

THE LAZY UNIVERSE

An Introduction to the Principle of Least Action

> Jennifer COOPERSMITH

THE LAZY UNIVERSE



Cornelius Lanczos From the archives of the Dublin Institute for Advanced Studies

THE LAZY UNIVERSE

An Introduction to the Principle of Least Action

by Jennifer Coopersmith





Great Clarendon Street, Oxford, OX2 6DP, United Kingdom

Oxford University Press is a department of the University of Oxford. It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide. Oxford is a registered trade mark of Oxford University Press in the UK and in certain other countries

© Jennifer Coopersmith 2017

The moral rights of the author have been asserted

First Edition published in 2017

Impression: 1

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, without the prior permission in writing of Oxford University Press, or as expressly permitted by law, by licence or under terms agreed with the appropriate reprographics rights organization. Enquiries concerning reproduction outside the scope of the above should be sent to the Rights Department, Oxford University Press, at the address above

You must not circulate this work in any other form and you must impose this same condition on any acquirer

Published in the United States of America by Oxford University Press 198 Madison Avenue, New York, NY 10016, United States of America

> British Library Cataloguing in Publication Data Data available

Library of Congress Control Number: 2016953473

ISBN 978-0-19-874304-0

Printed and bound by CPI Group (UK) Ltd, Croydon, CR0 4YY

Links to third party websites are provided by Oxford in good faith and for information only. Oxford disclaims any responsibility for the materials contained in any third party website referenced in this work. To Cornelius Lanczos (1893–1974) and Murray John Peake (1952–) "There's husbandry in heaven" ACT II, SCENE I MACBETH SHAKESPEARE

Preface

The mathematician and physicist, Cornelius Lanczos (1893–1974), wrote one of the best books of physics explanation that has ever been written: *The Variational Principles of Mechanics*.¹ It explains the true meaning and philosophical content of the Principle of Least Action. The present book is a shorter and simplified version of that classic – but it is an original interpretation (anything taken directly from Lanczos is attributed and referenced by page number); it is also a paean to Lanczos.

Who is the book for? It is for *you*! It concerns a principle that underpins the whole of physics, so how could it not be important to understand even a tiny bit of it? The greatest theories have one common feature - they always bring in wisdom far-surpassing their original remit, extending into almost every walk of life. Therefore, one should not limit one's area of study but rather follow the maxim: "The more you know, the more you *can* know; the more you understand, the more you *can* understand."² Moreover, as well as an increased understanding of the world, you will be privy to a rare reward - to "theories of excessive beauty".³

It is true that the reader is assumed to have a background in the physical sciences, however a layreader could also read this book, with profit and enjoyment, by skim-reading the mathematics, or by reading just the Introduction (Chapter 1), the Final Words (Chapter 9), and the historical and popular chapters (2 and 8). Furthermore, the book is not a textbook: there are many equations, but the aim is to *explain* - why these equations and not others, and what do the equations mean?⁴ The appendices are usually at a more advanced level and condensed in style, but they may be entirely passed over without loss of continuity in the main text. Apart from the invaluable asset of seeing how a problem is solved, the reason for so many appendices is twofold: to provide a

¹ Lanczos C, *The Variational Principles of Mechanics*, University of Toronto Press (1949). All page numbers will refer to the fourth edition, Dover Publications, Inc. New York (1970), and we shall write 'Lanczos, page x'.

- ² (author's maxim)
- ³ Lanczos, page 229.

⁴ This is in contrast to the approach in Synge and Griffith, Principles of Mechanics, 3rd edition, McGraw-Hill Book Company, Inc (1959), page 413, where the advice is: "Do not attempt to see a physical meaning in these [mathematical] operations; it will not help." compact resource for the physicist (for example, a physicist on a beach holiday); and for the layreader to know what subject-headings to follow up at a later stage, if so desired. Enrichment material has been included but often as optional reading (in small font), or in parentheses, or footnotes. Also, we have avoided detailing all the qualifications, exemptions, and special cases, in order to be able to say things simply, and make bold statements.

Terminology: on the authority of both Lanczos, and Richard Feynman,⁵ we refer to all the appropriate principles - Hamilton's Principle, Jacobi's Principle, Lagrange's Principle, and Maupertuis's Principle - as Principles of Least Action, as they are generically so. 'Variational Mechanics' means any physics problem that uses the Principle of Least Action. Also, the fact that we usually say 'least' as opposed to 'stationary' is explained in the main text - see Section 6.6.

