O'REILLY®

The New Quantum Era

AN OUTSIDER'S INTRODUCTION

SEBASTIAN HASSINGER



The New Quantum Era

An Outsider's Introduction

Sebastian Hassinger

Beijing • Boston • Farnham • Sebastopol • Tokyo O'REILLY®

The New Quantum Era

by Sebastian Hassinger

Copyright © 2024 Sebastian Hassinger. All rights reserved.

Printed in the United States of America.

Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472.

O'Reilly books may be purchased for educational, business, or sales promotional use. Online editions are also available for most titles (*https://oreilly.com*). For more information, contact our corporate/institutional sales department: 800-998-9938 or *corporate@oreilly.com*.

- Acquisitions Editor: Suzanne McQuade
- Development Editor: Shira Evans
- Production Editor: Jonathon Owen
- Copyeditor: nSight, Inc.
- Proofreader: Dwight Ramsey
- Interior Designer: Monica Kamsvaag

- Cover Designer: Susan Thompson
- Illustrator: Kate Dullea

- September 2024: First Edition

Revision History for the First Edition

- 2024-09-04: First Release

See *https://oreilly.com/catalog/errata.csp?isbn=9781098149413* for release details.

The O’Reilly logo is a registered trademark of O’Reilly Media, Inc. *The New Quantum Era*, the cover image, and related trade dress are trademarks of O’Reilly Media, Inc.

The views expressed in this work are those of the author and do not represent the publisher’s views. While the publisher and the author have used good faith efforts to ensure that the information and instructions contained in this work are accurate, the publisher and the author disclaim all responsibility for errors or omissions, including without limitation responsibility for damages resulting from the use of or reliance on this work. Use of the information and instructions contained in this work is at your own risk. If any

code samples or other technology this work contains or describes is subject to open source licenses or the intellectual property rights of others, it is your responsibility to ensure that your use thereof complies with such licenses and/or rights.
978-1-098-14936-9

Preface

Technologies are tools that human beings craft out of hard-won understanding of some aspect of our world, what British physicist David Deutsch calls explanatory knowledge. These tools have allowed us to attain capabilities of both a creative and a destructive nature unimaginable for any other form of life with which we are familiar. Explanatory knowledge has the power to reshape our existence in ways that convey maximum benefits to us, or to utterly destroy us. This terrible, marvelous gift is our central narrative—it provides abundance and marvels to the point of utter luxury, but the costs incurred over the centuries are coming due in the very near future. It seems entirely likely that this century will see a singular moment in which our fate will be determined. The weight of our debts will overwhelm us, exhausting our world and dooming us to whatever horrors lie beyond complete environmental collapse, or our toolmaking will surge ahead, maintaining the Anthropocene with a surplus available to service and eventually pay off our debt.

The scientific age gave us the crude tools of combustion and electricity to power the engines of the industrial era, and the first quantum age gave us the tools of the postindustrial

information age, where we reside today. This digital tool set may be enough to win our struggle against the destructive externalities of the machine of humanity, but it's not a foregone conclusion. In fact, the probability of our success is unknowable at this point in time, only to be revealed after that decisive moment coming in the future, after which we will fix our chances for survival at 100% or 0.

In anticipation of that fulcrum point, we are collectively in the process of mobilizing a historic, towering collection of resources with which we hope to stack the deck in our favor. The Biden Administration's Inflation Reduction Act was mostly an environmental policy in disguise, with subsidies for the decarbonization of the US economy, which in many cases have no effective cap, the legislative equivalent to spend "whatever it takes." Technologies such as solar panels, already the cheapest way to generate electricity, will play a vital, irreplaceable role.

However, our current technologies are rooted in the extractive, inefficient techniques gleaned from 19th-century scientific understanding. The chips that power our most powerful applications are power hungry and wasteful, and the same is true for all of our technologies, especially our most advanced ones. As we debate whether today's machine learning software will ever achieve a level of artificial general intelligence, or the

ability to generate explanatory knowledge comparable to human beings, it's worth noting that the most capable AI models today consume about 7 watt-hours per query. The human brain consumes 12 watts per hour, and who knows how many "queries" it handles in that time.

