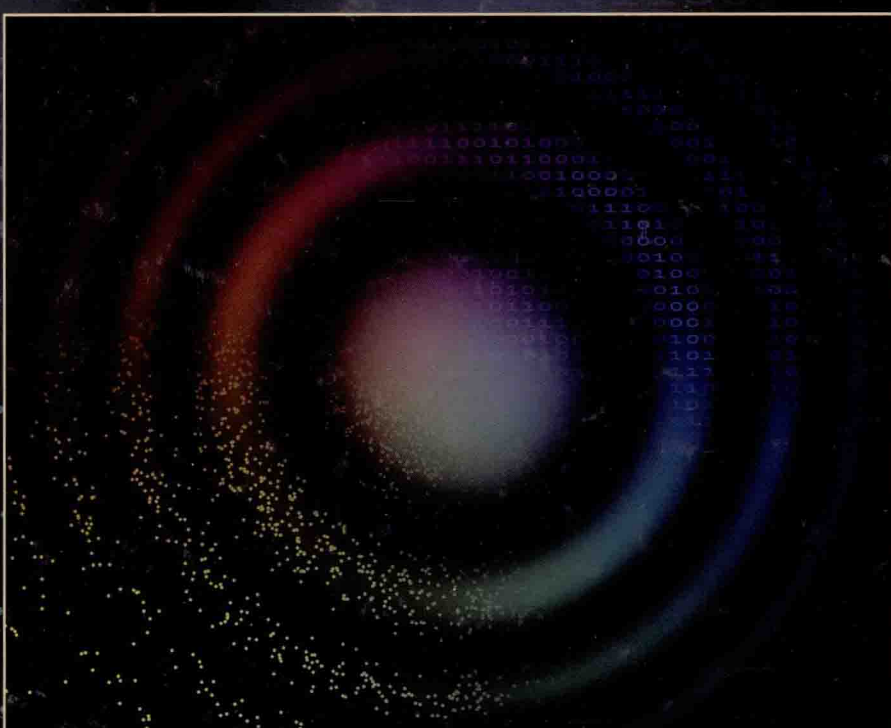
QUANTUM PROCESSES SYSTEMS, & INFORMATION



Benjamin Schumacher
Michael Westmoreland

Quantum Processes, Systems, and Information

BENJAMIN SCHUMACHER
Kenyon College
MICHAEL D. WESTMORELAND
Denison University

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore,
São Paulo, Delhi, Dubai, Tokyo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521875349

First published 2010

Printed in the United Kingdom at the University Press, Cambridge

A catalogue record for this publication is available from the British Library

Library of Congress Cataloging-in-Publication Data

Schumacher, Benjamin.
Quantum processes, systems, and information / Benjamin Schumacher, Michael Westmoreland.
p. cm.
ISBN 978-0-521-87534-9 (Hardback)
1. Quantum theory–Textbooks. I. Westmoreland, Michael D. II. Title.
QC174.12.S385 2010
530.12–dc22

2009039353

ISBN 978-0-521-87534-9 Hardback

Preface

The last two decades have seen the development of the new field of quantum information science, which analyzes how quantum systems may be used to store, transmit, and process information. This field encompasses a growing body of new insights into the basic properties of quantum systems and processes and sheds new light on the conceptual foundations of quantum theory. It has also inspired a great deal of contemporary research in optical, atomic, molecular, and solid state physics. Yet quantum information has so far had little impact on the way that quantum mechanics is taught.

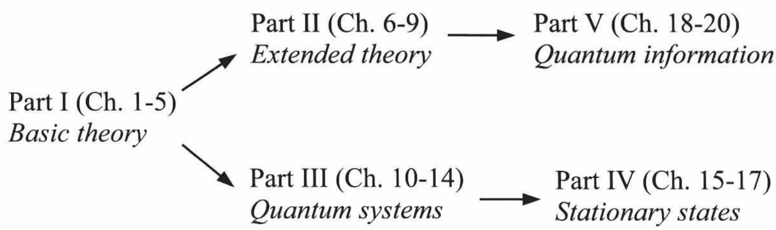
Quantum Processes, Systems, and Information is designed to be both an undergraduate textbook on quantum mechanics and an exploration of the physical meaning and significance of information. We do not regard these two aims as incompatible. In fact, we believe that attention to both subjects can lead to a deeper understanding of each. Therefore, the essential "story" of this book is very different from that found in most existing undergraduate textbooks.

