form a *stable attractor* of period 2.
- (d) What are the stable attractors of (6.5) for $r = 0.863$ and 0.88 ? What is the corresponding period? What are the stable attractors and corresponding periods for $r = 0.89, 0.891$, and 0.8922 ? □

Another way to determine the behavior of (6.5) is to plot the values of x as a function of r (see Figure 6.2). The iterated values of x are plotted after the initial transient behavior is discarded. Such a plot is generated by `BifurcateApp`. For each value of r , the first `ntransient` values of x are computed but not plotted. Then the next `nplot` values of x are plotted with the first half with the first half in one color and the second half in another. This process is repeated for a new value of r until the desired range of r values is reached. The magnitude of `nplot` should be at least as large as the longest period that you wish to observe. `BifurcateApp` extends `AbstractSimulation` rather than `AbstractCalculation` because the calculations can be time consuming. For this reason you might want to stop them before they are finished and reset some of the parameters.

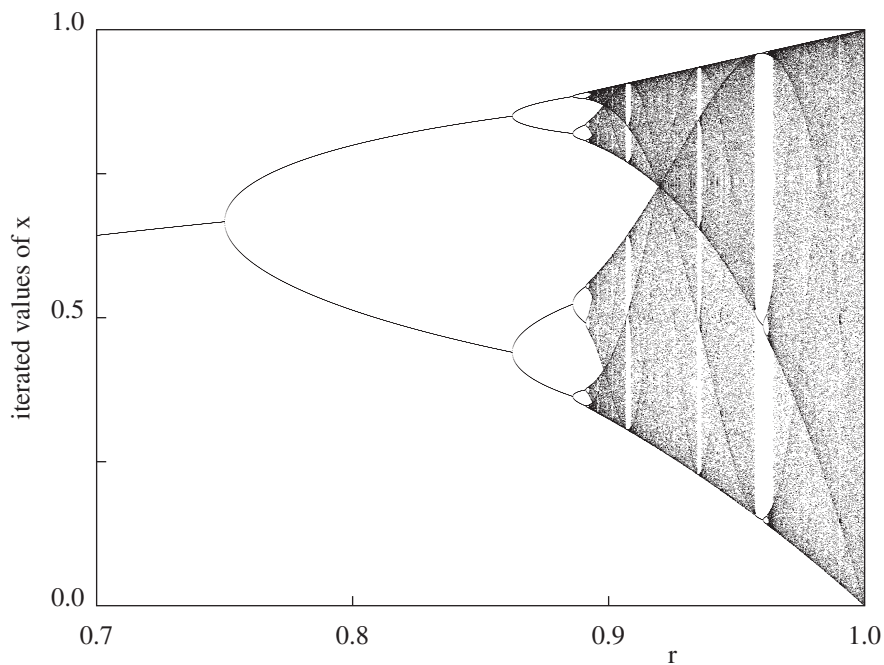


Figure 6.2: Bifurcation diagram of the logistic map. For each value of r , the iterated values of x_n are plotted after the first 1000 iterations are discarded. Note the transition from periodic to chaotic behavior and the narrow windows of periodic behavior within the region of chaos.

Listing 6.2: The `BifurcateApp` program generates a bifurcation plot of the logistic map

```
package org.opensourcephysics.sip.ch06;
import org.opensourcephysics.controls.*;
import org.opensourcephysics.frames.*;

public class BifurcateApp extends AbstractSimulation {
    double r;          // control parameter
    double dr;         // incremental change of r, suggest dr <= 0.01
    int ntransient;    // number of iterations not plotted
    int nplot;         // number of iterations plotted
    PlotFrame plotFrame = new PlotFrame("r", "x",
        "Bifurcation diagram");

    public BifurcateApp() {
        // small size gives better resolution
        plotFrame.setMarkerSize(0, 0);
        plotFrame.setMarkerSize(1, 0);
    }

    public void initialize() {
        plotFrame.clearData();
        r = control.getDouble("initial r");
        dr = control.getDouble("dr");
        ntransient = control.getInt("ntransient");
    }
}
```



```

        nplot = control.getInt("nplot");
    }

    public void doStep() {
        if(r<1.0) {
            double x = 0.5;
            for(int i = 0;i<ntransient;i++) {    // x values not plotted
                x = map(x, r);
            }
            // plot half the points in dataset zero
            for(int i = 0;i<nplot/2;i++) {
                x = map(x, r);
                // shows different x values for given value of r
                plotFrame.append(0, r, x);
            }
            // plot remaining points in dataset one
            for(int i = nplot/2+1;i<nplot;i++) {
                x = map(x, r);
                // dataset one has a different color
                plotFrame.append(1, r, x);
                i++;
            }
            r += dr;
        }
    }

    public void reset() {
        control.setValue("initial r", 0.2);
        control.setValue("dr", 0.005);
        control.setValue("ntransient", 200);
        control.setValue("nplot", 50);
    }

    double map(double x, double r) {
        return 4*r*x*(1-x);
    }

    public static void main(String[] args) {
        SimulationControl.createApp(new BifurcateApp());
    }
}

```

Problem 6.2. Qualitative features of the logistic map

- Use BifurcateApp to identify period 2, period 4, and period 8 behavior as can be seen in Figure 6.2. Choose $n_{\text{transient}} \geq 1000$. It might be necessary to “zoom in” on a portion of the plot. How many period doublings can you find?
- Change the scale so that you can follow the iterations of x from period 4 to period 16 behavior. How does the plot look on this scale in comparison to the original scale?
- Describe the shape of the trajectory near the bifurcations from period 2 to period 4, period 4 to period 8, etc. These bifurcations are frequently called *pitchfork bifurcations*. \square

The bifurcation diagram in Figure 6.2 indicates that the period doubling behavior ends at $r \approx 0.892$. This value of r is known very precisely and is given by $r = r_\infty = 0.892486417967\dots$. At $r = r_\infty$, the sequence of period doublings accumulates to a trajectory of infinite period. In Problem 6.3 we explore the behavior of the trajectories for $r > r_\infty$.

Problem 6.3. Chaotic behavior

- (a) For $r > r_\infty$, two initial conditions that are very close to one another can yield very different trajectories after a few iterations. As an example, choose $r = 0.91$ and consider $x_0 = 0.5$ and 0.5001 . How many iterations are necessary for the iterated values of x to differ by more than ten percent? What happens for $r = 0.88$ for the same choice of seeds?
- (b) The accuracy of floating point numbers retained on a digital computer is finite. To test the effect of the finite accuracy of your computer, choose $r = 0.91$ and $x_0 = 0.5$ and compute the trajectory for 200 iterations. Then modify your program so that after each iteration, the operation $x = x/10$ is followed by $x = 10*x$. This combination of operations truncates the last digit that your computer retains. Compute the trajectory again and compare your results. Do you find the same discrepancy for $r < r_\infty$?
- (c) What are the dynamical properties for $r = 0.958$? Can you find other windows of periodic behavior in the interval $r_\infty < r < 1$? □

6.3 Period Doubling

The results of the numerical experiments that we did in Section 6.2 probably have convinced you that the dynamical properties of a simple, nonlinear deterministic system can be quite complicated.

To gain more insight into how the dynamical behavior depends on r , we introduce a simple graphical method for iterating (6.5). In Figure 6.3 we show a graph of $f(x)$ versus x for $r = 0.7$. A diagonal line corresponding to $y = x$ intersects the curve $y = f(x)$ at the two fixed points $x^* = 0$ and $x^* = 9/14 \approx 0.642857$ [see (6.6b)]. If x_0 is not a fixed point, we can find the trajectory in the following way. Draw a vertical line from $(x = x_0, y = 0)$ to the intersection with the curve $y = f(x)$ at $(x_0, y_0 = f(x_0))$. Next draw a horizontal line from (x_0, y_0) to the intersection with the diagonal line at (y_0, y_0) . On this diagonal line $y = x$, and hence the value of x at this intersection is the first iteration $x_1 = y_0$. The second iteration x_2 can be found in the same way. From the point (x_1, y_0) , draw a vertical line to the intersection with the curve $y = f(x)$. Keep y fixed at $y = y_1 = f(x_1)$, and draw a horizontal line until it intersects the diagonal line; the value of x at this intersection is x_2 . Further iterations can be found by repeating this process.

This graphical method is illustrated in Figure 6.3 for $r = 0.7$ and $x_0 = 0.9$. If we begin with any x_0 (except $x_0 = 0$ and $x_0 = 1$), the iterations will converge to the fixed point $x^* \approx 0.643$. It would be a good idea to repeat the procedure shown in Figure 6.3 by hand. For $r = 0.7$, the fixed point is stable (an attractor of period 1). In contrast, no matter how close x_0 is to the fixed point at $x = 0$, the iterates diverge away from it, and this fixed point is unstable.

How can we explain the qualitative difference between the fixed point at $x = 0$ and at $x^* = 0.642857$ for $r = 0.7$? The local slope of the curve $y = f(x)$ determines the distance moved horizontally each time f is iterated. A slope steeper than 45° leads to a value of x further away from its initial value. Hence, the criterion for the stability of a fixed point is that the magnitude of the slope at the fixed point must be less than 45° . That is, if $|df(x)/dx|_{x=x^*}$ is less than unity, then x^* is stable; conversely, if $|df(x)/dx|_{x=x^*}$ is greater than unity, then x^* is unstable.

