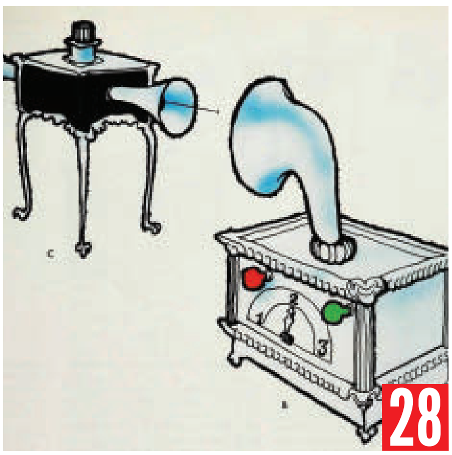
PHYSICS TODAY

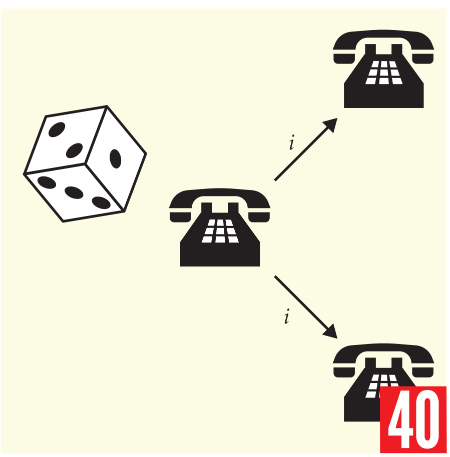
January 2025 • volume 78, number 1

A publication of the American Institute of Physics

SPECIAL ARCHIVAL ISSUE

OUR QUANTUM WORLD





PHYSICS TODAY

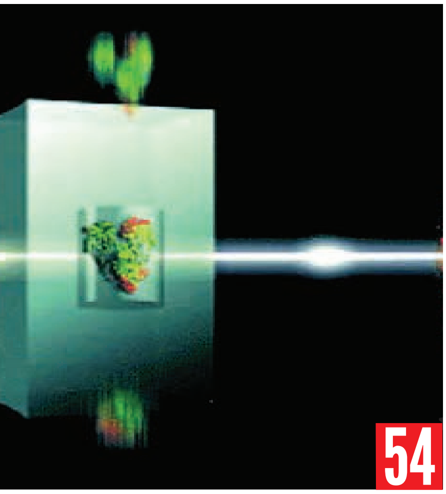
January 2025 | volume 78 number 1

SPECIAL ARCHIVAL ISSUE: OUR QUANTUM WORLD

ON THE COVER: Although it's natural to think of our world classically, quantum mechanics is all around us, and early debates about quantum theory, epitomized by Schrödinger's cat, have morphed into real-world quantum applications. In this special issue, we present articles from the PHYSICS TODAY archives that describe the birth of modern quantum mechanics, the emergence of several key concepts, and emerging applications that have the potential to transform both science and society. (Cover design by Three Ring Studio.)

PHYSICS TODAY

www.physicstoday.org



Address
American Institute of Physics
1 Physics Ellipse
College Park, MD 20740-3842
+1 301 209 3100
pteditors@aip.org

Physics Today (ISSN 0031-9228, coden PHTOAD) volume 78, number 1. Published monthly by the American Institute of Physics, 1305 Walt Whitman Rd, Suite 110, Melville, NY 11747-4300. Periodicals postage paid at Huntington Station, NY, and at additional mailing offices. POSTMASTER: Send address changes to Physics Today, American Institute of Physics, 1305 Walt Whitman Rd, Suite 110, Melville, NY 11747-4300. Views expressed in Physics Today and on its website are those of the authors and not necessarily those of AIP or any of its member societies.

Alliance for Audited Media

FROM THE EDITOR

Our quantum world

Richard J. Fitzgerald

Lasers. MRIs. Precision timekeeping. Solar cells. SI units of measure. High-contrast, high-efficiency display devices. Ultraprecise sensors. Optimized drug development. Secure communications. Most of us don't think about it, but we interact with quantum-enabled devices and applications on a regular basis, and that's only going to accelerate.

INTERNATIONAL YEAR OF
Quantum Science and Technology

The United Nations has declared 2025 to be the International Year of Quantum Science and Technology (IYQ). The timing is intentional: This year marks a century since what is traditionally considered to be the start of the "new quantum theory." (We'll have more about that timing in an upcoming issue of Physics Today.)

The goals of the yearlong event are broader than just recognizing the advances and impact that quantum science and technologies have had. As described in the story by Toni Feder on page 7, the focus will also be on raising awareness—among the public and policymakers—about the importance of quantum science and applications and their potential to help address the world's most pressing needs.

To kick off Physics Today's celebration of the IYQ, we present this special archival issue, in which we've pulled together several of our most enjoyable and informative quantum pieces. Most readers of Physics Today will have some familiarity with quantum mechanics but not necessarily with the history, the current state of the science, or the central concepts behind some of the most promising applications. We had a wealth of archival content to choose from, and filling those gaps was a prime goal of our selection process.

We present the articles in rough chronological order of their themes:

- I. I. Rabi, known for his work on molecular beams, was a graduate student at Columbia University in the 1920s and an eyewitness to the quantum revolution. In a transcription of a 1979 talk that Physics Today originally published in 2006, he shares his colorful recollections of that era.
- In a 1952 essay, Freeman Dyson, one of the main contributors to the development of quantum electrodynamics—the most precise, extensively tested theory in physics—draws parallels to the history of classical electrodynamics to convey "in simple words" the fruits of 25 years of development.
- John Bell's name is inseparable from discussions about the foundations of quantum mechanics. In a 2015 article, Reinhold Bertlmann recounts lively stories about working with Bell and explains why debates about those foundations still exist.
- In his memorably titled 1985 article "Is the Moon there when nobody looks? Reality and the quantum theory"—arguably Physics Today's most well-known article—David Mermin works through what "Bell-type" experiments say about the quantum nature of the world around us.
- Barbara Terhal, Michael Wolf, and Andrew Doherty in a 2003 feature describe how the irreducible quantum mechanical property of entanglement has emerged as an exploitable resource for such technologies as quantum teleportation, quantum communication, and, especially, quantum information processing.
- In their article from 2000, Daniel Gottesman and Hoi-Kwong Lo explain the principles underlying another emerging quantum technology: quantum cryptography.
- Lasing is an inherently quantum mechanical phenomenon. And as Philip Bucksbaum explains in an article from 2006, the ability of lasers to control and measure the quantum world is opening a wealth of new applications.
- In a 2021 article, Christine Middleton tours the layers of organization, operation, and abstraction that allow a user-friendly experience to emerge from the underlying qubits in a quantum computer.
- A 2014 Quick Study by Sheila Dwyer explores how the Heisenberg uncertainty principle can be exploited to improve the precision of quantum measurements—a technique that made possible the 2015 detection of gravitational waves.

The developments surveyed in these articles are snapshots in time of our understanding of the quantum world and of the advances then on the horizon. Collectively, they have helped shape today's frontier in quantum science. Fittingly, we end this special archival issue with a look at one direction of current research: macroscopic qubits.

Over the course of Physics Today's 77 years, we have published many more quantum-related articles than we have been able to include in this special archival issue. They are collected on our website at https://physicstoday.org/quantum.

Our celebration of the IYQ will continue throughout the year, as we bring you articles, Q&As, explainers, and other pieces that look both backward and forward. Like this editor's note and the story on page 7, each of our quantum-themed pieces will carry the IYQ logo. PT

2025 is the International Year of Quantum Science and Technology

Building awareness and inspiring a future workforce are two aims of the UN-designated quantum year.

Hands-on demonstrations of quantum entanglement, role-playing diplomacy games, continental-scale shindigs, and more activities for the International Year of Quantum Science and Technology (IYQ) are coming into focus. Last June, the United Nations declared 2025 the IYQ; since then, scientists, educators, and science lovers have been buzzing with ideas for how to celebrate the past century of quantum physics and its applications and look ahead to the next one.

The UN imprimatur lends visibility and legitimacy to efforts to raise awareness about quantum science and technology. It also comes with a commitment to the UN's 17 sustainable development goals—affordable and clean energy, quality education, and gender equality, to name a few. Many quantum-related activities are underway independent of the IYQ, says Enrica Porcari, head of CERN's IT department and a member of the IYQ steering committee. But the IYQ will "turbocharge" efforts, she says. "I think 2025 will see an explosion of events."

Quantum-based technologies are already ubiquitous, and many more applications in computing, communications, and sensing are on the horizon. "In physics, everyone understands how central quantum mechanics has become, but that's not the case for the public," says Paul Cadden-Zimansky, the physicist at Bard College who set the ball rolling that eventually resulted in the UN declaration and who is an IYQ global coordinator.

The IYQ can be called a success, Porcari says, if by the end of the year,

A DELEGATION headed by Joe Niemela (far right), a scientist emeritus at UNESCO and the Abdus Salam International Centre for Theoretical Physics, rallied support for the IYQ at the United Nations last April. The other delegates were, from left, Yanne Chembo (University of Maryland), American Physical Society president Young-Kee Kim (University of Chicago), and Ana María Cetto (National Autonomous University of Mexico). (Photo courtesy of Joe Niemela.)

people in quantum-underserved countries are saying, "I wouldn't miss this revolution."

Global events

The official IYQ launch is scheduled for 4–5 February at UNESCO's Paris headquarters. The event will introduce the year by focusing on the future of quantum science and technology, says Claudia Fracchiolla, head of public engagement at the American Physical Society, which is one of the six founding sponsors of the IYQ. The event, she says, will focus on questions like, What do policymakers need to think about? How will developments based on quantum physics benefit society? What education and workforce training are needed to prepare for the quantum revolution? What are the ethical considerations? Science ministers, Nobel Prize winners, educators, social scientists, and others will speak at the event.

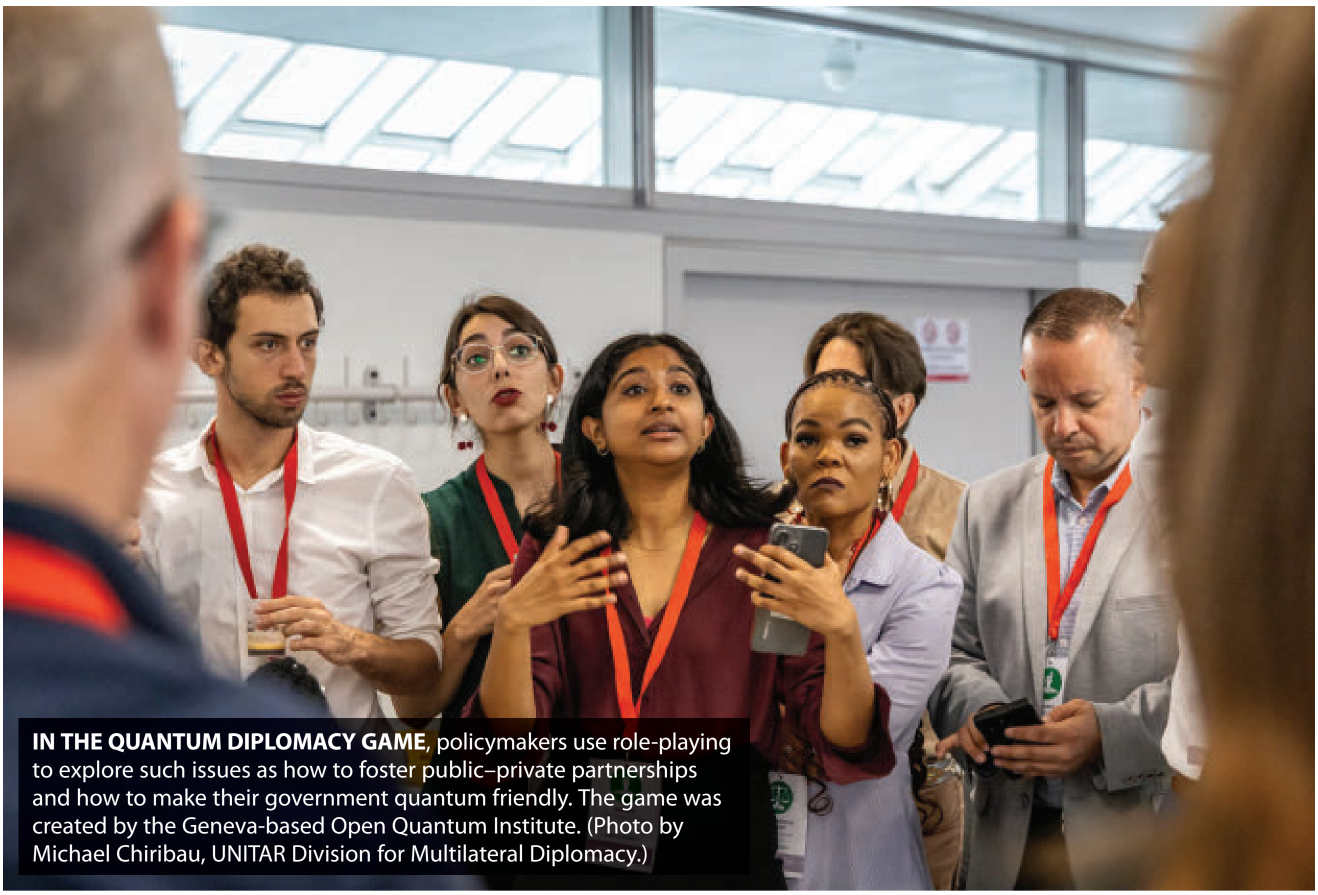
The IYQ sponsors, which have grown to include a couple dozen professional societies, foundations, universities, and companies from around the world, are planning a global event on each continent. Beyond that, the idea is to galvanize grassroots organization of activities large and small.

In March, the American Physical Society will host activities to celebrate the IYQ before and during its Global Physics Summit in Anaheim, California. Some activities, such as a quantum playground and treasure hunt, will be largely directed toward conference goers, but many will be public facing. Events will include dance and theater performances, art exhibits, an escape room, and a real-time demonstration of Bose–Einstein condensates being synthesized aboard the International Space Station.

One of the global events will likely take place in Ghana, which, along with Mexico, played a key role in bringing

INTERNATIONAL YEAR OF
Quantum Science
and Technology

the IYQ proposal to the UN. Riche-Mike Wellington, Ghana's focal person for the IYQ, says that training workshops and conferences, public awareness campaigns, and other activities are being planned in partnership with industry,

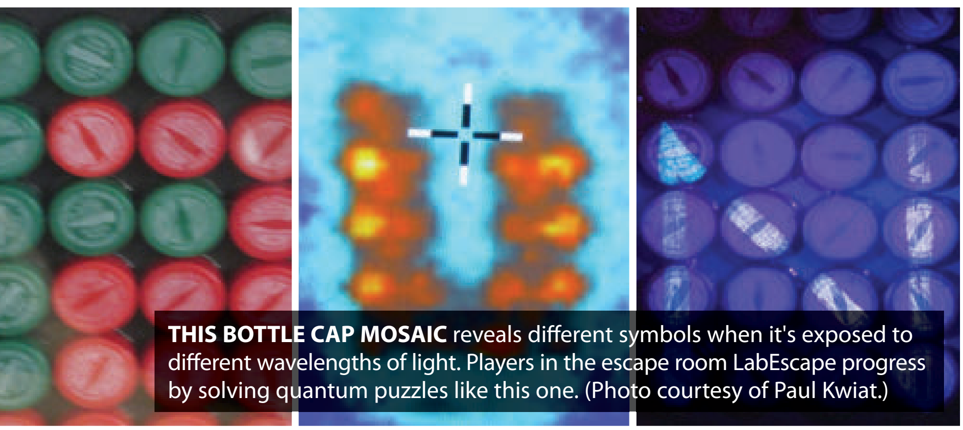


IN THE QUANTUM DIPLOMACY GAME, policymakers use role-playing to explore such issues as how to foster public–private partnerships and how to make their government quantum friendly. The game was created by the Geneva-based Open Quantum Institute. (Photo by Michael Chiribau, UNITAR Division for Multilateral Diplomacy.)

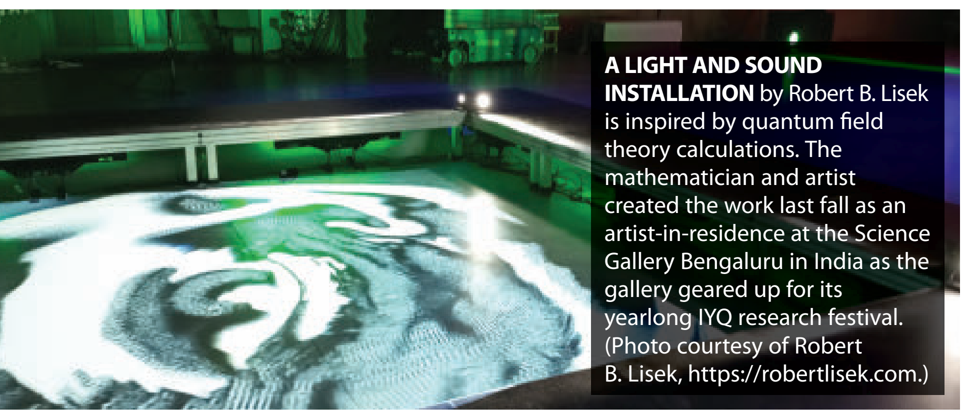
***PICOCOSMOS* (2024)**, a sculpture by artist Edy Fung, broadcasts sounds using data from a single-photon source at a Technical University of Berlin quantum communications lab. The sculpture was showcased in November by Studio Quantum and the Haus der Kulturen der Welt in Berlin at an event for artists and physicists that fostered discussions for IYQ events. (Photo courtesy of Edy Fung.)

THIS MOSAIC from the 1950s at the National Autonomous University of Mexico explores the past and present. The eastern wall (shown) portrays the contemporary world, with the atom taking center stage. (Photo by Miguel Zorrilla, General Directorate of Libraries and Digital Information Services, National Autonomous University of Mexico.)

IN THE MULTIMEDIA THEATER PIECE *The Art of Questionable Provenance*, scientists and artists explore the science of consciousness and the use of scientific forensics to analyze artwork. The University of Chicago's STAGE Center, which created and produced the show, is planning other projects for the IYQ. (Photo by Christopher Ash.)



THIS BOTTLE CAP MOSAIC reveals different symbols when it's exposed to different wavelengths of light. Players in the escape room LabEscape progress by solving quantum puzzles like this one. (Photo courtesy of Paul Kwiat.)



A LIGHT AND SOUND INSTALLATION by Robert B. Lisek is inspired by quantum field theory calculations. The mathematician and artist created the work last fall as an artist-in-residence at the Science Gallery Bengaluru in India as the gallery geared up for its yearlong IYQ research festival. (Photo courtesy of Robert B. Lisek, https://robertlisek.com.)

educators, and policymakers. The aim of IYQ activities, he says, is to "inspire future leaders and innovators in quantum science, driving economic growth and enhancing the quality of life for Ghanaians and Africans at large" and to bridge the "noticeable divide between the technologically rich North and the less-developed South."

Grassroots activities

In India, physics historian and museum director Jahnavi Phalkey is planning a yearlong quantum festival at Science Gallery Bengaluru. The preparations began last fall with a mathematician-artist who spent several weeks at the gallery creating quantum physics–inspired art. There will be installations, performances, and a beverage bar, called h-bar for Planck's constant. "The purpose is to create a sense of wonderment around quantum, not necessarily to explain it," says Phalkey. "It's to remind ourselves of the sheer beauty of what the mind is capable of."

People who have been involved in World Quantum Day, now in its fourth year, have a bit of a head start. The celebration has representatives in more than 60 countries. World Quantum Day is officially 14 April, but events take place on and around that date. Past activities have included explanatory video competitions for high school students, campaigns to translate "World Quantum Day" into many languages, museum talks that explore how quantum physics plays a role in people's day-to-day lives, and the creation of YouTube and other social media content.

Around the world, people at schools, museums, companies, and more are planning live and remote lectures, inviting students to intern in labs that do quantum-related research, hosting hackathons, and putting on events in which quantum science and art interact. If the UN-designated 2015 International Year of Light is anything to go by, expect upward of 13 000 events this year. Anyone can post an IYQ event or look up what's going on near them at https://quantum2025.org/en/event-resource.

Toni Feder PT

FROM THE ARCHIVES
AUGUST 2006

In 1979, **I. I. Rabi** (1898–1988) was the Emeritus University Professor of Physics at Columbia University in New York City. **Fraser Code** is an emeritus professor of physics at the University of Toronto in Canada.

Stories from the early days of quantum mechanics

Isidor Isaac Rabi
Transcribed and edited by R. Fraser Code

A colloquium delivered to the University of Toronto physics department on 5 April 1979 by the master of molecular beams offers a fresh look at an earlier era.

I have something in common with Ernest Rutherford, that distinguished physicist and professor at Canada's McGill University, who deplored the fact that, although a physicist, he got a Nobel Prize in chemistry. My career is the opposite. I started at Cornell as a chemist, and got a degree of bachelor of chemistry, which has since been discontinued. So I'm an orphan like the DeSoto, one of those cars that are no longer manufactured.

Anyway, after some years in which I tried various things that broadened my education but did not line my pocketbook, I went back to Cornell to study physical chemistry. But I'd taken all those courses so I said to myself "I'll study physics, and put the two together."

You know, that is somewhat like the person who wanted to study Chinese philosophy, so he looked up Chinese in the encyclopedia, and then he looked up philosophy, and finally tried to combine them.

But for me, when I started studying physics, I realized that the part of chemistry I liked was called physics. So that was the beginning of my career, and I entered the subject of physics more seriously around 1922.

Learning quantum mechanics in America

The year of 1922 was very significant. In fact, that whole time from the early twenties onward was a period of great ferment in physics, enormous ferment, all over the world—by which one means Denmark, England, France, but not the United States.

I remember one time when I was a graduate student at Cornell, sitting in the library amongst the students, just before the time when Professor [Arnold] Sommerfeld was to come and visit. And you could see one professor after another sneak in and take a look at Sommerfeld's book *Atombau und Spektrallinien* (*Atomic Structure and Spectral Lines*, Friedr. Vieweg & Sohn, 1919). That was all the exposure they had to the quantum theory. That was 1922 in America. By contrast, in Europe, quantum theory had been extant for quite a number of years. But in America, it had not yet achieved full recognition as something suitable for graduate study at Cornell, or for that matter at Columbia [where Rabi completed his PhD]. I'm not even sure that quantum theory was working very well here at Toronto in 1922!

Anyway, the faculty in America wasn't very much concerned with quantum physics, except experimentally. But at Columbia, a number of graduate students formed a weekly discussion group that we called a "Sunday soviet," by which I mean that we met every Sunday near 11 o'clock in the morning, and went on right through a Chinese dinner.

We learned a great deal just by ourselves. I'd recommend this method of learning to all the graduate students in this audience: If any of the faculty are deficient in some subject that interests you, just form a little soviet and do it on your own. As a matter of fact, it worked so well that when the Austrian physicist Erwin Schrödinger's paper first came out,[1] we read it and worked through all the equations.

Then, just as an exercise, Ralph Kronig and I decided to


COLUMBIA UNIVERSITY

Michael Pupin (1858–1935) at Columbia University, probably in the late 1920s. Pupin and I. I. Rabi were part of a small group at Columbia that was trying to figure out quantum mechanics in 1926. (Courtesy of AIP Emilio Segrè Visual Archives.)

do something with this new thing, Schrödinger's quantum theory. So we looked through [Max] Born's book[2] and found that the symmetrical top problem had not yet been done. So we sat down and, according to Schrödinger's prescription, formed the wave equation, separated the variables, got the angular momentum, as well as the various states, but then we ran into an equation that we didn't know how to solve.

And here's another lesson that I want you to hear from my own experience. Somehow or other after that Sunday soviet, I was sitting in the library reading the mathematical works of Carl Jacobi, who wrote beautifully in German. I understand that German is no longer required for graduate students here. Too bad, because in reading through that book, suddenly there appeared my equation—the one Kronig and I could not solve. It was the equation for the confluent hypergeometric series, which neither of us had ever heard of before. Using this reference, we were then able to solve the quantum mechanical problem of the symmetrical top molecule.[3]

But we did not have the faintest idea what the wavefunction ψ meant. It was a magical thing. What you got when you followed this prescription, as Schrödinger had done for the hydrogen atom, were the eigenvalues of the differential equation. These were the energy levels, which agreed with experiment. But we had no idea what the wavefunction was—what was this magic function ψ?

Of course, it became clear soon thereafter when Born[4] and others suggested that $|\psi|^2$, the absolute value of ψ squared, represented the probability density for finding that particular thing at that particular place. Suddenly the wave function ψ acquired a great meaning.

But it was so magical, that function ψ. You simply followed the formula, and out came real results. This was not a surprise. During the first period of its existence, quantum mechanics didn't predict anything that wasn't also predicted before by the old quantum mechanics plus that very magical abracadabra of the correspondence principle.

There were real artists at work on the correspondence principle. For example, they were able to deduce many things from the Kramers–Kronig formula, or from the Kramers–Heisenberg dispersion formula. The development of physical relationships from the correspondence principle was all done by artistry, by imagination, and from certain kinds of symmetry ideas. So the results that came out of quantum mechanics had to a large degree been previously anticipated from this correspondence principle.

But a very unfortunate thing happened to John Van Vleck, who wrote a remarkable book on the old quantum theory.[5] It was a wonderful book, a clear book, and he was a master. However, it was published and came out just at the time of the revolution in quantum mechanics. Unfortunately, it became obsolete almost on publication! The same was true with Wolfgang Pauli's first volume. When the revolution came, it all changed.

Now, it was the new quantum mechanics that was doing things and growing. Matrix mechanics, of course, was in many ways clearer, and in many ways more dense than Schrödinger's equation. But the matrix mechanics of Heisenberg used a different kind of mathematics.

Paul Dirac had been an engineer with a background in mechanics, rather than having been a physicist. So when he followed Heisenberg's first paper on matrix mechanics, he particularly noticed the commutation exchange relationships, and saw a certain parallel between Poisson brackets and the commutation exchange relationships. As a result, Dirac started his approach to matrix mechanics from that direction.

So that was a very great time because we could be the first to do something like the symmetrical top. And we were the first to do this important molecular problem, and just as graduate students! It was not for my dissertation, nor was it for Kronig's, but we did learn some quantum theory. While I was a graduate student at Columbia, there were no professors of theoretical physics. I was doing an experimental dissertation, and my supervisor was Professor Albert P. Wills.

In 1926, there was just our little group of serious thinkers, including Michael Pupin, sitting there trying to figure out Heisenberg's matrix mechanics. Schrödinger's formulation, of course, was our favorite. This was clear. It only required that you were familiar with differential equations, and it had a pictorial interpretation. In contrast, Heisenberg's approach involved matrices, which were not difficult but were messy. In addition, there was Heisenberg's use of abstract symbolism, which, of course, looked to us as the most mysterious of all.

And this shows how limited one can be if one is provincial. Because in the United States, as far as theoretical physics was concerned, we were provincial. Definitely provincial.

Visiting quantum physicists in Europe

So the time came when I had finished my dissertation.[6] But there were no jobs around in the US, so I got a small Barnard fellowship to go to Europe. It was $1500 a year for two years, not paying for transportation. And on this my wife and I went to Europe. Well of course, being an American, in many ways I was very naive. The first place I went was to Zürich, Switzerland, where I hoped to work with Professor Schrödinger.

Of course, I hadn't written a note beforehand to make arrangements to come. When I arrived in Zürich, I tried to find a pension [boarding house or small hotel] where I could stay. Afterwards, I went right down to the university, where there was a colloquium going on that afternoon. The man gave a fiery lecture, and I didn't understand a single word. I was very depressed, and I came out full of sorrow for what was going to happen to me. Here I had come all the way over to Europe from America, and now I felt very discouraged. So I looked around in the audience for somebody that I might know.

Well, I did find people in a very definite way. In 1927, the Russian revolution was about 10 years old. And Americans always wore white shirts, but with their collars attached. You could recognize an American anywhere that way. I looked around, and there at the colloquium was a man with a white shirt and collar attached.

He turned out to be Linus Pauling. I told him of my sorrow that I didn't understand what the lecturer was saying. He said "Don't worry, he was not talking German, he was talking Schweizerdeutsch," which was the local German dialect. I was very pleased to hear that. Later, Linus invited me to where he was staying and gave me a drink. I don't suppose you realize what this meant: In 1927, Prohibition was on in America and drink was a rare thing, especially when you had no money. He also recommended a good pension for me to stay at.

Well, the timing of my trip to Europe was not very good. I had just arrived in Zürich to visit with Schrödinger, and then Schrödinger left almost the same day. He'd gotten a good job in Berlin. But I was traveling lightly, except for a very heavy suitcase. So I went down to Munich to visit Sommerfeld. I arrived there, and just as I did in all these places, I came in and said, "My name is Rabi. I've come here to work." I hadn't written anything beforehand.

So there it was—Sommerfeld's office in Munich! I was shown to a room where some of his students worked, and there were Hans Bethe and Rudolf Peierls, who were graduate students at that time, and Albrecht Unsöld, who later became a well-known astrophysicist—that is, a theoretical astronomer. There were also two Americans who became very notable later. One was Edward Condon. You know the book, *The Theory of Atomic Spectra*, that he wrote with George Shortley (Cambridge U. Press, 1935), as well as Condon's other books. The other American was Howard P. Robertson, who was very well known in circles that deal with relativity. So we were the three Americans in Sommerfeld's group, who gave each other strength because we were worried that our German was not of the best quality. Every once in a while, Peierls and Bethe would go out in the hall and laugh, and we did have the suspicion that they were laughing at us.

Anyway, in the Germany of 1927, the working conditions for graduate students were very interesting in a way when compared to now. Once, Sommerfeld showed me around his offices. In the basement was one place where there was a closet with a board across, and a naked incandescent bulb over it. Right there was where Bethe worked. So there was nothing very much in the way of conveniences. I think there were only three graduate students actually working with Sommerfeld. But you can see their character somehow by their selection. Two of those three were Peierls and Bethe. I don't remember the third one.

Sommerfeld was a man with enormous dignity, a wonderful person. I was invited on Friday afternoons to the Englischer Garten to have tea with the Geheimrat [an honorary German title conferred on outstanding scientists]. It was very dignified.

Sommerfeld had a very large office, and then there was the office of his assistant, a man named Becker, and finally the place for his students. All the journals were in Sommerfeld's office. So if you wanted to look up something, you made your way to the assistant, who would then knock on the door of the Geheimrat, and then you walked in. Under those circumstances, you didn't look things up very much.

I am telling you these stories to show another way of life, which existed at that time, and to contrast it in a way from the one we have now. Of course, I don't know how it is since I finished working [in 1967]. For example, I don't know whether you need clearance [the need to make prior arrangements] at all to go from one place to another to work. I don't know whether you could come in and say, as a fresh-corked postdoc could say, "My name is Rabi. I've come to work here." The answer would probably be, "Who said your name *isn't* Rabi?" Well, it was a wonderful way to live, in a place like Germany. And as an American, you weren't part of it. You never expected to get a job there, so you were free.

In the fall, I left Munich intending to go first to England and then to Copenhagen. In England, I discovered that six marks—equivalent to six shillings—which carried me through the day in Germany, wouldn't quite give me a room in London. I saw financial disaster staring me in the face. So I went to Copenhagen.

Copenhagen, of course, was the mecca for everybody at that time who was interested in theoretical physics. Everything good came out of Copenhagen in one way or another. And so my wife and I went off. When we arrived in Copenhagen, I checked my bag, and my wife and I took our map and walked

Yoshio Nishina and Rabi in 1948. The two men wrote a paper together as part of Wolfgang Pauli's group in Hamburg, Germany, in 1927. (Courtesy of AIP Emilio Segrè Visual Archives.)

over to the Institute for Theoretical Physics [renamed the Niels Bohr Institute in 1965]. I rang the bell and said my usual spiel: "My name is Rabi. I've come to work." So the Institute's secretary gave me a key. I asked her for a suggestion on where we might stay, and she gave us a good one. I brought my wife and my bag there, and then came back.

This was September—a month of complete holiday. There was nobody around except the secretary and me. But there was something about Copenhagen that was in its walls, somehow or other. You couldn't be idle there. You just had to sit there and work, and try to think great thoughts. I recommend that you try it. It can be very frustrating.

In the course of time, several people were to appear. There was one gentleman with an enormous stutter. He tried to tell me his name, and I tried to help. And I said "Klein, Klein," as I knew Oskar Klein was Bohr's assistant, but when he came up with his name, it was Pascual Jordan, who later on became a professor and lecturer. And how he ever did it I don't know, except that he did not have this stutter when he had enough beer in him, or when he spoke English.