The University of La Trobe is thanked for granting me an honorary research associateship, and for the use of library and computing facilities. It was Joe Petrolito, Emeritus Professor of Engineering at La Trobe, Bendigo campus, who introduced me to the website of Edwin F Taylor, Senior Research Scientist Emeritus at the Massachusetts Institute of Technology (MIT). This led to a short but rewarding email correspondence.

At OUP, I thank Sönke Adlung and Ania Wronski. The use of LaTeX, Ubuntu, Gimp, Wikipedia, Google, Metapost, Scilab, and Mfpic are gratefully acknowledged. Mal Haysom and Deborah Peake are thanked for their help with the diagrams (Deborah drew the maze, the encumbered porcupine and Stevin's 'Wreath of Spheres'). George Rogers, librarian at DIAS, once again anticipated my needs and supplied superb images of Lanczos and of Hamilton, with no fuss. Most of all, I am grateful to Gerald Jay Sussman, Panasonic Professor of Electrical Engineering at MIT, who undertook a critical reading of the book prior to printing. The errors he identified have been corrected. Finally, I thank Murray Peake for helpful discussions on maths and physics, and for leading me to Lanczos in the first place.

If I have inspired any reader to seek out the work of the master, Cornelius Lanczos, then it will be 'mission accomplished'.

⁵ Feynman R P, Feynman's Lectures on Physics, Volume II, Chapter 19.

Contents

List of F	igures	xi
1	Introduction	1
2	Antecedents	12
3	Mathematics and physics preliminaries: of hills and plains and other things	31
4	The Principle of Virtual Work	59
5	D'Alembert's Principle	88
6	Lagrangian Mechanics	107
7	Hamiltonian Mechanics	143
8	The whole of physics	183
9	Final words	192
A1.1	Newton's Laws of Motion	197
A2.1	Portraits of the physicists	198
A3.1	Reversible displacements	201
A6.1	Worked examples in Lagrangian Mechanics	202
A6.2	Proof that <i>T</i> is a function of v^2	217
A6.3	Energy conservation and the homogeneity of time	222
A6.4	The method of Lagrange Multipliers	225
A6.5	Generalized Forces	228
A7.1	Hamilton's Transformation, examples	230
A7.2	Demonstration that the p_i s are independent coordinates	232
A7.3	Worked examples in Hamiltonian Mechanics	233
A7.4	Incompressibility of the phase fluid	237
A7.5	Energy conservation in extended phase space	238
A7.6	Link between the action, <i>S</i> , and the 'circulation'	241

Contents

A7.7	Transformation equations linking <i>p</i> and <i>q</i> via <i>S</i>	243
A7.8	Infinitesimal canonical transformations	244
A7.9	Perpendicularity of wavefronts and rays	248
A7.10	Problems solved using the Hamilton-Jacobi Equation	250
A7.11	Quasi refractive index in mechanics	255
A7.12	Einstein's link between Action and the de Broglie waves	256
Bibliogra Index	aphy and Further Reading	259 261

List of Figures

1.1	The suitors' puzzle.	2
2.1	Weights, connected by a cord, on inclined planes.	13
2.2	'Wreath of spheres' draped over inclined planes (after Stevin).	13
2.3	Huygens's canal-boat thought-experiment.	15
2.4	Brachystochrone, curve of 'swiftest descent' (schematic).	19
3.1	"Anamorphic Column" by Istvan Orosz.	43
3.2	The path of a cannon ball' (schematic).	49
3.3	Is the ladder stable?	52
3.4	The function, $f(x)$, and some variations, $f^{var}(x)$.	56
3.5	The distinction between df and δf .	57
4.1	Feynman's pivoting bar.	70
4.2	The 'black box'.	79
4.3	Knob and slider.	80
4.4	Stability of a ladder.	81
4.5	Two springs connected to a mass.	81
4.6	Spring loaded with a weight.	83
4.7	Soap bubble.	83
5.1	Connected masses, 'Half-Atwood' and 'Black Box'.	93
5.2	'Fictitious Forces', examples.	97
5.3	Newton's Bucket.	101
7.1	Nicolas Poussin, A Dance to the Music of Time.	142
7.2	Phase space: An oscillating spring.	158
7.3	Phase space: Simple pendulum.	159
7.4	Phase space: Spinning top.	160
7.5	Wavefront of common action in configuration space.	173
7.6	Huygens's wavelets.	176
7.7	Construction of the next wavefront.	177
A2.1	Portraits of the physicists.	199
A3.1	Washing-line seen through a window.	201
A6.1.1	Plane polar coordinates.	203
A6.1.2	Atwood's Machine.	205
A6.1.3	The 'Half-Atwood' Machine.	206
A6.1.4	Mass attached to a spring.	207
A6.1.5	Coordinates for the spherical pendulum.	210
A6.1.6	Projectile motion.	211