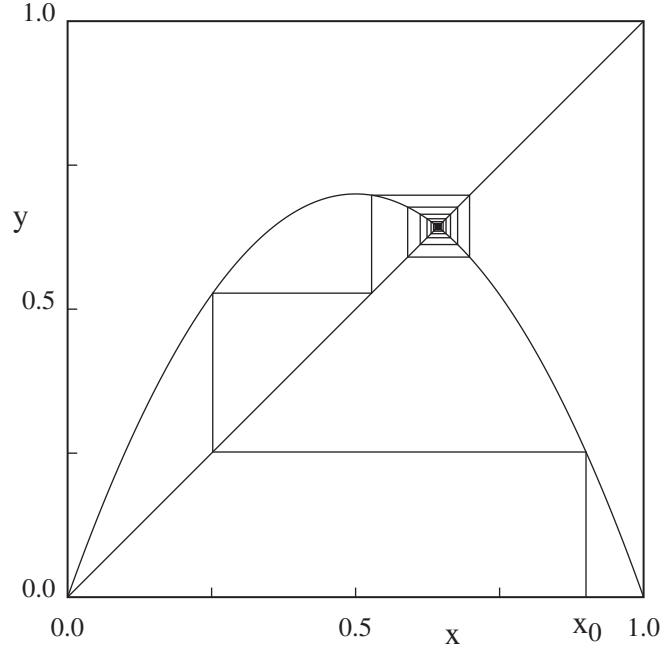


Figure 6.3: Graphical representation of the iteration of the logistic map (6.5) with $r = 0.7$ and $x_0 = 0.9$. Note that the graphical solution converges to the fixed point $x^* \approx 0.643$.

An inspection of $f(x)$ in Figure 6.3 shows that $x = 0$ is unstable because the slope of $f(x)$ at $x = 0$ is greater than unity. In contrast, the magnitude of the slope of $f(x)$ at $x = x^* \approx 0.643$ is less than unity, and this fixed point is stable. In Appendix 6A we show that

$$x^* = 0 \text{ is stable for } 0 < r < 1/4 \quad (6.6a)$$

and

$$x^* = 1 - \frac{1}{4r} \text{ is stable for } 1/4 < r < 3/4. \quad (6.6b)$$

Thus for $0 < r < 3/4$, the behavior after many iterations is known.

What happens if r is greater than $3/4$? We found in Section 6.2 that if r is slightly greater than $3/4$, the fixed point of f becomes unstable and bifurcates to a cycle of period 2. Now x returns to the same value after every second iteration, and the fixed points of $f(f(x))$ are the stable attractors of $f(x)$. In the following, we write $f^{(2)}(x) = f(f(x))$ and $f^{(n)}(x)$ for the n th iterate of $f(x)$. (Do not confuse $f^{(n)}(x)$ with the n th derivative of $f(x)$.) For example, the second iterate $f^{(2)}(x)$ is given by the fourth-order polynomial:

$$\begin{aligned} f^{(2)}(x) &= 4r[4rx(1-x)] - 4r[4rx(1-x)]^2 \\ &= 4r[4rx(1-x)][1 - 4rx(1-x)] \\ &= 16r^2x[-4rx^3 + 8rx^2 - (1+4r)x + 1]. \end{aligned} \quad (6.7)$$

What happens if we increase r still further? Eventually the magnitude of the slope of the fixed points of $f^{(2)}(x)$ exceeds unity, and the fixed points of $f^{(2)}(x)$ become unstable. Now

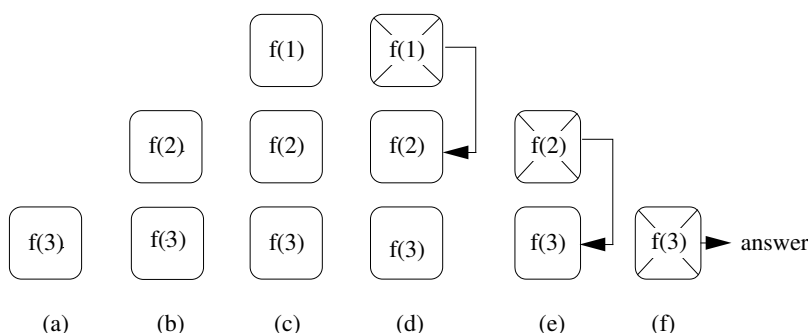


Figure 6.4: Example of the calculation of $f(0.4, 0.8, 3)$ using the recursive function defined in `GraphicalSolutionApp`. The number in each box is the value of the variable `iterate`. The computer executes code from left to right, and each box represents a copy of the function in the computer's memory. The input values $x = 0.4$ and $r = 0.8$, which are the same in each copy, are not shown. The arrows indicate when a copy is finished and its value is returned to one of the other copies. Notice that the first copy of the function $f(3)$ is the last one to finish. The value of $f(x, r, 3) = 0.7842$.

the cycle of f is period 4, and the fixed points of the fourth iterate $f^{(4)}(x) = f^{(2)}(f^{(2)}(x)) = f(f(f(f(x))))$ are stable. These fixed points also eventually become unstable, and we are led to the phenomena of *period doubling* that we observed in Problem 6.2.

`GraphicalSolutionApp` implements the graphical analysis of the iterations of $f(x)$. The n th-order iterates are defined in $f(x, r, \text{iterate})$, a *recursive* method. (The parameter `iterate` is 1, 2, and 4 for the functions $f(x)$, $f^{(2)}(x)$, and $f^{(4)}(x)$, respectively.) Recursion is an idea that is simple once you understand it, but it can be difficult to grasp initially. Although the method calls itself, the rules for method calls remain the same. Imagine that a recursive method is called. The computer then starts to execute the code in the method, but comes to another call of the same method as itself. At this point the computer stops executing the code of the original method, and makes an exact copy of the method with possibly different input parameters, and starts executing the code in the copy. There are now two possibilities. One is that the computer comes to the end of the copy without another recursive call. In that case the computer deletes the copy of the method and continues executing the code in the original method. The other possibility is that a recursive call is made in the copy, and a third copy is made of the method, and the code in the third copy is now executed. This process continues until the code in all the copies is executed. Every recursive method must have a possibility of reaching the end of the method; otherwise, the program will eventually crash.

To understand the method $f(x, r, \text{iterate})$, suppose we want to compute $f(0.4, 0.8, 3)$. First we write $f(0.4, 0.8, 3)$ as in Figure 6.4a. Follow the statements within the method until another call to $f(0.4, 0.8, \text{iterate})$ occurs. In this case, the call is to $f(0.4, 0.8, \text{iterate}-1)$ which equals $f(0.4, 0.8, 2)$. Write $f(0.4, 0.8, 2)$ above $f(0.4, 0.8, 3)$ (see Figure 6.4b). When you come to the end of the definition of the method, write down the value of f that is actually returned, and remove the method from the stack by crossing it out (see Figure 6.4d). This returned value for f equals y if $\text{iterate} > 1$, or it is the output of the method for $\text{iterate} = 1$. Continue deleting copies of f as they are finished, until there are no copies left on the paper. The final value of f is the value returned by the computer. Write a short program that defines

$f(x, r, \text{iterate})$ and prints the value of $f(0.4, 0.8, 3)$. Is the answer the same as your hand calculation?

Listing 6.3: GraphicalSolutionApp displays the graphical solution of the logistic map trajectory

```

package org.opensourcephysics.sip.ch06;
import org.opensourcephysics.controls.*;
import org.opensourcephysics.frames.PlotFrame;

public class GraphicalSolutionApp extends AbstractSimulation {
    PlotFrame plotFrame = new PlotFrame("iterations", "x",
                                         "graphical solution");
    double r;    // control parameter
    int iterate; // iterate of f(x)
    double x, y;
    double x0, y0;

    public GraphicalSolutionApp() {
        plotFrame.setPreferredMinMax(0, 1, 0, 1);
        plotFrame.setConnected(true);
        plotFrame.setXPointsLinked(true);
        // second argument indicates no marker
        plotFrame.setMarkerShape(2, 0);
    }

    public void reset() {
        control.setValue("r", 0.89);
        control.setValue("x", 0.2);
        plotFrame.setMarkerShape(0, 0);
        control.setAdjustableValue("iterate", 1);
    }

    public void initialize() {
        r = control.getDouble("r");
        x = control.getDouble("x");
        iterate = control.getInt("iterate");
        x0 = x;
        y0 = 0;
        clear();
    }

    public void startRunning() {
        if (iterate != control.getInt("iterate")) {
            iterate = control.getInt("iterate");
            clear();
        }
        r = control.getDouble("r");
    }

    public void doStep() {
        y = f(x, r, iterate);
        plotFrame.append(1, x0, y0);
        plotFrame.append(1, x0, y);
    }
}

```

```

        plotFrame.append(1, y, y);
        x = x0 = y0 = y;
        control.setValue("x", x);
    }

    void drawFunction() {
        int nplot = 200; // # of points at which function computed
        double delta = 1.0/nplot;
        double x = 0;
        double y = 0;
        for(int i = 0; i<=nplot; i++) {
            y = f(x, r, iterate);
            plotFrame.append(0, x, y);
            x += delta;
        }
    }

    void drawLine() { // draws line y = x
        for(double x = 0; x<1; x += 0.001) {
            plotFrame.append(2, x, x);
        }
    }

    public double f(double x, double r, int iterate) {
        if(iterate>1) {
            double y = f(x, r, iterate-1);
            return 4*r*y*(1-y);
        } else {
            return 4*r*x*(1-x);
        }
    }

    public void clear() {
        plotFrame.clearData();
        drawFunction();
        drawLine();
        plotFrame.repaint();
    }

    public static void main(String[] args) {
        SimulationControl control = SimulationControl.createApp(
            new GraphicalSolutionApp());
        control.addButton("clear", "Clear", "Clears the trajectory.");
    }
}

```

Problem 6.4. Qualitative properties of the fixed points

- (a) Use `GraphicalSolutionApp` to show graphically that there is a single stable fixed point of $f(x)$ for $r < 3/4$. It would be instructive to modify the program so that the value of the slope $df/dx|_{x=x_n}$ is shown as you step each iteration. At what value of r does the absolute value of this slope exceed unity? Let b_1 denote the value of r at which the fixed point of $f(x)$ bifurcates and becomes unstable. Verify that $b_1 = 0.75$.