Roughly speaking, the book is organized into five parts:

- Part I (Chapters 1–5) presents the basic outline of quantum theory, including a development of the essential ideas for simple "qubit" systems, a more general mathematical treatment, basic theorems about information and uncertainty, and an introduction to quantum dynamics.
- Part II (Chapters 6–9) extends the theory in several ways, discussing quantum entanglement, ideas of quantum information, density operators for mixed states, and dynamics and measurement on open systems.
- Part III (Chapters 10–14) uses the basic theory to discuss several specific quantum systems, including particles moving in one or more dimensions, systems with orbital or intrinsic angular momentum, harmonic oscillators and related systems, and systems containing many particles.
- Part IV (Chapters 15–17) deals with the stationary states of particles moving in 1-D and 3-D potentials, including variational and perturbation methods.
- Part V (Chapters 18–20) further develops the ideas of quantum information, examining quantum information processing, NMR systems, the meaning of classical and quantum entropy, and the idea of error correction.

These chapters are followed by Appendices on probability (Appendix A), Fourier series and Fourier transforms (Appendix B), Gaussian functions (Appendix C) and generalized quantum evolution (Appendix D).

Part I is the basis for all further work in the text. The remaining parts follow two quasi-independent tracks:



Thus, this book could be used as a text for either an upper-track or a lower-track style of course.[1]

We, however, strongly recommend including material from both tracks. This book is written from the conviction that a modern student of physics needs a broader set of concepts than conventional quantum mechanics textbooks now provide. Unitary time evolution, quantum entanglement, density operator methods, open systems, thermodynamics, concepts of communication, and information processing – all of these are at least as essential to the meaning of quantum theory as is solving the time-independent Schrödinger equation.

As we wrote this book, we had the benefit of useful and inspiring conversations with a great many colleagues and friends. Among these we wish particularly to express our gratitude to Charles Bennett, Herb Bernstein, Carl Caves, Chris Fuchs, Lucien Hardy, David Mermin, Michael Nielsen, and Bill Wootters. In a similar vein, we would also like to thank the other members of the (fondly remembered) Central Ohio Quantum Conspiracy: Michael Nathanson, Kat Christandl Gillen, and Lee Kennard. We have also received valuable input on the book from Matthew Neal and Ron Winters of Denison University and Ian Durham of St. Anselm College.

An early version of this book was used as an experimental textbook for a quantum mechanics course at Kenyon College, and the students in that course deserve their own thanks: Andrew Berger, Stephanie Hemmingson, John Hungerford, Lee Kennard, Joey Konieczny, Jeff Lanz, Max Lavrentovich, David Lenkner, Nikhil Nagendra, Alex Rantz, David Slochower, Jeremy Spater, Will Stanton, Adam Tassile, Chris Yorlano, and Matt Zaremsky.

Our faculty colleagues at both Kenyon College and Denison University have been wonderfully supportive throughout this project. One of us (MDW) is grateful to acknowledge a Robert C. Good Faculty Fellowship from Denison University. We also thank our editor at Cambridge University Press, Simon Capelin, for providing the initial impetus and for considerable patience and encouragement throughout.

[1] There are a few minor dependencies not indicated in this chart, but these can be easily accommodated in practice. The general discussion of composite systems in Section 6.1 is a useful preparation for work on many-particle systems in Chapter 14. The analysis of thermal states of a ladder system (Section 13.4) depends on the density operator formalism, but may be omitted if Chapter 8 has not been covered.

We are more grateful than we can readily express for the continuing love and support of our wives, Carol Schumacher and Bonnie Westmoreland. And finally, a word to our children, Barry, Patrick and Carolyn Westmoreland, and Sarah and Glynis Schumacher: This is what we have been so busy doing for the last few years. We hope you like it, because we are dedicating it to you.

Benjamin Schumacher
Department of Physics
Kenyon College

Michael D. Westmoreland
Department of Mathematics
and Computer Science
Denison University

Contents

1 Bits and quanta

1.1 Information and bits

On the evening of 18 April 1775, British troops garrisoned in Boston prepared to move west to the towns of Lexington and Concord to seize the weapons and capture the leaders of the rebellious American colonists. The colonists had anticipated such a move and prepared for it. However, there were two possible routes by which the British might leave the city: by land via Boston Neck, or directly across the water of Boston Harbor. The colonists had established a system of spies and couriers to carry the word ahead of the advancing troops, informing the colonial militias exactly when, and by what road, the British were coming.