In fact, throughout nature, we have found processes that do work at efficiencies that are unimaginable with our current explanatory knowledge. Plants use photosynthesis to create energy with only beneficial byproducts. Bacteria can fix nitrogen and eat plastic with no hazardous wastes. Science is slowly building our knowledge of these processes, in the hopes we can learn how to make abundance for ourselves without continuing to destroy our planet. In all cases, these natural processes at some level depend on quantum mechanics to work their magic.

The theory of quantum physics represents our best explanation of the operating system of our universe and in fact underpins the major components of our current digital age. Semiconductors used for microprocessors work because we understand and engineer their structures at a nanoscale where quantum mechanics describes behaviors of materials. Lasers were invented because of our understanding of photons and their quantum mechanical characteristics. However, quantum

physics' contribution is indirect. The material we make microprocessors with provides a substrate on which electrical circuits are drawn, and the work is done at a scale that is orders of magnitude larger than the subatomic. Billions of electrons flowing through conductive leads and billions of photons streaming along a fiber optic strand behave in much the same way as those flowing through a light switch or from a light bulb in your home, which is to say, classically, they do not take advantage of any of the unique behaviors of quantum mechanics.

This has been our state of affairs, but now, in the 21st century, we are seeing the dawn of a new quantum era. Decades of effort in the field of quantum mechanics has begun to translate into an entirely new set of engineering techniques that are being used to invent new technologies. These technologies, such as quantum sensing, quantum communication, and quantum computing, use quantum behaviors directly, leveraging them in the service of the desired application.

In December of 2017, I attended the Think Q Summit at IBM Thomas J. Watson Research Center, in Westchester County, New York. I worked for IBM at the time, though I was not affiliated with the research division. In fact, at that point I had no idea what quantum computing was, what it might be used for, or just

how mature it was—or was not. I was used to very early-stage emerging technologies, so I suppose I expected something like mobile computing in the ’90s, like a Motorola Envoy portable device in 1994, with 1 megabyte of RAM, a 4,800-bit-per-second dial-up modem, and weighing almost 2 pounds. A tantalizing proof of concept with a path to usefulness measured in years. The story of quantum computing that IBM presented at the summit was something else entirely.

Instead, on the auditorium stage, physicists were talking about future machines that were beginning to use quantum mechanical behaviors to carry out computation. Instead of marketing fluff, we were served up complex linear algebra formulas. In place of a product roadmap, we got a tour through the way IBM was making chips that manifested “artificial atoms,” which, when poked and prodded with microwave signals, would do math. Instead of applications, there were open-ended discussions about what they believed these machines might do, one day, if we got lucky. I found the most bleeding-edge technology, closer to fundamental scientific research than anything I’d seen before or will ever likely see again. I was hooked.

A few months later, I was part of the IBM quantum team, and though I left the company in 2021, I still work in the quantum

computing industry. I expect to finish my career working in this industry. Finding a way into this field I found so compelling was an intimidating challenge back in 2017. I am not a physicist, and my formal mathematics education ended with Calculus 2, which is the only class I ever failed, and I failed it in an utterly miserable way. It's a long story, but if I tell you I got nearly 100% in Cal 1 the first semester, you can infer that I had a very strange year as a freshman at University of British Columbia. So, with scant practical tools to understand the nitty-gritty of the linear algebra of quantum algorithms, I had to find another way to build up a competency. I fell back on the tools that I knew from my liberal arts and humanities training. I read. A lot. I read a lot that I didn't understand, really, but I grasped at the conceptual frameworks the authors used to convey the meaning of their work. I asked a lot of questions, often dumb questions, of the talented and very patient researchers I worked with at IBM and various universities.