- (b) Describe the trajectory of $f(x)$ for $r = 0.785$. Is the fixed point given by $x = 1 - 1/4r$ stable or unstable? What is the nature of the trajectory if $x_0 = 1 - 1/4r$? What is the period of $f(x)$ for all other choices of x_0 ? What are the values of the two-point attractor?
- (c) The function $f(x)$ is symmetrical about $x = 1/2$ where $f(x)$ is a maximum. What are the qualitative features of the second iterate $f^{(2)}(x)$ for $r = 0.785$? Is $f^{(2)}(x)$ symmetrical about $x = 1/2$? For what value of x does $f^{(2)}(x)$ have a minimum? Iterate $x_{n+1} = f^{(2)}(x_n)$ for $r = 0.785$ and find its two fixed points x_1^* and x_2^* . (Try $x_0 = 0.1$ and $x_0 = 0.3$.) Are the fixed points of $f^{(2)}(x)$ stable or unstable for this value of r ? How do these values of x_1^* and x_2^* compare with the values of the two-point attractor of $f(x)$? Verify that the slopes of $f^{(2)}(x)$ at x_1^* and x_2^* are equal.
- (d) Verify the following properties of the fixed points of $f^{(2)}(x)$. As r is increased, the fixed points of $f^{(2)}(x)$ move apart, and the slope of $f^{(2)}(x)$ at its fixed points decreases. What is the value of $r = s_2$ at which one of the two fixed points of $f^{(2)}$ equals $1/2$? What is the value of the other fixed point? What is the slope of $f^{(2)}(x)$ at $x = 1/2$? What is the slope at the other fixed point? As r is further increased, the slopes at the fixed points become negative. Finally at $r = b_2 \approx 0.8623$, the slopes at the two fixed points of $f^{(2)}(x)$ equal -1 , and the two fixed points of $f^{(2)}$ become unstable. (The exact value of b_2 is $b_2 = (1 + \sqrt{6})/4$.)
- (e) Show that for r slightly greater than b_2 , for example $r = 0.87$, there are four stable fixed points of $f^{(4)}(x)$. What is the value of $r = s_3$ when one of the fixed points equals $1/2$? What are the values of the three other fixed points at $r = s_3$?
- (f) Determine the value of $r = b_3$ at which the four fixed points of $f^{(4)}$ become unstable.
- (g) Choose $r = s_3$ and determine the number of iterations that are necessary for the trajectory to converge to period 4 behavior. How does this number of iterations change when neighboring values of r are considered? Choose several values of x_0 so that your results do not depend on the initial conditions. \square

Problem 6.5. Periodic windows in the chaotic regime

- (a) If you look closely at the bifurcation diagram in Figure 6.2, you will see that the range of chaotic behavior for $r > r_\infty$ is interrupted by intervals of periodic behavior. Magnify your bifurcation diagram so that you can look at the interval $0.957107 \leq r \leq 0.960375$, where a periodic trajectory of period 3 occurs. (Period 3 behavior starts at $r = (1 + \sqrt{8})/4$.) What happens to the trajectory for slightly larger r , for example, $r = 0.9604$?
- (b) Plot $f^{(3)}(x)$ versus x at $r = 0.96$, a value of r in the period 3 window. Draw the line $y = x$ and determine the intersections with $f^{(3)}(x)$. The stable fixed points satisfy the condition $x^* = f^{(3)}(x^*)$. Because $f^{(3)}(x)$ is an eighth-order polynomial, there are eight solutions (including $x = 0$). Find the intersections of $f^{(3)}(x)$ with $y = x$ and identify the three stable fixed points. What are the slopes of $f^{(3)}(x)$ at these points? Then decrease r to $r = 0.957107$, the (approximate) value of r below which the system is chaotic. Draw the line $y = x$ and determine the number of intersections with $f^{(3)}(x)$. Note that at this value of r , the curve $y = f^{(3)}(x)$ is tangent to the diagonal line at the three stable fixed points. For this reason, this type of transition is called a *tangent bifurcation*. Note that there is also an unstable point at $x \approx 0.76$.
- (c) Plot $x_{n+1} = f^{(3)}(x_n)$ versus n for $r = 0.9571$, a value of r just below the onset of period 3 behavior. How would you describe the behavior of the trajectory? This type of chaotic motion

k	b_k
1	0.750 000
2	0.862 372
3	0.886 023
4	0.891 102
5	0.892 190
6	0.892 423
7	0.892 473
8	0.892 484

Table 6.1: Values of the control parameter $r = b_k$ for the onset of the k th bifurcation. Six decimal places are shown.

is an example of *intermittency*; that is, nearly periodic behavior interrupted by occasional irregular bursts.

- (d) To understand the mechanism for the intermittent behavior, we need to “zoom in” on the values of x near the stable fixed points that you found in part (c). To do so change the arguments of the `setPreferredMinMax` method. You will see a narrow channel between the diagonal line $y = x$ and the plot of $f^{(3)}(x)$ near each fixed point. The trajectory can require many iterations to squeeze through the channel, and we see apparent period 3 behavior during this time. Eventually, the trajectory escapes from the channel and bounces around until it is again enters a channel at some unpredictable later time. \square

6.4 Universal Properties and Self-Similarity

In Sections 6.2 and 6.3 we found that the trajectory of the logistic map has remarkable properties as a function of the control parameter r . In particular, we found a sequence of period doublings accumulating in a chaotic trajectory of infinite period at $r = r_\infty$. For most values of $r > r_\infty$, the trajectory is very sensitive to the initial conditions. We also found “windows” of period 3, 6, 12, ... embedded in the range of chaotic behavior. How typical is this type of behavior? In the following, we will find further numerical evidence that the general behavior of the logistic map is independent of the details of the form (6.5) of $f(x)$.

You might have noticed that the range of r between successive bifurcations becomes smaller as the period increases (see Table 6.1). For example, $b_2 - b_1 = 0.112398$, $b_3 - b_2 = 0.023624$ and $b_4 - b_3 = 0.00508$. A good guess is that the decrease in $b_k - b_{k-1}$ is geometric; that is, the ratio $(b_k - b_{k-1})/(b_{k+1} - b_k)$ is a constant. You can check that this ratio is not exactly constant, but converges to a constant with increasing k . This behavior suggests that the sequence of values of b_k has a limit and follows a geometrical progression:

$$b_k \approx r_\infty - C\delta^{-k}, \quad (6.8)$$

where δ is known as the *Feigenbaum number* and C is a constant. From (6.8) it is easy to show that δ is given by the ratio

$$\delta = \lim_{k \rightarrow \infty} \frac{b_k - b_{k-1}}{b_{k+1} - b_k}. \quad (6.9)$$

2. Iterate the map with $r = r_0$ until x_n is within ϵ of $x(i)$. Then use (6.31) to determine r .
3. Turn off the control by setting $r = r_0$. □

Problem 6.14. Controlling chaos

- (a) Write a program that allows the user to turn the control on and off. The trajectory can be seen by plotting x_n versus n . The program should incorporate as input the desired unstable periodic trajectory $x(i)$, the period p , the value of r_0 , and the parameter ϵ .
- (b) Test your program with $r_0 = 0.95$ and the periods $p = 1, 5$, and 13 . Use $\epsilon = 0.02$.
- (c) Modify your program so that the values of r as well as the values of x_n are shown. How does r change if we vary ϵ ? Try $\epsilon = 0.05, 0.01$, and 0.005 .
- (d) Add a method to compute n_ϵ , the number of iterations necessary for the trajectory x_n to be within ϵ of $x(1)$ when the control is on. Find $\langle n_\epsilon \rangle$, the average value of n_ϵ , by starting with 100 random values of x_0 . Compute $\langle n_\epsilon \rangle$ as a function of ϵ for $\delta = 0.05, 0.005, 0.0005$, and 0.00005 . What is the functional dependence of $\langle n_\epsilon \rangle$ on ϵ ? □

6.7 Higher-Dimensional Models

So far we have discussed the logistic map as a mathematical model that has some remarkable properties and produces some interesting computer graphics. In this section we discuss some two- and three-dimensional systems that also might seem to have little to do with realistic physical systems. However, as we will see in Sections 6.8 and 6.9, similar behavior is found in realistic physical systems under the appropriate conditions.

We begin with a two-dimensional map and consider the sequence of points (x_n, y_n) generated by

$$x_{n+1} = y_n + 1 - ax_n^2 \tag{6.32a}$$

$$y_{n+1} = bx_n. \tag{6.32b}$$

The map (6.32) was proposed by Hénon who was motivated by the relevance of this dynamical system to the behavior of asteroids and satellites.