A6.1.7	Electrical and mechanical systems compared.	214
A6.1.8	One block sliding on another block.	215
A7.1	Coordinates for the spherical pendulum.	231

It would be wonderful if there was one principle, simple to state, that could account for every process in the physical universe. But there *is* such a principle, a surprisingly well-kept secret, that accounts for *almost* every physical process. It is a principle that is more powerful than Newton's ' $\mathbf{F} = m\mathbf{a}$ ', and a principle that doesn't have energy conservation as a requirement in every scenario. We know that Newtonian Mechanics must be replaced when speeds are very high, or the masses are tiny, or huge - but this 'new' principle still applies in these extreme regimes. How can one principle explain so much? The clue comes from the deep wisdom of the eighteenth-century French *philosophe*, Jean le Rond d'Alembert:¹

"L'univers, pour qui saurait l'embrasser d'un seul point de vue, ne serait, s'il est permis de le dire, qu'un fait unique et une grande vérité."

("If one could grasp the whole Universe from one viewpoint, it would appear, if it is permitted to say this, as a unique fact and a great truth.")

No one knows whether d'Alembert's beautiful claim is correct but one thing is certain, if we cannot find one universal viewpoint then we will not arrive at one universal truth. For our viewpoint to be universal it is not a question of us all looking at the view from the same hilltop, rather, it is a requirement that all viewpoints are equivalent, and that there is just *one* universal rule or law or algorithm that solves the problem. Our new principle achieves this - but it seems incredible that one simple 'algorithm' could cope with all the specificity, variety, and complexity across the whole of physics. To make this plausible, we consider the following fable.

¹ D'Alembert, J le R, Discours préliminaire de l'encyclopédie, 1751.



Figure 1.1 The suitors' puzzle.

There once was a King and he set a fiendish puzzle for prospective suitors who wanted to marry his daughter, the beautiful princess. The King had constructed a maze and the successful suitor had to provide the princess with instructions for collecting treasure from a casket. The young suitors had a few hours to look at a map of the maze and prepare their instructions. The King then looked at their answers and quickly whittled away the number of competitors to just two. These two suitors, Alfredo and Bruno, had very different approaches to the puzzle. Alfredo provided detailed instructions for every route, advising the princess to sight the tall column, visible from a distance over the hedges, to let her nose guide her to the fragrance of the frangipani tree, and also to listen out for the sound of bells chiming in the bell-tower, and water splashing at the fountain. Bruno came forwards with just a tiny scrap of paper on which it said, "Wherever you may find yourself, turn left at the next intersection. Eventually you will reach the casket." (The King had assured the contestants that there were no disconnected 'islands' within the maze.)

The King, a veritable sage, awarded the hand of his daughter in marriage to this second suitor.

The amazing thing about this suitor's instructions is that they are very simple to state and they are *universal* - they apply to any maze (although the maze must satisfy certain geometrical restrictions, for example, it cannot contain disconnected islands). There is also another

curious attribute of Bruno's algorithm - it is *local*. The princess need only ever look as far ahead as the very next step (while always keeping an eye open for the casket). Although Alfredo's instructions may be broken down into steps the method is not local in the true sense (there are references to distant features - the sound of the fountain, the smell of the frangipani, the sight and sound of the bells in the bell-tower).

This is only a story but it demonstrates some essential points. In our new principle the method is truly local, and this is never the case for Newtonian Mechanics, even when a path may be broken down into lots of tiny incremental steps. But most astounding of all we shall find that the ensuing equations are *invariant*, taking exactly the same form, no matter what the scenario or what coordinates are used. It is not even necessary that the setting is time-independent, or that the components are passive (so the maze could writhe and undulate with time, and the princess could affect the maze, say, by trimming the hedges as she passed by). The reason for this invariance is that we have at last found something absolute: it is not a universal timepiece, yardstick, or reference frame, for there are none, it is a *principle*, and one that applies across almost every area of physics. We introduce it by way of a brief historical aside.