Then, after a while, others showed up: great names in physics like Ivar Waller, Kronig (who had been there before me), and finally the great Professor Bohr came back from his vacation.

My arrival in Hamburg

And now I come to the beginning of the real story of my life, that is, the direction of my life. Bohr had had a very difficult summer, and his assistants thought that he had been overworked and that he should not have any people there except for Kronig, who had come earlier.

And here again a most fortunate thing happened. Without asking me, but making all the arrangements, they arranged for Yoshio Nishina and me to go to work with Pauli in Hamburg. This seemed disappointing at first, to go away from the center to a place like the University of Hamburg. But Hamburg actually was the greatest institution in the world for physics at that moment. Hamburg had Pauli; Walter Gordon [of the Klein–Gordon equation]; Wilhelm Lenz, who was in molecular theory, a brilliant man; and most of all, Otto Stern, in experiment. So there quite by accident, and partly against my will, I found myself in this very marvelous place. In addition, there was Ronald Fraser from Scotland, and John Taylor, who was an American. They had both done molecular beams before, and were working now with Stern. Pauli at that time, and this is toward the end of 1927, asked Nishina and me to write a paper with him.[7]

I became aware of the necessity for me to talk some English. This was a real physical necessity. The three of us English-speaking people there—Fraser and Taylor and I—formed a little group that I crowned "the three for we who were abroad." No matter what, you had to express yourself, and for me this was only possible in English. Shortly, I left Pauli's group. I had an idea about how to do an interesting experiment concerning the magnetic refraction of molecular beams and was invited by Otto Stern to do it in his laboratory at Hamburg.[8]

Remember, back at Columbia I said we were provincial. To show you the degree to which we were provincial—and by "we" I am talking about the United States, that land south of the Canadian border—in Germany they subscribed to the *Physical Review*, but waited until the end of the year to get their 12 issues at once, to save postage. It wasn't important enough to get each issue right away.

We—and here I mean Condon, Robertson, and others among my friends—felt that this was very humiliating and vowed we would change it. I must say that we did, because 10 years later the *Physical Review* was the leading journal in the world. It didn't take long. We came back and distributed ourselves among our various universities and began teaching students.

Teaching was just like raising fish—there were a lot of eggs, which we began to fertilize. And so we had this time bomb of emerging physicists. In America, we had numerous colleges and universities, the students were there, and they needed teachers. And we came back from Germany with the magic of quantum theory. Indeed, by the time World War II came, physicists could man all of the American research laboratories. We were able to recruit hundreds or thousands of people, people with a very sophisticated educational background. So it [the conversion of American physics from the provincial to the international] could be done.

And this is what frightened me so about the Russians when the first Sputnik was launched. I thought they were on to this trick of raising fish. But you can't do it unless you have a free society. This was done freely by the people themselves and was done without government support. There was no government money for physics before the war. But I'm getting ahead of my story.

The magical role of experiment

And now I begin the experimental part of my talk. It is about those great days, and how people saw marvelous things and didn't understand them.

It is well known that Stern and [Walther] Gerlach did a famous experiment that was intended to demonstrate space quantization. They passed a beam of silver atoms through an inhomogeneous magnetic field. When silver was evaporated, the atoms were supposed to have magnetic moments, which could be deflected by external magnetic field gradients. Since the atomic beam of silver had a Maxwell distribution of velocities, the beam would be deflected and broadened by the field gradients. Some would be deflected one way depending on their orientation, some the other way, and some not at all, if their orientation was perpendicular to the magnetic field.

Stern and Gerlach had a brilliant concept, and with very poor equipment they did the experiment. (See the article "Stern and Gerlach: How a bad cigar helped reorient atomic physics," Physics Today, December 2003, page 53.) And the experiment, as most of you have seen in elementary books, showed a split beam, plus and minus; some were deflected one way, some were deflected the other way. But what about the middle? What about the atoms that were perpendicular? [Rabi now refers to the old Bohr–Sommerfeld theory, in which ground-state silver had an erroneous orbital angular momentum ($L = 1$) and the electron's spin and g factor were yet to be discovered.] And the story at that time was that you assigned quantum numbers m_L that were equal to plus one, minus one, and zero. What about zero? There was no zero! Instead of that fact creating an enormous sensation, they just said, "Well, m_L equal to zero is missing," which was a great statement at that time, and nobody understood it.

Since there was no logical theory available, you could play it by ear; it seemed obvious that the zero state was missing. And to support the argument, they appealed to the theory of the Stark effect, in which the $m_L = 0$ orbit should hit the nucleus. So they said, "We can't have it hitting the nucleus, so we can say that the $m_L = 0$ quantum number is missing—you just don't have it." Now you begin to see why this strange experimental result was so useful. You didn't have to resort to these odd forms of chicanery about why the $m_L = 0$ state was missing. The whole point of the experiment was that they had seen atomic silver to have spin equal to one-half, and its orientation was either one way or the other. So it was right there in front of them, and because they had been so accustomed to glib talk, they didn't recognize it.

At that time, Stern was also doing experiments to show the wave nature of matter. First, he was scattering hydrogen atoms with a ruled surface, and then he successfully used another type of lattice. He showed that the scattering was associated with the de Broglie wavelength—not only for atoms, but also for molecules.

Now a molecule is not an atom, at least if you go back to the unsophisticated days. Once you have a de Broglie wavelength for a molecule with only two atoms, then why shouldn't a grand piano have a de Broglie wavelength? Any collection of things should scatter in this way. In fact, these scattering experiments were really demonstrating the wave nature of matter. Not just electron scattering, or even atomic scattering, but also molecular scattering was consistent with the same de Broglie relationship.

Later on [in 1933], pursuing the same idea, Stern and his collaborators measured the magnetic moment of the proton. This was done against the strong advice of his friend Pauli, among other theorists. They all said, "We know the moment of the proton, because we know the difference in mass between the proton and the electron, and we know the magnetic moment of the electron." Stern went ahead and did the experiment anyway, and, of course, all of those theorists were wrong.

Will physics ever come to an end?

I'm coming to the end of my talk, and I just want to tell you one more small story. I could go on telling stories, as you see, for a long, long time. But this is one story that you should take to heart.

I went with my mentor, Otto Stern, to visit the great Max Born, who was then at the very height of his glory, with his probabilistic interpretation of the wavefunction and so on. At that meeting, he told us very seriously that in six months' time, physics as we knew it would be over.

That was quite a blow! Born had an impressive personality, and he said this with a certain amount of reason because it was 1928, and Dirac had just given us his miraculous theory of the electron.[9] Making no assumptions other than relativistic invariance, Dirac derived the correct spin and magnetic moment of the electron. Everything that one wanted to know about the electron came without any extra assumptions beyond relativistic invariance. So this was a terrific achievement, of course. And Born apparently felt that it wouldn't take more than six months for these very bright boys around him to derive the spin and moment of the proton from a similar theory, and then it would be all over. As he explained, there would be a lot to do, of course, but physics as we knew it—more or less groping blindly around in our optimistic way, that portion of physics—would be behind us.

Well, I found Born's prediction very hard to believe. In fact, I couldn't actually let myself believe it. At my stage in life, I had far too much at stake. On the other hand, you will hear and see such predictions again as your careers develop. Most probably this will be particularly true for the graduate students and young people in the audience, because at every past period of synthesis in physics, the future looked closed.

In Newtonian times, physics was a closed book. There were central gravitational forces, and equations describing what they could do. People tried to come up with solutions to these equations, but some types of problems led them to invent other forces. And of course, along came Maxwell's theory of electromagnetism—all very beautiful, set, done, and apparently closed. But occasionally Nature does something strange, such as the photoelectric effect, which appeared just at the peak, the very triumph, of the Maxwell theory. It was uncovered first by accident during Heinrich Hertz's experiments on the detection of electromagnetic waves,[10] but he missed its significance and was unable to explain it. And so I have come to think that physics is a never-ending quest.

In closing, there is one other mystical thought that occurs to me. Now, in a day when we need all this big equipment for physics experiments, such as those vast accelerators that we have, I began to think: Will God reveal himself only to rich people? Would it really be true that you had to have a very wealthy country with a large population in order to get some basic information about how the universe is made? At this point I am a mystic, and I don't believe that only the rich and powerful can achieve true understanding. And I suppose it is up to you to prove me right.

Thank you. And I love questions.

Another view of things

One thing that I learned contains a tremendous amount of anthropology in just one sentence. One of Otto Stern's assistants was a man by the name of Fritz Knauer. One time I was telling Knauer that in my country you could travel from one place to another and you didn't have to register with the police—you just traveled freely. Knauer looked shocked at this, and he said to me, "You mean to say that you can live and die in America, and nobody cares?"

Now that may sound very funny to you, but it shows the other end of the telescope. Something that I thought was an awful imposition—registering with the police—was to him a great support. It takes quite a bit of training to live in a democratic country like America, it takes a lot of training indeed. Some people who came to America, such as Russian refugees, have been shocked to learn that they have to find a job by themselves.

Discussion

Jan van Kranendonk: A very down-to-earth question, perhaps. When you worked with Otto Stern, from what funds were the experimental apparatus supplied? How was this research work funded?

Rabi: That's a very good question. There was something, I think, called "der Notgemeinschaft der Deutschen Wissenschaft." Somebody might properly translate this, but it's the Society of Need for German Science, which got some money for grants, but I don't know whether it came from rich people or from the government. But the greater part of researchers' money, at least in some cases, came naturally from America. Didn't we beat the Germans in 1918? And now we had to pay!

The Rockefeller Foundation, and other foundations, supported students—people like Felix Bloch and Edward Teller. Many other people applied for and got Rockefeller fellowships and grants. They had equipment in the laboratories at Hamburg that we certainly didn't have at Columbia—and it was funded by American money. And very wisely, the Rockefeller Foundation was interested in getting good research and the best science for its money. And that was to be found in Germany at that time. That's where they spent it.

My eyes boggled when I saw all the equipment they had in Hamburg that I couldn't get in America. There were special kinds of vacuum pumps and other things. They had pumps which would cost $200 or $300, which was an enormous sum then. But when I came home and started doing research, I had to get pumps for $8. So you can see how research in Germany was funded: There was an enormous respect in the United States for German science, and an enormous feeling of inferiority for American science.

I think, as [J. Robert] Oppenheimer once expressed it, "We went to Germany, so to speak, on our hands and knees." But it took only a very short time, in the post–World War II period, for the whole flow to be reversed. In 1926 you couldn't get anywhere with English in Germany, because they didn't know any. I remember how surprised one German was to hear another German speak English. And if you wanted your research to be recognized, you would publish either in German or in the British journal *Nature*.

And you can compare that with today; English has almost become a universal language. But I would like to warn you: From 1927, the year that I was talking about, to 1937 or the beginning of the 1940s was only about 10 years, and during that time there was a reversal. And some of you who are very proud of not knowing any other language but English have got to learn some foreign languages. One other point about that: I know at Columbia they have also abolished the language requirements for the PhD. This is an enormous mistake.

If you want to read the originals of many important physics papers from the earlier part of the 20th century and most of the previous century, you won't be able to read them in English. Most of these original papers have not been translated into English, and you don't get the flavor of the original papers from textbooks. So I would suggest you take that very seriously to heart and learn some other languages. I don't know which, it's your guess . . . maybe Dutch [said with a kind smile toward van Kranendonk, referring to his slight accent].

Question: Could you elaborate further on how it was that you could appear, apparently unannounced, to work at the institute that you spoke about, and they knew that you would be acceptable? Is that what you intended to say?

Rabi: I was intending to show another period of time, when the world was simpler, and despite the first great World War, it still had that simplicity. A scholar could roam around and be accepted where he went. I didn't mean to put this to the test. But being a romantic, and an American, it didn't seem to me necessary to prearrange things. I mean that this favorable reception didn't surprise me. I just thought it was normal.

It is only when I look back on that time, especially with modern terms in mind, that I am surprised that nobody asked who funded me. At Hamburg, I had an idea for an experiment and I was invited to do it, and so I did it. But nobody asked me, "Are you funded?" No one at all. They gave me the equipment, and space, and so on. I had a marvelous time doing it.

We showed the Germans something that we called the "Amerikanische Arbeitsmethode," the American way of working. Usually the laboratory was opened strictly at 7am and then closed at 7pm—it was all so very un-American. We would come at 10am, and then, around 11 o'clock, the wives would come and make toast, crumpets, and so on while we went on doing our physics experiment. And we finished in very good time. It really worked. Also we were very happy while doing it. We'd have requests from the top floor of the building, "Would you please sing more quietly?" So it wasn't a time when you gritted your teeth and did an experiment. It was a joy all the time. That's the only way to do physics, I think.

Van Kranendonk: Perhaps I can ask a different question. You said that you were associated with Pauli, and I know that Pauli had a big reputation for being quite vicious. How did you find him? How did you like him and interact with him? Did you understand how he was when he worked?

Rabi: I have seen him being extremely vicious, as you say. I think I got along with him very well, but it was a result of a mistake that I made. Right after I came to Hamburg, I told him about some calculations I was making on the hydrogen molecule. And we had a misunderstanding between the Roman letter p and the Greek letter π [the latter is pronounced "pea" in both German and Greek]. When Pauli said "pea," I though he meant the Roman letter p [momentum], but he meant the number π. And so I said, my German being pretty poor, "Aber das ist Unsinn!" (That's nonsense!)

Nobody ever said that to Pauli. He rolled around and he said "Um . . . ist das Unsinn?" Somehow I had gotten in the first blow! But, you know, I was so upset by the way he did talk to people, until I saw that he was completely democratic—he talked the same way to Bohr. This was just Pauli's character, it was just Pauli's own way.

There was something called the "Pauli effect," which states that wherever Pauli went, misfortune followed. Not for Pauli, but for others.

Pauli had visited the astronomical observatory in Hamburg. The astronomers talked to him and then forgot about what they were doing, so the telescope hit the dome. Pauli caused things of that sort to happen. Stern would never let him into the laboratory. They were good friends, and Pauli would knock on the door and would usually want to borrow some money, and they would make their transaction right at the door.

I saw one of the most remarkable examples of the Pauli effect at a Physical Society meeting in Leipzig. News had come from America about the invention of talking pictures, and this local professor, I forget his name, was going to give a demonstration of them. The equipment was all set up, and when the assistant threw the switch . . . bang! bang! bang! came out of the loudspeaker, and then smoke. Pauli was beside himself. He shouted out, "My effect!" And they brought up another projector, and the same thing happened. Then they had a third one set up in a balcony above, where I suppose they used to have music of some sort. They connected that projector, and it worked, which showed the relationship between distance and the Pauli effect.

But the real explanation was given by Paul Ehrenfest. You see, Pauli was born in 1900, the beginning of the 20th century, which was just an illustration of the fact that misfortunes could never come up singly. The 20th century has been a terrible century. In terms of Pauli, misfortunes never did come singly.

Derek York: Do you know anything more about why Sommerfeld never received the Nobel Prize? If so, is there any inside story on this?

Rabi: I haven't heard any inside story about it, and I don't think anybody would have raised any objection if he had been given the prize. But you must remember that the Nobel Prize is given by a committee of the Swedish Academy, and they have their own idiosyncrasies. You know, there was a book published some 25 years ago about the various Nobel awards. It discussed many things, for example, about why didn't Dmitri Mendeleev get the Nobel Prize. It suggested some mistakes of the committee of the Swedish Academy. They were very human.

When the Nobel Prize was established, the choice of the awards was up to the Royal Swedish Academy, and they had very sincere doubts that they had the capacity to make such judgments. They felt they didn't have enough members that were *au courant* enough and mature enough to make good judgments. I must say that their early judgments were terrible. But they gave it to Albert Michelson, and they gave it to Pieter Zeeman. They really had a tremendous field to choose from, and I think that is what established the Nobel Prize with such prestige. In addition, the Nobel Prize is presented by the king and queen in royal fashion. All the Nobel recipients are able to live for a few days in the manner to which they would like to become accustomed.

Van Kranendonk: Well, perhaps on this note we should end, and may I then ask you to join me in thanking Professor Rabi for his visit, for his talk. And let's send him our best wishes.

REFERENCES

1. E. Schrödinger, *Ann. Phys. (Leipzig), Ser. 4* **79**, 361 (1926).
2. M. Born, *Vorlesungen über Atommechanik* (*The Mechanics of the Atom*), Julius Springer (1925).
3. R. de L. Kronig, I. I. Rabi, *Phys. Rev.* **29**, 262 (1927).
4. M. Born, *Z. Phys.* **38**, 803 (1926).
5. J. H. Van Vleck, *Quantum Principles and Line Spectra*, National Research Council (1926).
6. I. I. Rabi, *Phys. Rev.* **29**, 174 (1927).
7. Y. Nishina, I. I. Rabi, *Verh. Dtsch. Phys. Ges.* **9**, 6 (1928).
8. I. I. Rabi, *Nature* **123**, 163 (1929); *Z. Phys.* **54**, 190 (1929).
9. P. A. M. Dirac, *Proc. R. Soc. London A* **117**, 610 (1928); **118**, 351 (1928).
10. H. Hertz, *Ann. Phys. (Leipzig), Ser. 3* **33**, 983 (1887). PT

The online version of this article is linked to the complete and unedited transcript, which has considerably more material in it.

"There is one thing I would be glad to ask you. When a mathematician engaged in investigating physical actions and results has arrived at his conclusions, may they not be expressed in common language as fully, clearly, and definitely as in mathematical formulae? If so, would it not be a great boon to such as I to express them so?—translating them out of their hieroglyphics, that we might also work upon them by experiment."

From a letter of Faraday to Maxwell, 1857, quoted by Sir Lawrence Bragg, *Nature* **169**, 684 (1952).

QUANTUM

By ***F. J. Dyson***

HISTORICAL PARALLELS are never exact. Each development in science is something new and different from any which preceded it. Still it may be illuminating to discuss the progress that has recently been made in quantum electrodynamics, using the historical development of classical electrodynamics as a standard of comparison. So may we see our present knowledge and our present difficulties in their proper perspective. If Faraday's appeal quoted above had been more effectively answered in his day, might not electromagnetic waves have been discovered less than thirty years later? We cannot answer such a hypothetical question. But every theoretical physicist who reads Faraday's words will be uncomfortably aware that similar appeals are still being made and are still not being answered. This article attempts to express in simple words the results of our recent thinking in quantum electrodynamics, not fully, but clearly and definitely so far as that is possible.

First the meaning and scope of quantum electrodynamics must be defined. In our present state of ignorance we find it necessary to separate our ideas about the physical world into three compartments. In the first compartment we put our knowledge of nuclear structure, protons, neutrons, mesons, neutrinos, and the interactions of these particles with one another. In the second compartment we put theories of the large-scale structure and geometry of the universe, including Einstein's general theory of gravitation. In the third compartment we put our knowledge of all other phenomena, everything intermediate in scale between an atomic nucleus and a massive star. The third compartment includes the whole of classical mechanics, optics and electrodynamics, special relativity and extra-nuclear atomic physics. The convenience of these compartments is that they enable us to isolate the areas of our ignorance. The first two compartments are full of undigested experimental information, empirical rules, and mutually contradictory assumptions. These fields are only beginning to be explored and organized. On the other hand, the third compartment is unified by a logically consistent theory. We possess a set of mathematical equations which agree quantitatively, so far as is known, with all the wealth of accurate experimental data in this field. The equations consist of laws of motion for electrons, positrons, photons, and electromagnetic fields, incorporating the principles of quantum mechanics and of special relativity. This theory of the third compartment is what we mean by quantum electrodynamics.

QUANTUM ELECTRODYNAMICS occupies a unique position in contemporary physics. It is the only part of our science which has been completely reduced to a set of precise equations. It is the only field in which we can choose a hypothetical experiment and predict the result to five places of decimals, confident that the theory takes into account all the factors that are involved. Quantum electrodynamics gives us a complete description of what an electron does; therefore in a certain sense it gives us an understanding of what an electron is. It is only in quantum electrodynamics that our knowledge is so exact that we can feel we have some grasp of the nature of an elementary particle. That is the reason why theoretical physicists for the last thirty years have concen-

Freeman J. Dyson, theoretical physicist at Cornell University's Laboratory of Nuclear Studies, is one of the numerically small group of theorists who have contributed so heavily during the past few years to the mathematical development of quantum electrodynamics.

ELECTRODYNAMICS

trated their efforts so persistently on the electron. We must expect that the concepts, to which we have been led in our study of the electron, will later find their natural place in a more extended theory of elementary particles. Without these concepts and their mathematical expression in quantum electrodynamics, speculations concerning the nature of elementary particles would be mere guess-work.

The basic equations of quantum electrodynamics were formulated by Heisenberg, Pauli, and Dirac during the period from 1927 to 1929. Historically, they were the Maxwells of the new science. Just as Maxwell's equations in the thirty years after their discovery were triumphantly verified in experiment after experiment, so the equations of Heisenberg-Pauli-Dirac were tested during the 1930s and were found to give a correct account of all phenomena at that time accessible to exact measurement. In particular, all the complicated details of atomic spectra, and also the spectacular process of cascade multiplication of electrons and positrons observed in high-energy cosmic-ray showers, were shown to be in agreement with the theory.

Without stretching our analogies unduly, the historical parallelism between the development of classical and quantum electrodynamics can be pushed a great deal further, so as to include the events of the present day. After its initial successes, the Maxwell theory was found to have a perplexing feature. It predicted that the results of experiments should depend on the absolute velocity of the measuring instruments through space, the space being filled with an ether which provided an absolute frame of reference. It was one of the central features of Newtonian mechanics, on which Newton himself laid much stress, that no such observable effects of absolute velocities could exist. Thus the Maxwell theory, while not inconsistent with Newtonian mechanics, implied the abandonment of one of Newton's most cherished principles. Fortunately for Maxwell, the predicted effects of absolute velocity on measurable quantities were always of the order of the square of the ratio of the velocity to the velocity of light, and therefore too small to be detectable during his lifetime. So long as this was the case, it was possible to hold either of two opinions concerning these effects; either the effects would in time be discovered and the Newtonian principle would be disproved, or the effects would be shown to be absent and Maxwell's theory would have to be modified. Meanwhile, until the decision became experimentally possible, physicists could continue happily to believe in both Maxwell's and Newton's principles.

A STRANGELY SIMILAR evolution of ideas took place in quantum electrodynamics in the 1930s. It was early realized that the electromagnetic field around an electron carried with it energy, and that this energy possessed mass and inertia by virtue of Einstein's law of equivalence of mass and energy. The motion of an electron should thus be affected by some kind of dragging force resulting from the inertia of its own field. The effect of such a force[1] on the electron's motion

1. Strictly speaking, a reaction force is produced both by the field which the electron radiates away into space and by the field which the electron carries around and does not radiate. We use the words "field reaction" here to mean only the second of these two forces.

was called the "field reaction." As the theory was developed, two things gradually became clear. On the one hand, if calculations were made consistently ignoring the field reaction wherever it appeared, the results agreed perfectly with the experiments. On the other hand, when calculations including the field reaction were attempted, the results were always meaningless; the inertia of the electron's self-field turned out to be infinite and therefore the electron was predicted to behave like a particle of infinite mass. The physicists of that period were simply baffled by the situation. They had a theory which by every experimental test was shown to be correct. Yet its success depended on excluding from consideration the field reaction force, and excluding this force came close to denying the validity of Newton's law of the equality of action and reaction. If the electron can set up stresses in the electromagnetic field around it, how can these stresses be prevented from reacting back upon the motion of the electron?

Physicists were agreed upon one point. The experiments showed that the field reaction, if it existed, was too small to be detected by the techniques of that period. Trying to understand this fact, physicists split into two main opinions. One group held that the basic equations of the theory were correct, and that only the method of making calculations needed to be changed, so that the infinite reaction forces would be automatically omitted. The other group held that the basic equations of the theory should be modified in various ways so as to make the reaction forces finite. According to the first group the measured reaction force should be strictly zero; according to the second group it should be not zero but small. Neither group succeeded in making their arguments convincing; neither group had any physical model by which to justify their recommended procedures. Lacking an experimental test of these hypotheses, the majority of physicists continued to believe both in the general correctness of quantum electrodynamics and in the law of action and reaction. This unsatisfactory state of affairs persisted until the summer of 1947.

Both for the Maxwell theory and for quantum electrodynamics, the choice between contradictory alternatives was finally forced on theoretical physicists by a decisive experiment. The Michelson–Morley experiment of 1887 showed that Maxwell's predicted effects of the absolute velocity of measuring apparatus on the results of observations were nonexistent. The Lamb–Retherford experiment of 1947, a precise measurement of the fine-structure of the atomic hydrogen spectrum using the new technique of radio-frequency spectroscopy, showed that the field reaction force on an electron existed and produced a finite measurable displacement of the spectral lines. The physicists of the 1890s were thus faced with the necessity of reformulating the Maxwell theory, and those of the 1940s with the problem of reformulating quantum electrodynamics. In both cases, it was the experimental knowledge of what the results of the new theory ought to be which stimulated the efforts of the theorists and made a successful outcome possible.

It was Lorentz who created the new classical electrodynamics. The new theory was in fact not a departure from the Maxwell theory. It was a reinterpretation of the Maxwell theory, taking into account the fact that the electrical and mechanical properties of measuring instruments are not experimentally separable. In particular, the length of a solid object such as a measuring rod is determined by electrical forces between its constituent atoms, and other mechanical properties are in a similar way mixed up with electromagnetic effects. Lorentz observed that in any experiment in which the electrical effects of absolute motion through the ether should be detectable, there would also be effects of the same order of magnitude arising from effects of the motion on the mechanical properties of the apparatus. These mechanical effects would have to be included in any complete theory of the Michelson–Morley experiment. In particular, the effect of an "ether-wind" blowing lengthwise through the atoms of a measuring rod would be to diminish the length of the rod by a definite factor depending on the velocity. This special effect is called the "FitzGerald contraction" in honor of the man who first suggested it in 1893. Lorentz found that when all these effects of absolute velocity, electrical and mechanical, were taken correctly into account, they cancelled each other out exactly. The result of any measurement in any possible experiment would be independent of the absolute velocity, in agreement with the experience of Michelson and Morley. By reinterpreting the Maxwell theory in this way, Lorentz preserved the Newtonian principle of the unobservability of absolute velocities. This principle appeared in his theory as something of a miracle; the theory started with a real ether having a definite velocity relative to the measuring instruments; only at the end after long calculations it turned out that the ether velocity had no effect on the instruments' readings. Lorentz however was satisfied with his theory. It was ingenious and it gave the right answers to practical questions. What more could one want?

APPROPRIATELY the new quantum electrodynamics of 1947 originated with an idea proposed by Kramers, whose recent death is such a heavy loss to physics, and who happened to be the successor of Lorentz at Leiden. The mathematical formalism was later developed by Schwinger, Bethe, Tomonaga, and others. Kramers' idea was a simple one, and similar to that of Lorentz fifty years earlier. Kramers observed that the problematical inertial force on an electron due to the field reaction could under no circumstances be experimentally separated from the effects of the electron's ordinary mechanical inertia. The only observable inertia is the total inertia, the sum of the mechanical and the electrical effects. The physicists of the 1930s made the mistake of confusing the unobservable mechanical mass of an electron (let us call it m_0) with the observed mass of a free electron (let us call it m). For example, they calculated the field reaction inertia of an electron bound in a hydrogen atom, finding the result which we will call δm, an infinite quantity. They concluded that the total inertia of the bound electron should be $(m + \delta m)$, which is infinite since m is finite. This would be an infinite value for an observable quantity and would necessarily imply that the theory is wrong. However, as Kramers pointed out, the total inertia of the bound electron is not $(m + \delta m)$ but

$$m_0 + \delta m = m + \delta m - (m - m_0).$$

The quantity $(m - m_0)$ is by definition just the field inertia or the δm calculated for a free electron. For the observable total inertia to be finite, it is not necessary for δm to be finite. It is only necessary that the difference between the δm calculated for the bound electron and for a free electron be finite. Kramers suggested, and Schwinger afterwards verified, that this difference is in fact finite. This difference then represents the difference between the total inertia of a bound and a free electron, which is the quantity which is directly measured in the Lamb–Retherford experiment. After long and delicate calculations, it has recently been shown that the theoretical and experimental values of the difference agree to a phenomenally high degree of accuracy (at present about one part in a thousand, in an effect which was ten years ago beyond the limit of detection!).

The new quantum electrodynamics is, like the Lorentz electrodynamics, only a reinterpretation and not a departure from the older theory. It differs from the old theory only in this, that we now take consistently into account the effects of field reaction not only on the measured quantities but also on the standard mass m with which the measured quantities are compared. We can prove quite generally that when observable quantities are calculated and the results expressed in terms of the mass m instead of the unobservable m_0, the infinite expressions always cancel out and the results are finite. Further, the finite results have always turned out to agree with the experiments. A similar argument is also applied to the electronic charge. The measured charge on an electron, which we call e, is different from the quantity e_0 which appears in the starting equations of the theory, as a result of field reactions. If e is calculated in terms of e_0, the result involves infinities. But e_0 is an unobservable quantity, and measured quantities when expressed in terms of e are always finite. Therefore we have in the end a completely precise and workable theory. The starting equations contain the quantities m_0 and e_0 which are unobservable. When we make calculations of observable effects, we obtain expressions involving m_0 and e_0 together with infinite quantities, divergent integrals, and so forth. We have not to be afraid of the infinite quantities. We treat them as if they were ordinary numbers, and then at the end of the calculation, when everything is expressed in terms of the observed mass m and charge e, all the infinities drop out and the result is finite.

We are proud of our new quantum electrodynamics. Like the Lorentz theory, it is a triumph of ingenuity, and it succeeds in reconciling all the contradictions of the older theory without abandoning anything of value. It also shares with the Lorentz theory one other most striking feature. Namely, the whole success of the theory is based on an unexplained miracle. In the starting equations of the Lorentz theory there is a stationary ether. In quantum electrodynamics the starting equations involve the unobservable and mathematically meaningless symbols e_0 and m_0. In both cases there is a complicated mathematical cancellation, so that in calculations of observable quantities the final results are independent of either the ether velocity or of the meaningless symbols. Why these miraculous cancellations occur, the theories do not explain.

WE HAVE NOW brought our historical parallel down to the present moment. Can we extend it further still? The subsequent history of the Lorentz theory at least is well known. After Lorentz had worked for many years creating and perfecting his theory, Einstein appeared with the explanation of the miracle. He showed that all the consequences of the Lorentz theory could be deduced from a much simpler theory involving a new physical principle, the principle of special relativity. In the new theory there was no ether, no absolute velocities. Thus the absence of experimental effects of absolute velocities was assured from the beginning. The impossibility of detecting absolute motion in space was for Einstein the starting point, and everything else was derived from it. Einstein's theory did not substantially depart from the Lorentz theory in its predictions. Einstein simply turned the Lorentz theory upside down, so that the endpoint became the starting point and vice versa. After this inversion, all the satisfactory features of the Lorentz theory remained, and only the unobservable complications, the ether and the absolute velocities, vanished. Einstein's formulation of classical electrodynamics is so simple and complete that it still stands substantially as it did in 1905.

Can we hope for a similar revolution in quantum electrodynamics? It is my firm belief that we can. What we require is again to turn the theory upside down, so that its consequences remain unchanged while its principles are clarified. We need to find a way of starting the theory, so that the unobservable quantities e_0 and m_0 do not appear at all in the equations. That is, we need to describe an electron from the beginning, not as a mechanical particle plus an electromagnetic field, but as a unified whole. The new description should be based on a physical principle, similar to the principle of relativity, expressing just the impossibility of making an experimental separation of an electron into its mechanical and electrical parts. Only when we have such a description shall we understand the real reasons for the success of our present theory. To me it seems that this argument leads to a positive conclusion, that the unexplained success of the present theory is in itself a guarantee that a new and simpler description is waiting to be discovered. How long shall we have to wait for the discovery? This no one can guess. We must only be patient, and remember that the time scale of fundamental understanding is always slow. From Maxwell to Einstein was forty years, from Dirac to the present only twenty-five. PT

FROM THE ARCHIVES
JULY 2015

Reinhold Bertlmann is a professor of physics at the University of Vienna.

Magic moments with John Bell

Reinhold A. Bertlmann

John Bell, with whom I had a fruitful collaboration and warm friendship, is best known for his seminal work on the foundations of quantum physics, but he also made outstanding contributions to particle physics and accelerator physics.

John Stewart Bell and I met over tea in the common room of CERN's theory division. I had arrived a few weeks earlier, in April 1978, to work as an Austrian fellow. After one of the weekly theoretical seminars, the division held a welcome reception for all its newcomers. John was an impressive man, about 17 years older than me, with metal-rimmed glasses, red hair, and a beard. He asked about my research field, and when I replied, "quarkonium," he showed great interest. We immediately started a lively discussion in his office—the beginning of a fruitful collaboration and warm friendship.