Problem 6.15. The Hénon map

- (a) Write a program to iterate (6.32) for $a = 1.4$ and $b = 0.3$ and plot 10^4 iterations starting from $x_0 = 0, y_0 = 0$. Make sure you compute the new value of y using the old value of x and not the new value of x . Do not plot the initial transient. Look at the trajectory in the region defined by $|x| \leq 1.5$ and $|y| \leq 0.45$. Make a similar plot beginning from the second initial condition, $x_0 = 0.63135448, y_0 = 0.18940634$. Compare the shape of the two plots. Is the shape of the two curves independent of the initial conditions?
- (b) Increase the scale of your plot so that all points in the region $0.50 \leq x \leq 0.75$ and $0.15 \leq y \leq 0.21$ are shown. Begin from the second initial condition and increase the number of computed points to 10^5 . Then make another plot showing all points in the region $0.62 \leq x \leq 0.64$ and $0.185 \leq y \leq 0.191$. If time permits, make an additional enlargement and plot all points within the box defined by $0.6305 \leq x \leq 0.6325$ and $0.1889 \leq y \leq 0.1895$. You will

have to increase the number of computed points to order 10^6 . What is the structure of the curves within each box? Does the attractor appear to have a similar structure on smaller and smaller length scales? The region of points from which the points cannot escape is the basin of the Hénon attractor. The attractor is the set of points to which all points in the basin are attracted. That is, two trajectories that begin from different conditions will eventually lie on the attractor.

- (c) Determine if the system is chaotic; that is, sensitive to initial conditions. Start two points very close to each other and watch their trajectories for a fixed time. Choose different colors for the two trajectories.
- (d)* It is straightforward in principle to extend the method for computing the Lyapunov exponent that we used for a one-dimensional map to higher-dimensional maps. The idea is to linearize the difference (or differential) equations and replace dx_n by the corresponding vector quantity $d\mathbf{r}_n$. This generalization yields the Lyapunov exponent corresponding to the divergence along the fastest growing direction. If a system has f degrees of freedom, it has a set of f Lyapunov exponents. A method for computing all f exponents is discussed in Project 6.24. \square

One of the earliest indications of chaotic behavior was in an atmospheric model developed by Lorenz. His goal was to describe the motion of a fluid layer that is heated from below. The result is convective rolls, where the warm fluid at the bottom rises, cools off at the top, and then falls down later. Lorenz simplified the description by restricting the motion to two spatial dimensions. This situation has been realized experimentally and is known as a Rayleigh–Benard cell. The equations that Lorenz obtained are

$$\frac{dx}{dt} = -\sigma x + \sigma y \quad (6.33a)$$

$$\frac{dy}{dt} = -xz + rx - y \quad (6.33b)$$

$$\frac{dz}{dt} = xy - bz \quad (6.33c)$$

where x is a measure of the fluid flow velocity circulating around the cell, y is a measure of the temperature difference between the rising and falling fluid regions, and z is a measure of the difference in the temperature profile between the bottom and the top from the normal equilibrium temperature profile. The dimensionless parameters σ , r , and b are determined by various fluid properties, the size of the Rayleigh–Benard cell, and the temperature difference in the cell. Note that the variables x , y , and z have nothing to do with the spatial coordinates, but are measures of the state of the system. Although it is not expected that you will understand the relation of the Lorenz equations to convection, we have included these equations here to reinforce the idea that simple sets of equations can exhibit chaotic behavior.

LorenzApp displays the solution to (6.33) using the Open Source Physics 3D drawing framework and is available in the ch06 package. To make three-dimensional plots, we use the Display3DFrame class; the only argument of its constructor is the title for the plot. The following code fragment sets up the plot.

```
Display3DFrame frame = new Display3DFrame("Lorenz attractor");
Lorenz lorenz = new Lorenz();
frame.setPreferredMinMax(-15.0, 15.0, -15.0, 15.0, 0.0, 50.0);
frame.setDecorationType(VisualizationHints.DECORATION_AXES);
frame.addElement(lorenz); // lorenz is a 3D element
```

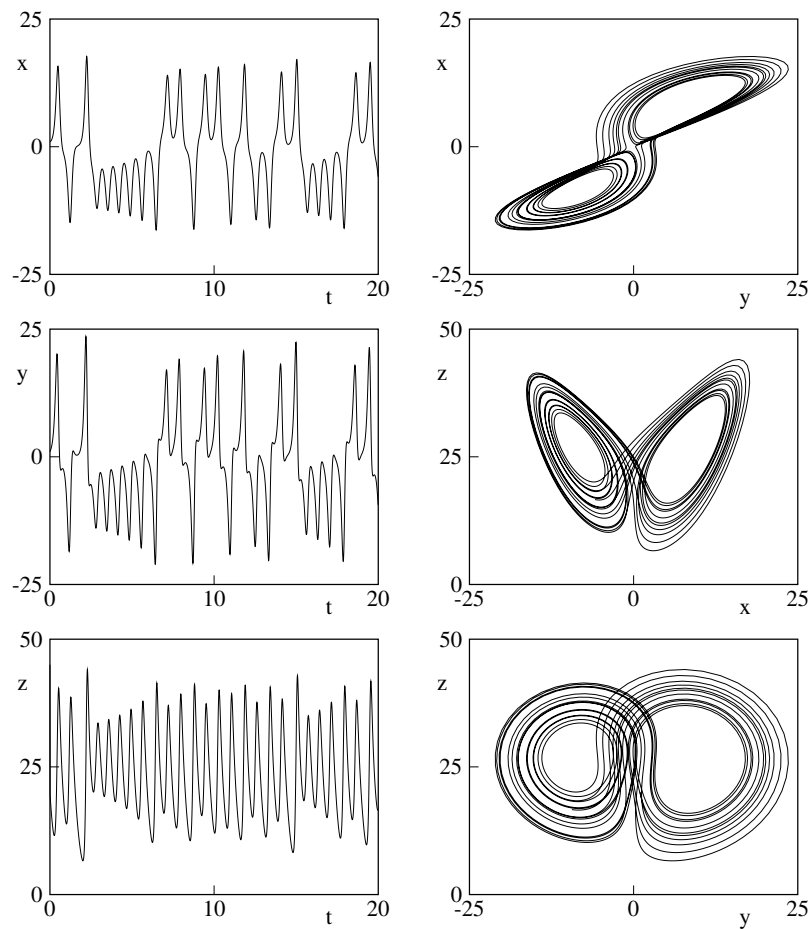


Figure 6.10: A trajectory of the Lorenz model with $\sigma = 10$, $b = 8/3$, and $r = 28$ and the initial condition $x_0 = 1, y_0 = 1, z_0 = 20$. A time interval of $t = 20$ is shown with points plotted at intervals of 0.01. The fourth-order Runge–Kutta algorithm was used with $\Delta t = 0.0025$.

Housekeeping methods such as `reset` and `initialize` are similar to methods in other simulations and are not shown.

The class `Lorenz` draws the attractor in the three-dimensional (x, y, z) space defined by (6.33). The state of the system is shown as a red ball in this 3D space, and the state's trajectory is shown as a trail. An easy way to show the time evolution is to extend the 3D Group class and create the ball and the trail inside the group. When points are added to the group, the trail is extended and the position of the ball is set. The `Lorenz` class imports `org.opensourcephysics.display3d.simple3d.*`. The ball and trail are then instantiated and added to the group as follows:

```
public class Lorenz extends Group implements ODE {
    ElementEllipsoid ball = new ElementEllipsoid();
    ElementTrail trail = new ElementTrail();
    addElement(trail);      // adds trace to Lorenz group
    addElement(ball);      // adds ball to Lorenz group
```

```
    ...
}
```

The properties of the `ball` and `trail` objects are set by

```
ball.setSizeXYZ(1, 1, 1); // sets size of ball in world coordinates
ball.getStyle().setFillColor(java.awt.Color.RED);
```

To plot each part of the trajectory through state space, we use the method `trail.addPoint(x, y, z)` to add to the trail and `ball.setXYZ(x, y, z)` to show the current state. The user can project onto two dimensions using the frame's menu or rotate the three-dimensional plot using the mouse because these capabilities are built into the frame. The `getRate` and `getState` methods model (6.33) by implementing the ODE interface.

Problem 6.16. The Lorenz model

- Use a Runge–Kutta algorithm such as RK4 or RK45 (see Appendix 3A) to obtain a numerical solution of the Lorenz equations (6.33). Generate three-dimensional plots using `Display3DFrame`. Explore the basin of the attractor with $\sigma = 10$, $b = 8/3$, and $r = 28$.
- Determine qualitatively the sensitivity to initial conditions. Start two points very close to each other and watch their trajectories for approximately 10^4 time steps.
- Let z_m denote the value of z where z is a relative maximum for the m th time. You can determine the value of z_m by finding the average of the two values of z when the right-hand side of (6.33) changes sign. Plot z_{m+1} versus z_m and describe what you find. This procedure is one way that a continuous system can be mapped onto a discrete map. What is the slope of the z_{m+1} versus z_m curve? Is its magnitude always greater than unity? If so, then this behavior is an indication of chaos. Why? \square

The application of the Lorenz equations to weather prediction has led to a popular metaphor known as the *butterfly effect*. This metaphor is made even more meaningful by inspection of Figure 6.10. The “butterfly effect” is often ascribed to Lorenz (see Hilborn). In a 1963 paper he remarked that:

“One meteorologist remarked that if the theory were correct, one flap of a seagull’s wings would be enough to alter the course of the weather forever.”

By 1972, the seagull had evolved into the more poetic butterfly and the title of his talk was “Predictability: Does the flap of a butterfly’s wings in Brazil set off a tornado in Texas?”

6.8 Forced Damped Pendulum

We now consider the dynamics of nonlinear systems described by classical mechanics. The general problem in classical mechanics is the determination of the positions and velocities of a system of particles subjected to certain forces. For example, we considered in Chapter 5 the celestial two-body problem and were able to predict the motion at any time. We will find that we cannot make long-time predictions for the trajectories of nonlinear classical systems when these systems exhibit chaos.