One of the most awe-inspiring developments in physics has been the shift from Newton's to Einstein's view of gravity. In Newton's Theory of Universal Gravitation, gravity is a force acting between bodies, however near or far. In Einstein's Theory of Gravitation (the Theory of General Relativity), the force is completely dispensed with. It is replaced by a patchwork of reference frames, sufficiently small, but seamlessly joined together, and, instead of responding to a force, the orbiting body now responds to geometry - the 'curvature' of 'space'. This is often represented heuristically by the image of, say, the Earth resting on a large two-dimensional surface, such as a trampoline, distorting this surface, and thereby affecting the trajectory of nearby small bodies, such as the Moon. These two theories - Newton's and Einstein's - are utterly different, and yet, amazingly, the experimental differences, for example, the predictions of the Moon's orbit, are practically nil. It turns out that Einstein's approach involves much more complicated calculations and, as we have just stated, barely any practical advantage - so why use it? The answer is that it has deeper explanatory power, it is applicable over a much greater range of problems, and it is philosophically more sound.

We have been talking just of gravitation, but the principle that we introduce explains not only gravitation but all kinds of problems in the physical sciences (statics and dynamics, optics, electricity and magnetism, quantum mechanics, physical chemistry, statistical mechanics, astronomy, materials science, hydrodynamics, quantum electrodynamics (QED), and so on); it also has deeper explanatory power, is applicable over a much greater range of problems, and is philosophically more sound. This principle is the Principle of Stationary Action (the PSA). It can be stated as:

The Principle of Stationary Action The physical system seeks out the 'flattest' region of 'space'.

This is equivalent to choosing the 'straightest' possible path, which (usually) translates as the path of *least* 'distance'. One more thing, whether considering the 'flatness of space' or the 'straightness of paths', or the 'least distance', only the 'space' *nearby* - that is to say, locally, - needs to be inspected.

The subtitle of this book is the Principle of *Least* Action (PLA), but the principle just given is the Principle of *Stationary* Action (PSA)? 'Stationary' is a mathematical term meaning 'at a flat point of 'space' ' but whether that flat point implies a least path requires a further investigation - therefore the PSA is the more general principle, and incorporates the PLA. However, we shall find that the more stringent condition, the PLA, *is* the one we need, and later on we'll switch to calling our principle: the Principle of Least Action. We'll explain this in a later chapter (Section 6.6, Chapter 6).

Einstein's Theory of Special Relativity, that preceded his theory of General Relativity, starts with two postulates: 1) the laws of physics take the same form in every reference frame,² 2) the speed of light in a vacuum is a constant. These two postulates are strikingly different: the first postulate (the Principle of Relativity) is philosophical in character - Einstein coached us into realizing that physics just couldn't be practised unless postulate 1) applied. On the other hand, postulate 2) appears to be empirical - the speed of light is constant, yes, but perhaps, in another Universe, it might have been variable? Similarly, in the case

² When we say 'reference frame' we shall automatically mean a 'valid reference frame'.

of the Principle of Stationary Action, the Principle has two postulates of very different natures. The first postulate, 1), we have already described that the system 'space' is as 'flat' as possible, locally. This appears reasonable but rather abstract and philosophical; it sounds more like geometry than physics, and we need to know - 'flat' with respect to what? This is where the second postulate comes in, the one that contains the physical input. Postulate 2) states that what is actually being flattened is a certain specific physical quantity - 'action'. This quantity has dimensions of energy \times time, or linear momentum \times distance or angular momentum \times angle, and so on. In one of its first incarnations, 'action' was given as ' $m \times v \times ds$ ', where m is the mass of a 'free' particle, v is its speed, and ds is a small distance along the particle's path. As we have to do with a postulate, we cannot justify the choice by deduction from even more elemental principles. Nevertheless, 'action' does seem like a worthy candidate for a telling physical quantity - it is a scalar (a pure magnitude, having no direction - therefore more likely to be an invariant), it 'spans the physical space' (nothing crucial is missed out), and it does so in the simplest way possible (*mvds* is postulated rather than, say, $m^2 v^3 d^4 s / dt^4$).