The vital message was delivered first by signal lamps hung in the steeple of Christ Church in Boston and observed by watchers over the harbor in Charlestown. As Henry Wadsworth Longfellow later wrote,

> One if by land, and two if by sea;
> And I on the opposite shore will be,
> Ready to ride and spread the alarm
> Through every Middlesex village and farm . . .

Two lamps: the British were crossing the harbor. A silversmith named Paul Revere, who had helped to organize the communication network, was dispatched on horseback to carry the news to Lexington. He stopped at houses all along the way and called out the local militia. By dawn on 19 April, the militiamen were facing the British on Lexington Common. The first battle of the American Revolutionary War had begun.

In the United States, Paul Revere[1] is remembered as a hero of the Revolutionary War, not for his later military career but for his "midnight ride" in 1775. Revere is famous to this day as a carrier of information.

Information is one of our central themes, and over the course of this book we will formalize the concept, generalize it, and subject it to extensive analysis. The story of Paul Revere illustrates several key ideas. What, after all, is information? We can give a heuristic definition that, though we will later stretch and generalize it almost beyond recognition, will serve to guide our discussion.

> *Information* is the ability to distinguish reliably between possible alternatives.

[1] William Dawes and Dr. Samuel Prescott, who accompanied Revere on the ride and helped spread the word, are almost forgotten – perhaps because they were never immortalized in verse by Longfellow.

Before Paul Revere's ride, the militiamen of Lexington could not tell whether the British were coming or not, or by what route. After Revere had reached them, they could distinguish which possibility was correct. They had gained *information*.

We can also distinguish between an abstract *message* and the physical *signal* that represents the message. The association between message and signal is called a *code*. For instance, here is the code used by the colonists for their church steeple signal.

Signal	**Message**
0 lamp	The British are not coming.
1 lamp	The British are coming by land.
2 lamps	The British are coming by sea.

Mathematically, the code is a function from a set of possible messages to a set of possible states of the physical system that will carry the message – in this case, the lamp configuration in the church tower.

There is no requirement that all possible signals are used in the code. (The Boston spies could have hung three lamps, or six, though their distant compatriots would have been rather perplexed.) On the other hand, the association between message and signal should be one-to-one, so that distinct messages are represented by distinct signals. Otherwise, it is not possible to deduce the message reliably from the signal.

Another very important point is that information can be *transformed* from one physical representation to another, so that the same message can be encoded into quite different physical signals. Paul Revere's message was not only represented by lamps in a church, but also by neural activity in his brain and then by patterns of sound waves as he cried, "The British are coming!"

Exercise 1.1 Identify at least seven distinct physical representations that this sentence has had from the time we wrote it to the time you read it.

The fact that the same message can be carried by very different signals is a fundamental truth about information.

This *transformability* of information allows us to simplify matters considerably, for we can always represent a message using signals of a standard type. The universal "currency" for information theory is the *bit*. The term *bit* is a generic term for a physical system with two possible distinguishable states. The states may be designated *no* and *yes*, *off* and *on*, or by the binary digits 0 and 1. These two states may be distinct voltage levels in an electrical device, two directions of magnetization in a small region of a computer disk, the presence or absence of a light pulse, two possible patterns of ink on a piece of paper, etc. All of these bits are *isomorphic*, in that information represented by one type of bit can be converted to another type of bit by a physical process.

A single bit has a limited "capacity" to represent information. We cannot store an entire book in one bit. The reason is that there are too many possible books (that is, possible messages) to be represented in a one-to-one manner by the two states 0 and 1. Somewhere in the code, we would inevitably have something like this:

Signal	Message
⋮	
0	*Alice in Wonderland* by Lewis Carroll
0	*The Guide for the Perplexed* by Moses Maimonides
⋮	

From a bit in the state 0, it would be impossible to choose the correct book in a reliable way.

To represent an entire book, we must use strings or sequences of many bits. If we have a string of n bits, then the number of distinct states available to us is

$$\#\text{ of states} = \underbrace{2 \times \cdots \times 2}_{n\text{ times}} = 2^n. \quad (1.1)$$

Exercise 1.2 A *byte* is a string of eight bits. How many possible states are there for one byte?