A familiar example of a nonlinear mechanical system is the simple pendulum (see Chapter 3). To make its dynamics more interesting, we assume that there is a linear damping term

SEMANAS 10 E 11

Chapter 7

Random Processes

Random processes are introduced in the context of several simple physical systems, including random walks on a lattice, polymers, and diffusion-controlled chemical reactions. The generation of random number sequences is also discussed.

7.1 Order to Disorder

In Chapter 6 we saw several examples of how, under certain conditions, the behavior of a non-linear deterministic system can appear to be random. In this chapter we will see some examples of how chance can generate statistically predictable outcomes. For example, we know that if we bet often on the outcome of a game for which the probability of winning is less than 50%, we will lose money eventually.

We first discuss an example that illustrates the tendency of systems of many particles to evolve to a well-defined state. Imagine a closed box that is divided into two parts of equal volume (see Figure 7.1). The left half contains a gas of N identical particles and the right half is initially empty. We then make a small hole in the partition between the two halves. What happens? We know that after some time, the average number of particles in each half of the box will become $N/2$, and we say that the system has reached equilibrium.

How can we simulate this process? One way is to give each particle an initial velocity and position and adopt a deterministic model of the motion of the particles. For example, we could assume that each particle moves in a straight line until it hits a wall of the box or another particle and undergoes an elastic collision. We will consider similar deterministic models in Chapter 8. Instead, we first simulate a probabilistic model based on a *random process*.

The basic assumptions of this model are that the motion of the particles is random and the particles do not interact with one another. Hence, the probability per unit time that a particle goes through the hole in the partition is the same for all N particles regardless of the number of particles in either half. We also assume that the size of the hole is such that only one particle can pass through at a time. We first model the motion of a particle passing through the hole by choosing one of the N particles at random and moving it to the other side. For visualization purposes, we will use arrays to specify the position of each particle. We then randomly generate an integer i between 0 and $N - 1$ and change the arrays appropriately. A more efficient Monte Carlo algorithm is discussed in Problem 7.2b. The tool we need to simulate this random process is a random number generator.

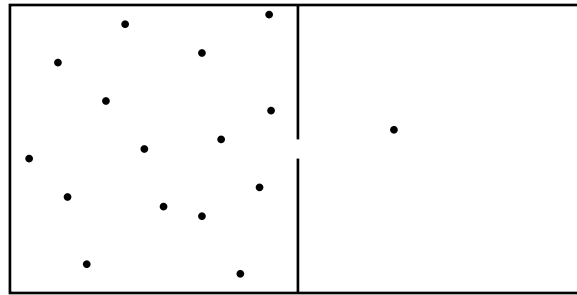


Figure 7.1: A box is divided into two equal halves by a partition. After a small hole is opened in the partition, one particle can pass through the hole per unit time.

It is counterintuitive that we can use a deterministic computer to generate sequences of random numbers. In Section 7.9 we discuss some of the methods for computing a set of numbers that appear statistically random but are in fact generated by a deterministic algorithm. These algorithms are sometimes called *pseudorandom number generators* to distinguish their output from intrinsically random physical processes, such as the time between clicks in a Geiger counter near a radioactive sample.

For the present we will be content to use the random number generator supplied with Java, although the random number generators included with various programming languages vary in quality. The method `Math.random()` produces a random number r that is uniformly distributed in the interval $0 \leq r < 1$. To generate a random integer i between 0 and $N - 1$, we write:

```
int i = (int)(N*Math.random());
```

The effect of the `(int)` cast is to eliminate the decimal digits from a floating point number. For example, `(int)(5.7) = 5`.

The algorithm for simulating the evolution of the model can be summarized by the following steps:

1. Use a random number generator to choose a particle at random.
2. Move this particle to the other side of the box.
3. Give the particle a random position on the new side of the box. This step is for visualization purposes only.
4. Increase the “time” by unity.

Note that this definition of time is arbitrary. Class `Box` implements this algorithm and class `BoxApp` plots the evolution of the number of particles on the left half of the box.

Listing 7.1: Class `Box` for the simulation of the approach to equilibrium.

```
package org.opensourcephysics.sip.ch07;
import java.awt.*;
import org.opensourcephysics.display.*;

public class Box implements Drawable {
    public double x[], y[];
```

```

public int N, nleft, time;

public void initialize() {
    // location of particles (for visualization purposes only)
    x = new double[N];
    y = new double[N];
    nleft = N;           // start with all particles on the left
    time = 0;
    for(int i = 0; i < N; i++) {
        x[i] = 0.5*Math.random(); // needed only for visualization
        y[i] = Math.random();
    }
}

public void step() {
    int i = (int) (Math.random()*N);
    if(x[i] < 0.5) {
        nleft--; // move to right
        x[i] = 0.5*(1+Math.random());
        y[i] = Math.random();
    } else {
        nleft++; // move to left
        x[i] = 0.5*Math.random();
        y[i] = Math.random();
    }
    time++;
}

public void draw(DrawingPanel panel, Graphics g) {
    if(x==null) {
        return;
    }
    int size = 2;
    // position of partition in middle of box
    int xMiddle = panel.xToPix(0.5);
    g.setColor(Color.black);
    g.drawLine(xMiddle, panel.yToPix(0), xMiddle, panel.yToPix(0.45));
    g.drawLine(xMiddle, panel.yToPix(0.55), xMiddle, panel.yToPix(1.0));
    g.setColor(Color.red);
    for(int i = 0; i < N; i++) {
        int xpix = panel.xToPix(x[i]);
        int ypix = panel.yToPix(y[i]);
        g.fillOval(xpix, ypix, size, size);
    }
}
}

```

Listing 7.2: Target class for plotting the approach to equilibrium.

```

package org.opensourcephysics.sip.ch07;
import org.opensourcephysics.controls.*;
import org.opensourcephysics.frames.*;

public class BoxApp extends AbstractSimulation {

```



```

Box box = new Box();
PlotFrame plotFrame = new PlotFrame("time", "number on left",
                                     "Box data");
DisplayFrame displayFrame = new DisplayFrame("Partitioned box");

public void initialize() {
    displayFrame.clearDrawables();
    displayFrame.addDrawable(box);
    box.N = control.getInt("Number of particles");
    box.initialize();
    plotFrame.clearData();
    displayFrame.setPreferredMinMax(0, 1, 0, 1);
}

public void doStep() {
    box.step();
    plotFrame.append(0, box.time, box.nleft);
}

public void reset() {
    // clicking resets erase positions of particles
    control.setValue("Number of particles", 64);
    plotFrame.clearData();
    enableStepsPerDisplay(true);
    setStepsPerDisplay(10);
}

public static void main(String[] args) {
    SimulationControl.createApp(new BoxApp());
}

```

How long does it take for the system to reach equilibrium? How does this time depend on the number of particles? After the system reaches equilibrium, what is the magnitude of the fluctuations? How do the fluctuations depend on the number of particles? Problems 7.2 and 7.3 explore such questions.

Exercise 7.1. Simple tests of operators and methods

- It is frequently quicker to write a short program to test how the operators and methods of a computer language work than to look them up in a manual or online. Write a test class to determine the values of $3/2$, $3.0/2.0$, $(\text{int})(3.0/2.0)$, $2/3$, and $(\text{int})(-3/2)$.
- Determine the behavior of `Math.round(arg)`, `Math.ceil(arg)`, and `Math rint(arg)`.
- Write a program to test whether the same sequence of random numbers appears each time the program is run if we use the method `Math.random()` to generate the sequence.
- Create an object from class `Random` and use the methods `setSeed(long seed)` and `nextDouble()`. Show that you obtain the same sequence of random numbers if the same seed is used. One reason to specify the seed rather than to choose it at random from the time (as is the default) is that it is convenient to use the same random number sequence when testing a program. Suppose that your program gives a strange result for a particular run. If you

notice a error in the program and change the program, you would want to use the same random number sequence to test whether your changes corrected the error. Another reason for specifying the seed is that another user could obtain the same results if you tell them the seed that you used. \square

Problem 7.2. Approach to equilibrium

- (a) Use BoxApp and Box and describe the nature of the evolution of n , the number of particles on the left side of the box. Choose the total number of particles N to be $N = 8, 16, 64, 400, 800$, and 3600 . Does the system reach equilibrium? What is your qualitative criterion for equilibrium? Does n , the number of particles on the left-hand side, change when the system is in equilibrium?
- (b) The algorithm we have used is needlessly cumbersome, because our only interest is the number of particles on each side. We used the positions only for visualization purposes. Because each particle has the same chance to go through the hole, the probability per unit time that a particle moves from left to right equals the number of particles on the left divided by the total number of particles, that is, $p = n/N$. Modify the program so that the following algorithm is implemented.
 - (i) Generate a random number r from a uniformly distributed set of random numbers in the interval $0 \leq r < 1$.
 - (ii) If $r \leq p = n/N$, move a particle from left to right, that is $n \rightarrow n - 1$; otherwise, $n \rightarrow n + 1$.
- (c) Does the time dependence of n appear to be deterministic for sufficiently large N ? What is the qualitative behavior of $n(t)$? Estimate the time for the system to reach equilibrium from the plots. How does this time depend on N ? \square

Problem 7.3. Equilibrium fluctuations

- (a) As a rough measure of the equilibrium fluctuations, visually estimate the deviation of $n(t)$ from $N/2$ for $N = 16, 64, 400, 800$, and 3600 ? Choose a time interval that is bigger than the time needed to reach equilibrium. How do your results for the deviation depend on N ?
- (b) A better measure of the equilibrium fluctuations is the mean square fluctuations Δn^2 , which is defined as

$$\Delta n^2 = \langle (n - \langle n \rangle)^2 \rangle = \langle n^2 \rangle - 2\langle n \rangle \langle n \rangle + \langle n \rangle^2 = \langle n^2 \rangle - 2\langle n \rangle^2 + \langle n \rangle^2 = \langle n^2 \rangle - \langle n \rangle^2. \quad (7.1)$$