D'Alembert's "one viewpoint" implies objectivity, and this is difficult to arrive at in everyday life where prejudice abounds. For example: we barely notice the reaction-force against the soles of our feet, or on our bottoms, that is present almost every minute of our lives (Einstein teaches us that we are thereby not in 'free-fall', and so do not serve as a 'natural' frame of reference); on the other hand, in the rapidly rotating 'gravitron' at the funfair, we feel pinned as if by a great weight but have no sensation of our spinning motion (we merely notice that we can barely nod or move our arms). When revisiting a park that we knew as a child, we find that it resembles a pocket handkerchief rather than a vast estate - is the slight increase in the height of our eyes the source of this change? No, it arises because we have undergone an enormous (nonlocal) translation in time, during which our brain has totally altered. We watch the water sloshing about in a neighbour's swimming pool perhaps they have installed a wave-generating machine? Upon closer inspection we find no machine, but realize that 'a giant hand' - an Earth tremor - is gently rocking the pool: therefore our initial assumption of an isolated system, defined by the edges of the pool, is wrong.

What helps us to achieve objectivity in physics (as opposed to everyday life) is the fact that we are bound by the strictures of mathematical tests. The PSA is centred on a mathematical test - a 'test of the flatness of 'space' ' - which is remarkable for its ability to winnow away the distracting observer-dependent features and so arrive at the true invariant laws. It succeeds in this because it involves an 'extremal' feature of the mathematical landscape (something like 'shortest route between Peshawar and Kabul'), and these features are unique in that they don't depend upon the type of map or even the units used (the route is shortest whether we use a Mercator's or a Peter's projection, and whether we measure in feet or in metres). Before the test can be applied, we have to define what we mean by 'space'. We follow the historical development, and return to our discussion of Newton.

You most likely know Newton's Laws of Motion³ but we want now to give a different perspective of them, emphasizing the philosophical assumptions. An implicit premise of Newton's Mechanics was the outstanding advance: 'space' and 'physics' are totally separate from each other. 'Physics' means forces, masses, and masses in motion. 'Space', following Descartes' invention of the coordinate system, means the three everyday space dimensions (commonly designated *x*, *y*, and *z*), and the time, *t*. All of *x*, *y*, *z*, and *t* are assumed independent of each other, and each goes on to infinity. Gone are the sixteenth century's tendencies, empathies, abhorrences, and vortices; and Newton's 'space' is empty, not full like Descartes'. (It is a void, which Newton does not abhor.)

Next after 'space' come particles - bodies with no internal structure but having an intrinsic property, mass. By Law (Newton's First), each 'free' particle is either at rest, or moves at constant speed and in a constant direction.

Finally, forces, F, are introduced, such as an attractive force between one particle and another. A force has one effect and one effect only - it causes a particle to accelerate. This is where mass plays its role, it determines how big the acceleration shall be, for a given force. (Apart from this, mass is inert - it doesn't depend on when or where the particle is, or on its state of motion.) All this is asserted in the Second Law, $F = ma.^4$ Another outstanding hypothesis was that for composite bodies (bodies made from many particles), or indeed for any complicated arrangement of particles, the net outcome could be obtained by 'summing' over the

³ See Appendix A1.1, Newton's Laws of Motion.

⁴ (But it could all have been so much more complicated; the force could have left the motion unchanged but caused the mass to swell, or it could have caused an acceleration not in line with **F**, or caused a third-order change in the position, and so on.)

influence of each particle considered on its own.⁵ Thus was born the idea of the 'rigid body', an extended body made up from separate particles, but which could itself be treated as if it were one single particle, with all its mass concentrated at one point.

325 years on from Newton's "*Principia*"⁶ it is hard to remain sufficiently impressed. As the philosopher, Schopenhauer, said, a theory passes from being rejected as ridiculous to being accepted but taken to be obvious.⁷ Consider Newton's use of 'acceleration'. It was already known, from Galileo's Principle of Relativity, that uniform motion is relative to the observer. Newton turned this around: all non-uniform, that is to say, *accelerated*, motion is *not* relative to the observer, it *can* be known absolutely. Acceleration with respect to what? Answer, with respect to 'space'. But as the accelerations are absolute, then Newton's 'space' is absolute. So we have arrived back at Newton's wonderful abstraction, an infinite 'space', an inert and absolute background to physical happenings.