The partner

Quarkonium, in analogy to positronium, designates a bound quark–antiquark system. Such states appear as narrow peaks in the energy spectra that are obtained after hadrons (particles containing quarks) interact; for that reason, quarkonium states are often called resonances. During the 1970s particle physicists discovered several such resonances, including the J/ψ, a bound state of charm and anticharm, and the Υ, a bound state of bottom and antibottom. The properties of those particles had to be understood, and so quarkonium states were a popular research field when John and I first got together.

At the time, physicists recognized that they could get pretty far considering just short-distance quark interactions. For instance, one could accurately predict the lifetimes of resonances.[1] John and I, however, wanted to understand the positions of the resonances; to do that, we had to include long-range interactions, which considerably upped the complexity of the calculations. For one thing, we had to consider interactions with and among gluons—particles analogous to photons—that convey the strong force that holds quarks together. That required us to go beyond perturbation theory and include the so-called gluon condensate: gluon fluctuations in the quantum chromodynamics vacuum.

Our approach was to approximate the full quantum field theory by something called potential theory, then a rather popular model. Within that framework, we succeeded in obtaining the ground-state energies of the J/ψ and Υ resonances[2] to within about 10%, though we were not able to construct a totally satisfactory bridge between the potential theory we used and the full-fledged quantum theory.[3] In carrying out our work, we had to make use of mathematical functions called moments. In view of the surprising success we achieved in obtaining the ground-state energies, we titled our paper "Magic moments."

Figure 1. Afternoon tea was a must when John Bell (right) and I (left) were working together. This shot was taken at John's home in 1980. (Photograph © Renate Bertlmann.)

I well remember one of our afternoon rituals. John, a true Irishman, always had to drink tea at four o'clock; figure 1 shows us checking out a sample at his home. We also practiced our ritual in the CERN cafeteria, where John always ordered *deux infusions verveine, s'il vous plaît*—two infusions of verbena, his favorite tea, for us to enjoy together. There, in a relaxed atmosphere, we talked about physics and philosophy. At times we were joined by my artist wife, Renate, and then the three of us had heated debates about modern art.

The particle physicist

John was a highly esteemed particle physicist who fascinated me with his extraordinary personality. I felt his fatherly kindness and admired his knowledge and wisdom. He had a deep understanding of quantum field theory and liked to illustrate his ideas with basic examples. He wrote several celebrated papers in particle physics, of which I'll mention just a few.

John's PhD thesis, submitted in the mid 1950s, included a fundamental paper, "Time reversal in field theory."[4] In that work he proved the so-called *CPT* theorem, where *C* is the charge conjugation operator, which replaces particles with antiparticles; *P* is the parity operator, which performs an inversion through the origin; and *T* is the time-reversal operation. The theorem states that any quantum field theory satisfying a small set of standard assumptions must be *CPT* symmetric. (For the record, the assumptions are that the theory is Lorentz invariant, local, and possesses a Hermitian Hamiltonian.) For many years all the credit went to Gerhart Lüders and Wolfgang Pauli, who proved the theorem a little bit before John did, but nowadays John is also rightly recognized.

John's most far-reaching contribution to particle physics was a paper called "A PCAC puzzle: $\pi^0 \to \gamma\gamma$ in the σ-model," written with Roman Jackiw, who was a postdoc at CERN at the time.[5] The "PCAC" in the title stands for "partially conserved axial current." The details aren't important here, but the idea is that the existence of a symmetry—the chiral symmetry that seemed to imply a conserved axial current in the limit that pions are massless—precluded the decay of the pion into two photons. The solution to the puzzle was that the very process of quantization can lead to the breakdown of a classical symmetry; when that happens, the quantum theory is said to be anomalous. Ultimately, the chiral-symmetry anomaly is responsible for the pion decay.

Stephen Adler helped to clarify the anomaly issue in a paper written independently of Bell and Jackiw's work.[6] Nowadays, the chiral anomaly is often referred to as the Adler-Bell-Jackiw anomaly. Further studies revealed anomalies to be not just a pathology of the quantization procedure but also keys to a deeper understanding of quantum field theory.[7] Anomalies are widespread in physical theories, including the standard model of particle physics and theories of gravitation.

Also worthy of mention is John's influential review "Weak interaction of kaons," coauthored with experimentalist Jack Steinberger, and the pioneering work on vector bosons and neutrino reactions that John wrote with his colleague Martinus Veltman.[8]

The accelerator physicist

After graduating from Queen's University Belfast in 1949 with two bachelor's degrees, John began his scientific career at the UK Atomic Energy Research Establishment at Harwell. There he met his future wife, Mary Ross, a reactor and accelerator physicist. She was working in the theoretical physics division, which was led by Klaus Fuchs, the well-known physicist who later got sentenced to prison because of his atomic espionage for the Soviet Union. In 1954 John and Mary were married and began to pursue their careers together.

Shortly after John came to Harwell, he and Mary were sent to the Telecommunications Research Establishment in

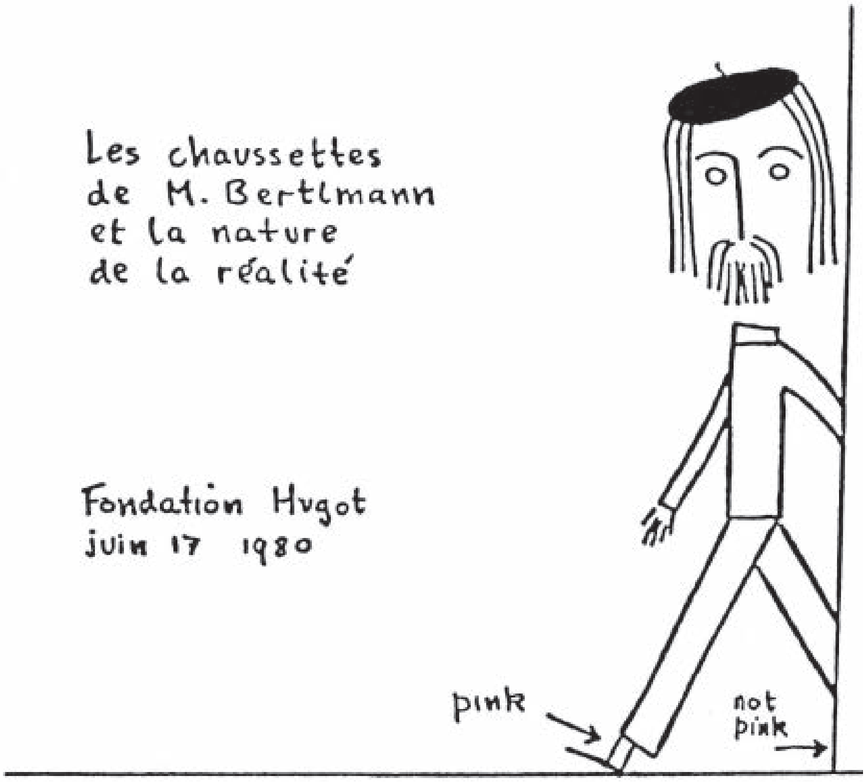


Figure 2. My socks were always of two different colors, as John Bell observed in this cartoon accompanying his paper "Bertlmann's socks and the nature of reality."[8,10] The paper, which addressed the difference between quantum and classical correlations, was based on a colloquium, "Conceptual Implications of Quantum Mechanics," organized by the Hugot Foundation of the Collège de France.

Malvern, where they stayed for about a year to work in William Walkinshaw's accelerator group. Walkinshaw highly appreciated John's abilities and noted that he "was a young man of high caliber who soon showed his independence on choice of project, with a special liking for particle dynamics. His mathematical talent was superb and elegant."[9]

Alone or in collaboration with Walkinshaw, John wrote several papers, mostly on how to focus a bunch of electrons or protons in a linear accelerator. In 1951 the whole accelerator group moved back to Harwell; soon after that, John turned to particle physics. By the end of the 1950s, he and Mary had become attracted to CERN, Europe's largest laboratory for basic science. The two moved there in 1960, John to be part of the theory division and Mary to join the accelerator research group.

During the 1980s John and Mary collaborated on accelerator work and wrote several papers together. One example is "Electron cooling in storage rings," in which they analyzed how changes in the electron velocity distribution would affect the electrons' ability to cool ion or proton beams in storage rings such as the Low Energy Antiproton Ring at CERN.[8] That paper was dedicated to Yuri Orlov, an accelerator physicist who was then imprisoned in the Soviet Union for his human rights activism and was freed later on. Such an act of solidarity was typical of the Bells.

A particularly attractive work, in my opinion, was Bell's combination of the Unruh effect of quantum field theory with accelerator physics. According to William Unruh, an observer who is uniformly accelerated through the electromagnetic vacuum will experience blackbody radiation with a temperature proportional to the acceleration. John's idea was to use electrons as the accelerated observers and the polarization of the electron beam as the thermometer that measures the temperature of the blackbody radiation. The result, published together with Jon Leinaas, a CERN fellow from Norway, was that the effect of the acceleration was small but measurable.[8]

I become famous

At CERN, John was a kind of oracle for particle physics, consulted by many colleagues who wanted to get his approval for their ideas. Of course, I had heard that he was also a leading figure in quantum mechanics—specifically, in quantum foundations. But nobody, either at CERN or anywhere else, could actually explain his foundational work to me. The standard answer was, "He discovered some relation whose consequence was that quantum mechanics turned out all right. But we knew that anyway, so don't worry." And I didn't. John, for his part, never mentioned his quantum work to me during the early years of our collaboration.

At the end of the summer of 1980, I returned for a while to my home institute, the University of Vienna. There was no internet then, and it was a common practice for physicists to send preprints of their work to all the main physics institutions in the world before their papers were published. Each week we in Vienna would exhibit the new incoming preprints on a special shelf.

One day I was sitting in our computer room with my computer cards, when my colleague Gerhard Ecker rushed in, waving a preprint in his hands. He shouted, "Reinhold, look, now you're famous!" I could hardly believe my eyes as I read and reread the title of a paper by John, "Bertlmann's socks and the nature of reality."[8,10] I was totally stunned. As I read the first page, my heart stood still. The paper begins

> The philosopher in the street, who has not suffered a course in quantum mechanics, is quite unimpressed by Einstein-Podolsky-Rosen [EPR] correlations. He can point to many examples of similar correlations in everyday life. The case of

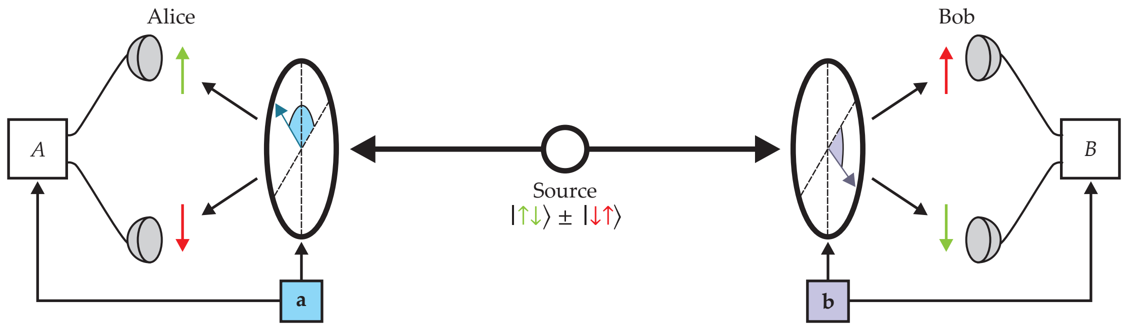


Figure 3. John Bell's famous inequality was derived for the setup illustrated here. A pair of spin-½ particles are prepared in a state of zero angular momentum, and each propagates freely in opposite directions to the measuring stations called Alice and Bob. Alice measures the spin in a direction **a** while Bob simultaneously measures in a direction **b**. In a hidden-variable theory, the measurement results are predetermined; the hidden variables might decree, for example, that if Alice measures her spin up, Bob will measure his down. (Adapted from R. A. Bertlmann, *J. Phys. A* **47**, 424007, 2014.)

> Bertlmann's socks is often cited. Dr. Bertlmann likes to wear two socks of different colours. Which colour he will have on a given foot on a given day is quite unpredictable. But when you see that the first sock is pink you can be already sure that the second sock will not be pink. Observation of the first, and experience of Bertlmann, gives immediate information about the second. There is no accounting for tastes, but apart from that there is no mystery here. And is not the EPR business [regarding quantum correlations] just the same?

John's paper included a cartoon (figure 2) that showed me with my odd socks; seeing it nearly knocked me down. It came so unexpectedly. I had no idea that John had noticed my habit of wearing socks of different colors—a habit I had cultivated since my early student days as my special 1960s-era protest. The article immediately pushed me into the quantum debate, and it thus really changed my life.

Now the time had come to understand why the "EPR business" was not just the same as "Bertlmann's socks" and to appreciate John's profound insight. I dove into his seminal works on hidden-variable theory and on Bell's inequality (see section 3 of reference 8) and his foundational quantum works.[10] I was impressed by John's clarity and depth of thought. From then on we had fruitful discussions about foundational issues; those interactions were a great fortune and honor for me. A new world had opened up—the universe of John Bell—and it has fascinated me ever since.

The critic of von Neumann

John was never satisfied with interpretations of quantum mechanics. Even as a student at Queen's University Belfast, he disliked the Copenhagen interpretation with its essential distinction between the quantum and classical worlds. He wondered where the quantum world stopped and the classical world began, and he wanted to get rid of the division.

When David Bohm published his reinterpretation of quantum theory as a deterministic, realistic theory with hidden variables,[11] his work was not appreciated by the physics community. Albert Einstein, for example, said that it "seems too cheap," and Wolfgang Pauli rejected it as "artificial metaphysics." John, however, was very much impressed and often remarked, "I saw the impossible thing done." For him, it was clear that in an appropriate reformulation of quantum theory, quantum particles would have definite properties governed by hidden variables. "Everything has definite properties," he would often say.

Hidden-variable theories take a set of observables $\{A, B, C, \ldots\}$ and assign to each individual system a set of eigenvalues $\{v(A, \lambda), v(B, \lambda), v(C, \lambda), \ldots\}$, one for each observable. Note that the assigned eigenvalues depend on the value of the hidden variable (or variables; there could be more than one) λ. For example, A, B, and C could be the x, y, and z components of an electron's spin in units of $\hbar/2$. Then, for a particular λ, $\{v(A), v(B), v(C)\}$ could be $\{+1, +1, -1\}$. Different members of an ensemble of states could have different assignments of the plus and minus signs according to their own individual λ; thus the hidden-variable theory must also provide a probability distribution for λ. When a quantum state—a state vector plus the specification of hidden variables—uniquely determines measurement outcomes, the state is said to be dispersion free.

In 1964 John started his investigation "On the problem of hidden variables in quantum mechanics"[10] by criticizing John von Neumann, who had given a proof that dispersion-free states, and thus hidden variables, are incompatible with quantum mechanics. What was the criticism? Consider three operators A, B, and C that satisfy $C = A + B$. If A and B commute, then the assigned eigenvalues must satisfy $v(C, \lambda) = v(A, \lambda) + v(B, \lambda)$.

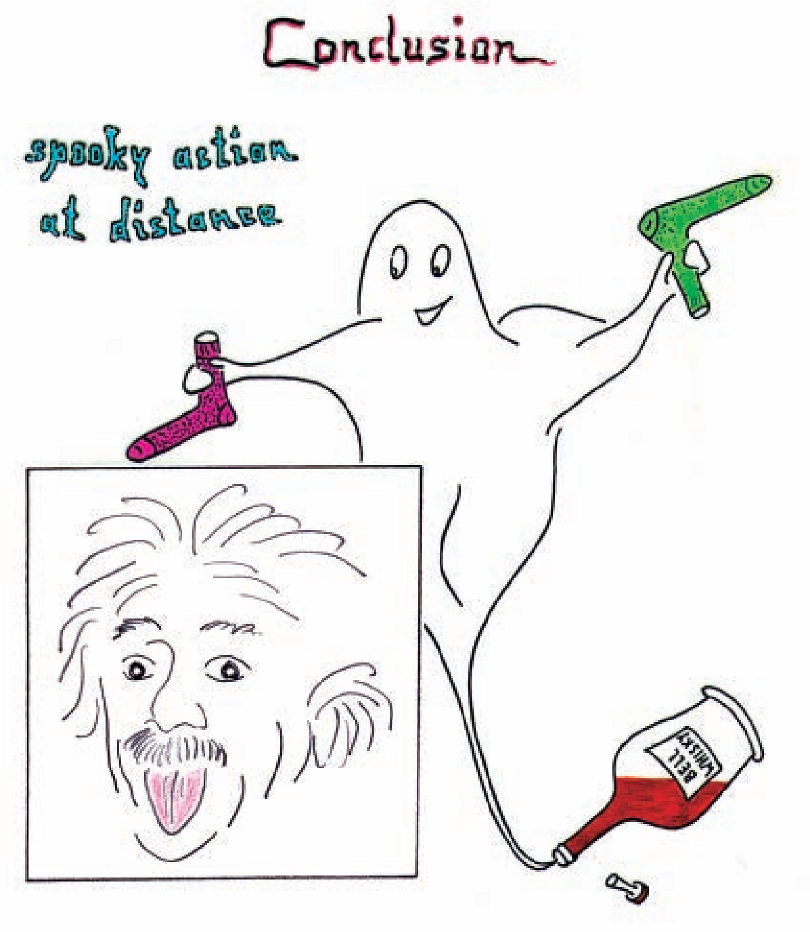


Figure 4. Real whisky bottles and spooky ghosts coexist in this cartoon that I drew to conclude a paper[17] dedicated to John Bell on the occasion of his 60th birthday.

Von Neumann, however, imposed the additivity property for noncommuting as well as commuting operators. "This is wrong," Bell grumbled, and before giving a general proof, he illustrated his dictum with the example of a spin measurement. Measuring the spin operator σ_x requires a suitably oriented Stern–Gerlach apparatus. The measurements of σ_y and $\sigma_x + \sigma_y$ require different orientations. Since the operators cannot be measured simultaneously, there is no necessity to impose additivity.

Thus John pointed to models for which results may depend on apparatus settings. Such models are called contextual, and they may agree with quantum mechanics. However, as demonstrated by the celebrated Kochen–Specker theorem, all noncontextual hidden-variable theories are indeed in conflict with quantum mechanics.[12]

The creator of Bell's theorem

At the end of his hidden-variable paper, John analyzed Bohm's reformulation more accurately. He discovered that according to Bohm's theory, in a system of two spin-½ particles—objects, like the electron, whose spin is $\hbar/2$—the behavior of one particle depends on the characteristics of the other, no matter how far apart the two particles are. He wondered, Was the dependence on remote characteristics just a defect of Bohm's particular hidden-variable model or would it hold more generally? Thus he was led to his seminal work "On the Einstein-Podolsky-Rosen paradox," which contained a proof that the result was general—the celebrated Bell inequality.[10]

John's profound discovery was that locality was incompatible with the statistical predictions of quantum mechanics. He proceeded from Bohm's spin version of the EPR paradox. As shown in figure 3, a pair of spin-½ particles in a spin singlet state (that is, the angular momentum of the pair is zero) propagates freely in opposite directions to measuring stations called Alice and Bob. Alice measures the spin in units of $\hbar/2$ along a direction **a** and obtains A; Bob measures along **b** and gets B. In a hidden-variable theory, the results are predetermined and specified by λ.

Assuming that A does not depend on Bob's measurement settings and B does not depend on Alice's—a condition now called Bell's locality hypothesis—the expectation value of the joint spin measurement of Alice and Bob is given by

$$E(\mathbf{a}, \mathbf{b}) = \int d\lambda\ \rho(\lambda)\ A(\mathbf{a}, \lambda) \cdot B(\mathbf{b}, \lambda).$$

Here the function $\rho(\lambda)$ represents a normalized distribution function for λ.

Alice's and Bob's spin measurements must satisfy $A(\mathbf{a}, \lambda) = \pm 1$ and $B(\mathbf{b}, \lambda) = \pm 1$. Given those relations, John was able to derive an inequality that must hold in *all* hidden-variable theories satisfying Bell's locality hypothesis: $1 + E(\mathbf{b}, \mathbf{c}) \geq |E(\mathbf{a}, \mathbf{b}) - E(\mathbf{a}, \mathbf{c})|$.

According to quantum mechanics, though, $E(\mathbf{a}, \mathbf{b}) = -\mathbf{a} \cdot \mathbf{b}$. Thus the quantum predictions violate Bell's inequality if, for example, **a**, **b**, and **c** lie in the same plane and are oriented, respectively, at 0°, 120°, and 60° relative to a common axis.

When I derived Bell's inequality for the first time, I was really impressed that it was possible to discriminate between all hidden-variable theories and quantum mechanics. How did John find his special combination of expectation values that contradicted quantum mechanics for certain sets of measurements? For me as a theorist the job was done. Nevertheless, experiment had to decide which was right, hidden-variable theory or quantum mechanics.

Classic experiments

The first to become interested in experimentally exploring Bell inequalities—nowadays there are several—was John

Clauser in the late 1960s. At that time, working in the field was a courageous act. Clauser relates, for example, how he once had an appointment with Richard Feynman to discuss an experimental EPR configuration for testing the predictions of quantum mechanics. Feynman immediately threw him out of the office saying, "Well, when you have found an error in quantum-theory's experimental predictions, come back then, and we can discuss *your* problem with it."[13] Fortunately, Clauser remained stubborn and, with Stuart Freedman, carried out the experiment in 1972. The outcome is well known; the results were in accord with quantum theory and in clear violation of a Bell inequality. Later experiments, notably by Edward Fry and Randall Thompson, confirmed the result.[14]

The 1980s saw a second generation of Bell experiments carried out, in particular by Alain Aspect and his group.[15] Aspect and colleagues worked with polarized photons, and their goal was to incorporate a fast-switch mechanism for the polarizers to exclude a possible mutual influence between the two observers Alice and Bob. Again, a Bell inequality was significantly violated, and again, experimental results agreed with the quantum mechanics predictions. In my opinion, the Aspect work was a turning point; the physics community began to realize that such explorations were getting at something essential. Research started into what is nowadays called quantum information and quantum communication, a flourishing field.

The third generation of Bell experiments commenced in the 1990s and has extended into the 21st century. It has taken advantage of new technologies such as spontaneous parametric down conversion, which is an effective way to create entangled photons. Anton Zeilinger and his group, in a landmark experiment, were able to ensure that the directions in which photon polarization was measured were set randomly and independently.[16] Fascinating experiments on quantum teleportation, quantum cryptography, and long-distance quantum communication followed.

A great puzzle

The essential ingredient in all Bell inequalities is Bell's locality hypothesis. So far, all experiments looking for violations in Bell inequalities have found them, so we have to conclude, along with John, that nature contains a nonlocality in its structure. That nonlocality disturbed John deeply, since for him it was equivalent to a breaking of Lorentz invariance—a feature he could hardly accept. He often remarked, "It's a great puzzle to me. Behind the scenes something is going faster than the speed of light."

John was totally convinced that realism is the proper position for a scientist. That is, he believed that experimental results are predetermined and not induced by the measurement process. In his analysis of EPR correlations, he did not so much assume reality as infer it. "It's a mystery," he said, "if looking at one sock makes the sock pink and the other one not-pink at the same time." He remained faithful to the hidden-variable program and was not discouraged by the outcome of the EPR–Bell experiments; rather, he found them puzzling. As he once remarked to me, "The situation is very intriguing that at the foundation of all that impressive success [of quantum mechanics] there are these great doubts."

At the end of his "Bertlmann's socks" paper, John again expressed his concern:

> It may be that we have to admit that causal influences *do* go faster than light. The role of Lorentz invariance in the completed theory would then be very problematic. An "ether" would be the cheapest solution. But the unobservability of this ether would be disturbing. So would the impossibility of "messages" faster than light.

I got back at John for "Bertlmann's socks" in a paper, "Bell's theorem and the nature of reality,"[17] that I dedicated to him in 1988 on the occasion of his 60th birthday. I sketched my conclusion in a cartoon, shown as figure 4. John, who strictly avoided alcohol, was very much amused by my illustration, since the spooky, nonlocal ghost emerged from a bottle of Bell's whisky, a brand that really did exist.

When I look back at my collaboration with John and remember his honest character and warm friendship, his deep and sharp intellect, and the knowledge I owe to him, I really feel privileged and thankful for the times I could spend with him. They were magic moments indeed.

I thank Renate Bertlmann for her company in all these years and for providing figure 1.

REFERENCES

1. J. S. Bell, R. A. Bertlmann, *Z. Phys. C* **4**, 11 (1980).
2. J. S. Bell, R. A. Bertlmann, *Nucl. Phys. B* **177**, 218 (1981); *Phys. Lett. B* **137**, 107 (1984); R. A. Bertlmann, *Nucl. Phys. B* **204**, 387 (1982).
3. J. S. Bell, R. A. Bertlmann, *Nucl. Phys. B* **187**, 285 (1981); R. A. Bertlmann, J. S. Bell, *Nucl. Phys. B* **227**, 435 (1983).
4. J. S. Bell, *Proc. R. Soc. London A* **231**, 479 (1955).
5. J. S. Bell, R. Jackiw, *Nuovo Cimento A* **60**, 47 (1969).
6. S. L. Adler, *Phys. Rev.* **177**, 2426 (1969).
7. R. A. Bertlmann, *Anomalies in Quantum Field Theory*, Oxford U. Press (2000).
8. M. Bell, K. Gottfried, M. Veltman, eds., *Quantum Mechanics, High Energy Physics and Accelerators: Selected Papers of John S. Bell*, World Scientific (1995).
9. P. G. Burke, I. C. Percival, *Biogr. Mem. Fellows R. Soc.* **45**, 1 (1999).
10. J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics*, 2nd ed., Cambridge U. Press (2004).
11. D. Bohm, *Phys. Rev.* **85**, 166 (1952); D. Bohm, *Phys. Rev.* **85**, 180 (1952).
12. See, for example, J. Conway, S. Kochen, in *Quantum [Un]speakables: From Bell to Quantum Information*, R. A. Bertlmann, A. Zeilinger, eds., Springer (2002), chap. 18.
13. J. F. Clauser, in ref. 12, p. 71.
14. E. S. Fry, T. Walther, in ref. 12, chap. 8.
15. A. Aspect, in ref. 12, chap. 9.
16. G. Weihs, in ref. 12, chap. 10.
17. R. A. Bertlmann, *Found. Phys.* **20**, 1191 (1990).

FROM THE ARCHIVES
APRIL 1985

David Mermin is director of the Laboratory of Atomic and Solid State Physics at Cornell University. A solid-state theorist, he has recently come up with some quasithoughts about quasicrystals. He is known to PHYSICS TODAY readers as the person who made "boojum" an internationally accepted scientific term. With N. W. Ashcroft, he is about to start updating the world's funniest solid-state physics text. He says he *is* bothered by Bell's theorem, but may have rocks in his head anyway.

Is the Moon there when nobody looks? Reality and the quantum theory

Einstein maintained that quantum metaphysics entails spooky actions at a distance; experiments have now shown that what bothered Einstein is not a debatable point but the observed behavior of the real world.

N. David Mermin

Quantum mechanics is magic[1]

In May 1935, Albert Einstein, Boris Podolsky and Nathan Rosen published[2] an argument that quantum mechanics fails to provide a complete description of physical reality. Today, 50 years later, the EPR paper and the theoretical and experimental work it inspired remain remarkable for the vivid illustration they provide of one of the most bizarre aspects of the world revealed to us by the quantum theory.

Einstein's talent for saying memorable things did him a disservice when he declared "God does not play dice," for it has been held ever since that the basis for his opposition to quantum mechanics was the claim that a fundamental understanding of the world can only be statistical. But the EPR paper, his most powerful attack on the quantum theory, focuses on quite a different aspect: the doctrine that physical properties have in general no objective reality independent of the act of observation. As Pascual Jordan put it[3]

> Observations not only disturb what has to be measured, they produce it. . . . We compel [the electron] to assume a definite position. . . . We ourselves produce the results of measurement.

Jordan's statement is something of a truism for contemporary physicists. Underlying it, we have all been taught, is the disruption of what is being measured by the act of measurement, made unavoidable by the existence of the quantum of action, which generally makes it impossible even in principle to construct probes that can yield the information classical intuition expects to be there.

Einstein didn't like this. He wanted things out there to have properties, whether or not they were measured:[4]

> We often discussed his notions on objective reality. I recall that during one walk Einstein suddenly stopped, turned to me and asked whether I really believed that the moon exists only when I look at it.

The EPR paper describes a situation ingeniously contrived to force the quantum theory into asserting that properties in a space-time region **B** are the result of an act of measurement in another space-time region **A**, so far from **B** that there is no possibility of the measurement in **A** exerting an influence on region **B** by any known dynamical mechanism. Under these conditions, Einstein maintained that the properties in **A** must have existed all along.