The brackets $\langle \dots \rangle$ denote an average taken after the system has reached equilibrium. The relative magnitude of the fluctuations is $\Delta n / \langle n \rangle$. Modify your program so that averages are taken after equilibrium has been reached. Run for a time that is long enough to obtain meaningful results. Compute the mean square fluctuations Δn^2 for the same values of N considered in part (a). How do the relative fluctuations, $\Delta n / \langle n \rangle$, depend on N ? (You might find it helpful to see how averages are computed in Listings 7.3 and 7.4.) \square

From Problem 7.2 we see that $n(t)$ decreases in time from its initial value to its equilibrium value in an almost deterministic manner if $N \gg 1$. It is instructive to derive the time dependence of $n(t)$ to show explicitly how chance can generate deterministic behavior. If there are

$n(t)$ particles on the left side after t moves, then the change in $\langle n \rangle(t)$ in the time interval Δt is given by

$$\Delta \langle n \rangle = \left[\frac{-\langle n \rangle(t)}{N} + \frac{N - \langle n \rangle(t)}{N} \right] \Delta t. \quad (7.2)$$

(We defined the time so that the time interval $\Delta t = 1$ in our simulations.) What is the meaning of the two terms in (7.2)? If we treat $\langle n \rangle$ and t as continuous variables and take the limit $\Delta t \rightarrow 0$, we have

$$\frac{\Delta \langle n \rangle}{\Delta t} \rightarrow \frac{d \langle n \rangle}{dt} = 1 - \frac{2 \langle n \rangle(t)}{N}. \quad (7.3)$$

The solution of the differential equation (7.3) is

$$\langle n \rangle(t) = \frac{N}{2} [1 + e^{-2t/N}] \quad (7.4)$$

where we have used the initial condition $\langle n \rangle(t = 0) = N$. Note that $\langle n \rangle(t)$ decays exponentially to its equilibrium value $N/2$. How does this form (7.4) compare to your simulation results for various values of N ? We can define a *relaxation time* τ as the time it takes the difference $[\langle n \rangle(t) - N/2]$ to decrease to $1/e$ of its initial value. How does τ depend on N ? Does this prediction for τ agree with your results from Problem 7.2?

***Problem 7.4.** A simple modification

Modify your program so that each side of the box is chosen with equal probability. One particle is then moved from the side chosen to the other side. If the side chosen does not have a particle, then no particle is moved during this time interval. Do you expect that the system behaves in the same way as before? Do the simulation starting with all the particles on the left side of the box and choose $N = 800$. Do not keep track of the positions of the particles. Compare the behavior of $n(t)$ with the behavior of $n(t)$ found in Problem 7.3. How do the values of $\langle n \rangle$ and Δn compare? \square

The probabilistic method discussed on page 198 for simulating the approach to equilibrium is an example of a *Monte Carlo* algorithm, that is, the random sampling of the most probable outcomes. An alternative method is to use *exact enumeration* and determine all the possibilities at each time interval. For example, suppose that $N = 8$ and $n(t = 0) = 8$. At $t = 1$ the only possibility is $n = 7$ and $n' = 1$. Hence, $P(n = 7, t = 1) = 1$, and all other probabilities are zero. At $t = 2$ one of the seven particles on the left can move to the right, or the one particle on the right can move to the left. Because the first possibility can occur in seven different ways, we have the nonzero probabilities, $P(n = 6, t = 2) = 7/8$ and $P(n = 8, t = 2) = 1/8$. Hence, at $t = 2$ the average number of particles on the left side of the box is

$$\langle n(t = 2) \rangle = 6P(n = 6, t = 2) + 8P(n = 8, t = 2) = \frac{1}{8} [6 \times 7 + 8 \times 1] = 6.25. \quad (7.5)$$

Is this exact result consistent with what you found in Problem 7.2? In this example N is small, and we could continue the enumeration of all the possibilities indefinitely. However, for larger N the number of possibilities becomes very large after a few time intervals, and we need to use Monte Carlo methods to sample the most probable outcomes.

7.2 Random Walks

In Section 7.1 we considered the random motion of many particles in a box, but we did not care about their positions—all we needed to know was the number of particles on each side.

Suppose that we want to characterize the motion of a dust particle in the atmosphere. We know that as a given dust particle collides with molecules in the atmosphere, it changes its direction frequently, and its motion appears to be random. A simple model for the trajectory of a dust particle in the atmosphere is based on the assumption that the particle moves in any direction with equal probability. Such a model is an example of a *random walk*.

The original statement of a random walk was formulated in the context of a drunken sailor. If a drunkard begins at a lamp post and takes N steps of equal length in random directions, how far will the drunkard be from the lamp post? We will find that the mean square displacement of a random walker, for example, a dust particle or a drunkard, grows linearly with time. This result and its relation to diffusion leads to many applications that might seem to be unrelated to the original drunken sailor problem.

We first consider an idealized example of a random walker that can move only along a line. Suppose that the walker begins at $x = 0$ and that each step is of equal length a . At each time interval the walker has a probability p of a step to the right and a probability $q = 1 - p$ of a step to the left. The direction of each step is independent of the preceding one. After N steps the displacement x of the walker from the origin is given by

$$x_N = \sum_{i=1}^N s_i \quad (7.6)$$

where $s_i = \pm a$. For $p = 1/2$ we can generate one walk of N steps by flipping a coin N times and increasing x by a each time the coin is heads and decreasing x by a each time the coin is tails.

We expect that if we average over a sufficient number of walks of N steps, then the average of x_N , denoted by $\langle x_N \rangle$, would be $(p - q)Na$. We can derive this result by writing $\langle x \rangle = \sum_{i=1}^N \langle s_i \rangle = N\langle s \rangle$, because the average of the sum is the sum of the averages, and the average of each step is the same. We have $\langle s \rangle = p(a) + q(-a) = (p - q)a$, and the result follows. For simplicity, we will frequently drop the subscript N and write $\langle x \rangle$.

Because $\langle x \rangle = 0$ for $p = 1/2$, we need a better measure of the extent of the walk. One measure is the displacement squared:

$$x_N^2 = \left[\sum_{i=1}^N s_i \right]^2. \quad (7.7)$$

For $p \neq 1/2$ it is convenient to consider the mean square net displacement Δx^2 defined as

$$\Delta x^2 \equiv \langle (x - \langle x \rangle)^2 \rangle \quad (7.8a)$$

$$= \langle x^2 \rangle - \langle x \rangle^2. \quad (7.8b)$$

(We have written Δx^2 rather than $\langle \Delta x^2 \rangle$ for simplicity.) To determine Δx^2 , we write (7.8a) as the sum of two terms:

$$\Delta x^2 = \left\langle \sum_{i=1}^N \Delta_i \sum_{j=1}^N \Delta_j \right\rangle = \left\langle \sum_{i=1}^N \Delta_i^2 \right\rangle + \left\langle \sum_{i \neq j=1}^N \Delta_i \Delta_j \right\rangle \quad (7.9)$$

where $\Delta_i = s_i - \langle s \rangle$. The first sum on the right-hand side of (7.9) includes terms for which $i = j$; the second sum is over i and j with $i \neq j$. Because each step is independent, we have $\langle \Delta_i \Delta_j \rangle = \langle \Delta_i \rangle \langle \Delta_j \rangle$. This term is zero because $\langle \Delta_i \rangle = 0$ for any i . The first term in (7.9) equals $N\langle \Delta^2 \rangle = N[\langle s^2 \rangle - \langle s \rangle^2] = N[a^2 - (p - q)^2 a^2] = 4pqNa^2$, where we have used the fact that $p + q = 1$. Hence

$$\Delta x^2 = 4pqNa^2 \quad (\text{analytic result}). \quad (7.10)$$

We can gain more insight into the nature of random walks by doing a Monte Carlo simulation, that is, by using a computer to “flip coins.” The implementation of the random walk algorithm is simple, for example,

```

if (p < Math.random()) {
    x++;
}
else {
    x--;
}

```

Clearly we have to sample many N step walks because, in general, each walk will give a different outcome. We need to do a Monte Carlo simulation many times and average over the results to obtain meaningful averages. Each N -step walk is called a *trial*. How do we know how many trials to use? The simple answer is to average over more and more trials until the average results don’t change within the desired accuracy. The more sophisticated answer is to do an error analysis similar to what we do in measurements in the laboratory. Such an analysis is discussed in Section 11.4.

The more difficult parts of a program to simulate random walks are associated with book-keeping. The walker takes a total of N steps in each trial, and the net displacement x is computed after every step. Our convention will be to use a variable name ending in Accumulator (or Accum) to denote a variable that accumulates the value of some variable. In Listing 7.3 we provide two classes to be used to simulate random walks.