With the PSA, every tenet of Newton's Mechanics is challenged: the absolutes of Newton are avoided as every measure - whether it be a position, a speed, a direction, a time, and so on - is always defined with reference to something *within* the given system; action-at-a-distance does not occur - a global picture is built up, piece by piece, but by antennae which are sensitive only to conditions locally; the axes of 'space' no longer extend to infinity, are not necessarily independent of each other, and not necessarily independent of the masses within; and Newton's modular approach, building up complexity from more and more complicated arrangements of particles, each taken singly, is replaced by a holistic 'systems' approach.

Let's explain this 'systems' approach by analogy. Bertrand Russell (philosopher and mathematician) quipped that the activities of mankind amounted to the redistribution of matter within $\pm 0.2\%$ of the Earth, at its surface (this was before the era of space travel). To check Russell's claim, we could exhaustively track the motion of every single person, throughout recorded history, and note what masses they were carrying and where they deposited them; or, we could estimate the

 $^{^5\,}$ (Again, it could have been more complicated, the force might have been cast anew for, say, each trio of particles.)

⁶ Newton, Isaac, Philosophiae naturalis principia mathematica (The Mathematical Principles of Natural Philosophy), 1687.

⁷ (There is then the third phase, when the theory is again rejected.)

matter-content of cities built, crops grown, monuments constructed, bodies buried, and so on. In the second, 'systems', approach, we have lost the simplicity of basic elements (a person, their movements, what they are carrying) and instead have more abstract concepts relating to the whole system (cities, roads, pyramids, etc.). Some totally new possibilities arise ('the deceased') that were not catered for (!) in the first approach. We end up with a static tally of the mass distribution.

Another, more dynamic, example is given by the description of a football match. In the modular description we have only 'players' and 'a football'; in the other approach, the 'systems-view', we have 'defence-position', 'attack', 'tackling', 'dribbling the ball', 'goal-kick', and so on. One counter-intuitive aspect of this systems-view is that we appear to have lost that quintessential feature of motion - its directionality. However, we soon realize that it is not lost but embedded in the whole-system structures (for example, 'goal-kick' has no absolute direction (say, 30° West) yet it conveys all the directional information required, and makes reference only to features within the system - the goal posts).

Returning now to the PSA, the method is as follows: (i) Instead of particles we have individual components of the system. These are chosen in a system-specific way. (They can be billiard balls, atoms, planets, lever arm, pendulum, capacitor, and so on, as the given problem demands); (ii) Instead of forces there are 'scalar structure functions', what in a previous life we have called the energy functions (the kinetic energy, and the potential energy); (iii) we identify all the independent 'motions' that the system can undergo. These 'motions' are the physical changes that happen naturally and that characterize the given system - what in a later life we shall call the 'degrees of freedom'. (For example, the planet orbits, the lever arm rotates, swings swing, and roundabouts turn.) (iv) We come finally to the application of a principle, a principle that requires an exploration of the 'space' in which the physical problem occurs. Knowing all the 'motions', we then choose an *alternative* set of 'motions' that could occur. These motions are hypothetical - we have hypothesized them - however we are not free to hypothesize anything we like: the 'motions' must all be in the same given 'space' (each system has its own 'system-space'), occur in the given time-window, and they must be 'nearby'. Now these 'motions' imply certain amounts of kinetic energy and potential energy consumed or generated in the given time. From these energies we compute a certain quantity - the total action - used up

in the given time. In short, we determine the total hypothetical action for this choice of hypothetical motions. We then continue the exploration and consider another choice of hypothetical motions and again determine the consequential hypothetical action. And so on. The principle then asserts that of all choices for hypothetical motions, *the actual motions are those which make the change-in-action-between-choices come out to zero*. More evocatively, the system finely adjusts itself, via the actual motions (acting in concert but instant by instant taking their marching orders from the scalar structure functions) in just such a way that the action used is least. That these subtly different versions - the italicized one and the evocative one - are the same will emerge during the course of this book.

Be reassured: these ideas are new and many and abstract; there is no way they can be understood in one go. It is useful to collect them together in one place, but not possible to convey all the nuances in a single paragraph. For example, we shall later on discover that sometimes the 'space' exploration is made explicitly by us (in the method known as the Principle of Virtual Work) and sometimes the mathematics takes care of it (in the Variational or Lagrangian Mechanics). Also, there is sometimes an elision made between the 'motions' as 'degrees of freedom' and the 'motions' as hypothetical 'variations'.