The fuzzy beginnings of the understanding of the science was augmented by my experience in emerging technology from the start of my professional career, establishing an internet service provider in Vancouver the year the NSF's Acceptable Use Policy first allowed commercialization of network access. Many of the patterns discernible in the early stages of the quantum computing market bore resemblances to those early days of the

internet, or of the web, mobile computing, open source, big data, and the like. Nothing matched exactly, it never does, but while history never repeats itself, it certainly rhymes, and picking out the meter and guessing what comes next in a new emerging tech can be helpful even when one doesn't have deep specialized knowledge. It has helped as I built my understanding of how the quantum phenomena are being shaped into system architectures, as the value chain is beginning to form, as the complete quantum computing stack begins to take shape. There is a vast amount that I do not know and do not understand about quantum computing, but I feel confident that my guesses are not horribly wrong.

Gathering this knowledge, I have found that my fascination with quantum physics and this particular moment when its science is being turned into technology has only grown, has never diminished. I love the history of physics, the yielding of classical theory to the counterintuitive and often upsetting quantum theory. I love the history of information theory and the dawning of the digital age. My understanding of both has grown and cross-pollinated, mirroring the field of quantum information technology itself. At some point, I was inspired to write this book.

This book is not a textbook or a how-to of any kind. It will not teach the reader how to program quantum computers. It contains virtually no math and no code. My hope is that the reader will have a better understanding of what quantum computing is and what this moment in its development is about. This stage of quantum computing's development is, in some ways, mimicking the path taken by digital computing in the middle of the 20th century, or at least it's rhyming with it. The first functioning devices are being created in a bespoke fashion, with each one an idiosyncratic attempt to find a path toward the usefulness and scalability of the modern microprocessor.

In the midst of this uncertainty and enormous complexity, there can be general confusion and even, sometimes, attempts to leverage confusion as a tool in itself, to gain investment, or revenue, or reputation. This book is an attempt to provide explanatory knowledge to the reader, to help interpret marketing and scientific claims when you encounter them. Mostly I hope your sense of awe and wonder grow a little as you gain a greater appreciation for just how audacious our goals for this new quantum era are.

O’Reilly Online Learning

NOTE

For more than 40 years, *O’Reilly Media* has provided technology and business training, knowledge, and insight to help companies succeed.

Our unique network of experts and innovators share their knowledge and expertise through books, articles, and our online learning platform. O’Reilly’s online learning platform gives you on-demand access to live training courses, in-depth learning paths, interactive coding environments, and a vast collection of text and video from O’Reilly and 200+ other publishers. For more information, visit *https://oreilly.com*.

How to Contact Us

Please address comments and questions concerning this book to the publisher:

- O’Reilly Media, Inc.
- 1005 Gravenstein Highway North
- Sebastopol, CA 95472

- 800-889-8969 (in the United States or Canada)
- 707-827-7019 (international or local)
- 707-829-0104 (fax)
- *support@oreilly.com*
- *https://oreilly.com/about/contact.html*

We have a web page for this book, where we list errata, examples, and any additional information. You can access this page at *https://oreil.ly/new-quantum-era*.

For news and information about our books and courses, visit *https://oreilly.com*.

Find us on LinkedIn: *https://linkedin.com/company/oreilly-media*.

Watch us on YouTube: *https://youtube.com/oreillymedia*.

Acknowledgments

This book is the product of many fascinating conversations with brilliant, patient scientists without whom I wouldn't understand any of this stuff. Thanks to Pat Gumann, Chris Lirakis, Dimitri Maslov, John Preskill, Grant Salton, Peter Shor, Scott Aaronson, Sebastian Weidt, Charles Bennett, Yasunobu

Nakamura, Dave Bacon, Ken Brown, Brian DeMarco, Susannah Glickman, Lin Lin, and so many others.

Special thanks to those who read early drafts and provided invaluable feedback: Gilad Ben-Shach, Pedro Lopes, Gabe Paré, and Emily Edwards. Thanks to O'Reilly and especially Mike Loukides, with whom I first shared the idea to write a book at Foo Camp in 2020; Shira Evans, who has been an absolute rock; and Zan McQuade, whose enthusiasm is unending and always appreciated.