Suppose that there are M possible messages. If the number of possible signals is at least as large as M, then we can find a code in which the message can be reliably inferred from the signal. We can thus determine whether the message can be represented by n bits.

- If $M > 2^n$, then n bits are not enough.
- If $M \leq 2^n$, then n bits are enough.

The number n of bits necessary to represent a given message is a way of measuring "how much information" is in the message. This is a very practical sort of measure, since it tells us what resources (bits) are necessary to perform a particular task (represent the message faithfully).

We define the *entropy* H of our message to be

$$H = \log M, \quad (1.2)$$

where M is the number of possible messages. (From now on, unless we otherwise indicate, "log" will denote the logarithms with base 2: $\log \equiv \log_2$.) The entropy H is a measure of the information content of the message. From our discussion above, we see that if $n < H$, n bits will not be enough to represent the message faithfully. On the other hand, if $n \geq H$, then n bits will be enough. Thus, H measures the number of bits that the message requires.[2]

We can think of H as a measure of *uncertainty* – that is, of how much we do not know before we get the message. It is also a measure of how our uncertainty is reduced when we identify which message is the right one. In other words, before we receive and decode the signal, there are M possibilities and $H = \log M$. Afterward, we have uniquely identified the right message, and the entropy is now $H' = \log 1 = 0$.

We have called H the "entropy," which is the name used for H by Claude Shannon in his pioneering work on the mathematical theory of information. The name harkens back to

[2] Anticipating later developments, we should note here that our definition of H implicitly assumes that the M possible messages are all *equally likely*. To cope with more general situations, we will need a more general expression for the entropy. However, Eq. 1.2 will do for our present purposes.

thermodynamics, and for very good reason. Ludwig Boltzmann showed that if a macroscopic system has W possible microscopic states, then the thermodynamic entropy S_θ is

$$S_\theta = k_B \ln W, \tag{1.3}$$

where $k_B = 1.38 \times 10^{-23}$ J/K, called *Boltzmann's constant.* This famous relation (which is inscribed on Boltzmann's tomb in Vienna) can be viewed as a fundamental link between information and thermodynamics. Up to an overall constant factor, the thermodynamic entropy S_θ is just a measure of our uncertainty of the microstate of the system.

Exercise 1.3 A liter of air under ordinary conditions has a thermodynamic entropy of about 5 J/K. How many bits would be necessary to represent the microstate of a liter of air?

We will have more to say about the connection between information and thermodynamics later on.

Suppose A and B are two messages, having M_A and M_B possible values respectively. The two messages taken together form a joint message that we denote AB. If A and B are independent of each other, every combination of A and B values is a possible joint message, and so $M_{AB} = M_A M_B$. In this case the entropy is additive:

$$H(AB) = H(A) + H(B). \tag{1.4}$$

On the other hand, if the messages are not independent, it may be that some combinations of A and B are not allowed, so that the joint entropy $H(AB)$ may be less than $H(A) + H(B)$.

Exercise 1.4 Suppose A and B each have 16 possible values. What is the joint entropy $H(AB)$ (a) if the messages are independent, and (b) if B is known to be an exact copy of A?

How much information was contained in the message sent from the Christ Church steeple in 1775? There were three possible messages, and so $H = \log 3 \approx 1.58$. This means that one bit would not suffice to represent the message, but two bits would be more than enough. That much is clear; but can we give a more exact meaning to H? Does it make sense to say that a message contains 1.58 bits of information?

Suppose our message is a decimal digit, which can take on values 0 through 9. There are ten possible values for this message, so the entropy is $H = \log 10 \approx 3.32$. We shall need at least four bits to represent the digit. But imagine that our task is to encode, not just a single digit, but a whole sequence of independent digits. We could simply set aside four bits per digit, but we can do better by considering groups (or *blocks*) of three digits. Each group has $10^3 = 1000$ possible values, and so has an entropy of $3 \log 10 = 9.97$. Therefore we can encode three digits in ten bits, using (on average) $10/3 \approx 3.33$ bits per digit. This is more efficient, and is very close to using $\log 10$ bits per digit.

Exercise 1.5 Devise a binary code for triples of digits (as described above) and use your code to represent the first dozen digits of π.