Did we write 'hypothetical'? Yes, this is the *piece de resistance*: the 'system-space' is a *virtual* abstract space, and this is what finally enables us to achieve the required objectivity (the actual physical space could have this, that, or the other observer-bias, whereas the virtual abstract space is neutral).

Here is a summary of the main virtues of the PSA.

- (i) It does the job.
- (ii) As forces play no part in the method then 'forces-of-constraint' also play no part. Incredible but true.
- (iii) Better understanding of physics. We can now have 'cat' and 'mouse' instead of only 'particles' and 'particle-particle interactions'. But a 'cat' is more than the sum of its 'particles'.
- (iv) No hard and fast distinction between 'active' and 'passive' components. (Just as a river carves out the river-bed, and the river-bed determines the path of the river, so the 'curvature of space' affects the paths of bodies, and moving bodies affect the 'curvature of space'.)

- (v) Philosophically superior: local (and this is all we ever detect experimentally); no pre-existing empty 'space' (just take the world as it is and then examine it⁸); the system is what's important.
- (vi) Global View. Even when 'space' is flat locally, 'curvature' can still arise globally - by patching together smaller regions with the requirement that there is no 'puckering at the seams'.
- (vii) The PSA gives prominence to energy and to the whole system. Kinetic energy is shown to be a more fundamental and primitive concept than force. Also the dichotomy between kinetic energy and potential energy is explained.
- (viii) The PSA reveals a deep connection between symmetry and conserved properties. (Newton's Mechanics does not lead naturally to any conservation laws except for one, the conservation of linear momentum.)
 - (ix) In Newtonian Mechanics no attempt is given to show how robust the solution is (how it changes following small changes in the starting conditions) or to give ball-park estimates. The PSA, via Hamilton's Mechanics, does address these questions.
 - (x) Amazing unity of approach across almost the whole of physics. (Almost? The exceptions will be discussed in due course.)

Apart from the fact that the PSA is the keystone of physics, and therefore an indispensable tool for the professional engineer or physical scientist, there are two other reasons why we would like to awaken an appreciation of it, even a non-mathematical appreciation. The first is a pragmatic reason. We all know, roughly speaking, what space, time, mechanics, quantum mechanics, matter, and energy are about. These ideas have passed into the public domain. It would be inefficient to start from scratch in our science classes, and not even incorporate advances made during the seventeenth and eighteenth centuries. Somehow the PSA has got missed out - it is time to correct this. The second reason is aesthetic. The beauty of physics does not reside only in the beauty of the night sky, a rainbow, or a sunset. It resides even more in the interior logic, the principles which reach across vast areas of the physical

⁸ This is more correct, especially when we remember that all our observations really have been carried out in the presence of large gravitating masses. Even where experiments are carried out in remote regions, the results still need to be brought back to Earth - in our present state of evolution.

world, unifying them into a self-consistent whole, and with the most economical set of starting premises. (One could call it the 'Aha' feeling.)

Alice (from Lewis Carroll's *Alice's Adventures in Wonderland*) tried so hard to get through the door in order to see the exquisitely beautiful garden beyond. Consider this garden as a metaphor for physics: it's true that without mathematics we shall never be able to wander freely around this garden, but it would be wonderful indeed if we could just be lifted up to peer at it through the keyhole.

Antecedents

A new principle rarely arises completely out of the blue, there are usually vague presentiments in the air. Some precursors to the Principle of Least Action are described (we don't attempt to give an exhaustive history).

Simon Stevin (1548–1620)

Stevin (Stevinus) was a 'geometer' from Bruges in Flanders. The discovery of which he was most proud concerns the condition for static equilibrium of weights on inclined planes (see Figure 2.1):

What relation must hold between the masses and the lengths of the triangle in order for the system to be in equilibrium? Our first guess might be to 'resolve the forces' into components, then balance the horizontal components to zero, and, finally, equate the vertical components to the weights. Stevin, in Antwerp in 1588, attacked the problem in a different way - (he was working 200 years before vectors, and 99 years before Newton's forces).

The first remarkable thing Stevin did was to bring (hypothetical) motion into this static set-up. He recast the problem and imagined a chain of spherical masses that could circulate around the triangle, with the lower part of the chain hanging freely. Stevin idealized this arrangement: the chain could move in either direction, without friction, and would never get stuck on the pointy bits (imagine a little frictionless pulley at the apex of the triangle). Also, the mass had to be *uniformly* distributed along the chain - think of a bead chain like the ones used today to open and close Venetian blinds. Stevin argued that if the section of chain from A to X pulls more than the section from X to B then the chain will over-balance to the left. It will circulate anticlockwise (excuse the anachronism) until the original portion AX then occupies the arc Y to A. The entire length of chain that hangs freely below the little table, between A and B, can be ignored - by symmetry it is always in balance



Figure 2.1 Weights, connected by a cord, on inclined planes.