Spooky actions at a distance

Many of his simplest and most explicit statements of this position can be found in Einstein's correspondence with Max Born.[5] Throughout the book (which sometimes reads like a Nabokov novel), Born, pained by Einstein's distaste for the statistical character of the quantum theory, repeatedly fails, both in his letters and in his later commentary on the correspondence, to understand what is really bothering Einstein. Einstein tries over and over again, without success, to make himself clear. In March 1948, for example, he writes:

> That which really exists in **B** should . . . not depend on what

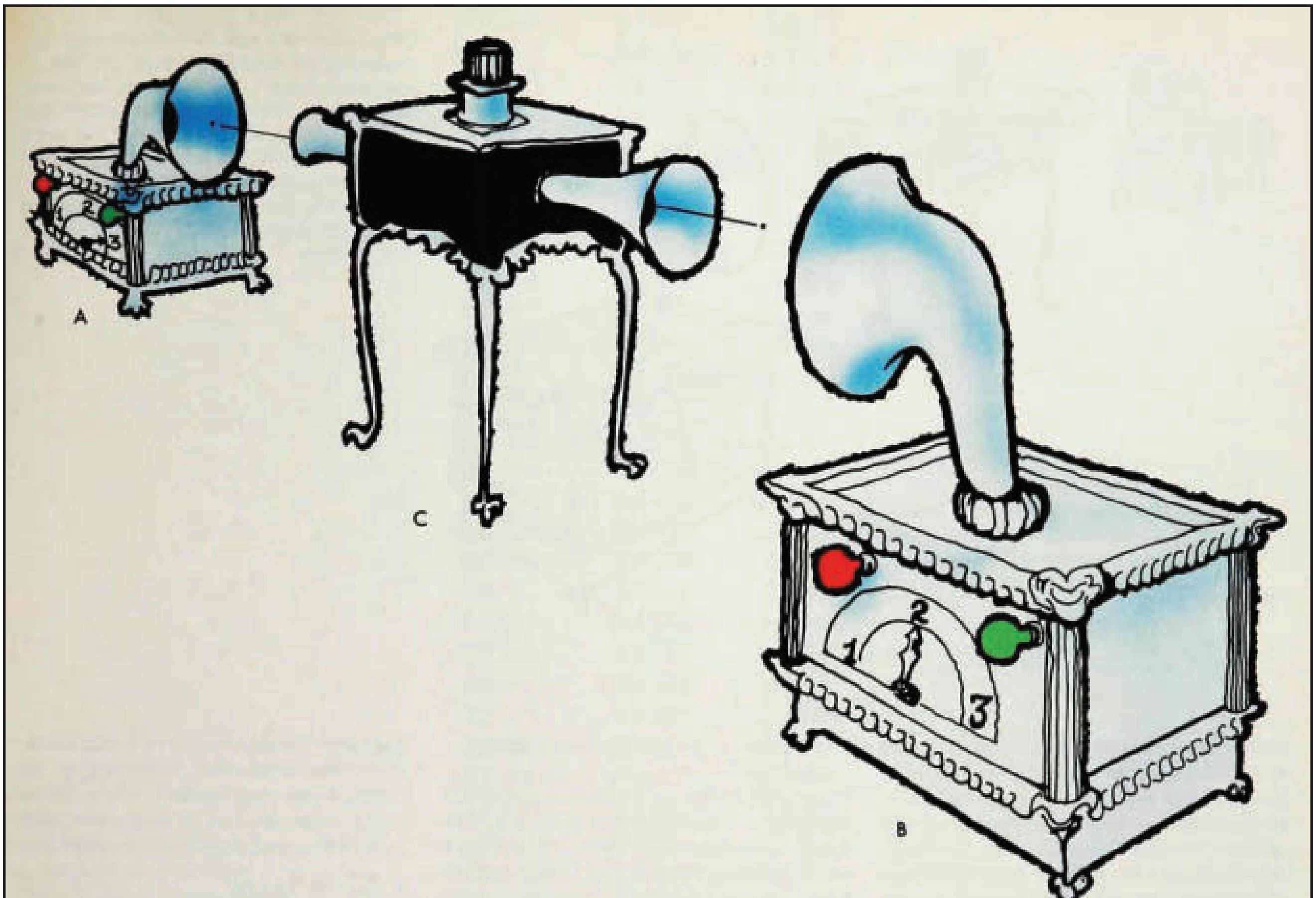
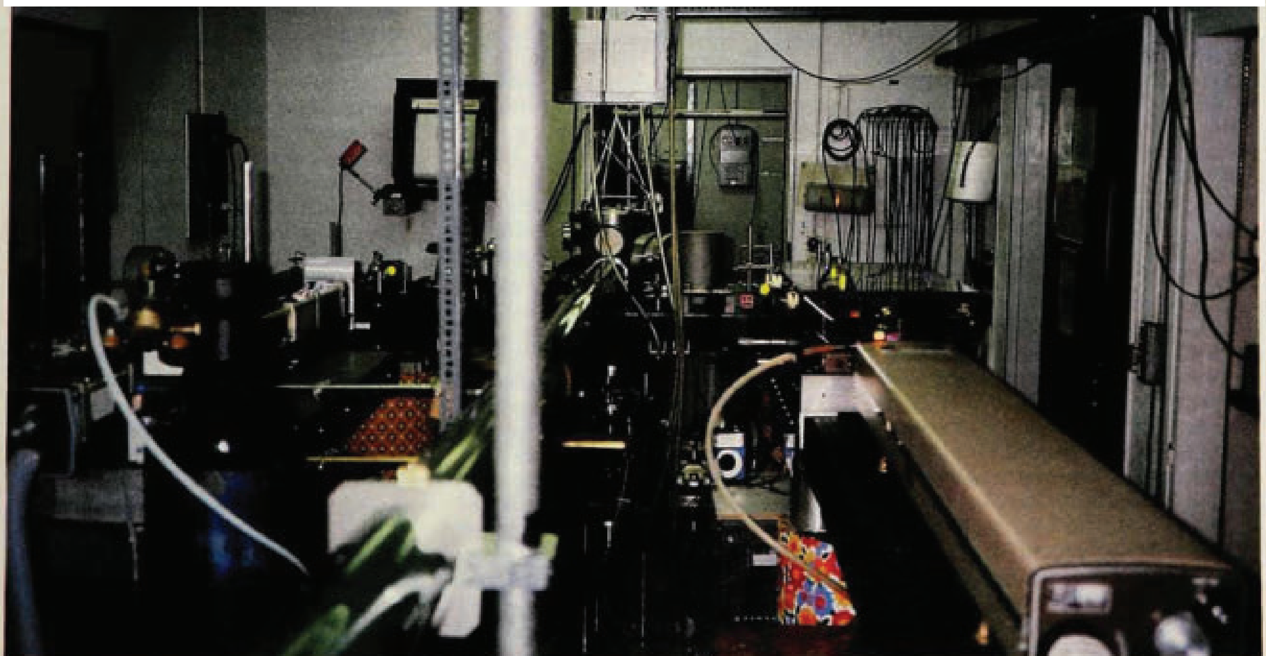


Figure 1. An EPR apparatus. The experimental setup consists of two detectors, **A** and **B**, and a source of something ("particles" or whatever) **C**. To start a run, the experimenter pushes the button on **C**; something passes from **C** to both detectors. Shortly after the button is pushed each detector flashes one of its lights. Putting a brick between the source and one of the detectors prevents that detector from flashing, and moving the detectors farther away from the source increases the delay between when the button is pushed and when the lights flash. The switch settings on the detectors vary randomly from one run to another. Note that there are no connections between the three parts of the apparatus, other than via whatever it is that passes from **C** to **A** and **B**. The photo below shows a realization of such an experiment in the laboratory of Alain Aspect in Orsay, France. In the center of the lab is a vacuum chamber where individual calcium atoms are excited by the two lasers visible in the picture. The re-emitted photons travel 6 meters through the pipes to be detected by a two-channel polarizer.



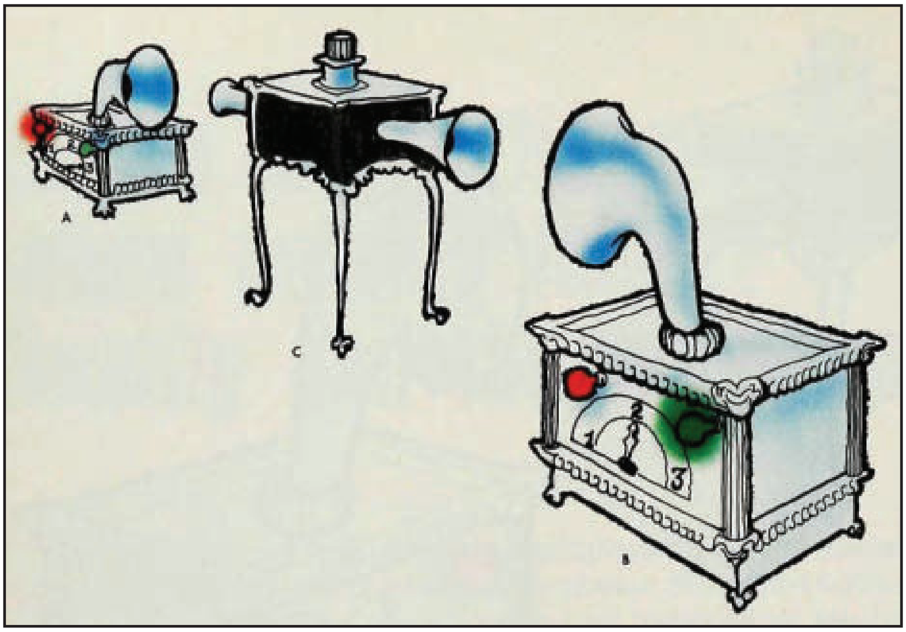


Figure 2. The result of a run. Shortly after the experimenter pushed the button on the source in figure 1 , the detectors flash one lamp each. The experimenter records the switch settings and the colors of the lamps and then repeats the experiment. Here, for example, the record reads 32RG—the switches are in positions 3 and 2 and the lamps flashed R and G, respectively.

> kind of measurement is carried out in part of space **A**; it should also be independent of whether or not any measurement at all is carried out in space **A**. If one adheres to this program, one can hardly consider the quantum-theoretical description as a complete representation of the physically real. If one tries to do so in spite of this, one has to assume that the physically real in **B** suffers a sudden change as a result of a measurement in **A**. My instinct for physics bristles at this.

Or, in March 1947,

> I cannot seriously believe in [the quantum theory] because it cannot be reconciled with the idea that physics should represent a reality in time and space, free from spooky actions at a distance.

The "spooky actions at a distance" (*spukhafte Fernwirkungen*) are the acquisition of a definite value of a property by the system in region **B** by virtue of the measurement carried out in region **A**. The EPR paper presents a wavefunction that describes two correlated particles, localized in regions **A** and **B**, far apart. In this particular two-particle state one can learn (in the sense of being able to predict with certainty the result of a subsequent measurement) either the position or the momentum of the particle in region **B** as a result of measuring the corresponding property of the particle in region **A**. If "that which really exists" in region **B** does not depend on what kind of measurement is carried out in region **A**, then the particle in region **B** must have had both a definite position and a definite momentum all along.

Because the quantum theory is intrinsically incapable of assigning values to both quantities at once, it must provide an incomplete description of the physically real. Unless, of course, one asserts that it is only by virtue of the position (or momentum) measurement in **A** that the particle in **B** acquires its position (or momentum): spooky actions at a distance.

At a dramatic moment Pauli appears in the *Born–Einstein Letters*, writing Born from Princeton in 1954 with his famous tact on display:

> Einstein gave me your manuscript to read; he was *not at all* annoyed with you, but only said you were a person who will not listen. This agrees with the impression I have formed myself insofar as I was unable to recognize Einstein whenever you talked about him in either your letter or your manuscript. It seemed to me as if you had erected some dummy Einstein for yourself, which you then knocked down with great pomp. In particular, Einstein does not consider the concept of "determinism" to be as fundamental as it is frequently held to be (as he told me emphatically many times). . . . In the same way, he *disputes* that he uses as criterion for the admissibility of a theory the question: "Is it rigorously deterministic?"

Pauli goes on to state the real nature of Einstein's "philosophical prejudice" to Born, emphasizing that "Einstein's point of departure is 'realistic' rather than 'deterministic.'" According to Pauli the proper grounds for challenging Einstein's view are simply that

> One should no more rack one's brain about the problem of whether something one cannot know anything about exists all the same, than about the ancient question of how many angels are able to sit on the point of a needle. But

33RR	33GG	13GR
12GR	31GR	23GR
33GG	12GG	22RR
21GR	21GR	11RR
21RR	33GG	21GR
22RR	21RR	21RR
33GG	12GR	23GG
11GG	22RR	32GR
23RR	13RG	33RR
32GR	12RG	33GG
12GR	23GG	33GG
12RG	11GG	23GR
11GG	13RG	21GR
31RG	21RG	12RR
12RG	33RR	32GR
13GR	32GR	32GR
22GG	32GG	33GG
12RG	33GG	31RG
12GR	21RR	13RR
22GG	12RG	13RG
23GR	22GG	32RG
33RR	11GG	31GR
33GG	23GR	23RR
31RG	22RR	33RR
31RR	11GG	13GR
33RR	32GR	11GG
32RG	13RG	31GR
31RG	13GR	31RG
11RR	23GG	13GR
23GR	33RR	23RG
12GG	31GR	31GG
11GG	13RG	23RG
13RG	23RR	21RR
31RG	12GR	23RG
23GR	31RG	11GG
31GR	32RG	22GG
23RG	21GR	11GG
22RR	22GG	11GG
12GR	22RR	21RG
32GR	13RR	11RR
22RR	21GG	12RG
12GG	23GR	23GR
33RR	22GG	32GR
11RR	22GG	21GG
23GG	31GG	21RG
23GG	13GR	13RG
33RR	21GR	13RG
23GR	33RR	13RG
21GG	23RR	13GR
13GR	22RR	23RG
33GG	12RR	22GG
11GG	23RG	11RR
12RR	23RG	31RG
12GG	32GR	23RR
31GG	31RG	23RG
32RG	22GG	11RR
21GR	11GG	32RG
22GG	11GG	32GR

Figure 3. Data produced by the apparatus of figure 1. This is a fragment of an enormous set of data generated by many, many runs: Each entry shows the switch settings and the colors of the lights that flashed for a run. The switch settings are changed randomly from run to run.

it seems to me that Einstein's questions are ultimately always of this kind.

Faced with spooky actions at a distance, Einstein preferred to believe that things one cannot know anything about (such as the momentum of a particle with a definite position) do exist all the same. In April 1948 he wrote to Born:

> Those physicists who regard the descriptive methods of quantum mechanics as definitive in principle would . . . drop the requirement for the independent existence of the physical reality present in different parts of space; they would be justified in pointing out that the quantum theory nowhere makes explicit use of this requirement. I admit this, but would point out: when I consider the physical phenomena known to me, and especially those which are being so successfully encompassed by quantum mechanics, I still cannot find any fact anywhere which would make it appear likely that [the] requirement will have to be abandoned. I am therefore inclined to believe that the description of quantum mechanics . . . has to be regarded as an incomplete and indirect description of reality. . . .

A fact is found

The theoretical answer to this challenge to provide "any fact anywhere" was given in 1964 by John S. Bell, in a famous paper[6] in the short-lived journal *Physics*. Using a *gedanken* experiment invented[7] by David Bohm, in which "properties one cannot know anything about" (the simultaneous values of the spin of a particle along several distinct directions) are required to exist by the EPR line of reasoning, Bell showed ("Bell's theorem") that the nonexistence of these properties is a direct consequence of the quantitative numerical predictions of the quantum theory. The conclusion is quite independent of whether or not one believes that the quantum theory offers a complete description of physical reality. If the data in such an experiment are in agreement with the numerical predictions of the quantum theory, then Einstein's philosophical position has to be wrong.

In the last few years, in a beautiful series of experiments, Alain Aspect and his collaborators at the University of Paris's Institute of Theoretical and Applied Optics in Orsay provided[8] the experimental answer to Einstein's challenge by performing a version of the EPR experiment under conditions in which Bell's type of analysis applied. They showed that the quantum-theoretic predictions were indeed obeyed. Thirty years after Einstein's challenge, a fact—not a metaphysical doctrine—was provided to refute him.

Attitudes toward this particular 50-year sequence of intellectual history and scientific discovery vary widely.[9] From the very start Bohr certainly took it seriously. Léon Rosenfeld describes[10] the impact of the EPR argument:

> This onslaught came down upon us as a bolt from the blue. Its effect on Bohr

12GR	31GR	23GR
33GG	12GG	**22RR**
21GR	21GR	**11RR**
21RR	**33GG**	21GR
22RR	21RR	21RR
33GG	12GR	23GG
11GG	**22RR**	32GR
23RR	13RG	**33RR**
32GR	12RG	**33GG**
12GR	23GG	**33GG**
12RG	**11GG**	23GR
11GG	13RG	21GR
31RG	21RG	12RR
12RG	**33RR**	32GR
13GR	32GR	32GR
22GG	32GG	**33GG**
12RG	**33GG**	31RG
12GR	21RR	13RR
22GG	12RG	13RG
23GR	**22GG**	32RG
33RR	**11GG**	31GR
33GG	23GR	23RR
31RG	**22RR**	**33RR**
31RR	**11GG**	13GR
33RR	32GR	**11GG**
32RG	13RG	31GR
31RG	13GR	31RG
11RR	23GG	13GR
23GR	**33RR**	23RG
12GG	31GR	31GG
11GG	13RG	23RG
13RG	23RR	21RR
31RG	12GR	23RG
23GR	31RG	**11GG**
31GR	32RG	**22GG**
23RG	21GR	**11GG**
22RR	**22GG**	11GG
12GR	**22RR**	21RG
32GR	13RR	11RR
22RR	21GG	12RG
12GG	23GR	23GR
33RR	**22GG**	32GR
11RR	**22GG**	21GG
23GG	31GG	21RG
23GG	13GR	13RG
33RR	21GR	13RG
23GR	**33RR**	13RG
21GG	23RR	13GR
13GR	**22RR**	23RG
33GG	12RR	**22GG**
11GG	23RG	**11RR**
12RR	23RG	31RG
12GG	32GR	23RR
31GG	31RG	23RG
32RG	**22GG**	**11RR**
21GR	**11GG**	32RG
22GG	**11GG**	32GR
22RR	21RG	13GG

Figure 4. Switches set the same: the data of figure 3, but highlighted to pick out those runs in which both detectors had the same switch settings as they flashed. Note that in such runs the lights always flash the same colors.

> was remarkable. . . . A new worry could not have come at a less propitious time. Yet, as soon as Bohr had heard my report of Einstein's argument, everything else was abandoned.

Bell's contribution has become celebrated in what might be called semipopular culture. We read, for example, in *The Dancing Wu Li Masters* that[11]

> Some physicists are convinced that [Bell's theorem] is the most important single work, perhaps, in the history of physics.

And indeed, Henry Stapp, a particle theorist at Berkeley, writes that[12]

> Bell's theorem is the most profound discovery of science.

At the other end of the spectrum, Abraham Pais, in his recent biography of Einstein, writes[13] of the EPR article—that "bolt from the blue," the basis for "the most profound discovery of science":

> The only part of this article which will ultimately survive, I believe, is a phrase ["No reasonable definition of reality could be expected to permit this"] which so poignantly summarizes Einstein's views on quantum mechanics in his later years.

I think it is fair to say that more physicists would side with Pais than with Stapp, but between the majority position of near indifference and the minority position of wild extravagance is an attitude I would characterize as balanced. This was expressed to me most succinctly by a distinguished Princeton physicist on the occasion of my asking how he thought Einstein would have reacted to Bell's theorem. He said that Einstein would have gone home and thought about it hard for several weeks—that he couldn't guess what he would then have said, except that it would have been extremely interesting. He was sure that Einstein would have been very bothered by Bell's theorem. Then he added,

> Anybody who's not bothered by Bell's theorem has to have rocks in his head.

To this moderate point of view I would only add the observation that contemporary physicists come in two varieties. Type 1 physicists are bothered by EPR and Bell's theorem. Type 2 (the majority) are not, but one has to distinguish two subvarieties. Type 2a physicists explain why they are not bothered. Their explanations tend either to miss the point entirely (like Born's to Einstein) or to contain physical assertions that can be shown to be false. Type 2b are not bothered and refuse to explain why. Their position is unassailable. (There is a variant of type 2b who say that Bohr straightened out[14] the whole business. but refuse to explain how.)

A *gedanken* demonstration

To enable you to test which category you belong to, I shall describe, in black-box

12GR	31GR	23GR
33GG	12GG	22RR
21GR	21GR	11RR
21RR	33GG	21GR
22RR	21RR	21RR
33GG	12GR	23GG
11GG	22RR	32GR
23RR	13RG	33RR
32GR	12RG	33GG
12GR	23GG	33GG
12RG	11GG	23GR
11GG	13RG	21GR
31RG	21RG	12RR
12RG	33RR	32GR
13GR	32GR	32GR
22GG	32GG	33GG
12RG	33GG	31RG
12GR	21RR	13RR
22GG	12RG	13RG
23GR	22GG	32RG
33RR	11GG	31GR
33GG	23GR	23RR
31RG	22RR	33RR
31RR	11GG	13GR
33RR	32GR	11GG
32RG	13RG	31GR
31RG	13GR	31RG
11RR	23GG	13GR
23GR	33RR	23RG
12GG	31GR	31GG
11GG	13RG	23RG
13RG	23RR	21RR
31RG	12GR	23RG
23GR	31RG	11GG
31GR	32RG	22GG
23RG	21GR	11GG
22RR	22GG	11GG
12GR	22RR	21RG
32GR	13RR	11RR
22RR	21GG	12RG
12GG	23GR	23GR
33RR	22GG	32GR
11RR	22GG	21GG
23GG	31GG	21RG
23GG	13GR	13RG
33RR	21GR	13RG
23GR	33RR	13RG
21GG	23RR	13GR
13GR	22RR	23RG
33GG	12RR	22GG
11GG	23RG	11RR
12RR	23RG	31RG
12GG	32GR	23RR
31GG	31RG	23RG
32RG	22GG	11RR
21GR	11GG	32RG
22GG	11GG	32GR
22RR	21RG	13GG

Figure 5. Switches set any way: the data of figure 3, but highlighted to emphasize only the colors of the lights that flashed in each run, no matter how the switches were set when the lights flashed. Note that the pattern of colors is completely random.

terms, a very simple version of Bell's *gedanken* experiment, deferring to the very end any reference whatever either to the underlying mechanism that makes the gadget work or to the quantum-theoretic analysis that accounts for the data. Perhaps this backwards way of proceeding will make it easier for you to lay aside your quantum theoretic prejudices and decide afresh whether what I describe is or is not strange.[15]

What I have in mind is a simple *gedanken* demonstration. The apparatus comes in three pieces. Two of them (**A** and **B**) function as detectors. They are far apart from each other (in the analogous Aspect experiments over 10 meters apart). Each detector has a switch that can be set to one of three positions; each detector responds to an event by flashing either a red light or a green one. The third piece (**C**), midway between **A** and **B**, functions as a source. (See figure 1.)

There are no connections between the pieces—no mechanical connections, no electromagnetic connections, nor any other known kinds of relevant connections. (I promise that when you learn what is inside the black boxes you will agree that there are no connections.) The detectors are thus incapable of signaling to each other or to the source via any known mechanism, and with the exception of the "particles" described below, the source has no way of signaling to the detectors. The demonstration proceeds as follows:

The switch of each detector is independently and randomly set to one of its three positions, and a button is pushed on the source; a little after that, each detector flashes either red or green. The setting of the switches and the colors that flash are recorded, and then the whole thing is repeated over and over again.

The data consist of a pair of numbers and a pair of colors for each run. A run, for example, in which **A** was set to 3, **B** was set to 2, **A** flashed red, and **B** flashed green, would be recorded as "32RG," as shown in figure 2.

Because there are no built-in connections between the source **C** and the detectors **A** and **B**, the link between the pressing of the button and the flashing of the light on a detector can only be provided by the passage of something (which we shall call a "particle," though you can call it anything you like) between the source and that detector. This can easily be tested; for example, by putting a brick between the source and a detector. In subsequent runs, that detector will not flash. When the brick is removed, everything works as before.

Typical data from a large number of runs are shown in figure 3. There are just two relevant features:

▶ If one examines only those runs in which the switches have the same setting (figure 4), then one finds that the lights always flash the same colors.

▶ If one examines all runs, without any regard to how the switches are set (figure 5), then one finds that the pattern of flashing is completely random. In particular, half the time the lights flash the same colors, and half the time different colors.

That is all there is to the *gedanken* demonstration.

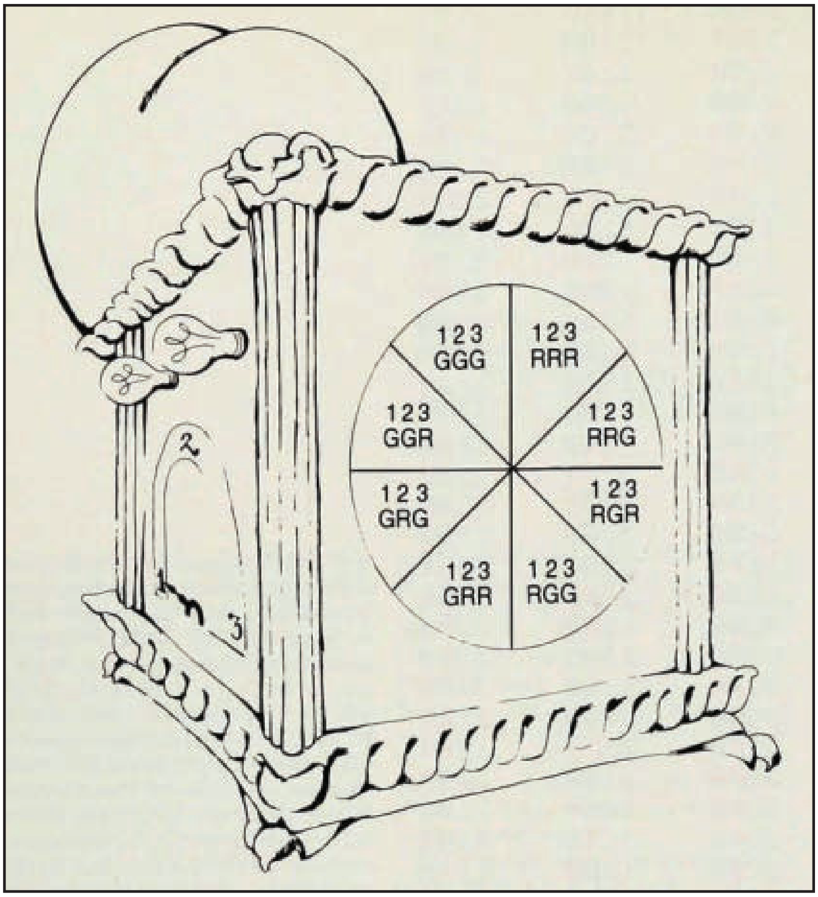


Figure 6. Model of a detector to produce data like those in figure 4. Particles from the source fall with equal probability into any of the eight bins; for each bin the color flashed depends on the switch as indicated on the back of the box.

Should you be bothered by these data unless you have rocks in your head?

How could it work?

Consider only those runs in which the switches had the same setting when the particles went through the detectors. In all such runs the detectors flash the same colors. If they could communicate, it would be child's play to make the detectors flash the same colors when their switches had the same setting, but they are completely unconnected. Nor can they have been preprogrammed always to flash the same colors, regardless of what is going on, because the detectors are observed to flash different colors in at least some of those runs in which their switches are differently set, and the switch settings are independent random events.

How, then, are we to account for the first feature of the data? No problem at all. Born, in fact, in a letter of May 1948, offers[5] such an explanation to Einstein:

> It seems to me that your axiom of the "independence of spatially separated objects **A** and **B**" is not as convincing as you make out. It does not take into account the fact of coherence; objects far apart in space which have a common origin need not be independent. I believe that this cannot be denied and simply has to be accepted. Dirac has based his whole book on this.

In our case the detectors are triggered by particles that have a common origin at the source **C**. It is then easy to dream up any number of explanations for the first feature of the data.

Suppose, for example, that what each particle encounters as it enters its detector is a target (figure 6) divided into eight regions, labeled RRR, RRG, RGR, RGG, GRR, GRG, GGR, and GGG. Suppose each detector is wired so that if a particle lands in the GRG bin, the detector flips into a mode in which the light flashes G if the switch is set to 1, R if it is set to 2, and G if it is set to 3; RGG leads to a mode with R for 1 and G for 2 and 3, and so on. We can then easily account for the fact that the lights always flash the same colors when the switches have the same settings by assuming that in each run the source always fires its particles into bins with the same labels.

Evidently this is not the only way. One could imagine that particles come in eight varieties: cubes, spheres, tetrahedra, ... All settings produce R when a cube is detected, a sphere results in R for settings 1 and 2, G for setting 3, and so forth. The first feature of the data is then accounted for if the two particles produced by the source in each run are always both of the same variety.

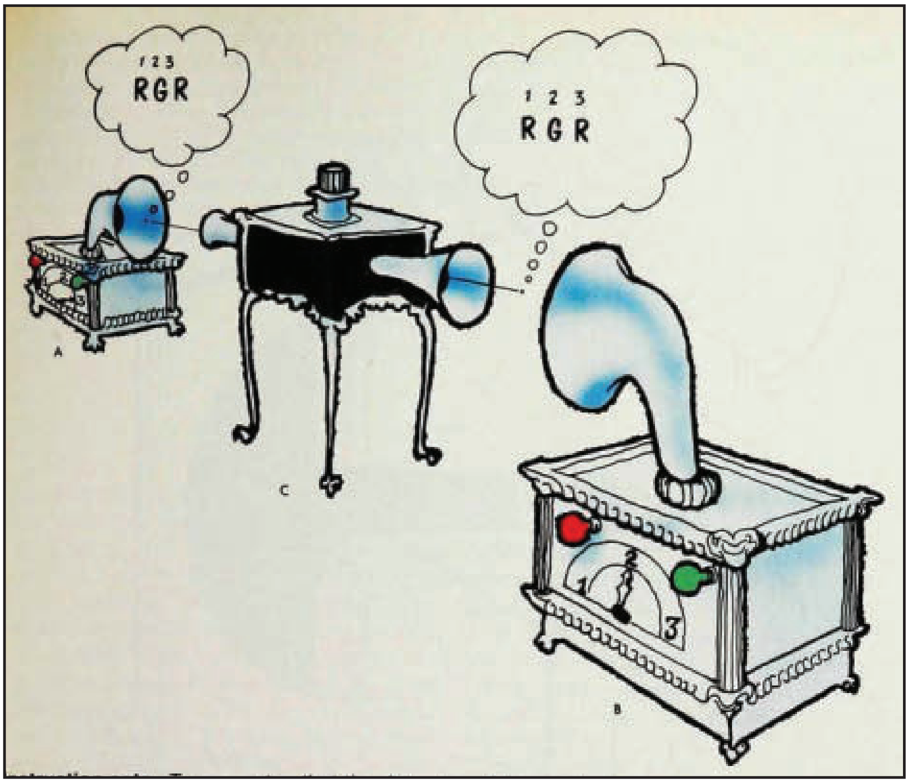


Figure 7. Instruction sets. To guarantee that the detectors of figure 6 flash the same color when the switches are set the same, the two particles must in one way or another carry instruction sets specifying how their detectors are to flash for each possible switch setting. The results of any one run reveal nothing about the instructions beyond the actual data; so in this case, for example, the first instruction (1R) is "something one cannot know anything about," and I've only guessed at it, assuming that "it exists all the same."

Common to all such explanations is the requirement that each particle should, in one way or another, carry to its detector a set of instructions for how it is to flash for *each* of the three possible switch settings, and that in *any* run of the experiment both particles should carry the same instruction sets:

▶ A set of instructions that covers *each* of the three possible settings is required because there is no communication between the source and the detectors other than the particles themselves. In runs in which the switches have the same setting, the particles cannot know whether that setting will be 11, 22, or 33. For the detectors always to flash the same colors when the switches have the same setting, the particles must carry instructions that specify colors for each of the three possibilities.

▶ The absence of communication between source and detectors also requires that the particles carry such instruction sets in *every* run of the experiment—even those in which the switches end up with different settings—because the particles always have to be prepared: Any run may turn out to be one in which the switches end up with the same settings.

This generic explanation is pictured schematically in figure 7.

Alas, this explanation—the only one, I maintain, that someone not steeped in quantum mechanics will ever be able to come up with (though it is an entertaining game to challenge people to try)—is untenable. It is inconsistent with the second feature of the data: There is no conceivable way to assign such instruction sets to the particles from one run to the next that can account for the fact that in all runs taken together, without regard to how the switches are set, the same colors flash half the time.

Pause to note that we are about to show that "something one cannot know anything about"—the third entry in an instruction set—cannot exist. For even if instruction sets did exist, one could never learn more than two of the three entries (revealed in those runs where the switches ended up with two different settings). Here is the argument.

Consider a particular instruction set, for example, RRG. Should both particles be issued the instruction set RRG, then the detectors will flash the same colors when the switches are set to 11, 22, 33, 12, or 21; they will flash different colors for 13, 31, 23, or 32. Because the switches at each detector are set randomly and independently, each of these nine cases is equally likely, so the instructions set RRG will result in the same colors flashing $5/9$ of the time.

Evidently the same conclusion holds for the sets RGR, GRR, GGR, GRG and RGG, because the argument uses only the fact that one color appears twice and the other once. All six such instruction sets also result in the same colors flashing $5/9$ of the time.

But the only instruction sets left are RRR and GGG, and these each result in the same colors flashing all of the time.

Therefore if instruction sets exist, the same colors will flash in at least $5/9$ of all the runs, regardless of how the instruction sets are distributed from one run of the demonstration to the next. This is

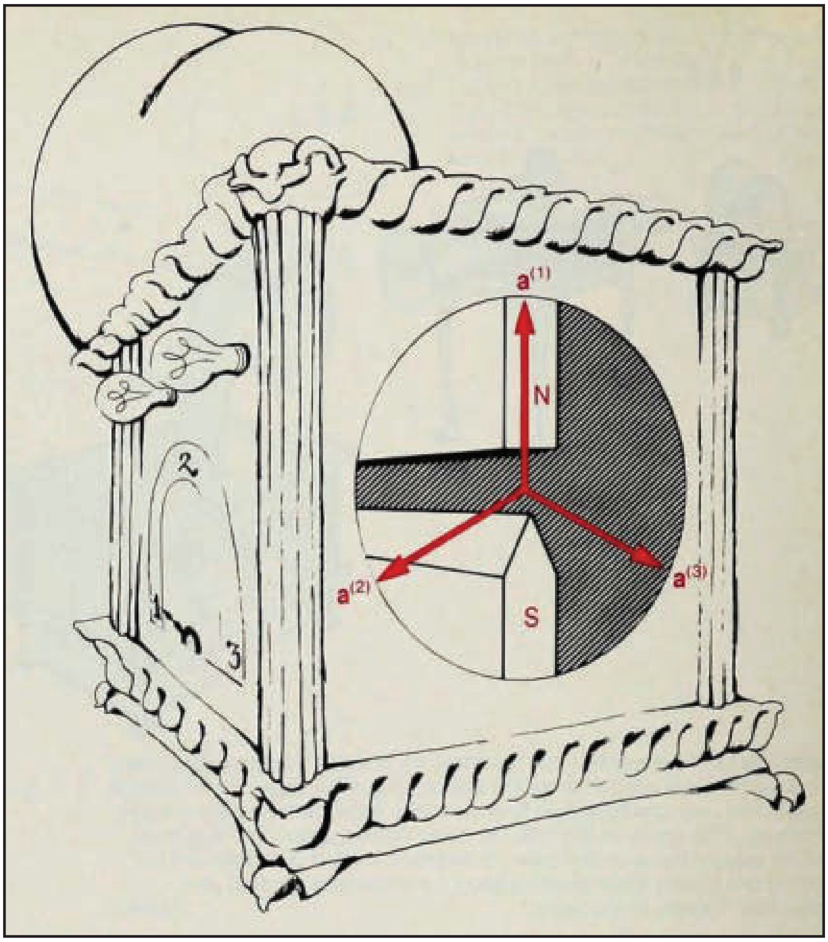


Figure 8. A realization of the detector to produce the data of figure 3. The particles have a magnetic moment and can be separated into "spin up" and "spin down" particles by the Stern–Gerlach magnet inside the detector. Setting the switch to positions 1, 2, or 3 rotates the north pole of the magnet along the coplanar unit vectors $\mathbf{a}^{(1)}$, $\mathbf{a}^{(2)}$, or $\mathbf{a}^{(3)}$, separated by 120°. The vector sum of the three unit vectors is, of course, zero. The switch positions on the two detectors correspond to the same orientations of the magnetic field. One detector flashes red for spin up, green for spin down; the other uses the opposite color convention.