This motivates the following argument. Consider a message having an entropy H. If we have a long sequence of independent messages of this type, we can group them into blocks and encode the blocks into bits. If the blocks have n messages, each block has an entropy nH. Let N be the minimum number of bits needed to represent a message block. This will be the smallest integer that is at least as big as nH, and so

$$nH \leq N < nH + 1. \quad (1.5)$$

Calculating the number of bits required on a "per message" basis, we are using $K = N/n$ bits per message, and

$$H \leq K < H + \frac{1}{n}. \quad (1.6)$$

If we consider very large message blocks, $n \gg 1$ and so $1/n$ is very small. The two ends of the inequality chain squeeze together, and for large blocks we will use almost exactly H bits per message to represent the information. Therefore, if we encode our messages "wholesale," the entropy H precisely measures the number of bits per message that we need.

Exercise 1.6 Consider a type of message that has three possible values (like the message of the colonial spies in Boston). Calculate the minimum number of bits required to encode blocks of 2, 3, 5, 10, or 100 such messages. In each case, also calculate the number of bits used per message.

Things become more complicated in the presence of noise. *Noise* is a general term for any process that prevents a signal from being transferred and read unambiguously. For example, imagine that there had been fog on Boston Harbor on that April night in 1775. In a heavy fog, the church steeple might not have been visible at all from Charlestown, and no information would have been conveyed. In a lighter mist, the observers might have been able to see that there were lamps in the steeple, but not been able to count them. They would then have known that the British troops were on the move, but not which way they were going. A part of the information would have been transmitted successfully, but not all.

It is possible to formalize this notion of partial information. Before any communication takes place, there are M possible messages and the entropy is $H = \log M$. Afterward, we have reduced the number of possible messages from M to M', but because of noise $M' > 1$. The amount of information conveyed in this process is defined to be

$$H - H' = \log \frac{M}{M'}. \quad (1.7)$$

Exercise 1.7 A friend is thinking of a number between 1 and 20 (inclusive). She tells you that the number is prime. How much information has she given you?

The concept of information is fundamental in scientific fields ranging from molecular biology to economics, not to mention computer science, statistics, and various branches of engineering. It is also, as we will see, an important unifying idea in physics.

1.2 Wave-particle duality

Since the 17th Century, there have been two basic theories about the physical nature of light. Isaac Newton believed that light is composed of huge numbers of particle-like "corpuscles." Christiaan Huygens favored the idea that light is a wave phenomenon, a moving periodic disturbance analogous to sound. Both theories explain the obvious facts about light, though in different ways. For example, we observe that two beams of light can pass through one another without affecting each other. In the Newtonian corpuscle theory, this simply means that the light particles do not interact with each other. In the Huygensian wave theory, it implies that light waves obey the *principle of superposition*: the total light wave is simply the sum of the waves of the two individual beams.

To take another example, we observe that the shadows of solid objects have sharp edges. This is easily explained by the Newtonian theory, since the light particles move in straight lines through empty space. On the other hand, this observation seems at first to be a fatal blow to the wave theory, because waves moving past an obstacle should spread out in the space beyond. However, if the wavelength of light were very short, then this spreading might be too small to notice. For over a hundred years, the known experimental facts about light were not sufficient to settle whether light was a particle phenomenon or a wave phenomenon, and both theories had many adherents.

Then, in 1801, Thomas Young performed a crucial experiment in which Huygens's wave theory was decisively vindicated. This was the famous two-slit experiment.

Suppose that a beam of monochromatic light shines on a barrier with a single narrow opening, or "slit." The light that passes through the slit falls on a screen some distance away. We observe that the light makes a small smudge on the screen. (For thin slits, this smudge of light actually gets wider when the slit is made narrower, and on either side of the main smudge there are several much dimmer smudges. These facts are already difficult to explain without the wave theory, but we will skip this point for now.)

Light passing through another slit elsewhere in the barrier will make a similar smudge centered on a different point. But suppose two nearby slits are both open at once. If we imagine that light is simply a stream of non-interacting Newtonian corpuscles, we would expect to see a somewhat broader and brighter smudge of light, the result of the two corpuscle-showers from the individual slits.

But what happens in fact (as Young observed) is that the region of overlap of the two smudges shows a pattern of light and dark bands called *interference fringes*, see Fig. 1.1.