My infinite adoration and appreciation for Ayann, whose beautiful mind keeps me endlessly stimulated and curious, and whose love and depthless confidence in me kept me buoyed from the very beginning.

Chapter 1. Quantum Quinceañera

> *It's easy for quantum computing to sound like a fanciful theorist's dream…But just extrapolating Moore's Law… somewhere around the year 2020 or 2030 is when we hit one bit per atom…So if we really want to keep getting faster… Quantum weirdness is essentially the only resource we have that's still untapped for computing.*
>
> —Isaac Chuang and Neil Gershenfeld

Google's blog post "What Our Quantum Computing Milestone Means" was written by its CEO, Sundar Pichai, so it was clearly meant to be a big deal. That week in October 2019, it was accompanied by several articles and an op-ed in the *New York Times*, an article and a peer-reviewed paper in the prestigious journal *Nature*, and stories in several other tech outlets. Google had gone to great lengths to generate attention for its accomplishment. It was not, however, immediately apparent just what those claims meant. Pichai called it the "hello world" moment for the field of quantum computing, something Google had been working at for 13 years. Google had now accomplished this feat on its device named Sycamore, seen in Figure 1-1. But to the general public, to the tech journalists

covering it, and to the vast majority of people, it was far from clear just what Google had accomplished. Pichai wrote in the blog post that his researchers had achieved “quantum supremacy,” and that meant that they had “used a quantum computer to solve a problem that would take a classical computer an impractically long amount of time [to accomplish].” This, the reader was informed, was a “milestone in our effort to harness the principles of quantum mechanics to solve computational problems.”