Bell's theorem (also known as Bell's inequality) for the *gedanken* demonstration.

But in the actual *gedanken* demonstration the same colors flash only ½ the time. The data described above violate this Bell's inequality, and therefore there can be no instruction sets.

If you don't already know how the trick is done, may I urge you, before reading how the *gedanken* demonstration works, to try to invent some other explanation for the first feature of the data that does not introduce connections between the three parts of the apparatus or prove to be incompatible with the second feature.

One way to do it

Here is one way to make such a device:

Let the source produce two particles of spin ½ in the singlet state, flying apart toward the two detectors. (Granted, this is not all that easy to do, but in the Orsay experiments described below, the same effect is achieved with correlated photons.) Each detector contains a Stern–Gerlach magnet, oriented along one of three directions ($\mathbf{a}^{(1)}$, $\mathbf{a}^{(2)}$, or $\mathbf{a}^{(3)}$), perpendicular to the line of flight of the particles, and separated by 120°, as indicated in figure 8. The three settings of the switch determine which orientation is used. The light on one detector flashes red or green, depending on whether the particle is deflected toward the north (spin up) or south (spin down) pole of the magnet as it passes between them; the other detector uses the opposite color convention.

That's it. Clearly there are no connections between the source and the detectors or between the two detectors. We can nevertheless account for the data as follows:

When the switches have the same setting, the spins of both particles are measured along the same direction, so the lights will always flash the same colors if the measurements along the same direction always yield opposite values. But this is an immediate consequence of the structure of the spin singlet state, which has the form

$$|\psi\rangle = (1/\sqrt{2})[|+-\rangle - |-+\rangle] \qquad (1)$$

independent of the direction of the spin quantization axis, and therefore yields $+-$ or $-+$ with equal probability, but never $++$ or $--$, whenever the two spins are measured along any common direction.

To establish the second feature of the data, note that the product $m_1 m_2$ of the results of the two spin measurements (each of which can have the values $+½$ or $-½$) will have the value $-¼$ when the lights flash the same colors and +¼ when they flash different colors. We must therefore show that the product vanishes when averaged over all the nine distinct pairs of orientations the two Stern–Gerlach magnets can have. For a given pair of orientations, $\mathbf{a}^{(i)}$ and $\mathbf{a}^{(j)}$, the mean value of this product is just the expectation value in the state ψ of the corresponding product of (commuting) hermitian observables $\mathbf{a}^{(i)} \cdot \mathbf{S}^{(1)}$ and $\mathbf{a}^{(j)} \cdot \mathbf{S}^{(2)}$. Thus the second feature of the data requires:

$$0 = \textstyle\sum_{ij} \langle \psi | [\mathbf{a}^{(i)} \cdot \mathbf{S}^{(1)}][\mathbf{a}^{(j)} \cdot \mathbf{S}^{(2)}] | \psi \rangle \qquad (2)$$

But equation 2 is an immediate consequence of the linearity of quantum mechanics, which lets one take the sums inside the matrix element, and the fact that the three unit vectors around an equilateral triangle sum to zero:

$$\textstyle\sum_i \mathbf{a}^{(i)} = \sum_j \mathbf{a}^{(j)} = 0 \qquad (3)$$

This completely accounts for the data. It also unmasks the *gedanken* demonstration as a simple embellishment of Bohm's version of the EPR experiment. If we kept only runs in which the switches had the same setting, we would have precisely the Bohm–EPR experiment. The assertion that instruction sets exist is then blatant quantum-theoretic nonsense, for it amounts to the insistence that each particle has stamped on it in advance the outcome of the measurements of three different spin components corresponding to noncommuting observables $\mathbf{S} \cdot \mathbf{a}^{(i)}$, $i = 1, 2, 3$. According to EPR, this is merely a limitation of the quantum-theoretic formalism, because instruction sets are the only way to account for the first feature of the data.

Bell's analysis adds to the discussion those runs in which the switches have different settings, extracts the second feature of the data as a further elementary prediction of quantum mechanics, and demonstrates that any set of data exhibiting this feature is incompatible with the existence of the instruction sets apparently required by the first feature, quite independently of the formalism used to explain the data, and quite independently of any doctrines of quantum theology.

The experiments

The experiments of Aspect and his colleagues at Orsay confirm that the quantum-theoretic predictions for this experiment are in fact realized, and that the conditions for observing the results of the experiment can in fact be achieved. (A distinguished colleague once told me that the answer to the EPR paradox was that correlations in the singlet state could never be maintained over macroscopic distances—that anything, even the passage of a cosmic ray in the next room, would disrupt the correlations enough to destroy the effect.)

In these experiments the two spin-½ particles are replaced by a pair of photons and the spin measurements become polarization measurements. The photon pairs are emitted by calcium atoms in a radiative cascade after suitable pumping by lasers. Because the initial and final atomic states have $J = 0$, quantum theory predicts (and experiment confirms) that the photons will be found to have the same polarizations (lights flashing the same colors in the analogous *gedanken* experiment) if they are measured along the same direction—feature number 1. But if the polarizations are measured at 120° angles, then theory predicts (and experiment confirms) that they will be the same only a quarter of the time [$\frac{1}{4} = \cos^2(120°)$]. This is precisely what is needed to produce the statistics of feature number 2 of the *gedanken* demonstration: The randomly set switches end up with the same setting (same polarizations measured) ⅓ of the time, so in all runs the same colors will flash $\frac{1}{3} \times 1 + \frac{2}{3} \times (\frac{1}{4}) = \frac{1}{2}$ the time. The people in Orsay were interested in a somewhat modified version of Bell's argument in which the angles of greatest interest were multiples of 22.5°, but they collected data for many different angles, and, except for EPR specialists, the conceptual differences between the two cases are minor.[16]

There are some remarkable features to these experiments. The two polarization analyzers were placed as far as 13 meters apart without producing any noticeable change in the results, thereby closing the loophole that the strange quantum correlations might somehow diminish as the distance between regions **A** and **B** grew to macroscopic proportions. At such separations it is hard to imagine that a polarization measurement of photon #1 could, in any ordinary sense of the term, "disturb" photon #2. Indeed, at these large separations, a hypothetical disturbance originating when one photon passed through its analyzer could only reach the other analyzer in time to affect the outcome of the second polarization measurement if it traveled at a superluminal velocity.

In the third paper of the Orsay group's series, bizarre conspiracy theories are dealt a blow by an ingenious mechanism for rapidly switching the directions along which the polarizations of each photon are measured. Each photon passes to its detector through a volume of water that supports an ultrasonic standing wave. Depending on the instantaneous amplitude of the wave, the photon either passes directly into a polarizer with one orientation or is Bragg reflected into another with a different orientation. The standing waves that determine the choice of orientation at each detector are independently driven and have frequencies so high that several cycles take place during the light travel time from one detector to the other. (This corresponds to a refinement of the *gedanken* demonstration in which, to be absolutely safe, the switches are not given their random settings until *after* the particles have departed from their common source.)

What does it mean?

What is one to make of all this? Are there "spooky actions at a distance"? A few years ago I received the text of a letter from the executive director of a California thinktank to the Under Secretary of Defense for Research and Engineering, alerting him to the EPR correlations:

> If in fact we can control the faster-than-light nonlocal effect, it would be possible . . . to make an untappable and unjammable command-control-communication system at very high bit rates for use in the submarine fleet. The important point is that since there is no ordinary electromagnetic signal linking the encoder with the decoder in such a hypothetical system, there is nothing for the enemy to tap or jam. The enemy would have to have actual possession of the "black box" decoder to intercept the message, whose reliability would not depend on separation from the encoder nor on ocean or weather conditions. . . .

Heady stuff indeed! But just what is this nonlocal effect? Using the language of the *gedanken* demonstration, let us talk about the "*N*-color" of a particle (*N* can be 1, 2, or 3) as the color (red or green) of the light that flashes when the particle passes through a detector with its switch set to *N*. Because instruction sets cannot exist, we know that a particle cannot at the same time carry a definite 1-color, 2-color and 3-color to its detector. On the other hand, for any particular *N* (say 3), we can determine the 3-color of the particle heading for detector **A** before it gets there by arranging things so that the other particle first reaches detector **B**, where its 3-color is measured. If the particle at **B** was 3-colored red, the particle at **A** will turn out to be 3-colored red, and green at **B** means green at **A**.

Three questions now arise:

▶ Did the particle at **A** have its 3-color prior to the measurement of the 3-color of the particle at **B**? The answer cannot be yes, because, *prior* to the measurement of the 3-color at **B**, it is altogether possible that the roll of the dice at **B** or the whim of the **B**-operator will result in the 2-color or the 1-color being measured at **B** instead. Barring the most paranoid of conspiracy theories, "prior to the measurement of the 3-color at **B**" is indistinguishable from "prior to the measurement of the 2- (or 1-) color at **B**." If the 3-color already existed, so also must the 2- and 1-colors have existed. But instruction sets (which consist of a specification of the 1-, 2-, and 3-colors) do not exist.

▶ Is the particle at **A** 3-colored red *after* the measurement at **B** shows the color red? The answer is surely yes, because under these circumstances it is invariably a particle that will cause the detector at **A** to flash red.

▶ Was something (the value of its 3-color) transmitted to the particle at **A** as a result of the measurement at **B**?

Orthodox quantum metaphysicians would, I believe, say no, nothing has changed at **A** as the result of the measurement at **B**; what has changed is our knowledge of the particle at **A**. (Somewhat more spookily, they might object to the naive classical assumption of localizability or separability implicit in the phrases "at **A**" and "at **B**.") This seems very sensible and very reassuring: *N*-color does not characterize the particle at all, but only what we know about the particle. But does that last sentence sound as good when "particle" is changed to "photon" and "*N*-color" to "polarization"? And does it really help you to stop wondering why the lights always flash the same colors when the switches have the same settings?

What is clear is that if there is spooky action at a distance, then, like other spooks, it is absolutely useless except for its effect, benign or otherwise, on our state of mind. For the statistical pattern of red and green flashes at detector **A** is entirely random, however the switch is set at detector **B**. Whether the particles arriving at **A** all come with definite 3-colors (because the switch at **B** was stuck at 3) or definite 2-colors (because the switch was stuck at 2) or no colors at all (because there was a brick in front of the detector at **B**)—all this has absolutely no effect on the statistical distribution of colors observed at **A**. The manifestation of this "action at a distance" is revealed only through a comparison of the data independently gathered at **A** and at **B**.

This is a most curious state of affairs, and while it is wrong to suggest that EPR correlations will replace sonar, it seems to me something is lost by ignoring them or shrugging them off. The EPR experiment is as close to magic as any physical phenomenon I know of, and magic should be enjoyed. Whether there is physics to be learned by pondering it is less clear. The most elegant answer I have found[17] to this last question comes from one of the great philosophers of our time, whose view of the matter I have taken the liberty of quoting in the form of the poetry it surely is:

We always have had a great deal of difficulty
in understanding the world view
that quantum mechanics represents.

At least I do,
because I'm an old enough man
that I haven't got to the point
that this stuff is obvious to me.

Okay, I still get nervous with it. . . .

You know how it always is,
every new idea,
it takes a generation or two
until it becomes obvious
that there's no real problem. . . .

I cannot define the real problem,
therefore I suspect there's no real problem,
but I'm not sure
there's no real problem.

Nobody in the 50 years since Einstein, Podolsky and Rosen has ever put it better than that.

Some of the views expressed above were developed in the course of occasional technical studies of EPR correlations supported by the National Science Foundation under grant No. DMR 83-14625.

REFERENCES

1. D. M. Greenberger, discussion remarks at the Symposium on Fundamental Questions in Quantum Mechanics, SUNY at Albany, April 1984.
2. A. Einstein, B. Podolsky, N. Rosen, *Phys. Rev.* **47**, 777 (1935).
3. Quoted by M. Jammer, *The Philosophy of Quantum Mechanics: The Interpretations of Quantum Mechanics in Historical Perspective*, Wiley-Interscience (1974), p. 161.
4. A. Pais, *Rev. Mod. Phys.* **51**, 863 (1979).
5. A. Einstein, M. Born, H. Born, *The Born–Einstein Letters: Correspondence Between Albert Einstein and Max and Hedwig Born from 1916 to 1955*, Walker (1971).
6. J. S. Bell, *Physics* **1**, 195 (1964).
7. D. Bohm, *Quantum Theory*, Prentice-Hall (1951), pp. 614–619.
8. A. Aspect, P. Grangier, G. Roger, *Phys. Rev. Lett.* **47**, 460 (1981); A. Aspect, P. Grangier, G. Roger, *Phys. Rev. Lett.* **49**, 91 (1982); A. Aspect, J. Dalibard, G. Roger, *Phys. Rev. Lett.* **49**, 1804 (1982).
9. For a discussion of the views of today's physicists toward the meaning of the quantum theory, see the interesting and provocative essay "Cognitive repression in contemporary physics" by E. F. Keller, *Am. J. Phys.* **47**, 718 (1979).
10. L. Rosenfeld, in *Niels Bohr: His Life and Work as Seen by His Friends and Colleagues*, S. Rozental, ed., North-Holland (1967), pp. 114–136.
11. G. Zukav, *The Dancing Wu Li Masters: An Overview of the New Physics*, Morrow (1979), p. 282. On the same page it is also said that "Bell's theorem is a mathematical construct which, as such, is indecipherable to the nonmathematician," a view that I hope the rest of this article will dispel.
12. H. P. Stapp, *Nuovo Cimento B* **40**, 191 (1977).
13. A. Pais, *"Subtle Is the Lord . . .": The Science and the Life of Albert Einstein*, Oxford U. Press (1982), p. 456.
14. N. Bohr, *Phys. Rev.* **48**, 696 (1935).
15. What follows is a somewhat refined version of an argument I published a few years ago in *Am. J. Phys.* **49**, 940 (1981), incorporating some improvements suggested by Richard Friedberg. For other elementary treatments see J. S. Bell's beautiful essay, "Bertlmann's socks and the nature of reality," *J. Phys. Colloq.* **42**, C2-41 (1981), B. d'Espagnat's article in the November 1979 *Scientific American*, or d'Espagnat's recent book, *In Search of Reality*, Springer-Verlag (1983).
16. For a survey of other attempts to realize the EPR experiment, and the variants of Bell's original argument used to interpret experimental tests, see J. F. Clauser, A. Shimony, *Rep. Prog. Phys.* **41**, 1881 (1978).
17. R. P. Feynman, *Int. J. Theor. Phys.* **21**, 467 (1982), p. 471. PT

FROM THE ARCHIVES
APRIL 2003

Barbara Terhal is a research staff member in the physical sciences department at the IBM Corp's T. J. Watson Research Center in Yorktown Heights, New York. **Michael Wolf** is a physics PhD student at the Technical University at Brunswick in Brunswick, Germany. **Andrew Doherty** is a postdoctoral scholar in the quantum optics group at the California Institute of Technology.

Quantum entanglement: A modern perspective

It's not your grandfather's quantum mechanics. Today, researchers treat entanglement as a physical resource: Quantum information can now be measured, mixed, distilled, concentrated, and diluted.

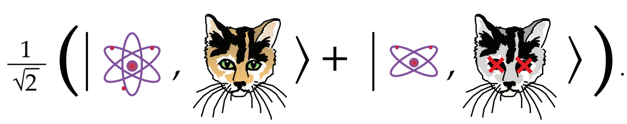
Barbara M. Terhal, Michael M. Wolf, and Andrew C. Doherty

"If two separated bodies, each by itself known maximally, enter a situation in which they influence each other, and separate again, then there occurs regularly that which I have [just] called entanglement of our knowledge of the two bodies."
—Erwin Schrödinger (translation by J. D. Trimmer)

Erwin Schrödinger coined the word *entanglement* in 1935 in a three-part paper[1] on the "present situation in quantum mechanics." His article was prompted by Albert Einstein, Boris Podolsky, and Nathan Rosen's now celebrated EPR paper that had raised fundamental questions about quantum mechanics earlier that year.

Einstein and his coauthors had recognized that quantum theory allows very particular correlations to exist between two physically distant parts of a quantum system; those correlations make it possible to predict the result of a measurement on one part of a system by looking at the distant part. On that basis, the EPR paper argued that the distant predicted quantity should have a definite value even *before* being measured if the theory were to claim completeness and respect locality. However, because quantum mechanics disallows such definite values prior to measuring, the EPR authors concluded that, from a classical perspective, quantum theory must be incomplete.

Schrödinger's 1935 perspective comes closer to the modern view: The wavefunction or state vector gives us all the information that we can have about a quantum system. About entangled quantum states, he wrote, "The whole is in a definite state, the parts taken individually are not,"[1] which we now understand as the essence of pure-state entanglement. In that same 1935 article, Schrödinger also introduced his famous cat as an extreme illustration of entanglement: A cat physically isolated in a box with a decaying atom and vial of cyanide represents a quantum state having macroscopic degrees of freedom. If the atom were to decay and trigger the release of cyanide, the cat would die. The quantum-mechanical description of the system is a coherent superposition of one state in which the atom is still excited and the cat alive, and another state in which the atom has decayed and the cat is dead:



The isolated cat-trigger-atom-cyanide system as a whole is in a definite entangled state, even though the cat itself exists as a probabilistic mixture of being alive or dead.

For the three decades following the 1935 articles, the debate about entanglement and the "EPR dilemma"—how to make sense of the presumably nonlocal effect one particle's measurement has on another—was philosophical in nature, and for many physicists it was nothing more than that. The 1964 publication[2] by John Bell (pictured in figure 1) changed that situation dramatically. Bell derived correlation inequalities that can be violated in quantum mechanics but have to be satisfied within every model that is local and complete—so-called local

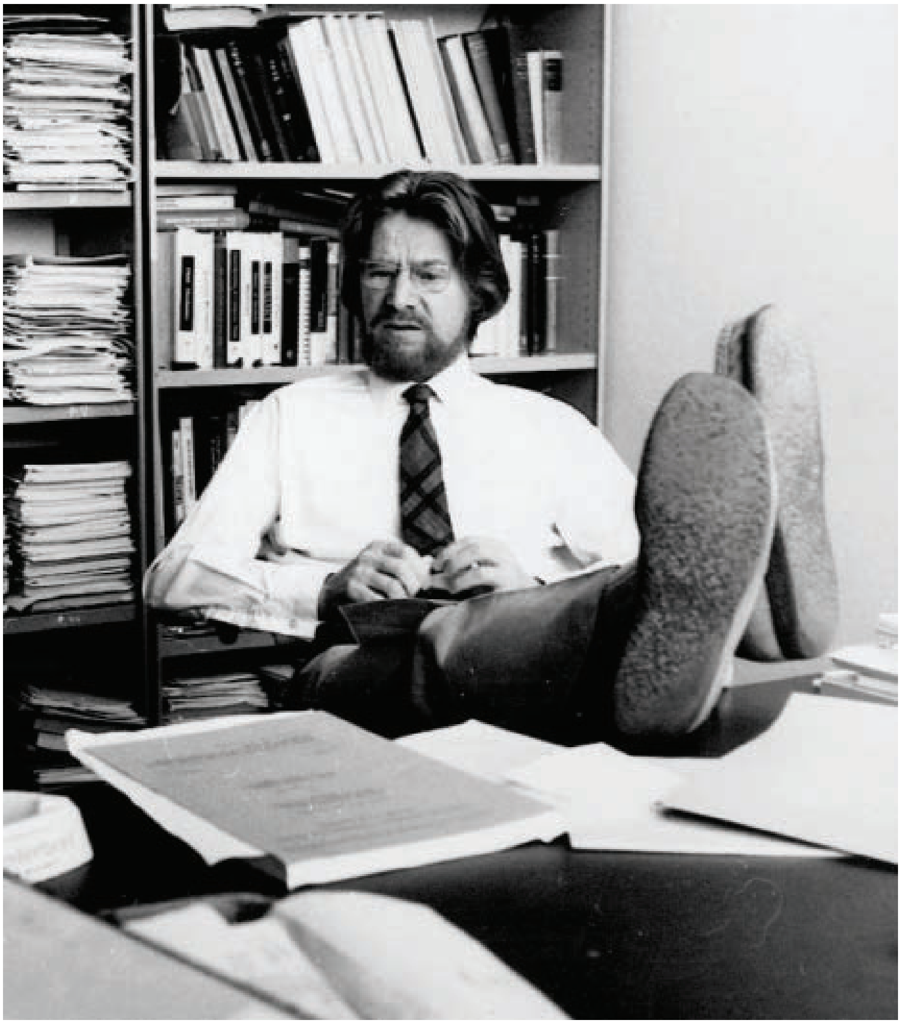


Figure 1. John Bell in repose. His seminal work clarified the difference between correlations generated by entanglement and correlations in local hidden-variable models. Nowadays, quantum information theorists exploit this difference to create advantages that communication protocols using entanglement have over classical ones.

hidden-variable models. Bell's work made it possible to test whether local hidden-variable models can account for observed physical phenomena. Early and ongoing recent experiments[3] showing violations of such Bell inequalities have invalidated local hidden-variable models and lend support to the quantum-mechanical view of nature. In particular, an observed violation of a Bell inequality demonstrates the presence of entanglement in a quantum system.

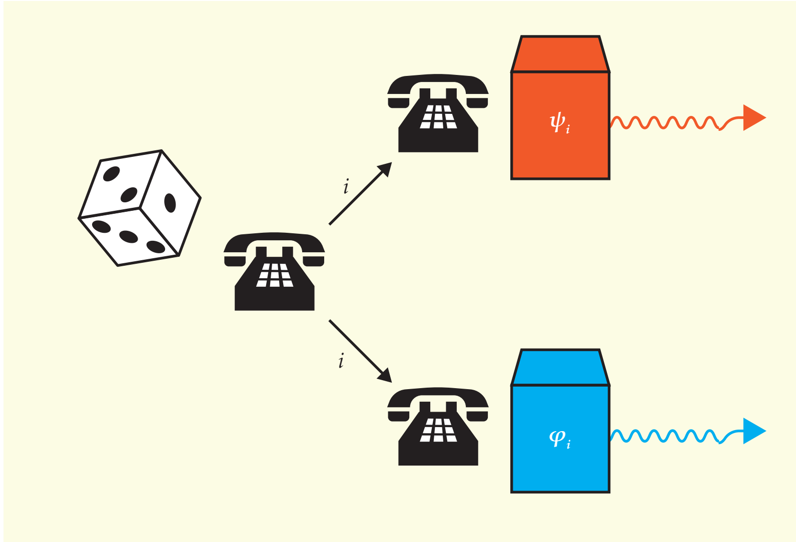
In 1995, Peter Shor at AT&T Research discovered that, for certain problems, computation with quantum states instead of classical bits can result in tremendous savings in computation time.[4] He found a polynomial-time quantum algorithm that solves the problem of finding prime factors of a large integer. To date, no classical polynomial-time algorithm for this problem exists.

Shor's breakthrough generated an avalanche of interest in quantum computation and quantum information theory. In this context, a modern theory of entanglement has begun to emerge: Researchers now treat entanglement not simply as a paradoxical feature of quantum mechanics, but as a physical resource for quantum-information processing and computation. A whole zoo of various kinds of pure and mixed entangled states may be prepared—well beyond the simple pure-state superpositions that Schrödinger envisioned. And those mixed entangled states may be measured, distilled, concentrated, diluted, and manipulated. A surprisingly rich picture of entanglement is now taking shape.

Entanglement for the 21st century

The discovery of quantum teleportation by IBM researcher Charles Bennett and five collaborators in 1993 marks the starting point of the modern view. In quantum teleportation (see the article by Charles Bennett in PHYSICS TODAY, October 1995, page 24), an experimentalist, Alice, wishes to send an unknown state $|s\rangle = \alpha|0\rangle + \beta|1\rangle$ of a two-level quantum system to another experimentalist, Bob, in a distant laboratory. The two-level system could refer, for example, to the polarization of a single photon, the electronic excitation of an effective two-level atom, or the nuclear magnetic spin of a hydrogen atom. Alice and Bob do not have the means of directly transmitting the quantum system from one place to another (for photons, this could be the case when using a high-loss optical fiber), but let us imagine that they do share an entangled

Figure 2. Classically correlated, or separable, quantum states are generated when Alice (red) and Bob (blue) locally prepare quantum states ψ_i and φ_i depending on the result *i* of a classical random number generator. If the correlations in a bipartite quantum state *cannot* be produced by such a procedure, then the state is considered entangled.



state. Consider the case in which Alice and Bob each have one spin of a shared singlet state of two spin-½ particles $|\Psi^-\rangle = 1/\sqrt{2}(|\uparrow,\downarrow\rangle - |\downarrow,\uparrow\rangle)$, also called an EPR pair. Alice can transmit her spin $|s\rangle$ to Bob by performing a certain joint measurement on her spin state $|s\rangle$ and her half of the EPR pair. She tells Bob the result of her measurement and, depending on her information, Bob rotates his half of the EPR pair to obtain the state $|s\rangle$. The teleportation protocol demonstrates that the resources of classical communication and the sharing of prior EPR entanglement are sufficient to transmit an unknown spin state $|s\rangle$. (For the experimental realization, see Physics Today, February 1998, page 18.)

The spin-singlet EPR state that Alice and Bob share in quantum teleportation is called a maximally entangled state. Even though the two spins together constitute a definite pure state, each spin state is maximally undetermined or mixed when considered separately. In mathematical terms, Alice's local density matrix—obtained by tracing over Bob's spin degrees of freedom, $\mathrm{Tr}_B(|\Psi^-\rangle\langle\Psi^-|)$—has equal probability for spin up and spin down. In keeping with Schrödinger's understanding of entanglement, one measures the amount of entanglement in a general pure state φ in terms of the lack of information about its local parts. The von Neumann entropy $S(\rho) = -\mathrm{Tr}(\rho \log\rho)$ is used as a measure of that information. In other words, the entropy of entanglement E of the pure state φ is equal to the von Neumann entropy of, say, Alice's density matrix $\rho = \mathrm{Tr}_B|\varphi\rangle\langle\varphi|$.

Mixed entanglement

In the quantum teleportation scenario, we imagined, unrealistically, that Alice and Bob shared an EPR pair free of noise or decoherence. More generally, Alice and Bob have quantum systems that interact directly or through another mediating quantum system—like Rydberg atoms in a laser cavity that interact via photons, or two ions in an ion trap that interact through phonon modes of the trap.[5] A related example of interest in quantum computation is an array of interconnected ion traps, each holding a small number of ions that are coupled by traveling photons or by ions that are moved between the traps.[6] The interaction, or "quantum link," between a pair of systems is subject to noise or decoherence through photon loss or heating of the phonons, for instance. For simplicity, assume that Alice and Bob's local operations on the quantum systems—operations on the ions in a single trap, say—are perfect, and their exchange of classical information is also perfectly noise free. That idealization enables one to measure the strength of the quantum link between the systems.

An essential question is, Given unavoidable noise levels, is it possible to establish a strong quantum link—a set of pure EPR pairs, in other words—between two systems? If it is, then the noise is weak enough to permit the error-free exchange of quantum information between the systems, since the teleportation through the generated EPR pairs will be error free. That capability may come at a certain cost, determined by the amount of noisy interaction required to generate an EPR pair. If it is not possible to generate EPR pairs, that decoherence in the system imposes a fundamental limitation on our ability to perform quantum information processing.

The possibility of generating shared EPR entanglement in noisy environments is not only of interest in entanglement theory, but is crucial for the realization of long-distance quantum communication[7] and possibly large-scale quantum computation. For example, it was recently shown[8] that fault-tolerant quantum computation can be achieved in the presence of very high noise levels in the interaction link—a link can have an error rate of two-thirds—between quantum systems that are "small" in a particular sense, if one assumes that local quantum processing on each end is (almost) error free.

Pure quantum states have their entanglement quantified fairly intuitively by considering the degree of local "mixedness" or entropy. However, mixtures of entangled and unentangled states are murkier: Recognizing which mixtures are still entangled may be difficult. So, just what physical systems can we call "entangled"? An operational description—expressing entanglement in terms of its negation—is helpful. Suppose that Alice and Bob, working in their distant labs, each receive the same random number over the phone. Depending on the random number, each of them locally prepares a certain quantum state. The physical state of their whole system, expressed as a density matrix, typically exhibits correlations between the two systems. However, those correlations would be classical, since they arise from classical random numbers. A quantum state that can be prepared in this way *over the phone* is called "unentangled" or separable, and such a state can be mathematically expressed as a mixture of unentangled pure states (see figure 2). Conversely, a state is "entangled" if it cannot be prepared over the phone, but requires coherent interaction between the two systems or the transmission of superpositions of quantum states.

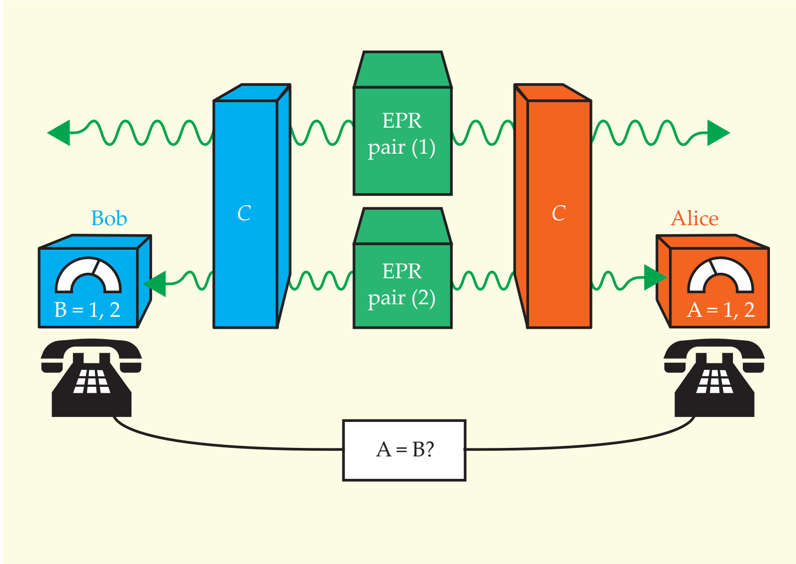


Figure 3. Entanglement distillation—the conversion of many noisy less-entangled states into fewer, more-entangled ones. Imagine two Einstein-Podolsky-Rosen pairs that pick up noise when their parts are transmitted to Alice and Bob. Assume that the noisy states are still entangled. Alice and Bob can use the following protocol to increase the entanglement: (i) each of them applies a controlled-shift operation C to the states sent to them; the shift operation acts on the upper green system (1) and the lower green system (2). For i and $j = 0,1$, $C|i\rangle_1 \otimes |j\rangle_2 = |i\rangle_1 \otimes |i \oplus j\rangle_2$, where $\oplus$ means addition modulo 2. (ii) Each measures the lower EPR(2) pair in the $\{|0\rangle, |1\rangle\}$ basis and they compare their results. If the outcomes are the same (checked over the phone), the entanglement in the first EPR pair will have increased. The various ways of iterating the procedure to distill more entangled states are known as recurrence protocols[9] or entanglement pumping.[8]

Measures of noisy entanglement

For mixed states, it is harder to establish a good measure of entanglement, since such a measure has to distinguish between entropy arising from classical correlations in the state—a state of thermal equilibrium, for example—and local entropy due to purely quantum correlations. Two measures of entanglement that have explicit physical meaning in the processing of quantum information have emerged from the quantum-link notion just described: the entanglement cost $E(\rho)$ of a quantum state and the distillable entanglement $D(\rho)$ of a quantum state, first defined in reference 9.

Assume that Alice and Bob have created, using their noisy link, many (n) shared copies of an entangled quantum state ρ; we denote such a collection as $\rho^{\otimes n}$. To distill some EPR pairs from those copies, Alice and Bob perform several rounds of local, error-free operations to their parts of the copies and communicate their measurements (or other classical data) to each other. Such a protocol is called entanglement distillation; figure 3 illustrates one round of such a scheme. The aim is to produce fewer states that are, however, more entangled than the initial ones. Ideally, the protocol produces nearly perfect maximally entangled EPR pairs in the limit of a large number of input states $\rho^{\otimes n}$ with $n \to \infty$. The distillable entanglement $D(\rho)$ is then the number of such EPR pairs that can be extracted per copy of ρ in this asymptotic limit.