Listing 7.3: Listing of Walker class.

```

package org.opensourcephysics.sip.ch07;
public class Walker {
    // accumulated data on displacement of walkers, index is time
    int xAccum[], xSquaredAccum[];
    int N; // maximum number of steps
    double p; // probability of step to the right
    int position; // position of walker

    public void initialize() {
        xAccum = new int[N+1];
        xSquaredAccum = new int[N+1];
    }

    public void step() {
        position = 0;
        for(int t = 0; t < N; t++) {
            if(Math.random() < p) {
                position++;
            } else {
                position--;
            }
            // determine displacement of walker after each step
            xAccum[t+1] += position;
            xSquaredAccum[t+1] += position*position;
        }
    }
}

```

Listing 7.4: Target class for random walk simulation.

```

package org.opensourcephysics.sip.ch07;
import org.opensourcephysics.controls.*;
import org.opensourcephysics.frames.*;

public class WalkerApp extends AbstractSimulation {

    Walker walker = new Walker();
    PlotFrame plotFrame = new PlotFrame("time", "<x>,<x^2>", "Averages");
    HistogramFrame distribution =
        new HistogramFrame("x", "H(x)", "Histogram");
    int trials; // number of trials

    public WalkerApp() {
        plotFrame.setXYColumnNames(0, "t", "<x>");
        plotFrame.setXYColumnNames(1, "t", "<x^2>");
    }

    public void initialize() {
        walker.p = control.getDouble("Probability p of step to right");
        walker.N = control.getInt("Number of steps N");
        walker.initialize();
        trials = 0;
    }

    public void doStep() {
        trials++;
        walker.step();
        distribution.append(walker.position);
        distribution.setMessage("trials = "+trials);
    }

    public void stopRunning() {
        plotFrame.clearData();
        for(int t = 0; t<=walker.N; t++) {
            double xbar = walker.xAccum[t]*1.0/trials;
            double x2bar = walker.xSquaredAccum[t]*1.0/trials;
            plotFrame.append(0, 1.0*t, xbar);
            plotFrame.append(1, 1.0*t, x2bar - xbar*xbar);
        }
        plotFrame.repaint();
    }

    public void reset() {
        control.setValue("Probability p of step to right", 0.5);
        control.setValue("Number of steps N", 100);
    }

    public static void main(String[] args) {
        SimulationControl.createApp(new WalkerApp());
    }
}

```

Problem 7.5. Random walks in one dimension

- (a) In class `Walker` the steps are of unit length so that $a = 1$. Use `Walker` and `WalkerApp` to estimate the number of trials needed to obtain Δx^2 for $N = 20$ and $p = 1/2$ with an accuracy of approximately 5%. Compare your result for Δx^2 to the exact answer in (7.10). Approximately how many trials do you need to obtain the same relative accuracy for $N = 100$?
- (b) Is $\langle x \rangle$ exactly zero in your simulations? Explain the difference between the analytic result and the results of your simulations. Note that we have used the same notation $\langle \dots \rangle$ to denote the exact average calculated analytically and the approximate average computed by averaging over many trials. The distinction between the two averages should be clear from the context.
- (c) How do your results for $\langle x \rangle$ and Δx^2 change for $p \neq q$? Choose $p = 0.7$ and determine the N dependence of $\langle x \rangle$ and Δx^2 .
- (d)* Determine Δx^2 for $N = 1$ to $N = 5$ by enumerating all the possible walks. For simplicity, choose $p = 1/2$ so that $\langle x \rangle = 0$. For $N = 1$ there are two possible walks: one step to the right and one step to the left. In both cases $x^2 = 1$, and hence $\langle x_1^2 \rangle = 1$. For $N = 2$ there are four possible walks with the same probability: (i) two steps to the right, (ii) two steps to the left, (iii) first step to the right and second step to the left, and (iv) first step to the left and second step to the right. The value of x_2^2 for these walks is 4, 4, 0, and 0, respectively, and hence $\langle x_2^2 \rangle = (4 + 4 + 0 + 0)/4 = 2$. Write a program that enumerates all the possible walks of a given number of steps and compute the various averages of interest exactly. \square

The class `WalkerApp` displays the distribution of values of the displacement x after N steps. One way of determining the number of times that the variable x has a certain value would be to define a one-dimensional array, `probability`, and let

```
probability[x] += 1;
```

In this case because x takes only integer values, the array index of `probability` is the same as x itself. However, the above statement does not work in Java because x can be negative as well as positive. What we need is a way of mapping the value x to a bin or index number. The `HistogramFrame` class, which is part of the Open Source Physics display package, does this mapping automatically using the Java `Hashtable` class. In simple data structures, data is accessed by an index that indicates the location of the data in the data structure. `Hashtable` data is accessed by a *key*, which in our case is the value of x . A hashing function converts the key to an index. The `append` method of the `HistogramFrame` class takes a value, finds the index using a hashing function, and then increments the data associated with that key. The `HistogramFrame` class also draws itself.

The `HistogramFrame` class is very useful for taking a quick look at the distribution of values in a data set. You do not need to know how to group the data into bins or the range of values of the data. The default bin width is unity, but the bin width can be set using the `setBinWidth` method. See `WalkerApp` for an example of the use of the `HistogramFrame` class. Frequently, we wish to use the histogram data to compute other quantities. You can collect the data using the `Data Table` menu item in `HistogramFrame` and copy the data to a file. Another option is to include additional code in your program to analyze the data. The following statements assume that a `HistogramFrame` object called `histogram` has been created and data entered into it.

```

// creates array entries of data from histogram
java.util.Map.Entry[] entries = histogram.entries();
for (int i = 0, length = entries.length; i < length; i++) {
    // gets bin number
    Integer binNumber = (Integer) entries[i].getKey();
    // gets number of occurrences for bin number i
    Double occurrences = (Double) entries[i].getValue();
    // gets value of left edge of bin
    double value = histogram.getLeftMostBinPosition(binNumber.intValue());
    // sets value to middle of bin
    value += 0.5*histogram.getBinWidth();
    // convert from Double class to double data type
    double number = occurrences.doubleValue();
    // use value and number in your analysis
}

```

Problem 7.6. Nature of the probability distribution

- Compute $P_N(x)$, the probability that the displacement of the walker from the origin is x after N steps. What is the difference between the histogram, that is, the number of occurrences, and the probability? Consider $N = 10$ and $N = 40$ and at least 1000 trials. Does the qualitative form of $P_N(x)$ change as the number of trials increases? What is the approximate width of $P_N(x)$ and the value of $P_N(x)$ at its maximum for each value of N ?
- What is the approximate shape of the envelope of $P_N(x)$? Does the shape change as N is increased?
- Fit the envelope $P_N(x)$ for sufficiently large N to the continuous function

$$C \frac{1}{\sqrt{2\pi\Delta x^2}} e^{-(x-\langle x \rangle)^2/2\Delta x^2}. \quad (7.11)$$

The form of (7.11) is the standard form of the Gaussian distribution with $C = 1$. The easiest way to do this fit is to plot your results for $P_N(x)$ and the form (7.11) on the same graph using your results for $\langle x \rangle$ and Δx^2 as input parameters. Visually choose the constant C to obtain a reasonable fit. What are the possible values of x for a given value of N ? What is the minimum difference between these values? How does this difference compare to your value for C ? \square

Problem 7.7. More random walks in one dimension

- Suppose that the probability of a step to the right is $p = 0.7$. Compute $\langle x \rangle$ and Δx^2 for $N = 4, 8, 16$, and 32 . What is the interpretation of $\langle x \rangle$ in this case? What is the qualitative dependence of Δx^2 on N ?
- An interesting property of random walks is the mean number $\langle D_N \rangle$ of *distinct* lattice sites visited during the course of an N step walk. Do a Monte Carlo simulation of $\langle D_N \rangle$ and determine its N dependence. \square

We can consider either a large number of successive walks as in Problem 7.7 or a large number of noninteracting walkers moving at the same time as in Problem 7.8.

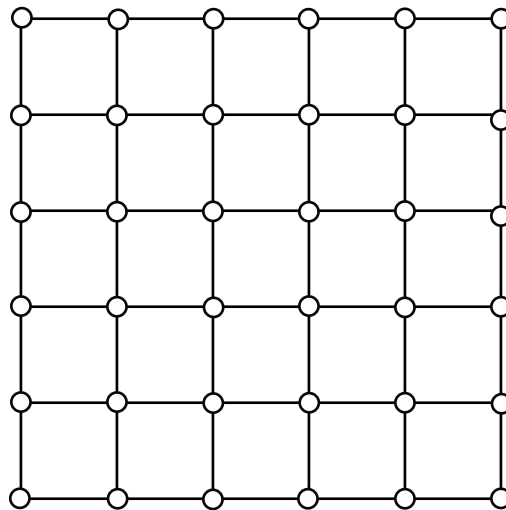


Figure 7.2: An example of a 6×6 square lattice. Note that each site or node has four nearest neighbors.

Problem 7.8. A random walk in two dimensions

- Consider a collection of walkers initially at the origin of a square lattice (see Figure 7.2). At each unit of time, each of the walkers moves at random with equal probability in one of the four possible directions. Create a drawable class, `Walker2D`, which contains the positions of M walkers moving in two dimensions and draws their location, and modify `WalkerApp`. Unlike `WalkerApp`, this new class need not specify the maximum number of steps. Instead, the number of walkers should be specified.
- Run your application with the number of walkers $M \geq 1000$ and allow the walkers to take at least 500 steps. If each walker represents a bee, what is the qualitative nature of the shape of the swarm of bees? Describe the qualitative nature of the surface of the swarm as a function of the number of steps N . Is the surface jagged or smooth?
- Compute the quantities $\langle x \rangle$, $\langle y \rangle$, Δx^2 , and Δy^2 as a function of N . The average is over the M walkers. Also compute the mean square displacement R^2 given by

$$R^2 = \langle x^2 \rangle - \langle x \rangle^2 + \langle y^2 \rangle - \langle y \rangle^2 = \Delta x^2 + \Delta y^2. \quad (7.12)$$

What is the dependence of each quantity on N ? (As before, we will frequently write R^2 instead of R_N^2 .)