This is really strange. Consider a point on the screen in the middle of one of the dark fringes. When either one of the slits is open, some light does fall on this point. But when both slits are open, the spot is dark. In other words, we can *decrease* the intensity of light at some points by *increasing* the amount of light that passes through the barrier.

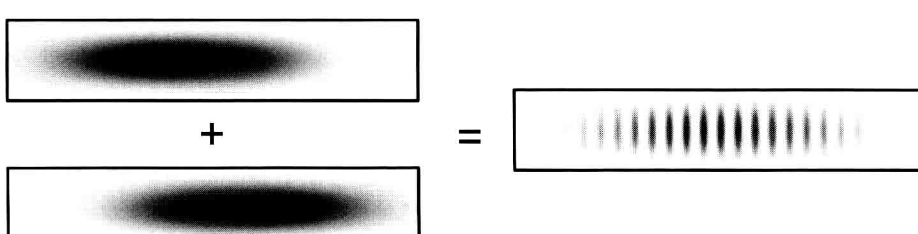


Fig. 1.1 The light patterns from two single slits combine to form a pattern of interference fringes. (For clarity on the printed page, the negative of the pattern is shown; more ink means higher intensity.)

The situation is no less peculiar for the bright fringes. Take a point in the middle of one of these. When either slit is opened, the intensity of light at the point has some value I. But with both slits open, instead of an intensity $2I$ (as we might have expected), we see an intensity of $4I$! The *average* of the intensity over the light and dark fringes is indeed $2I$, but the pattern of light on the screen is less uniform than a particle theory of light would suggest.

Young realized that this curious behavior could easily be explained by the wave theory of light. Waves emerge from each of the two slits, and the combined wave at the screen is just the sum of the two disturbances. Denote by $\phi(\vec{r}, t)$ the quantity that describes the wave in space and time. In sound waves, for example, the "wave function" ϕ describes variations in air pressure. The two slits individually produce waves ϕ_1 and ϕ_2, and by the principle of superposition the two slits together produce a combined wave $\phi = \phi_1 + \phi_2$.

Two further points complete the picture. First we note that ϕ can take on either positive or negative values. By analogy to surface waves on water, the places where ϕ is greatest are called the wave "crests," while the places where ϕ is least (most negative) are called the wave "troughs." Second, the observed intensity of the wave at any place is related to the square of the magnitude of the wave function there: $I \propto |\phi|^2$.

At some points on the screen, the two partial waves ϕ_1 and ϕ_2 are "out of phase," so that a crest of ϕ_1 is coincident with a trough of ϕ_2 and vice versa. At these points, the waves cancel each other out, and $|\phi|^2$ is small. This phenomenon is called *destructive interference* and is responsible for the dark fringes.

At certain other points on the screen, the two partial waves ϕ_1 and ϕ_2 are "in phase," by which we mean that their crests and troughs arrive synchronously. When ϕ_1 is positive, so is ϕ_2, and so on. The partial waves reinforce each other, and $|\phi|^2$ is large. This phenomenon, *constructive interference*, is responsible for the bright fringes.

At intermediate points, ϕ_1 and ϕ_2 neither exactly reinforce one another nor exactly cancel, so the resulting intensity has an intermediate value.

Exercise 1.8 In the two slit experiment, in a particular region of the screen the light from a single slit has an intensity I, but when two slits are open, the intensity ranges over the interference fringes from 0 to $4I$. Explain this in terms of ϕ_1 and ϕ_2.

Young was able to use two-slit interference to determine the wavelength λ of light, which does turn out to be quite small. (For green light, λ is only 500 nm.) Later in the 19th Century, James Clerk Maxwell put the wave theory of light on a firm foundation by showing that light is a travelling disturbance of electric and magnetic fields – an *electromagnetic wave*.

But the wave theory of light was not the last word. In the first years of the 20th Century, Max Planck and Albert Einstein realized that the interactions of light with matter can only be explained by assuming that the energy of light is carried by vast numbers of discrete light *quanta* later called *photons*. These photons are like particles in that each has a specific discrete energy E and momentum p, related to the wave properties of frequency f and wavelength λ:

$$\begin{aligned} E &= hf, \\ p &= \frac{h}{\lambda}, \end{aligned} \tag{1.8}$$