The reverse process also has physical meaning. What is the smallest number k of EPR pairs that Alice and Bob initially need to create a set of n copies of ρ for $n \to \infty$ by local error-free operations? This asymptotic ratio k/n is the second measure of entanglement, the entanglement cost $E(\rho)$.

Reversible and irreversible manipulation

Attentive readers may have noticed a quirk in our notation: The formalism uses the same symbol E to denote both the entanglement cost for general states and the entropy of entanglement for pure states. The notation coincidence is harmless since the creation cost of a pure state equals the local entropy of entanglement E. Furthermore, for a pure state φ, it turns out that $E(\varphi) = D(\varphi)$ (see box 1 on page 44). Physically, this means that the process of entanglement dilution—converting EPR pairs into lesser entangled pure states φ—can be reversed without loss of entanglement. The reverse process is called entanglement concentration and it produces $D(\varphi)n = E(\varphi)n$ EPR pairs from an initial supply of n states φ.

For mixed states, D is believed to be generically less than E, which implies that the preparation of mixed states

Box 1. The law of large numbers and interconvertible entanglement

Suppose one generates a bit string of length k by k realizations of a binary random variable that takes the value 1 with probability p and the value 0 with probability $1 - p$. By the law of large numbers, among the k-bit strings there exist typical strings that have a high probability of occurring—ones in which approximately $pk + O(\sqrt{k})$ bits are 1 and $(1 - p)k$ bits are 0, for instance—and atypical strings, the string of all zeros, for example. The key to understanding the protocols of pure state entanglement concentration and dilution[18] is this typicality of sequences.

Suppose Alice and Bob would like to convert some shared entangled states $\varphi^{\otimes k}$ with $|\varphi\rangle = \sqrt{p}|11\rangle + \sqrt{1-p}|00\rangle$ to a smaller supply of Einstein-Podolsky-Rosen (EPR) pairs Ψ^-. In other words, suppose they wish to *concentrate* their entanglement in fewer qubits. Alice and Bob will each do a local measurement that counts the number of ones in a bit string (but not which bits are ones). With high probability—approaching 1 as $k \to \infty$—they both have pk as their measurement outcome, indicating that pk bits out of k are one. With that outcome, Alice and Bob will have obtained a quantum state whose local density matrix has eigenvalues that are all equal which number approximately

$$\binom{k}{pk} \approx 2^{kH(p) - O(\sqrt{k})} = 2^{kH(\varphi) - O(\sqrt{k})}.$$

Here, $H(p)$ is the Shannon entropy of the distribution $(p, 1 - p)$. Thus Alice and Bob can make a local change of basis (a unitary rotation) and truncate the dimension of the space to 2^n and obtain $n \approx kE(\varphi) - O(\sqrt{k})$ EPR pairs.

In the reverse process of dilution, one converts n EPR pairs into k states φ by quantum teleporting an approximation φ_k to $\varphi^{\otimes k}$ from Alice to Bob using the EPR pairs. In the local spectrum of the state $\varphi^{\otimes k}$, there exist typical eigenstates, with approximately pk bits equal to 1 and $(1 - p)k$ bits equal to 0, and atypical eigenstates. The approximation φ_k is obtained from $\varphi^{\otimes k}$ by truncating the local spectrum to the eigenstates that are in this *typical* subspace. The dimension of this typical subspace is $2^{kH(p) + O(\sqrt{k})}$ and therefore the state φ_k can be teleported using $n \approx kE(\varphi) + O(\sqrt{k})$ EPR pairs. In the limit of large k, the conversion ratios k/n of the dilution and concentration protocols will be the same and thus prove the asymptotic reversibility of the processes.

from EPR pairs is a process involving an irreversible loss of entanglement. Curiously, the $D < E$ conjecture has only been proven for some special classes of mixed states.[10]

In 1998, the Horodecki family of Gdańsk, Poland (father Ryszard and sons Paweł and Michał), identified a class of entangled states that exhibit an extreme form of irreversibility. They proved that no entanglement can be distilled ($D = 0$) from these "bound entangled states."[11] And for a large set of states from that class, irreversibility was established by proving that entanglement is required to prepare the states $E > 0$.

Consider the metaphor illustrated in figure 4. If EPR pairs were nodes connected by lines or strands that represent quantum correlations between particles, then one could think of mixed entanglement as entanglement in which the strands are simply mixed up. The mixing may make it hard to reconstruct which particle of Alice is entangled with which particle of Bob. Cutting a few strands reduces the clutter, but every line cut represents an EPR pair lost (compare this process with the distillation protocol in figure 3). Bound entangled states are those mixtures that are so thoroughly mixed up that every single line has to be cut to remove the noise or clutter from the system. But, when every line is cut, no entanglement remains to be distilled.

"Black holes" of quantum information

Because the modern theory of entanglement treats quantum states as physical resources for processing information, one might consider them hierarchically. A simple and ideal world would have only two classes of quantum states: unentangled, classically correlated states that are useless as a resource in quantum teleportation and don't violate any Bell inequalities, and entangled states whose distillation rate D measures their usefulness in quantum teleportation. If the distillation rate D is nonzero, one can distill from such states some EPR pairs, known to violate Bell inequalities.

Bound entanglement tells us that life is not so simple. Bound entangled states are costly ($E > 0$), but useless in various quantum-information-processing protocols like teleportation. Furthermore, there is evidence that bound entangled states do not violate any Bell inequalities.

In those two senses, bound entangled states are the "black holes" of quantum information theory. Entanglement goes in but is impossible to recover. And like black holes in the theory of gravitation, bound entangled states test the limits of our understanding and puzzle us by their intrinsic irreversibility.

Bound entanglement and partial transposition

In what sense are bound states so thoroughly mixed up that no entanglement at all can be extracted? Bound entangled states behave intrinsically differently from every other entangled state: They remain physical under the *unphysical* operation of partial transposition.

Researchers realized that they could characterize entanglement in terms of how states behave under certain unphysical operations.[12] In 1996, Asher Peres at the Technion–Israel Institute of Technology in Haifa, Israel, noted that matrix transposition is just such an unphysical operation when applied to entangled states. Taking the transpose of a system's density matrix produces another density matrix—a physically valid result. And taking the transpose of, say, Bob's part of an unentangled state $\psi_A \otimes \psi_B$ yields another physically valid quantum state, since each part of the quantum state can transform separately; ψ_A is not changed, and the density matrix of ψ_B is transposed. But when applied to part of a pure entangled state, matrix transposition produces an unphysical result. (For details, see box 2 on page 46.)

Peres conjectured that partial transposition was the defining criterion for entanglement. In other words, all entangled states—pure or mixed—should map onto unphysical states by partial matrix transposition, and all unentangled states will remain physical under the same operation.

Remarkably, the truth of that conjecture depends on the dimension of the underlying Hilbert spaces or phase spaces. If one considers the state of two spin-½ particles, the polarization degrees of freedom of two laser beams, or two modes of a light field having a Gaussian Wigner function, then, indeed, all entangled states map onto unphysical states by partial transposition. However, for two spin-one (or higher-dimensional system) particles or a Gaussian light field with at least two modes for both Alice and Bob, that is no longer true in general; there exist entangled mixed states that pass the "partial transpose" test and have therefore lost an essential property of entanglement.

The loss of that property is precisely what the Horodecki family showed would lead to a zero distillation rate D. Entangled states that pass the partial transpose test are the bound entangled states in which the entanglement is forever locked or "bound" inside.

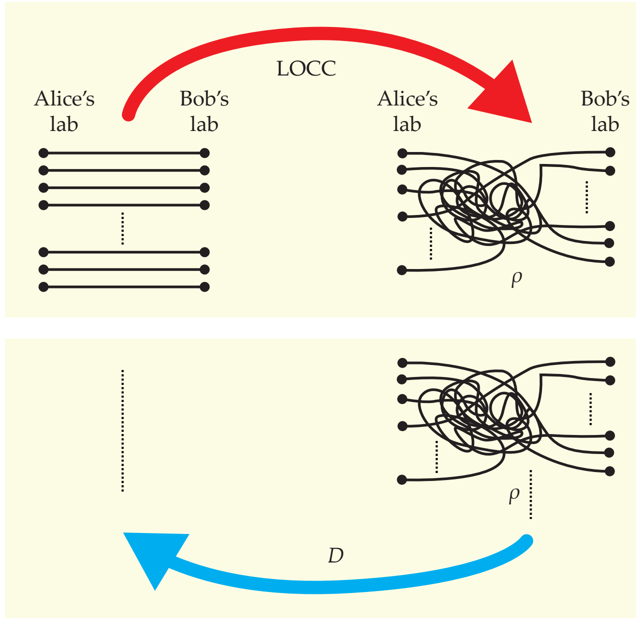


Figure 4. Irreversibility in noisy entanglement. An entangled EPR pair is represented by a single line or strand connecting two nodes or particles, one each in Alice and Bob's labs. The red arrow signifies the creation of some mixed entanglement from the single strands by local operations on the particles (and classical communication, on the phone, say); the process is abbreviated LOCC. One state ρ that has five particles for both Alice and Bob is created. The entanglement cost is the number of EPR pairs that is needed per single noisy state ρ, in this case 7⁄1 because Alice and Bob began with seven EPR pairs. But how does one reverse the process and extract some single strands—EPR pairs—from the noisy mixtures? The distillation rate D is the number of EPR pairs that can be extracted per noisy state ρ. *Bound entangled* mixtures are those that are so thoroughly mixed up that there are no means to extract any single strands. In other words, for a bound entangled state the blue arrow representing the distillation rate D is zero.

Entanglement witnesses

Given that entanglement can be such a subtle property of quantum states, just how can one distinguish between entangled and unentangled states? A violation of a Bell inequality has been the traditional telltale sign of entanglement in a quantum system. Examples of such experiments[3] used pairs of entangled photons created from nonlinear optical processes, especially parametric down-conversion; the polarization degrees of freedom of the emitted photons carried entanglement. Alice and Bob checked for a Bell inequality violation by using local analyzers to measure the polarization of the photons along various angles.

Unfortunately, many quantum states, including the set of bound entangled states, are not known to violate any Bell inequality. And considering the existing limitations on experimental control of quantum systems, experimentalists prefer to check for entanglement using the fewest possible local measurements. The theoretical framework of an entanglement witness, of which a Bell inequality is a particular example,[13] addresses those two issues. The defining property of an entanglement witness W is that its expectation value with respect to any unentangled state ρ is always nonnegative, $\mathrm{Tr}(W\rho) \geq 0$. At the same time, there exist entangled states σ for which $\mathrm{Tr}(W\sigma) < 0$. Measuring W on a quantum state σ and finding a negative expectation value thus establishes the entanglement of σ. The good news is that there is an entanglement witness for every entangled state; given an experimental means, any entanglement, bound or otherwise, can be detected. The bad news is that entanglement witnesses are nonlocal observables. Nevertheless, one can measure the expectation value of W by measuring the expectation value of a number of local observables W_i, such that $W = \sum_i W_i$. Research is under way to determine the minimal number of local measurements for a given witness.[14]

Bell's communication advantages

Given the framework of entanglement witnesses, what is special about Bell inequalities? Although they can be considered a type of entanglement witness, Bell inequalities do not, strictly speaking, test for entanglement but for a departure from local hidden-variable theories. Interpreted as such, Bell inequalities have taken on a whole new life in quantum-communication science. Researchers consider remote parties who have to carry out a certain task with minimal communication between them. One compares the amount of communication necessary if those parties are given shared random bits (that can be viewed as local hidden variables) or an

Box 2. Partial matrix transposition and time reversal

Matrix transposition on density matrices is closely related to the operation of time reversal—represented by an antiunitary operation—in quantum mechanics. The time-reversal operation reverses the momenta, including angular momenta and spin, of a quantum system. It is possible to represent the operation by complex conjugation that maps the momentum operator $\hat{p} = -id/dx$ onto $\hat{p} = id/dx$. Applied to Hermitian density matrices, complex conjugation is identical to matrix transposition $T : \rho \to \rho^T$ in a given basis. When applying this operation on an entire density matrix ρ, one obtains another valid density matrix $\rho^T = \rho^*$ with nonnegative eigenvalues. But when the transposition operation is applied "partially" to half of a joint system—the maximally entangled state $|\Phi\rangle_{AB} = 1/\sqrt{2}(|00\rangle + |11\rangle)$, for example—then one may no longer end up with a valid quantum state. Indeed, transposition in the $\{|0\rangle, |1\rangle\}$ basis on Bob's half of the state Φ_{AB} (and the identity operation I_A on Alice's half) gives $(I_A \otimes T)(|\Phi\rangle\langle\Phi|)$ =

$$\left(I_A \otimes T\right)\frac{1}{2}\begin{pmatrix}1 & 0 & 0 & 1\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 1 & 0 & 0 & 1\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1\end{pmatrix},$$

a matrix that has a negative eigenvalue, and is therefore unphysical. The relevance of partial transposition for detecting entanglement in a quantum state was first noted by Asher Peres in 1996. He observed that any unentangled state remains unentangled under partial transposition, because a product state $|\varphi_A\rangle \otimes |\varphi_B\rangle$ is mapped onto another product state $|\varphi_A\rangle \otimes |\varphi_B{}^*\rangle$ by transposition of Bob's system.

entangled quantum state. Sharing entangled states leads to savings in communication precisely because the correlations in quantum states cannot always be adequately described by local hidden-variable theories[15] (see the article by Andrew M. Steane and Wim van Dam, in Physics Today, February 2000, page 35).

What lies beyond

The efforts of the quantum information theorists over the past eight years would come to little if the theory were not supplemented by an ability to create and manipulate entanglement in the lab. There is a rapidly growing list of physical systems—optical and atomic systems especially—in which it is possible to prepare various kinds of entangled states. As discussed previously, the use of photonic degrees of freedom, such as polarization or momentum, has been a long-time favorite way to create entanglement.[3] Entangled states consisting of the quadrature observables of different modes of light have been prepared in optical parametric oscillators and optical fibers.[16] Entanglement in the states of motion of the valence electrons[5] of trapped ions or of Rydberg atoms in cavity quantum electrodynamics has involved up to four different atoms. Another promising avenue is the recently observed entanglement of large ensembles of atoms.[17]

This short review showcases just a few striking facets of the modern theory of entanglement. Most notably, entanglement shared between more than two subsystems is outside our scope here. The broader study of entanglement between many subsystems may lead the field to better understand the role of large-scale entanglement in quantum computation or quantum many-body systems.

We have focused on the role of entanglement in the transmission of quantum information. Entanglement also proves useful, however, when the goal is to transmit classical information as efficiently as possible. Researchers are studying many measures of mixed entanglement beyond the two most prominent measures discussed in this review. As for bound entanglement, there is some evidence that it may have a role to play as "helper" entanglement, useless by itself, but useful when combined with other sources of entanglement. For entanglement-theory overview articles that highlight the field, see volume 1 of *Quantum Information and Computation* (July 2001).

REFERENCES

1. E. Schrödinger, *Naturwissenschaften* **23**, 807 (1935); **23**, 823 (1935); **23**, 844 (1935). English trans. in *Proc. Am. Philos. Soc*. **124**, 323 (1980).
2. J. S. Bell, *Physics* **1**, 195 (1964).
3. A. Aspect, J. Dalibard, G. Roger, *Phys. Rev. Lett*. **49**, 1804 (1982); W. Tittel, G. Weihs, *Quantum Inf. Comput*. **1**, 3 (2001).
4. P. W. Shor, *SIAM J. Comput*. **26**, 1484 (1997). An earlier version is P. W. Shor, *Proceedings 35th Annual Symposium on Foundations of Computer Science*, IEEE Computer Society Press (1994), p. 124.
5. C. A. Sackett, *Quantum Inf. Comput*. **1**, 57 (2001); J. M. Raimond, M. Brune, S. Haroche, *Rev. Mod. Phys*. **73**, 565 (2001).
6. D. J. Wineland et al., *Philos. Trans. R. Soc. London Ser. A* **361**, 1349 (2003).
7. H.-J. Briegel, W. Dür, J. I. Cirac, P. Zoller, *Phys. Rev. Lett*. **81**, 5932 (1998); W. Dür, H.-J. Briegel, J. I. Cirac, P. Zoller, *Phys. Rev. A* **59**, 169 (1999).
8. W. Dür, H.-J. Briegel, *Phys. Rev. Lett*. **90**, 067901 (2003).
9. C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, W. K. Wootters, *Phys. Rev. A* **54**, 3824 (1996).
10. G. Vidal, W. Dür, J. I. Cirac, *Phys. Rev. Lett*. **89**, 027901 (2002); K. G. H. Vollbrecht, R. F. Werner, M. M. Wolf, *Phys. Rev. A* **69**, 062304 (2004).
11. M. Horodecki, P. Horodecki, R. Horodecki, *Phys. Rev. Lett*. **80**, 5239 (1998).
12. M. Horodecki, P. Horodecki, R. Horodecki, *Phys. Lett. A* **223**, 1 (1996).
13. B. M. Terhal, *Theor. Comput. Sci*. **287**, 313 (2002).
14. O. Gühne et al., *J. Mod. Opt*. **50**, 1079 (2002).
15. Č. Brukner, M. Żukowski, J.-W. Pan, A. Zeilinger, *Phys. Rev. Lett*. **92**, 127901 (2004).
16. Z. Y. Ou, S. F. Pereira, H. J. Kimble, K. C. Peng, *Phys. Rev. Lett*. **68**, 3663 (1992); A. Furusawa, J. L. Sørensen, S. L. Braunstein, C. A. Fuchs, H. J. Kimble, E. S. Polzik, *Science* **282**, 706 (1998); C. Silberhorn, P. K. Lam, O. Weiß, F. König, N. Korolkova, G. Leuchs, *Phys. Rev. Lett*. **86**, 4267 (2001).
17. B. Julsgaard, A. Kozhekin, E. S. Polzik, *Nature* **413**, 400 (2001).
18. C. H. Bennett, H. J. Bernstein, S. Popescu, B. Schumacher, *Phys. Rev. A* **53**, 2046 (1996).

PT

FROM THE ARCHIVES
NOVEMBER 2000

Daniel Gottesman is a postdoctoral fellow in the theory group at Microsoft Corp in Richmond, Washington. **Hoi-Kwong Lo** is chief scientist and senior vice president for research and development at MagiQ Technologies Inc (http://www.magiqtech.com) in New York, New York.

FROM QUANTUM CHEATING TO QUANTUM SECURITY

Daniel Gottesman and Hoi-Kwong Lo

For thousands of years, code-makers and code-breakers have been competing for supremacy. Their arsenals may soon include a powerful new weapon: quantum mechanics.

Cryptography—the art of code-making—has a long history of military and diplomatic applications, dating back to the Babylonians. In World War II, the Allies' feat of breaking the legendary German code Enigma contributed greatly to their final victory. Nowadays, cryptography is becoming increasingly important in commercial applications for electronic business. Sensitive data such as credit card numbers and personal identification numbers (PINs) are routinely transmitted in encrypted form. Quantum mechanics is a new tool for both code-breakers and code-makers in their eternal arms race. It has the potential to revolutionize cryptography both by creating perfectly secure codes and by breaking standard encryption schemes.

The best-known application of cryptography is secure communication,[1] illustrated in figure 1. Suppose Alice would like to send a message to Bob, but there is an eavesdropper, Eve, who is wiretapping the channel. To prevent Eve from knowing the message, Alice may perform encryption—that is, transform the message to something that is unintelligible to Eve—during the communication. On receiving the message, Bob inverts the transformation and recovers the message.

Bob's advantage over Eve lies in his knowledge of a secret, commonly called the key, that he shares with Alice. The key tells him how to decode the message. Consider this example (in the style of Cold War espionage thrillers):

> *The rumble of Soviet tanks shook the Prague hotel room (number 117) as secret agent John Blond finished decoding his orders from his superior, N. He tore the used page from the codebook and immediately burned it with his lighter.*

Blond is using a perfectly unbreakable cipher, a "one-time pad." The secret codebook allows N and Blond to share a long secret binary string—the key—before Blond leaves on his mission. Whenever N would like to send a message to Blond, she first converts it to binary. She then takes the exclusive-OR (XOR) between each bit of the message and the corresponding key bit to generate the encrypted message, which is transmitted over a public channel. An enemy can intercept the encrypted message, but without the key, it is incomprehensible gibberish, offering no clue to the contents of the original message. On the other hand, Blond, by looking up the key in the codebook, can recover the original message by taking the XOR between the encrypted message and the key. Blond immediately burns the used page of the codebook to prevent it from falling into enemy hands in the future.

Key distribution problem

> *John Blond finally snapped shut the codebook and sighed. He had been on duty in Czechoslovakia for so long that his codebook was getting thin. He knew his days in Prague would soon be over: N would have to recall him before he used up his whole codebook. Blond recalled master cryptographer R's remonstration: "This is no joking matter, double-one seven.* Never *reuse the one-time pad."*

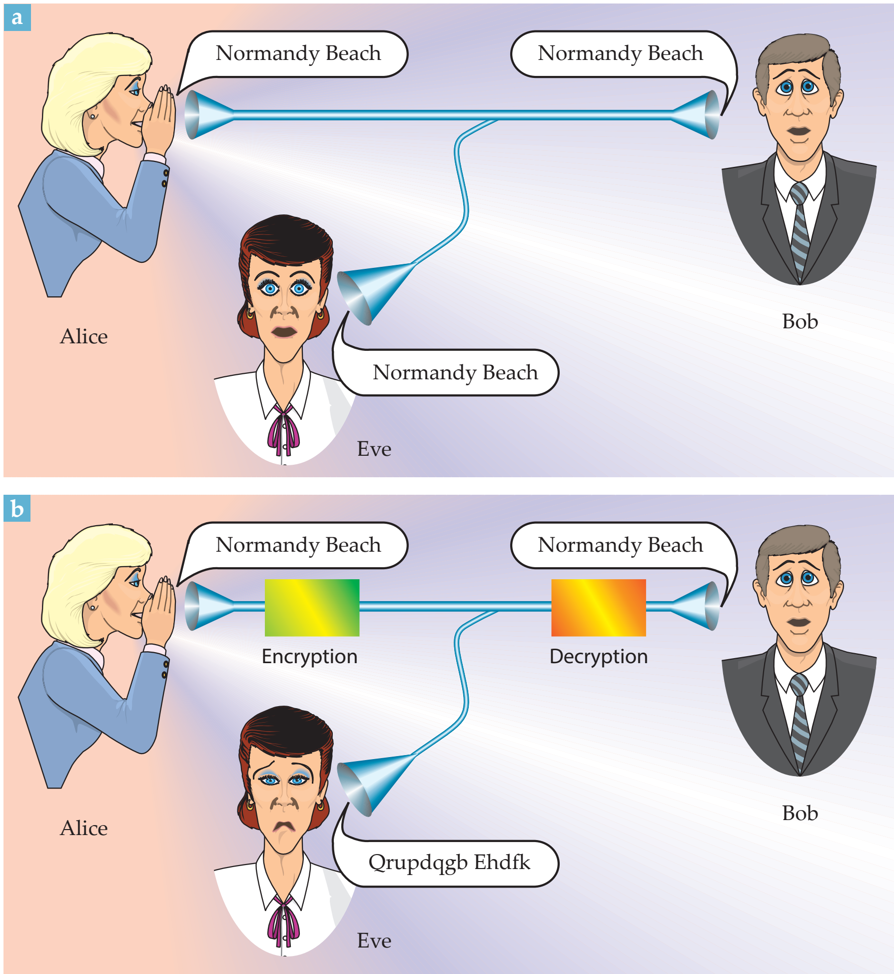


FIGURE 1. COMMUNICATION security. **(a)** Alice sends a message to Bob through a communication channel, but an eavesdropper, Eve, is wiretapping. **(b)** A message is encrypted by Alice using an encryption key. The encrypted message, the *ciphertext*, is now unintelligible to Eve. Bob, who has the same key as Alice, can decrypt the ciphertext and recover the original message. (The code used in this figure is not very secure. Try breaking it yourself; the solution is at the end of the article.)

R was serious for a good reason. The reuse of keys by the Soviet Union (due to the manufacturer's accidental duplication of one-time pad pages) enabled US cryptanalysts to unmask the atomic spy Klaus Fuchs in 1949.[2] When the key for a one-time pad is used more than once, enemy cryptanalysts have the opportunity to look for patterns in the encrypted messages that might reveal the key. Nevertheless, excellent cryptosystems (known as symmetric cryptosystems) that reuse the key have been developed. The longer the key, the more secure the system. For instance, a widely used system is the Data Encryption Standard (DES), which has a key length of 56 bits. No method substantially more efficient than trying all 2^{56} values of the key is known for breaking DES. It is still conceivable, however, that some yet unknown clever algorithm could defeat DES and its cousins.

For top-secret applications, therefore, the one-time pad is preferable. Blond's predicament illustrates the drawback of the one-time pad: When the secret key is used up, the code cannot be used until the sender and receiver get together to share a new secret key. Sending a courier with a new codebook into the Prague Spring is a dangerous and unreliable business. Even if the courier arrives, Blond and N can never be sure that the codebook was not copied during its journey.

This issue is known as the "key distribution problem." A possible solution is public key cryptography. Instead of a single long key shared between the sender and receiver, public key cryptography uses two sorts of keys: one public key, which is known to the world, and one private key, known only to the receiver. Anyone with the public key can send secret messages, but only someone who knows the private key can read them. The important defining feature of public key cryptography is that, even knowing the encryption key, there is no known *computationally efficient* way of working out what the decryption key really is. As an example, the security of the best-known public key cryptosystem, RSA, relies on the difficulty of factoring large integers (see figure 2).

Public key cryptography can be used for another important task: digital signatures. A digital signature exchanges the role of the keys used in public key cryptography: The private key is used to generate a signature and the public key is used to verify it. Only someone with the private key could have created the signature.

Quantum code-breaking

Both DES and RSA rely on an unproven assumption: There is no fast algorithm to determine the secret key. For instance, RSA is believed to be secure because mathematicians throughout the world have worked very hard to break it, steadily producing modest improvements in factoring algorithms, but without groundbreaking success. With only modest increases in key size, users of RSA can easily keep ahead even of the exponential growth in computing power over the years.

Quantum mechanics changed this. In 1994, Peter Shor of AT&T Laboratories invented a *quantum* algorithm for efficient factoring of large numbers.[3] The state of a quantum

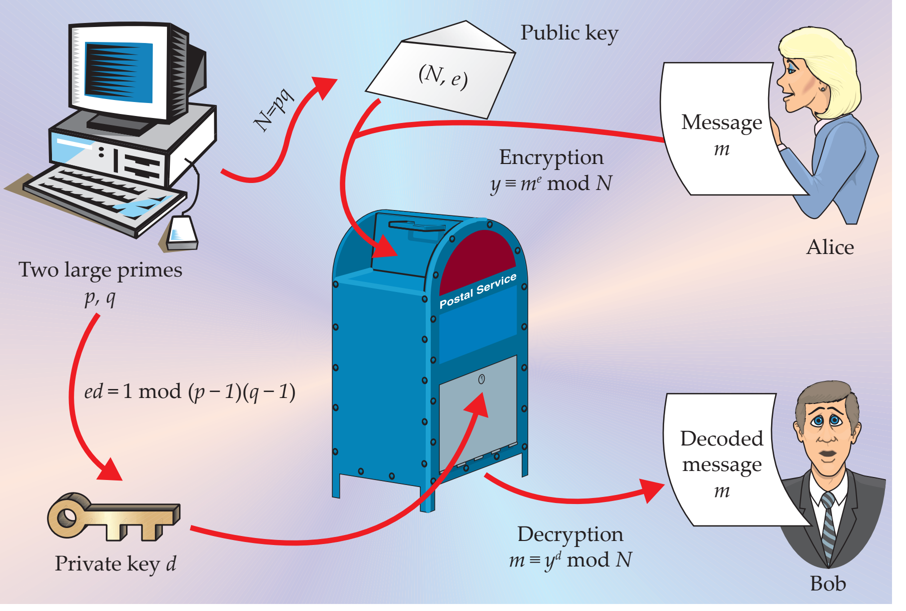


FIGURE 2. THE RSA PUBLIC KEY cryptosystem. The best-known public key system is called RSA, after its inventors Ronald Rivest, Adi Shamir, and Leonard Adleman. It is based on modular arithmetic over a large base N that is the product of two large primes p and q. If x is relatively prime to N, the Euler–Fermat theorem tells us that $x^r \equiv 1 \bmod N$, where $r = (p-1)(q-1)$. The public key is a pair of numbers (N, e), and the private key is d, with $ed \equiv 1 \bmod r$ (that is, $ed = kr + 1$ for some integer k). To encrypt a message m, the sender (Alice) computes $y \equiv m^e \bmod N$. To decrypt the message y, the receiver (Bob) computes $y^d \bmod N \equiv m^{ed} \bmod N \equiv m$. For this step, Bob has to know the private key d. Anyone can send Bob an encrypted message, but only Bob can decrypt it.

computer is a superposition of exponentially many basis states, each of which corresponds to a state of a classical computer of the same size. By taking advantage of interference and entanglement in this system, a quantum computer can perform in a reasonable time some tasks that would take ridiculously long on a classical computer. Shor's discovery propelled the then-obscure subject of quantum computing into a dynamic and rapidly developing field, and stimulated scores of experiments and proposals aimed toward building quantum computers.

Another remarkable discovery was made by Lov Grover of Bell Laboratories, Lucent Technologies, who in 1996 invented a quantum searching algorithm[4] (see PHYSICS TODAY, October 1997, page 19). To find one particular item among N objects requires checking $O(N)$ items classically. With Grover's algorithm, a quantum computer need only look up items $O(\sqrt{N})$ times. It can be used to radically speed up the exhaustive key search of DES (that is, trying all 2^{56} possibilities).

If a quantum computer is ever constructed in the future, much of conventional cryptography will fall apart! To provide the same security, the key lengths of symmetric schemes like DES would have to be doubled due to Grover's algorithm. The most commonly used public key schemes are RSA and others based on discrete logarithms or elliptic curves; Shor's algorithm breaks all of them. Even if it is decades until a sufficiently large quantum computer can be built, this is a matter of current concern: Some data, such as nuclear weapons designs, will still need to remain secret, and it is important that today's secret messages cannot be decoded tomorrow.

Quantum code-making

Even if DES and RSA do fall apart, the one-time pad remains a perfectly unbreakable cipher even against a quantum computer. However, as previously discussed, it has a serious catch: the key distribution problem. It presupposes that Alice and Bob share a key that is secret and as long as the message. There is no way to guarantee that in practice. Trusted couriers can be bribed or even intercepted without their knowledge. More generally, classical signals are distinguishable. An eavesdropper can reliably read the signals without changing them. Therefore, in classical physics there is nothing, in principle, to prevent an eavesdropper from wiretapping the key distribution channel passively.

Fortunately, quantum mechanics helps to make codes as well as break them.[5] (See also Charles Bennett's article, "Quantum information and computation," PHYSICS TODAY, October 1995, page 24.) The Heisenberg uncertainty principle dictates that it is fundamentally impossible to know the exact values of complementary variables such as a particle's momentum and its position. This apparent limitation imposed by quantum mechanics can be a powerful tool in catching eavesdroppers. The central idea is to use nonorthogonal quantum states to encode information. More concretely, the essence of quantum cryptography can be understood in a single question: Given a single photon in one of four possible polarizations (↔, ↕, ⤢, or ⤡), can one determine its polarization with certainty? Surprisingly, the answer is no. The rectilinear basis (↔ and ↕) and the diagonal basis (⤢ and ⤡) are incompatible, so the Heisenberg uncertainty principle forbids us from simultaneously measuring both. More generally, experiments distinguishing nonorthogonal states, even if only partially reliable, will disturb the states.