- Estimate R^2 for $N = 8, 16, 32$, and 64 by averaging over a large number of walkers for each value of N . Assume that $R = \sqrt{R^2}$ has the asymptotic N dependence:

$$R \sim N^\nu \quad (N \gg 1), \quad (7.13)$$

and estimate the exponent ν from a log-log plot of R^2 versus N . We will see in Chapter 13 that the exponent $1/\nu$ is related to how a random walk fills space. If $\nu \approx 1/2$, estimate the magnitude of the self-diffusion coefficient D from the relation $R^2 = 4DN$.

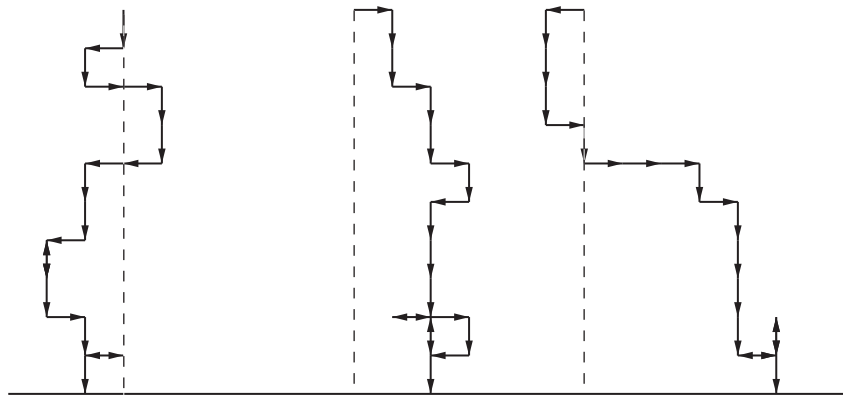


Figure 7.3: Examples of the random path of a raindrop to the ground. The step probabilities are given in Problem 7.9. The walker starts at $x = 0$, $y = h$.

- (e) Do a Monte Carlo simulation of R^2 on a triangular lattice (see Figure 8.5) and estimate ν . Can you conclude that the exponent ν is independent of the symmetry of the lattice? Does D depend on the symmetry of the lattice? If so, give a qualitative explanation for this dependence.
- (f)* Enumerate all the random walks on a square lattice for $N = 4$ and obtain exact results for $\langle x \rangle$, $\langle y \rangle$, and R^2 . Assume that all four directions are equally probable. Verify your program by comparing the Monte Carlo and exact enumeration results. \square

Problem 7.9. The fall of a rain drop

Consider a random walk that starts at a site a distance $y = h$ above a horizontal line (see Figure 7.3). If the probability of a step down is greater than the probability of a step up, we expect that the walker will eventually reach a site on the horizontal line. This walk is a simple model of the fall of a rain drop in the presence of a random swirling breeze. Do a Monte Carlo simulation to determine the mean time τ for the walker to reach any site on the line $y = 0$ and find the functional dependence of τ on h . Is it possible to define a velocity in the vertical direction? Because the walker does not always move vertically, it suffers a net displacement x in the horizontal direction. How does Δx^2 depend on h and τ ? Reasonable values for the step probabilities are 0.1, 0.6, 0.15, 0.15, corresponding to up, down, right, and left, respectively. \square

7.3 Modified Random Walks

So far we have considered random walks on one- and two-dimensional lattices where the walker has no “memory” of the previous step. What happens if the walker remembers the nature of the previous steps? What happens if there are multiple random walkers, with the condition that no double occupancy is allowed? We explore these and other variations of the simple random walk in this section. All these variations have applications to physical systems, but the applications are more difficult to understand than the models themselves.

The fall of a raindrop considered in Problem 7.9 is an example of a *restricted* random walk, that is, a walk in the presence of a boundary. In the following problem, we discuss in a more

general context the effects of various types of restrictions or boundaries on random walks. Other examples of a restricted random walk are given in Problems 7.17 and 7.23.

Problem 7.10. Restricted random walks

- (a) Consider a one-dimensional lattice with trap sites at $x = 0$ and $x = L$ ($L > 0$). A walker begins at site x_0 ($0 < x_0 < L$) and takes unit steps to the left and right with equal probability. When the walker arrives at a trap site, it can no longer move. Do a Monte Carlo simulation and verify that the mean number of steps τ for the particle to be trapped (*the mean first passage time*) is given by

$$\tau = (2D)^{-1}x_0(L - x_0) \quad (7.14)$$

where D is the self-diffusion coefficient in the absence of traps [see (7.29)].

- (b) Random walk models in the presence of traps have had an important role in condensed matter physics. For example, consider the following idealized model of energy transport in solids. The solid is represented as a lattice with two types of sites: hosts and traps. An incident photon is absorbed at a host site and excites the host molecule or atom. The excitation energy or *exciton* is transferred at random to one of the host's nearest neighbors, and the original excited molecule returns to its ground state. In this way the exciton wanders through the lattice until it reaches a trap site at which a chemical reaction occurs. A simple version of this energy transport model is given by a one-dimensional lattice with traps placed on a periodic sublattice. Because the traps are placed at regular intervals, we can replace the random walk on an infinite lattice by a random walk on a circular ring. Consider a lattice of N host or nontrapping sites and one trap site. If a walker has an equal probability of starting from any host site and an equal probability of a step to each nearest neighbor site, what is the N dependence of the mean survival time τ (the mean number of steps taken before a trap site is reached)? Use the results of part (a) rather than doing a simulation.
- (c) Consider a one-dimensional lattice with reflecting sites at $x = -L$ and $x = L$. For example, if a walker reaches the reflecting site at $x = L$, it is reflected at the next step to $x = L - 1$. At $t = 0$ the walker starts at $x = 0$ and steps with equal probability to the left and right. Write a Monte Carlo program to determine $P_N(x)$, the probability that the walker is at site x after N steps. Compare the form of $P_N(x)$ with and without the presence of the reflecting sites. Can you distinguish the two probability distributions if N is the order of L ? At what value of N can you first distinguish the two distributions? \square

Problem 7.11. A persistent random walk

- (a) In a persistent random walk, the *transition* or jump probability depends on the previous step. Consider a walk on a one-dimensional lattice, and suppose that step $N - 1$ has been made. Then step N is made in the same direction with probability α ; a step in the opposite direction occurs with probability $1 - \alpha$. Write a program to do a Monte Carlo simulation of the persistent random walk in one dimension. Estimate $\langle x \rangle$, Δx^2 , and $P_N(x)$. Note that it is necessary to specify both the initial position and an initial direction of the walker. What is the $\alpha = 1/2$ limit of the persistent random walk?
- (b) Consider $\alpha = 0.25$ and $\alpha = 0.75$ and determine Δx^2 for $N = 8, 64, 256$, and 512 . Assume that $\Delta x^2 \sim N^{2\nu}$ for large N , and estimate the value of ν from a log-log plot of Δx^2 versus N

for large N . Does v depend on α ? If $v \approx 1/2$, determine the self-diffusion coefficient D for $\alpha = 0.25$ and 0.75 . In general, D is given by

$$D = \frac{1}{2d} \lim_{N \rightarrow \infty} \frac{\Delta x^2}{N} \quad (7.15)$$

where d is the dimension of space. That is, D is given by the asymptotic behavior of the mean square displacement. (For the simple random walk considered in Section 7.2, $\Delta x^2 \propto N$ for all N .) Give a physical argument for why $D(\alpha \neq 0.5)$ is greater (smaller) than $D(\alpha = 0.5)$.

- (c) You might have expected that the persistent random walk yields a nonzero value for $\langle x \rangle$. Verify that $\langle x \rangle = 0$, and explain why this result is exact. How does the persistent random walk differ from the biased random walk for which $p \neq q$?
- (d) A persistent random walk can be considered as an example of a *multistate* walk in which the state of the walk is defined by the last transition. The walker is in one of two states; at each step the probabilities of remaining in the same state or switching states are α and $1 - \alpha$, respectively. One of the earliest applications of a two-state random walk was to the study of diffusion in a chromatographic column. Suppose that a molecule in a chromatographic column can be either in a mobile phase (constant velocity v) or in a trapped phase (zero velocity). Instead of each step changing the position by ± 1 , the position at each step changes by $+v$ or 0 . A quantity of experimental interest is the probability $P_N(x)$ that a molecule has moved a distance x in N steps. Choose $v = 1$ and $\alpha = 0.75$ and determine the qualitative behavior of $P_N(x)$. \square

Problem 7.12. Synchronized random walks

- (a) Randomly place two walkers on a one-dimensional lattice of L sites, so that both walkers are not at the same site. At each time step randomly choose whether the walkers move to the left or to the right. Both walkers move in the same direction. If a walker cannot move in the chosen direction because it is at a boundary, then this walker remains at the same site for this time step. A trial ends when both walkers are at the same site. Write a program to determine the mean time and the mean square fluctuations of the time for two walkers to reach the same site. This model is relevant to a method of doing cryptography using neural networks (see Rutter et al.).
- (b) Change your program so that you use biased random walkers for which $p \neq q$. How does this change affect your results? \square

Problem 7.13. Random walk on a continuum

One of the first continuum models of a random walk was proposed by Rayleigh in 1919. In this model the length a of each step is a random variable and the direction of each step is uniformly random. In this case the variable of interest is R , the distance of the walker from the origin after N steps. The model is known as the freely jointed chain in polymer physics (see Section 7.7) in which case R is the end-to-end distance of the polymer. For simplicity, we first consider a walker in two dimensions with steps of equal (unit) length at a random angle.

- (a) Write a Monte Carlo program to compute $\langle R \rangle$ and determine its dependence on N .
- (b) Because R is a continuous variable, we need to compute $p_N(R)\Delta R$, the probability that R is between R and $R + \Delta R$ after N steps. The quantity $p_N(R)$ is the *probability density*. Because