Figure 2.2 'Wreath of spheres' draped over inclined planes (after Stevin).

with itself. However, the replacement section of chain now lying along AX will again over-balance that new section now lying along XB. In other words, the new state, after circulation, is identical to the initial state. But this will always be true (the new state, after circulation, will always be identical to the initial state), and so the chain will always over-balance, and will keep circulating for ever.

But a continuously circulating chain, without an engine to drive it, is absurd. This is the second remarkable step taken by Stevin - he considered 'perpetual motion' as self-evidently absurd, and used this as the basis for an argument (a *reductio ad absurdum* argument).¹ To avoid the absurd outcome then the starting premise - section AX over-balances section XB - must be wrong; in other words, the initial distribution of masses, with the chain draping itself smoothly over the surfaces, must already be in equilibrium. This is a powerful proof, as it applies for an infinity of different starting states (different inclined planes).

Stevin realized he had discovered an eternal truth, and he displayed a diagram of his 'Wreath of Spheres' (*clootcrans*) as the frontispiece of his book on mechanics. The punchline, from our point of view, is that: without the use of forces, considering only gentle movements of the chain in harmony with the constraints (motion within certain surfaces), Stevin showed that, with respect to these motions, nothing changes, and equilibrium is maintained. In fact (going beyond Stevin's own interpretation), the proof can be made to sound like pure geometry: for equilibrium, it is only necessary that the length of chain along an inclined plane is equal to the length of that inclined plane. Moreover, as any bumpy surface can be thought of as lots of tiny triangles, we have only to lie a uniform chain on this surface, with no bunching up or stretching out, and it *will* be in equilibrium.

Christiaan Huygens (1629–95)

Huygens, in Paris in 1656, tried to understand billiard-ball-type collisions but without the use of forces (this was some thirty years before Newton's *Principia*, but billiards was a game that had been played since the fifteenth century). Then, as now, scientists were motivated by the thrill of showing that a great name was wrong. Descartes was the towering authority in the middle of the seventeenth century, and we shall see that three scientists (natural philosophers) in our story found errors in Descartes' work - Fermat, Huygens, and Leibniz.

Back to Huygens. In his rules of collision, Descartes had claimed that a small body, hitting a larger body at rest, would never be able to shift it. Huygens knew that this couldn't possibly be right - it didn't agree with

¹ This is probably the first time this had ever been done - that is, the use of the impossibility of perpetual motion as the basis of an argument.

Antecedents

experiment.² To show that Descartes was wrong, he used a remarkable demonstration. (Huygens might have been a bit apprehensive about contradicting the Cartesian view - Descartes had been a regular visitor to the Amsterdam home of Huygens's parents, and his tutor was horrified to hear of the young Huygens' 'heresy'.)

Huygens's remarkable demonstration consisted in putting the Principle of the Relativity of Motion (due to Galileo) to *quantitative* use - the first time this had ever been done.³ He imagined a 'billiard-ball collision' viewed simultaneously from two vantage points - from a smoothly coasting canal-boat, and from the canal-bank (undoubtedly his childhood in Amsterdam was an influence). Here is a picture of this thought-experiment, taken from the frontispiece of Huygens's treatise.

It's not clear from this picture whether the experiment was carried out on the boat or on the canal-bank. It doesn't matter, so let's say it happened on the boat. On the boat, the large mass hits a stationary smaller mass. Now we are free to choose the speed of the canal-boat



Figure 2.3 Huygens's canal-boat thought experiment⁴ (from *De Motu corpore ex percussione*, 1656)⁵.

² Descartes also knew that experiment was contradicted, but he had the perfect fudge - see Coopersmith, J. *Energy, the Subtle Concept*, Revised Edition, Oxford University Press, (2015) - hereafter referred to as Coopersmith, *EtSC*.

³ (to the author's knowledge)

⁴ The figure shows balls of the same size, yet we can imagine that one is more massive than the other.

⁵ Huygens C, De Motu corpore ex percussione, 1656, published by Martin Nijhoff NV.