The key distribution problem can be partially solved by quantum mechanics using the idea of quantum key distribution (QKD). The first and best-known protocol, usually called "BB84" because it was published in 1984 by Charles Bennett and Gilles Brassard,[6] is described in the box on page 51. In a

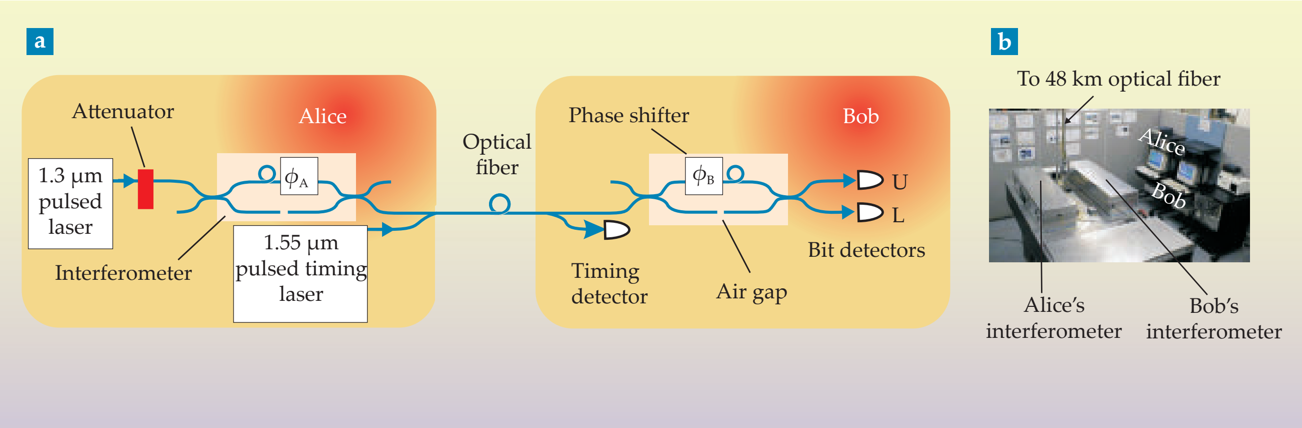


FIGURE 3. EXPERIMENTAL QUANTUM KEY DISTRIBUTION. (a) Schematic of the experiment at Los Alamos[8] that implements the protocol known as B92 (see the box on page 51) over 48 km of optical fiber. A laser with a wavelength of 1.3 μm, attenuated to approximate a single-photon source, is the source of the key bits. Its output is passed through Alice's interferometer. The two nonorthogonal quantum states used in the B92 protocol are realized as two possible settings for the phase delay ϕ_A in one branch of the interferometer. To measure the state, Bob passes the photon through his interferometer, adding one of two possible phase shifts ϕ_B, and detects the photon in one of the two bit detectors. A bright pulse from a second laser tells Bob when to expect a photon from Alice. Air gaps in both interferometers allow Alice and Bob to tweak the optical path lengths to keep properly synchronized. **(b)** The actual setup of the experiment. The two boxes in the foreground are the interferometers, connected to each other only through 48 km of optical fiber. (Figure courtesy of Richard Hughes.)

prototypical QKD protocol, Alice sends some nonorthogonal quantum states to Bob, who makes some measurements. Then, by talking on the phone (which need not be secure), they decide if Eve has tampered with the quantum states. If not, they have a shared key that is guaranteed to be secret. Note that Alice and Bob must share some authentication information to begin with; otherwise, Bob has no way to know that the person on the phone is really Alice, and not a clever mimic. The key generated by QKD can subsequently be used for both encryption and authentication, thus achieving two major goals in cryptography.

Experimental QKD

QKD is an active experimental subject. The first working prototype, constructed in 1989 at IBM in Yorktown Heights, New York, transmitted quantum signals over 32 cm of open air.[7] Since then, various groups—including those led by Paul Townsend at the British Telecommunications Photonics Technology Research Centre (now part of Corning), Jim Franson of Johns Hopkins University, Nicolas Gisin and Hugo Zbinden of the University of Geneva, and Richard Hughes of Los Alamos National Laboratory—have made important contributions. A primary focus has been a series of impressive experiments over commercial optical fibers. The world record distance for QKD,[8] at the time of writing, is about 50 km. One of the long-distance experiments, performed at Los Alamos, is depicted in figure 3.

Most experiments to date have used variants of either the BB84 or B92 schemes (see the box), although recently three groups—one led by Paul Kwiat of Los Alamos, Gisin and Zbinden's group at Geneva, and a collaboration led by Anton Zeilinger of the University of Vienna and Harald Weinfurter of the University of Munich—have independently implemented protocols based on entangled pairs of particles, also known as Einstein-Podolsky-Rosen or EPR states. In the BB84 and B92 schemes, typically a single-photon source is simulated using attenuated coherent states—on average, only a fraction of a photon is actually sent. With additional losses in the fiber, very few arriving laser pulses actually contain a photon. This low yield does not interfere much with key distribution, however, since only the photons that reach Bob are used in the protocol. The key is generally encoded in either the polarization or the phase of the photon. Error rates in the photons actually received are usually a few percent.

For commercial applications in, say, a local area network environment, it is useful for a quantum cryptographic system to be integrated into a passive multiuser optical fiber network and its equipment to be miniaturized. Townsend's group has done much work in this area.[9] For point-to-point applications, the Geneva group has devised a so-called "plug and play" system that automatically compensates for polarization fluctuations.[10] Such systems might someday convey secret information between government agencies around Washington, DC, or connect bank branches within a city.

QKD has also been performed in open air,[11] during daylight, with a current range of about 1.6 km. Ambitious schemes to perform a ground-to-satellite QKD experiment have been proposed. If successful, quantum cryptography may be used to ensure the security of command control of satellites from control centers on the ground.

Future experiments will aim to make QKD more reliable, to integrate it with today's communications infrastructure, and to increase the distance and rate of key generation.

Another ambitious goal is to produce a quantum repeater using techniques of quantum error correction. Such an accomplishment will require substantial technical breakthroughs, but would allow key distribution over arbitrarily long distances.

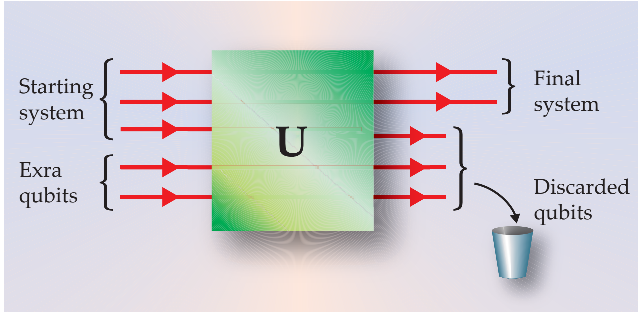


FIGURE 4. THE CHURCH OF THE LARGER HILBERT SPACE. In cryptography—and other areas—the quantum mechanical description of explicitly quantum aspects (such as single-photon polarizations) can be expanded to include other parts as well, including measurements and random number generation. This alternative treatment consists of three steps. First, the original quantum system—which might consist of two-level quantum bits (called "qubits"), for example—is augmented with an additional system. In this expanded Hilbert space, all operations are unitary and can be combined into a single quantum mechanical step (here denoted by "U"). Part of the output of the transformation is thrown away, leaving only the final quantum system of interest. Using quantum mechanics to simulate classical computations and working with pure quantum states allows the most generalized treatment of a problem and simplifies the task of determining whether a given protocol is secure. Describing a protocol in the Church of the Larger Hilbert Space does not change the protocol in any way; it merely provides a new and sometimes simpler way of looking at the system.

Is QKD secure?

While experiments in QKD forged ahead, the theory developed more slowly. A clever Eve can adopt many possible strategies to fool Alice and Bob, including subtle quantum attacks entangling all of the particles sent by Alice. Taking all possibilities into account, as well as the effects of realistic imperfections in Alice and Bob's apparatus and channel, has been difficult. A long series of partial results has appeared over the years, addressing restricted sets of strategies by Eve,[12] but only in the past few years have complete proofs appeared.

One class of proofs, by Dominic Mayers[13] and subsequently by others, including Eli Biham and collaborators and Michael Ben-Or,[14] attacks the problem directly and proves that the standard BB84 protocol is secure. Another approach, by one of us (HKL) and H. F. Chau,[15] proves the security of a new QKD protocol that uses quantum error-correcting codes.[5] (For more on quantum error correction, see John Preskill, "Battling decoherence: The fault-tolerant quantum computer," Physics Today, June 1999, page 24.) This approach allows one to apply *classical* probability theory to tackle a *quantum* problem directly. It works because the relevant observables all commute with each other. While conceptually simpler, this protocol requires a quantum computer to implement. The two approaches have been unified by Peter Shor and John Preskill,[16] who showed that a quantum error-correcting protocol could be modified to become BB84 without compromising its security.

The proof of the security of QKD is a fine theoretical result, but it does not mean that a real QKD system would be secure.[17] Some known and unknown security loopholes might prove to be fatal. Apparently minor quirks of a system can sometimes provide a lever for an eavesdropper to break the encryption. For instance, instead of producing a single photon, a laser may produce two; Eve can keep one and give the other to Bob. She can then learn what polarization Alice sent without revealing her presence. There are various possible solutions to this particular problem; it is the unanticipated flaws that present the greatest security hazard. Ultimately, we cannot have confidence that a real-life quantum

The BB84 protocol

In the best-known quantum key distribution (QKD) scheme, BB84, Alice sends Bob a sequence of photons, each independently prepared in one of four polarizations (↔, ↕, ⤢, or ⤡). For each photon, Bob randomly picks one of the two (rectilinear and diagonal) bases to perform a measurement. He keeps the measurement outcome secret. Now Alice and Bob publicly compare their bases. They keep only the polarization data for which they measured in the same basis. In the absence of errors and eavesdropping by Eve, these data should agree.

To test for tampering, they now choose a random subset of the remaining polarization data, which they publicly announce. From there they can compute the error rate (that is, the fraction of data for which their values disagree). If the error rate is unreasonably high—above, say, 10%—they throw away all the data (and perhaps try again later). If the error rate is acceptably small, they perform error correction and also "privacy amplification" to distill a shorter string that will act as the secret key. These steps essentially ensure that their keys agree, are random, and are unknown to Eve.

Other QKD schemes have also been proposed. For example, Artur Ekert of the University of Oxford suggested one based on quantum mechanically correlated (that is, entangled) photons, using Bell inequalities as a check of security. In 1992, Charles Bennett of IBM proposed a simple QKD scheme, called B92, that uses only two nonorthogonal states.

cryptographic system is secure until it has withstood attacks from determined real-life adversaries. Traditionally, breaking cryptographic protocols has been considered to be as important as making them—the protocols that survive are more likely to be truly secure. The same standard will have to be applied to QKD.

Post–Cold War applications

There are many problems beyond secure communication that can be addressed by cryptography.

> *Alice and Bob are considering going on a date, but neither is willing to admit their interest unless the other is also interested. How can they decide whether or not to date without letting slip any unnecessary information?*

This dating problem can be phrased as the problem of computing a function $f(a, b) = ab$, where a and b are single bits held respectively by Alice and Bob (0 = not interested, 1 = interested). Problems like this can be solved classically using variants of public key cryptography, which we know might be rendered insecure by quantum computers. By exchanging quantum states, can Alice and Bob solve the above dating problem with absolute security?

There are many possible functions f that two people might wish to compute together, too many to consider each of them individually. Instead, cryptographers rely on a suite of primitive operations that can be combined to build more complex functions. One important protocol is called bit commitment, and it is the electronic equivalent of a locked box. Alice chooses a bit, 0 or 1, and writes it on a piece of paper, which she deposits in the box. She gives the box to Bob but keeps the key. She cannot change what she wrote, and without the key, Bob cannot open the box. But at some later point, Alice can give Bob the key and reveal her bit. By itself, bit commitment is useful mostly for debunking professional psychics, but it serves as a useful building block for more interesting functions.

Consider the following bit commitment scheme[6] proposed by Bennett and Brassard: If Alice wishes to commit to a 0, she sends Bob a polarized photon in the rectilinear basis; if she wishes to commit to a 1, she sends Bob a polarized photon in the diagonal basis. In either case, Alice flips a coin to decide which of the two polarizations to send. Bob has no way to tell which basis Alice used; no matter which bases Alice and he choose, Bob would measure a random value. But when Alice unveils her bit, telling Bob which of the four states she sent, Bob can measure in the appropriate basis to verify that Alice is telling the truth. If she lies about which basis she used, Bob has a 50% chance of finding out. If the protocol is repeated many times, Alice's chance of successfully cheating is abysmally small.

This protocol is secure against a classical cheater, who does not have much ability to store and manipulate quantum states. But as Bennett and Brassard recognized, a *quantum* cheater can break the protocol. Suppose that instead of picking a specific state and sending it to Bob, Alice creates an entangled pair of photons, $(|{\leftrightarrow}{\updownarrow}\rangle - |{\updownarrow}{\leftrightarrow}\rangle)/\sqrt{2}$ (an EPR pair), and sends the second photon to Bob, keeping the first one. She stores the quantum state of the first photon and delays measuring it. Suppose that when the time comes for Alice to open the commitment, she decides she would like the committed bit to read 0, which requires her to specify a state in the rectilinear basis. Because of the entanglement, Alice knows that if she and Bob measure in the same basis, they will get opposite results. Therefore, she can measure her photon in the rectilinear basis and tell Bob he has the opposite polarization, and she will always be right.

If Alice instead wishes the committed bit to read 1, she needs a state in the diagonal basis. But $(|{\leftrightarrow}{\updownarrow}\rangle - |{\updownarrow}{\leftrightarrow}\rangle)/\sqrt{2} \equiv (|{\swarrow\!\!\nearrow}{\nwarrow\!\!\searrow}\rangle - |{\nwarrow\!\!\searrow}{\swarrow\!\!\nearrow}\rangle)/\sqrt{2}$. So Alice can measure her particle in the diagonal basis and again be sure that Bob's measurement outcome will be opposite to hers. Quantum cheating allows Alice to change her mind at the last minute without being caught by Bob, thus totally defeating the purpose of bit commitment.

Nonetheless, more sophisticated schemes for quantum bit commitment were proposed, and for a long time were believed to be secure. Eventually, the bubble burst and it was shown that the above quantum cheating strategy, which uses EPR nonlocality and delayed measurements, can be generalized to break all two-party quantum bit commitment schemes.[18] If Alice and Bob hold one of two pure quantum states that are indistinguishable to Bob, then Alice, acting unilaterally, can change one to the other. Therefore, the two basic requirements of bit commitment—that Bob does not know the bit and that Alice cannot change it—are fundamentally incompatible with quantum mechanics.

The strength of the proof lies in its generality. The idea is to treat the whole system as if it were quantum mechanical, extending the part that was originally quantum to include any dice, measuring devices, and classical computations that appear in the protocol. From this point of view, the original protocol is equivalent to a purely quantum one, with some of the output being thrown in the trash (see figure 4). Note that throwing something away can never help a cheater, so we might as well assume that the state shared by Alice and Bob is the pure quantum state that is completely determined by the protocol. That assumption substantially reduces the complexity of the problem. It is not difficult to show that when Alice and Bob hold a pure state, quantum bit commitment is impossible.

Following the fall of quantum bit commitment, other important basic quantum cryptographic protocols have also been proven to be insecure by one of us (HKL), thus leaving the field in a shambles. What is left?

Some potential applications in cryptography are too similar to bit commitment and cannot be done at all quantum mechanically. Others have more modest goals and *can* be solved by quantum protocols. For instance, Lior Goldenberg, Lev Vaidman, and Stephen Wiesner of Tel Aviv University have proposed a method of "quantum gambling," in which a cheater must pay a large fine if caught. The majority lie in a middle ground—we do not know whether they can be solved. The dating problem is an example. Many approaches to it tread too near bit commitment and are doomed to failure, but it's possible there are others, as yet undiscovered, that do not.

Physics today, cryptology tomorrow

Quantum computers are still on the drawing boards, and quantum cryptographic systems are only prototypes. Still, there are a number of reasons for thinking about quantum cryptology today. Unlike other cryptosystems, the security of QKD is based on fundamental principles of quantum mechanics, rather than unproven computational assumptions. QKD eliminates the great threat of unanticipated advances in algorithms and hardware breaking a widely used cryptosystem. Small-scale QKD systems are well within the capabilities of today's technology, and commercial systems could be available within a few years (although whether such systems are widely adopted depends on many nonacademic factors, including cost).

Furthermore, grappling with the problems posed by quantum protocols can give us insight into more general questions about quantum mechanical systems in many fields of physics. For instance, one reason it is hard to analyze protocols and attacks is that they frequently involve a combination of quantum and classical behaviors. In considering bit commitment, though, it was possible to replace classical parts of the protocol with a quantum description, an approach that is useful for many problems inside and outside the field of quantum cryptography. This fully quantum treatment is sometimes called the Church of the Larger Hilbert Space, following John Smolin of IBM. All quantum operations, including measurements, are unitary when considered as acting on a larger Hilbert space (figure 4).

Finally, quantum mechanics changes the world of cryptology, and it is important to know what the new terrain will look like to decide on cryptographic standards that may last for decades. In a world where quantum computers and communication are commonplace, today's most widespread public key cryptosystems would no longer work; in the worst case, perhaps *no* public key cryptosystem will work. If so, symmetric cryptosystems and QKD would partially fill the gap, allowing secure communication. Unfortunately, digital signatures would fail as well, meaning important communications would need to be notarized by a trusted third party.

Of course, QKD and symmetric cryptosystems are not useful in situations in which Alice and Bob have never met. Solving this problem would probably require a quantum cryptographic center, which could verify the identity of both of them. The center would have to be known and trusted by both Alice and Bob.

Problems beyond secret communication and digital signatures are a mixed bag. Many, such as bit commitment and perhaps the dating problem, would be impossible, whereas others, such as quantum gambling, could be carried out with complete security.

This is just one of a number of possible futures. Perhaps some new or existing public key cryptosystems will survive quantum computation, or perhaps new public key systems will be developed that can only run on a quantum computer. Perhaps quantum computers will always remain difficult to build (we believe that this is unlikely), and public key cryptography will remain widespread, despite its potential flaws. Only time will tell who benefits more from quantum cryptology: the code-makers or the code-breakers.

Decoding the message in figure 1

The code is a "Caesar's cipher," in which each letter is shifted by a fixed number of places in the alphabet. In this case, the shift is three places.

REFERENCES

1. B. Schneier, *Applied Cryptography*, 2nd ed., Wiley (1996).
2. C. J. Phillips, "What Made Venona Possible?" in *Venona: Soviet Espionage and the American Response 1939–1957*, R. L. Benson, M. Warner, eds., National Security Agency, Central Intelligence Agency (1996), p. xv.
3. P. W. Shor, in *Proceedings 35th Annual Symposium on Foundations of Computer Science*, S. Goldwasser, ed., IEEE Computer Science Press (1994), p. 124.
4. L. K. Grover, in *Proceedings of the 28th Annual ACM Symposium on the Theory of Computing*, ACM Press (1996), p. 212.
5. H.-K. Lo, S. Popescu, T. Spiller, eds., *Introduction to Quantum Computation and Information*, World Scientific (1998).
6. C. H. Bennett, G. Brassard, in *Proceedings of the International Conference on Computers, Systems, and Signal Processing*, IEEE Press (1984), p. 175. The first paper on quantum cryptography was written by Stephen Wiesner around 1970, but it remained unpublished until 1983: S. Wiesner, *Sigact News* **15**(1), 78 (1983).
7. C. H. Bennett et al., *J. Cryptology* **5**, 3 (1992).
8. R. J. Hughes, G. L. Morgan, C. G. Peterson, *J. Mod. Opt.* **47**, 533 (2000); P. D. Townsend, *Opt. Fiber Technol.* **4**, 345 (1998) and references therein.
9. P. D. Townsend, *Nature* **385**, 47 (1997).
10. A. Muller et al., *Appl. Phys. Lett.* **70**, 793 (1997).
11. B. C. Jacobs, J. D. Franson, *Opt. Lett.* **21**, 1854 (1996); W. T. Buttler et al., *Phys. Rev. Lett.* **84**, 5652 (2000).
12. For a review, see C. H. Bennett, P. W. Shor, *Science* **284**, 747 (1999) and references therein.
13. D. Mayers, *J. Assoc. Comput. Mach.* **48**, 351 (2001). A preliminary version appeared in D. Mayers, *Advances in Cryptology—Crypto '96*, N. Koblitz, ed., Springer (1996), p. 343.
14. See, for example, E. Biham et al., in *Proceedings of the 32nd Annual ACM Symposium on the Theory of Computing*, ACM Press (2000), p. 715.
15. H.-K. Lo, H. F. Chau, *Science* **283**, 2050 (1999). This paper built on earlier work on quantum privacy amplification: D. Deutsch et al., *Phys. Rev. Lett.* **77**, 2818 (1996); **80**, 2022 (1998) (erratum).
16. P. W. Shor, J. Preskill, *Phys. Rev. Lett.* **85**, 441 (2000).
17. G. Brassard, N. Lütkenhaus, T. Mor, B. Sanders, *Phys. Rev. Lett.* **85**, 1330 (2000).
18. D. Mayers, *Phys. Rev. Lett.* **78**, 3414 (1997); H.-K. Lo, H. F. Chau, *Phys. Rev. Lett.* **78**, 3410 (1997). PT

Philip Bucksbaum is professor of physics and applied physics at Stanford University in Stanford, California, and at the Stanford Synchrotron Radiation Laboratory in Menlo Park. He directs the university's PULSE (Photon Ultrafast Laser Science and Engineering) Center.

Photon science and quantum control

Philip H. Bucksbaum

Recent advances in laser technology have hastened developments in other fields—precision measurement, atomic cooling, gravitational-wave sensing, quantum computing, cryptography, and many more. Like the laser itself, those fields may transform society.

In the past decade, no fewer than four Nobel Prizes—one in chemistry and three in physics—were awarded for work done on the science of atoms and molecules interacting with laser light. The remarkable efforts that led to those prizes mark the latest stage in the gradual transformation of lasers, still less than a half-century old, from sources of directed photons for spectroscopy to something more: tools for the control of the quantum world. New lasers are reaching previously unknown regimes of intensity, stability, wavelength, and pulse duration. And those developments are driving new cross-disciplinary research in atomic, solid-state, and x-ray physics; quantum optics; physical chemistry; and laser engineering.

The 10 laureates who shared the four Nobel Prizes in the past 10 years have helped shape the new discipline of laser-driven quantum control. Nine come from the field of atomic, molecular, and optical physics. A recent National Research Council interim report[1] describes some of the research opportunities in the AMO field and points to anticipated advances in six areas: precision measurements, ultracold matter, ultra-high-intensity and short-wavelength lasers, ultrafast control, nanophotonics, and quantum information science. Keeping in mind Yogi Berra's caveat, "It's tough to make predictions, especially about the future," I describe here some of the opportunities in those areas and try to convey the excitement accompanying them and the rapid growth in photon science generally; the growth appears likely to continue for years.

Clocks and lights

What time is it, really? The most recent Nobel Prize in Physics was awarded in part to John Hall and Theodor Hänsch for advances in one of the oldest subjects in physics: measuring the passage of time (see PHYSICS TODAY, December 2005, page 19). At the most fundamental level, physicists still don't know what time is, although we surely know how to quantify it more precisely than any other physical property. And things that we can compare to our most accurate clocks—the spin rate of a pulsar, for example, or the frequency of an atomic transition many light-years away—may be the sources of new discoveries. The source of the next great discovery in physics is mere speculation, of course; but the remarkable improvement in atomic-clock precision is a fact, and improvements to the state of the art continue.

Ultrafast pulsed-laser sources, developed in the past decade for such applications as optical digital communication and the investigation of transient phenomena, have found new uses because of their special spectral properties: The pulses in those lasers contain an optical-frequency comb that stretches from the near-UV to the near-IR wavelength range. The frequency comb enables direct and precise conversion between optical frequencies and the microwave frequencies of atomic clocks. We can now literally count the optical-frequency (10^{15} Hz) waves and thereby measure optical-frequency ratios more precisely than ever.

Time reversal—through the looking glass. A clock appears to run backwards, or counterclockwise, when viewed in a mirror. But, of course, it doesn't run more slowly. The mirror's backward minute is precisely the same duration as a forward minute. But what if a physical process went backwards? Could we even tell? That is not just a whimsical remark, but a serious question about the fundamental forces of nature. We have known for years that neutral K mesons created in high-energy collisions display a tiny bit of time-reversal difference, or asymmetry, but we don't yet know why. We don't even know whether ordinary matter has the same property, and we have very few ways to seek the answer.

The measurement of atomic electric-dipole moments (EDMs) could provide clues (see the article by Norval Fortson, Patrick Sandars, and Steve Barr in Physics Today, June 2003, page 33). EDMs in atoms cannot exist in a perfect time-symmetric world. Indeed, no one has ever observed a permanent electric-dipole moment in an atom, even though with today's instrumental sensitivity, a relative charge displacement between an atom's electrons and nucleus as small as a trillionth the width of the nucleus would be detectable. Nevertheless, some of the most promising theories that offer explanations for particle-physics time-reversal violations also predict atomic EDMs not much smaller than the present limit. Similar advancements in precision measurements can also search for matter–antimatter *CPT* violations—that is, violations to symmetry under the combined operation of charge conjugation, parity, and time reversal—or even violations of Einstein's famous principle of relativity at levels far more sensitive than ever before possible.

Position sensing—where are we? Laser navigation gyros are not new. They are optical interferometers that detect motion by measuring changes in the relative length of two optical paths. The same operating principle is behind gravitational-wave observatories such as the Laser Interferometer Gravitational-Wave Observatory, which are pushing the concept of relative length to extreme limits: LIGO can measure length changes as small as a hundredth of a proton diameter over the length of a football field. A future space-based gravitational-wave observatory named LISA (Laser Interferometer in Space Antenna) would be even more sensitive.

Cold and fast

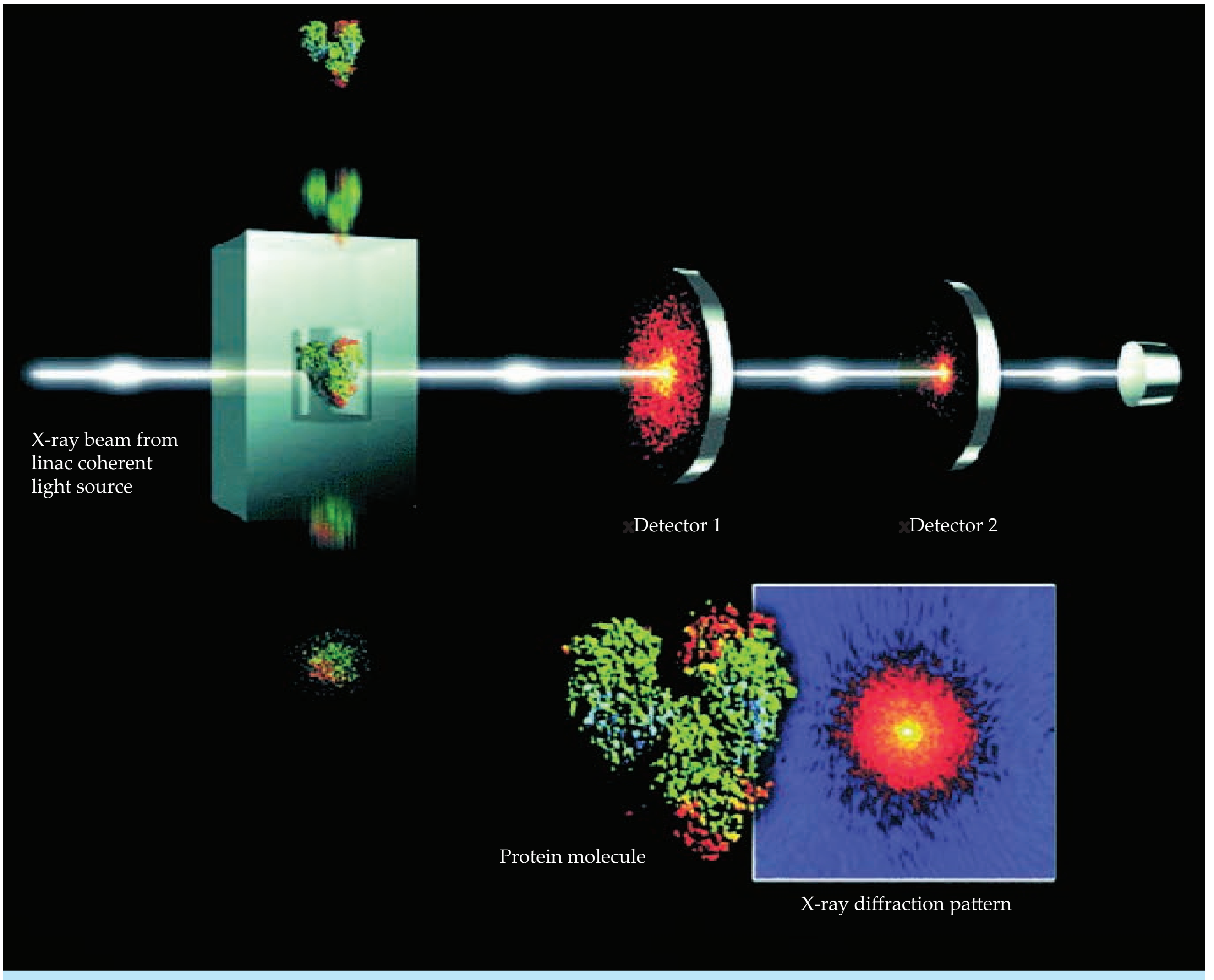
Where is the coldest place in the universe? Boulder, Colorado. That's not just the punchline of a joke about weather on the Front Range. The low-temperature record for any macroscopic object is held by atomic Bose–Einstein condensates at about one billionth of 1 K, far colder than the 2.7-K cosmic background temperature of deep space.

Scientists investigating ultracold atoms captured two of the past decade's Nobel Prizes, and the field itself was the most spectacularly successful research area in atomic, molecular, and optical physics over the period. Meanwhile, ultracold atoms are beginning to have important research and technological applications. For example, ultracold atoms form the inner workings of the highest-accuracy atomic clocks, such as the NIST-F1 atomic fountain clock, which are partly responsible for the advances in precision time measurements.

Quantum interferometers. Interferometry and ultracold gases come together in a revolutionary combination to improve the sensitivity of interferometers by replacing the light waves with quantum matter waves. Matter-wave interferometers could provide huge improvements in the accuracy of navigational systems and systems that measure changes in local gravity. Right now, for example, gravimeters based on laser interferometers that suspend mirrors as test masses are used to explore local gravity anomalies on Earth. The technology has found applications in oil exploration and other endeavors. In quantum matter-wave interferometers, the acceleration of gravity is applied not to the mirror but to the wave itself. These types of interferometers would so greatly extend the sensitivity of gravimeters that it would be possible to detect tiny underground nonuniformities in Earth's gravity from airborne laboratories.

Ultracold gases are a new arena for quantum control as well. They have been used to simulate some of the quantum many-body physics in condensed matter systems. They can be confined to one or two dimensions and can be controlled to mimic periodic structures of crystalline solids. The approach has led to a new research field of quantum simulators, which is beginning to attract attention from the solid-state physics community.

The next superpower. Lasers, already the brightest lights on Earth, will reach peak powers beyond a quadrillion watts in the next decade. That's more power concentrated in the peak of a high-powered laser pulse—if only for a few femtoseconds—than is consumed by all the nations on Earth. That capability will bring to the laboratory bench the plasma conditions that exist in stellar interiors, and high-energy-density physics is poised to make great advances because of it. High-powered lasers also produce very large field gradients that can be used to accelerate particles; that



The linac coherent light source, a free-electron laser under construction at SLAC, will emit femtosecond pulses a billion times brighter than any other existing x-ray source once it begins operating in 2009. The schematic here pictures the diffraction from a protein molecule that falls through the beam. Scientists will merge a series of diffraction patterns of the molecule in many different positions. The resulting three-dimensional reconstruction will reveal the structures of proteins that cannot be crystallized and studied any other way.

approach to particle acceleration will be yet another active research field in the coming decade.

A new kind of high-powered laser is about to switch on for the first time: The x-ray free-electron lasers under construction in the US, Europe, and Asia, which will be about a billion times brighter than any other source operating in the x-ray region, represent a merging of the most advanced technical capabilities of high-energy particle accelerators and x-ray light and laser sources. They derive their energy from relativistic electrons compressed to femtosecond bunches in linear accelerators. The x-ray bursts from such lasers are expected to be brief enough to capture motion on the atomic scale in molecules and bright enough to record an image of a biological molecule like a virus or a protein (see the figure). The first x-ray FEL is scheduled to start operations at SLAC in 2009. International teams have already assembled to plan research on these revolutionary machines.

Ultra-ultrafast pulses. Ordinary molecules at room temperature rotate in picoseconds; they vibrate and collide in femtoseconds. Thus, much can be learned from femtosecond lasers that excite or probe matter. The 1999 Nobel Prize in Chemistry recognized achievements in this fast-moving field (see PHYSICS TODAY, December 1999, page 19). Subpicosecond lasers are now commercially available and ultrafast pump–probe techniques have become routine. I've already mentioned the contributions of ultrafast lasers to precision measurements, but there is much more to the rapidly expanding field.

One of the most intriguing challenges for the future is to push for still-shorter pulse durations. In the past few years new sources have produced pulses shorter than one femtosecond—an achievement that heralds the age of attosecond science. Just as femtosecond pulses are ideal tools for exploring atomic motion in molecules, attosecond pulses go one step further and can be used to explore electron motion within atoms. Electron motion creates and destroys bonds, the physical basis for chemistry and materials science. The new capabilities are likely to produce new physical insights.

Subfemtosecond pulses are already being used in atomic physics. The rearrangement of electrons in atoms following the excitation of a core-level electron is known to take place at very short time scales, often under one femtosecond. The inaugural experiments with attosecond pulses observed that process three years ago (see Physics Today, April 2003, page 27).

Learning from the quantum world. New advances in optical pulse shaping enable the generation of light pulses whose shape, polarization, intensity, and frequency can all be controlled at will. Such total control of light can be translated into near-total control of the quantum state of a molecule. Many examples now exist of mode-selective chemistry, in which optical pulses are tailored to push a chemical reaction to favor one product or another, simply by changing the pulse shape. The search for pulse shapes that can control reactions to favor the rare over the common can proceed via computer control in learning feedback systems. Those systems are capable of producing hundreds of different pulse shapes per second, performing similar experiments with each one, and analyzing and ranking the results (see the article by Ian Walmsley and Herschel Rabitz, Physics Today, August 2003, page 43).

There are many more examples in which lasers are used to control the quantum world. Quantum computing, photonic crystals, quantum cryptography, and negative-index materials are each new, rapidly growing fields with tremendous potential to expand science. Some of those areas may even transform society, just as the laser itself has done. Certainly the new research areas that explore control of the quantum world are experiencing a decade of rapid progress. As Yogi said, predictions about the future may be difficult; but the general prediction that much rich research in quantum control lies ahead seems a safe bet.

This essay is adapted from a talk given at the 75th-anniversary celebration of the American Institute of Physics in May 2006.

REFERENCE

1. Committee on AMO 2010, National Research Council, *Controlling the Quantum World of Atoms, Molecules, and Photons: An Interim Report*, National Academies Press (2005). PT

FROM THE ARCHIVES
MARCH 2021

Christine Middleton is a senior editor at PHYSICS TODAY.

What's under the hood of a quantum computer?

Christine Middleton

Many layers lie between everyday users and the delicate, error-prone hardware they manipulate.

When most people sit down at their computers to work, they're thinking about all the things they need to get done; far from mind is any consideration of how their keystrokes and mouse clicks are translated into logic operations and electrical signals. That separation between hardware and user interface is the product of decades of development. Now quantum computer developers are navigating similar terrain.

The quantum computing stack is everything that lies between a user and the physical qubits. The stack needs to perform essential functions; for instance, it must facilitate user interaction, turn inputs into hardware manipulation, and correct for numerous error sources. (For more about quantum architectures, see the article by Anne Matsuura, Sonika Johri, and Justin Hogaboam, PHYSICS TODAY, March 2019, page 40.) There's no one right way to divide those tasks into discrete levels, though, and researchers and technology companies are still pursuing different visions for future quantum architectures.

On page 28 of PHYSICS TODAY's March 2021 issue, Harrison Ball, Michael Biercuk, and Michael Hush present the quantum computing stack proposed by Q-CTRL, the quantum technology company founded by Biercuk. The authors explain in detail how the functionality of a quantum firmware layer—one component of a quantum computer—is critical for managing qubit errors. Here we explain what happens in the rest of the layers of a quantum computer.

Qubit hardware

Classical computers store information as bits that each take a value of 0 or 1. Underlying those bits are field-effect transistors that act as switches; each can take a value of either 0 or 1 depending on whether the switch is on or off. At the most basic level, everything a computer does—save information, execute calculations, run programs—is just manipulating the values of those billions of bits with small electrical voltages.

Quantum computers instead rely on qubits that can be in one of two states, $|0\rangle$ or $|1\rangle$, or a linear superposition of those two states, $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, in which the coefficients α and β are related to the probability of finding the qubit in each state.

Why is it useful for qubits to exist in a superposition of states? It comes down to how much information you can store in n independent bits compared with the same number of qubits that are linked through entanglement—a phenomenon that cannot be described by classical physics.

Each classical bit requires only one value to describe whether it's on or off, so n bits represent n binary digits. At first glance it may seem like qubits would have $2n$ numbers akin to those binary digits because each has two coefficients, α and β. But the advantage can be even bigger than that; describing a quantum state made of n qubits can require up to 2^n coefficients.

Consider, for example, a three-qubit system. Each qubit can be in the state $|0\rangle$ or $|1\rangle$, so there are eight possible states that the system could be measured in—and eight coefficients

A REFRIGERATION SYSTEM houses an IBM Q System One quantum computer. (Photo from IBM.)

describing the probability of each state. The more qubits in a system, the bigger the informational advantage over classical bits. Taking advantage of that huge computation space is no mean feat, though; writing algorithms that benefit from qubit properties is a challenge because, although computations may manipulate 2^n parameters, they output just n values—the final qubit states. (For more on that, see the section on quantum algorithms on page 61.)

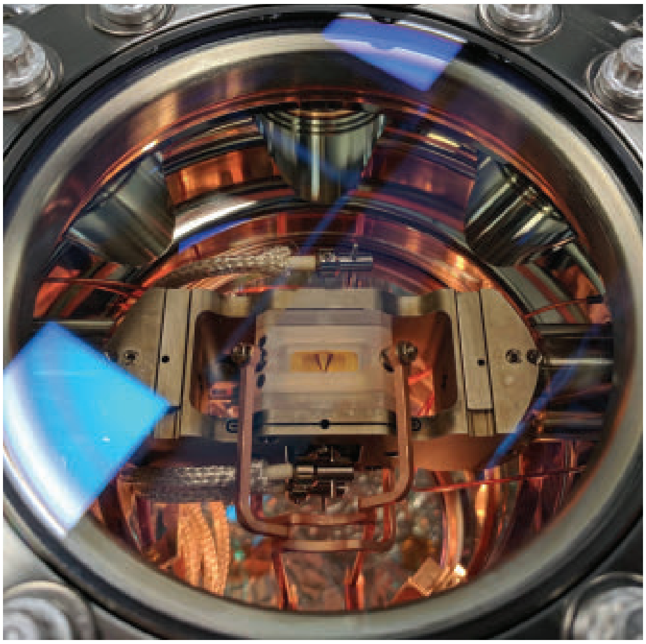
Whereas classical computing has largely settled on one type of bit hardware, qubits still come in many varieties. Any two-level quantum system—a nuclear spin, a photon's polarization, or a quantum dot's spin, to name a few—can be used as a qubit. The usefulness of a particular system, however, depends on things such as how easily the qubits are to manipulate and entangle, how long they remain in desired quantum states, and how prone they are to having their states destroyed by outside noise.

One popular example of qubit hardware implementation is trapped-ion qubits. In those designs, charged particles are confined by electromagnetic traps, and a valence electron moving between two states acts as the qubit. Hyperfine transitions in neutral atoms can serve the same function (see the article by David Weiss and Mark Saffman, Physics Today, July 2017, page 44), as can electron spin-flips in quantum dots (see the article by Lieven Vandersypen and Mark Eriksson, Physics Today, August 2019, page 38).

Some of the most well-known quantum computers, including those from IBM and Google, rely on superconducting transmon qubits. Transmons are superconducting islands of charge in which the difference between $|0\rangle$ and $|1\rangle$ is the presence of an additional Cooper pair of bound electrons.

Quantum firmware

Qubits are prone to errors. All sorts of environmental factors—thermal fluctuations, electromagnetic radiation, magnetic fields—can knock a qubit out of its intended state. That degradation of information is known as decoherence and can occur in a fraction of a second. Despite the use of refrigeration to reduce thermal fluctuations, decoherence eventually creeps in and produces hardware errors, like accidentally flipping a qubit's state from $|0\rangle$ to $|1\rangle$. (The commonly used refrigeration systems, like the one shown above from IBM, are what many people picture when they imagine a quantum computer.) The number of operations that can be performed with a qubit is limited by the qubit's decoherence time. Moreover, every set of qubit hardware has its own unique deviations from ideal



GOLD ELECTRODES produce a trap for charged particles in this ion-trap quantum computer. The electrodes are structured to permit microwave and laser-beam access. The entire system is housed in an ultrahigh-vacuum chamber. (Photo by Michael J. Biercuk, University of Sydney.)

performance (see the article by Ian Walmsley and Herschel Rabitz, PHYSICS TODAY, August 2003, page 43).

But higher levels in the quantum computing stack can't be expected to account for such system-to-system variation; a programmer needs to be able to request that an operation be performed without knowing about the underlying hardware's quirks. (Imagine if every computer required personalized software!)

Quantum firmware creates a virtualized version of the qubit hardware for the higher levels of the computing stack. It is focused on all the low-level quantum control tasks that can be used to stabilize the hardware and mitigate errors. For instance, it uses information about the hardware to autonomously define error-resistant versions of the RF or microwave pulses that act on the qubits to execute quantum logic operations.

Although quantum firmware alone doesn't solve the problem of hardware errors, it is particularly efficient at suppressing slow drifts in hardware parameters such as a qubit's resonant frequency; those drifts are a dirty secret of quantum computing hardware. That capability makes firmware a strong complement to quantum error correction protocols that are better suited to dealing with stochastic errors.

For more on the quantum firmware layer, see the PHYSICS TODAY article by Ball, Biercuk, and Hush referred to earlier.

Hardware-aware quantum compiler

In classical computers, compilers take higher-level instructions for tasks that need to be completed and translate those instructions into a series of operations that are performed using the underlying hardware. The same thing happens in a quantum computer.

The hardware-aware quantum compiler, also known as a transpiler, is responsible for figuring out how to complete a set of logic operations in a manner that accounts for the physical connections between qubits. Although physical qubits can't easily be moved, the states of two qubits can be swapped for an effective rearrangement. The transpiler works out how to implement an arbitrary operation between qubits given the hardware constraints, such as which qubits are directly connected to each other. It also decides which qubits to use for each operation—for instance, if a particular qubit is known to be faulty, information might need to be routed around it.

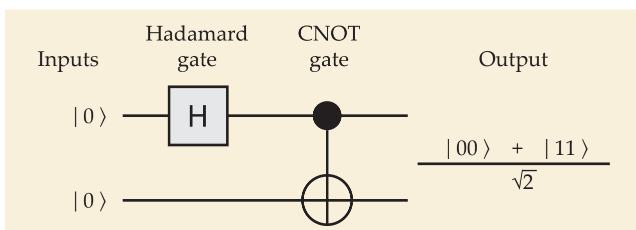
In the current era of quantum computing, the hardware-aware compiler is the only compiler. As such, it bears the additional responsibility of reducing the number of quantum logic operations needed to execute an algorithm. Optimizing qubit usage in that way allows a task to be completed as quickly as possible, which is important given the short lives of qubit states.

In the future, when quantum error correction is routinely used, some of this responsibility will be borne by higher-level logical-layer compilation. The lower-level compiler will be tasked with translating logical-qubit operations into their constituent physical-qubit manipulations.

Quantum error correction

Even with quantum firmware, errors inevitably arise from both decoherence and imperfect qubit manipulation. Quantum error correction (QEC) is designed to detect and fix those errors. It works by smearing information across many qubits in a way that protects against individual qubit failures. Each error-correcting group of physical qubits makes up a single logical qubit that can then be used in a quantum circuit. Amazingly, logical qubits can be designed such that even as the underlying qubit states decohere, the logical qubit state persists, in principle indefinitely.

Once a logical qubit is encoded, a complex algorithm is used to identify errors and apply corrections in a way that



TWO QUBITS start in pure $|0\rangle$ states. A Hadamard gate acts on the first qubit and puts it in a superposition of states $|0\rangle$ and $|1\rangle$ with an equal probability of finding the qubit in each state. The two-qubit CNOT gate flips the target qubit (⊕) to $|1\rangle$ only if the control qubit (•) is in state $|1\rangle$, thereby producing the entangled output state shown. Bell states are used in, for example, quantum cryptography (see the article by Marcos Curty, Koji Azuma, and Hoi-Kwong Lo, PHYSICS TODAY, March 2021, page 36).

doesn't lose the encoded information. (Measuring the qubits directly would destroy their quantum states.) A simple implementation uses redundancy to provide protection; even if one of the qubits ends up in the wrong state, the probability that they're all wrong is lower.

Correcting qubit errors with QEC is inherently resource intensive—some current schemes use tens of physical qubits per logical block—and will likely require more qubits than are available in existing devices to provide any real benefit. Accordingly, QEC is more important in the long term than it is for current machines. Quantum firmware aims to reduce the burden on QEC routines by dealing with more predictable noise, thereby improving QEC's resource efficiency.

Logical-level compilation and circuit optimization

A quantum circuit is a map of the sequential logic gates that are applied to a series of qubits to run an algorithm. A simple example of a circuit that entangles two qubits in a Bell state is shown on the previous page.

The initial qubit states are on the left, the final states on the right, and between them a series of gates that indicate the operations performed on each qubit. The qubits represented in the circuit aren't physical qubits; rather, they're abstract objects known as logical qubits. One logical qubit may be realized using many interacting physical qubits whose hardware errors are mitigated by QEC.

A single algorithm can be represented by multiple logically equivalent circuits, and the goal of circuit optimization is to find the one requiring the fewest operations or timesteps. Executing fewer operations enables the algorithm to run faster—an important goal for any quantum computer, whether or not it is using QEC.

Quantum algorithms and applications

Quantum algorithms play the same role as classical algorithms: They provide step-by-step instructions for completing a computational task.

Although a regular algorithm could in principle be run on a quantum computer, a true quantum algorithm takes advantage of the underlying hardware's quantum nature. For example, manipulating one qubit in a quantum computer affects the entire n-qubit state and each of the 2^n coefficients needed to describe it, effectively doing that many operations in parallel. However, it's not quite parallel computing. When the final qubit states are measured, each is either a 0 or a 1; the algorithm outputs only n values rather than all 2^n coefficients. (For more on quantum computation, see, for example, the article by Charles Bennett, Physics Today, October 1995, page 24.)

Given that measurement limitation, truly taking advantage of a quantum computer's huge computational space is tricky. The entire field of quantum algorithm development is devoted to figuring out how to efficiently leverage that resource. Some problems, like factorizing prime numbers, are known to be sped up by quantum algorithms. That speedup is reflected in the number of steps the algorithm must go through to arrive at an answer. Whereas the number of steps a conventional computer requires to factor a prime number scales exponentially with the size of the number, the number of steps for a quantum computer scales only polynomially. Quantum Fourier transforms are also significantly faster than their classical counterparts. Other tasks, such as playing chess, garner little to no benefit from quantum algorithms because the number of steps needed would still grow too quickly with the complexity of the problem.

```
from qiskit import QuantumRegister, QuantumCircuit, Aer, execute

q = QuantumRegister(1)
hello_qubit = QuantumCircuit(q)

hello_qubit.iden(q[0])

job = execute(hello_qubit, S_simulator)
result = job.result()
result.get_statevector()
```

```
array([1.+0.j, 0.+0.j])
```

THIS SHORT QISKIT ALGORITHM, akin to a "Hello, World!" program, initializes one qubit in the state $|1\rangle$. (Image from D. Koch, L. Wessing, P. M. Alsing, http://arxiv.org/abs/1903.04359.)

A variational quantum algorithm is a compromise between classical and quantum ones. It breaks up a computation into a small quantum component and a larger classical optimization problem and therefore requires a much smaller quantum computer than, say, the quantum Fourier transform. Such algorithms are promising for solving problems in finance, logistics, and chemistry.

User interface, QAAS, and operating system

Most people who want to use quantum computers aren't going to build or even buy one—at least not anytime soon. To facilitate access to the limited existing quantum computing resources, companies have put together cloud-based infrastructures that allow remote operation. As in a classical computer, the highest level of the quantum computing stack provides the interface that users interact with.

Amazon Braket, Microsoft Azure Quantum, and Rigetti Quantum Cloud Services are examples of quantum-as-a-service (QAAS) offerings. However, those companies aren't necessarily providing access to their own quantum computers; rather, they connect users and computers. For example, Amazon Braket can connect users to resources from D-Wave, Rigetti, and IonQ. That approach makes quantum computers similar to other managed, cloud-based computational resources, such as graphical processing units.

The above services can be used to write code using high-level programming languages. The resulting algorithms probably wouldn't look particularly exotic to someone with programming experience. For example, the open-source software development kits Ocean (from D-Wave), Qiskit (from IBM), and Forest (from Rigetti) support the programming language Python. Languages specifically designed for quantum computing include Quantum Computation Language (QCL), which resembles C, and Q Language, which works as an extension in C++. The code defines a sequence of operations that constitute a logical algorithm. PT

QUICK STUDY

FROM THE ARCHIVES
NOVEMBER 2014

Sheila Dwyer is a postdoctoral research fellow at the LIGO Hanford Observatory in Richland, Washington.

Squeezing quantum noise

Sheila Dwyer

You can't beat the Heisenberg uncertainty principle, but you can engineer systems so that most of the uncertainty is in the variable of your choice. Doing so can improve the precision of delicate measurements.

Most of the time, an imperfect measurement technique can be blamed for any deviation from the actual value of the variable you are trying to observe. However, in the probabilistic world of quantum mechanics, the observable properties of a physical system are truly uncertain; identical measurements on the same particle will result in different values even if each individual measurement is perfect. The Heisenberg uncertainty principle states that fundamental physics will limit how small the uncertainty in given pairs of observables can be.

The best-known uncertainty relation places a minimum on the product of the uncertainties (designated by Δ) in position x and momentum p—namely, $\Delta x\, \Delta p \geq \hbar/2$. Uncertainty relations are an aspect of quantum mechanics that is disconcertingly nonclassical. If, for example, a particle exists in a very specific location, its momentum must be highly uncertain and vice versa. Squeezed states are a class of quantum states that exemplify that kind of behavior, with a small uncertainty in one observable and, therefore, a large uncertainty in another.

Noisy vacuum

In addition to position and momentum, many other pairs of observables—for instance, the polarizations of light or the spin components of particles—satisfy uncertainty relations. For light, which can be treated as a quantum harmonic oscillator, the roles of position and momentum may be taken on by a pair of unitless observables, X_1 and X_2, known as quadratures. The uncertainties of those operators, ΔX_1 and ΔX_2, are governed by the uncertainty relation $\Delta X_1\, \Delta X_2 \geq 1$. In some situations, X_1 and X_2 correspond to the amplitude and phase of the electric field, and their uncertainties represent the amplitude noise and phase noise of that field.

The lowest energy state of a harmonic oscillator, called the ground or vacuum state, also has the minimum uncertainty allowed by the Heisenberg principle. Panel a of the figure, which represents the ground state of the electric field, shows the probability of measuring specific values of X_1 and X_2. As is characteristic of the ground state, measurement uncertainty is equally distributed between the two variables.

Such vacuum fluctuations exist everywhere, even in places that in a classical world would be totally dark. And they exist with every possible frequency, polarization, and direction of propagation. Their energy and field strengths are tiny, but their presence has several important physical consequences, including spontaneous emission and Casimir forces, and they limit the precision of sensitive measurements.

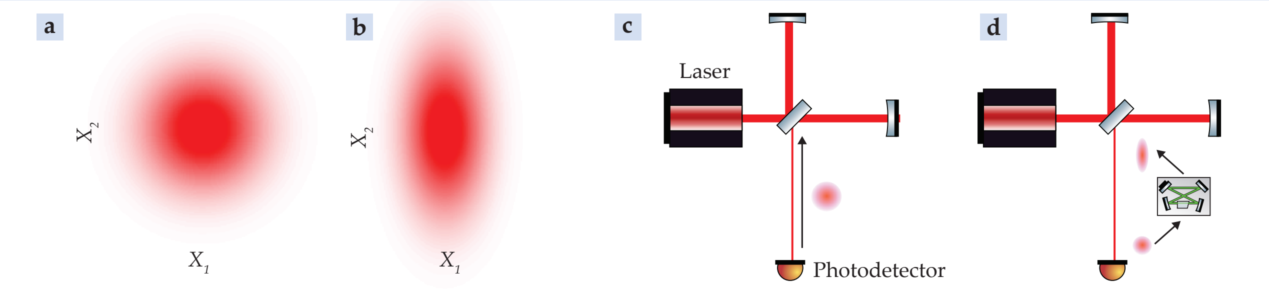
Although the uncertainty principle places a minimum on the product of the uncertainty in pairs of observables, it doesn't place any restrictions on the uncertainty in either observable alone. You can think of the uncertainty product pictorially as the area of the dark red bull's-eye in panel a of the figure. For a commonly used class of states called Gaussian states, the minimum-product requirement means that the area for the state must be at least as large as the area for the ground state. However, a state can be squeezed as in panel b to reduce the uncertainty in one observable, provided a larger uncertainty in the conjugate variable preserves or increases the area. Today researchers are using squeezed states to improve some of the most delicate measurements ever made.

So how do you actually squeeze a state? The key to any squeezing technique is to create correlations between normally independent fluctuations; such correlations can lead to the reduction of noise. Today the most widely used methods for generating squeezed states of light rely on parametric down conversion, a process in which one photon is converted into two lower-frequency photons whose phases are correlated. Squeezed states of light have also been created with optomechanical systems in which the mechanical response of a resonator to radiation pressure is used to create correlations between the amplitude and phase noise of light. Uncertainty in the direction of atomic spins can also be squeezed via measurement of light that interacts with the atoms in an optical cavity. Indeed, spin squeezing has been realized in ensembles of cold atoms and in Bose–Einstein condensates, an accomplishment that could improve the stability of atomic clocks and the performance of atom interferometers.

Detecting spacetime ripples

Gravitational waves—whether from supernovae, spinning neutron stars, or the inspirals and coalescence of compact-object binaries—distort spacetime. As a result, waves with frequencies below 10 kHz will change the length of any object they pass through. But the alterations are minuscule: The kilometer-scale interferometers currently under construction to detect gravitational waves, Advanced LIGO (Laser Interferometer Gravitational-Wave Observatory) and Advanced Virgo, will need to measure changes in their arm lengths of roughly 10^{-20} m, five orders of magnitude less than the width of a proton! The large instruments, with their 40-kg mirrors and multikilometer-long arms, are not the sort of systems normally expected to exhibit quantum behavior. However, the displacements they measure are so small that the uncertainties imposed by quantum mechanics limit their performance.

All the major Earth-based gravitational-wave interferometers are variations on the Michelson interferometer; panel c of the figure gives a schematic diagram of the device. A laser



Squeezed light and interferometry. The quantum state of light can be depicted by probability distributions such as shown in **(a)** and **(b)**. The so-called quadrature variables X_1 and X_2 here describe the electric field. Panel a gives the distribution of "vacuum fluctuations" for the ground state of the electromagnetic field; panel b gives the distribution for squeezed light. **(c)** The Michelson interferometer is the basis for Earth-based interferometers designed for detecting gravitational waves. The difference in arm lengths due to the passage of a gravitational wave is measured by monitoring the intensity of light on the photodetector shown at the bottom of the schematic. Vacuum fluctuations symbolized by the target shape enter an interferometer from the unused port (thin, red line) and cause quantum noise. **(d)** Reflections off a nonlinear cavity (inset) convert the vacuum fluctuations that would normally enter the interferometer to squeezed-vacuum fluctuations; the result is reduced quantum noise and improved measurement precision.

beam is sent down two orthogonal arms by a beamsplitter and reflected back toward the beamsplitter, where the light from the two arms interferes constructively or destructively depending on the relative length of the arms. By measuring the power at a photodetector, one can make sensitive measurements of changes in the arm lengths.

Since vacuum fluctuations propagate everywhere, they enter a Michelson interferometer from the unused port of the beamsplitter, where the photodetector is located. In 1981 Carlton Caves explained how those vacuum fluctuations cause the two types of quantum noise that limit the performance of gravitational-wave detectors: quantum radiation pressure noise, which results from fluctuations in the momentum imparted to the interferometer mirrors when light reflects off them, and shot noise, due to fluctuations in the amplitude of light arriving at the photodetector. Those distinct types of noise can be attributed to the uncertainties in X_1 and X_2. Caves suggested that the performance of a gravitational-wave detector could be improved by substituting squeezed states for the vacuum fluctuations that enter from the dark port of the interferometer.

During the past decade, members of the LIGO scientific collaboration have created sources of squeezed vacuum states suitable for integration into gravitational-wave detectors. Two gravitational-wave detectors have already used squeezed-state injection to improve their sensitivity: the GEO600 detector near Hanover, Germany, in 2010 and the LIGO detector in Washington State in 2011 (see Physics Today, November 2011, page 11). As depicted in panel d of the figure, the vacuum fluctuations that would normally enter the interferometer are first reflected off a nonlinear cavity that converts the ground-state vacuum fluctuations to squeezed vacuum fluctuations. In both the GEO and LIGO experiments, the squeezing reduced shot noise and increased the quantum radiation pressure noise; still, the quantum radiation pressure noise remained well below other limiting noise sources in the two interferometers.

For the Advanced LIGO interferometers, however, quantum radiation pressure noise will dominate in the astrophysically important 10- to 30-Hz band, so injection of squeezed states that reduce shot noise would degrade the interferometers' low-frequency sensitivity. For that reason, the Advanced LIGO instruments will include filter cavities that reduce the level of squeezing at low frequencies and preserve the high-frequency squeezing. More than 30 years after Caves made his proposal, squeezing combined with suitable filter cavities has emerged as the most practical way for Advanced LIGO to improve its sensitivity.

From thought experiment to practical tool

Squeezed states were first considered almost a century ago as theoretical constructs illustrating one of the difficult nonclassical concepts of quantum mechanics: the uncertainty principle. Once scientists created those states in the lab, they used them to test fundamental ideas of quantum mechanics. Now squeezed states are becoming a tool to improve precision measurements, to search for signals from distant astrophysical events, and to demonstrate quantum teleportation and quantum cryptography. The coming years may well see the implementation of new types of squeezed states, new methods for generating squeezed states, and further applications of squeezing to solve novel problems.

I thank Lisa Barsotti, Kim Burtnyk, and Nergis Mavalvala for detailed comments on an earlier draft, and I acknowledge with appreciation NSF's contribution toward the construction of the Laser Interferometer Gravitational-Wave Observatory.

ADDITIONAL RESOURCES

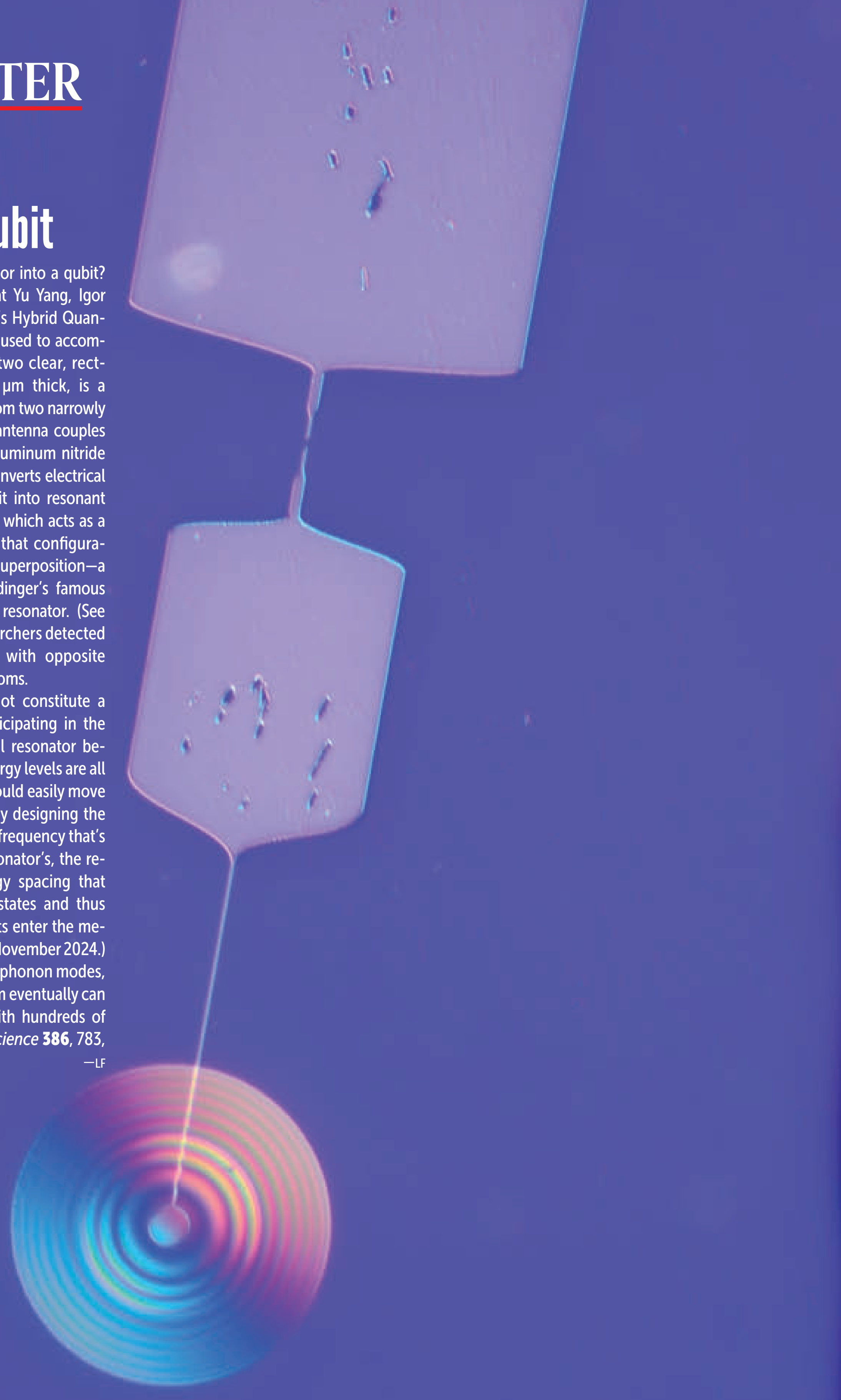
- C. M. Caves, "Quantum-mechanical noise in an interferometer," *Phys. Rev. D* **23**, 1693 (1981).
- N. Bigelow, "Quantum engineering: Squeezing entanglement," *Nature* **409**, 27 (2001).
- J. Abadie et al. (LIGO collaboration), "A gravitational wave observatory operating beyond the quantum shot-noise limit," *Nat. Phys.* **7**, 962 (2011).
- N. Mavalvala, T. Corbitt, "Vibrating membrane puts a squeeze on light," *Physics* **6**, 95 (2013).
- J. Aasi et al. (LIGO collaboration), "Enhanced sensitivity of the LIGO gravitational wave detector by using squeezed states of light," *Nat. Photon.* **7**, 613 (2013).
- Special issue, "Squeezed Light: From Inspiration to Application," *LIGO Magazine*, issue 3 (September 2013). PT

BACK SCATTER

A macroscopic qubit

How do you turn a mechanical resonator into a qubit? This micrograph shows the system that Yu Yang, Igor Kladarić, and colleagues in ETH Zürich's Hybrid Quantum Systems Group, led by Yiwen Chu, used to accomplish that task. Sandwiched between two clear, rectangular sapphire crystals, each 400 μm thick, is a superconducting qubit that is formed from two narrowly separated rectangles of aluminum. An antenna couples the qubit to a dome of piezoelectric aluminum nitride (at bottom, 400 μm in diameter) that converts electrical signals from the superconducting qubit into resonant vibrations in the upper sapphire crystal, which acts as a mechanical resonator. The team used that configuration in 2023 to generate a quantum superposition—a so-called cat state, after Erwin Schrödinger's famous thought experiment—in a mechanical resonator. (See Physics Today, July 2023, page 16.) Researchers detected two oscillations, or phonon modes, with opposite phases in the upper sapphire crystal's atoms.

But superposed states alone do not constitute a qubit, which has only two states participating in the superposition. Because the mechanical resonator behaves like a harmonic oscillator, the energy levels are all evenly spaced. As a result, the system could easily move between multiple phonon states. But by designing the superconducting qubit with a resonant frequency that's slightly offset from the mechanical resonator's, the researchers induced variations in energy spacing that enabled them to isolate two energy states and thus make the resonator a qubit. (See "Qubits enter the mechanical world," Physics Today online, 19 November 2024.) A single resonator can host hundreds of phonon modes, and the researchers hope that the system eventually can be used to build a quantum circuit with hundreds of qubits on just one chip. (Y. Yang et al., *Science* **386**, 783, 2024; image courtesy of Yu Yang.) —LF



TO SUBMIT CANDIDATE IMAGES FOR BACK SCATTER VISIT https://contact.physicstoday.org.