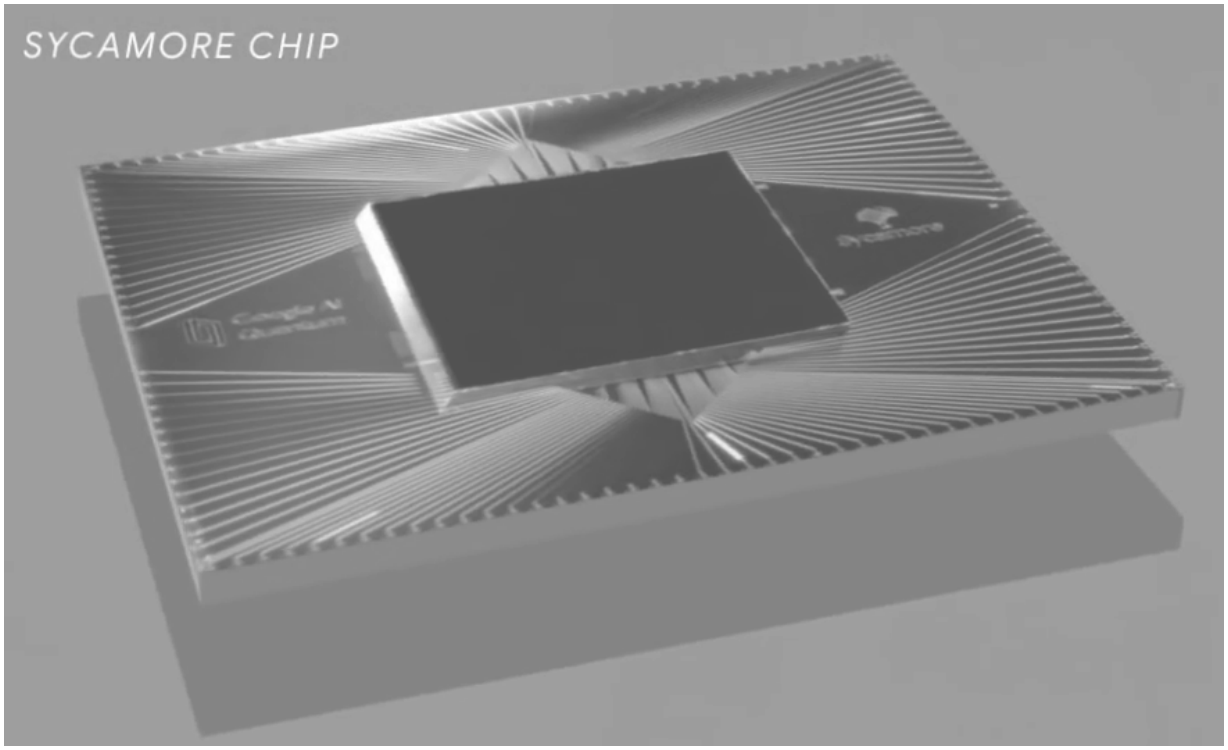


Figure 1-1. The Google Sycamore quantum processor used in the supremacy experiment Wikimedia Commons, October 23, 2019

Journalists struggled to get to grips with the implications of what was being announced. First things first, what was meant by a “classical” computer? The term is borrowed from physics, where the behaviors at the subatomic scale are referred to as quantum, and above that scale physical behaviors are considered “classical.” So your laptop, your phone, even the largest supercomputer in the world—all are computing classically. This begged the question, what did it mean to build computers using the “principles of quantum mechanics”? In the blog post, Pichai had offered the following as a starting point:

> *A bit in a classical computer can store information as a 0 or 1. A quantum bit—or qubit—can be both 0 and 1 at the same time, a property called superposition. So if you have two quantum bits, there are four possible states that you can put in superposition, and those grow exponentially. With 333 qubits there are 2^333, or 1.7x10^100—a Googol—computational states you can put in superposition, allowing a quantum computer to simultaneously explore a rich space of many possible solutions to a problem.*

A “1” followed by 100 zeros is called a googol and is the source of the company’s name. Including the Easter egg in the form of a very large number, as cute as it was, was not particularly helpful for building understanding. A common characteristic of

attempts to explain quantum computing seems to be to utilize analogies that seem straightforward but don't end up being very useful for building understanding. For example, Pichai's statement that quantum computers can "simultaneously explore…many possible solutions" is entirely inaccurate. And yet these false analogies persist to this day, as even well-informed and competent professionals such as Pichai struggle with the inadequacy of our language to explain what is *actually* going on.

Cade Metz, a journalist writing in the *New York Times* the same day, side-stepped these challenges, writing merely that the technology "relies on the mind-bending ways some objects act at the subatomic level or when exposed to extreme cold." Metz's colleague at the *New York Times*, David Yaffe-Bellany, also weighed in on the same day, grappling with the topic a bit more deeply. Whereas classical computers represented information as "bits," which can only assume the values of "1" or "0" and support computation based on that Boolean representation of two states called true and false, quantum computers used "qubits." Qubits, a portmanteau of "quantum" and "bits," Yaffe-Bellany wrote, can be 1 and 0 simultaneously, and therefore could hold an exponentially larger amount of information when compared to an equal number of classical bits.

The definition of qubit is accurate, but Yaffe-Bellany's explanation of superposition is, like Pichai's, another great example of an attempt to explain a quantum mechanical concept that drives physicists crazy: its inaccuracies leave an entirely incorrect impression. Quantum physicists will tell you that superposition doesn't mean that the qubit has both values at the same time but instead has a probabilistic combination of both, which doesn't exactly help shed light on the matter. There has been a struggle, largely unsuccessful, to come to grips with the "spooky" behaviors at the quantum scale, and the press coverage of the Google event didn't fare any better. In *IEEE Spectrum*, qubits were likened to "marbles" instead of coins, with their hemispheres labeled "zero" and "one," rolling around, while bits are coins that lie flat with only one value visible. A metaphor that physicists find particularly galling is that quantum computers evaluate "all possible answers at once" before settling on the correct answer, something they decidedly do not do. However, coverage of quantum computing continues to fall back on these flawed explanations.

The coverage around the Google supremacy experiment, having failed to create metaphors that created more clarity, turned to a framing that was more comfortable—industry competition. Metz's piece noted that "many of the tech industry's biggest names, including Microsoft, Intel and IBM as well as Google, are

jockeying for a position in quantum computing." In fact, the reactions to the supremacy announcement had also featured some uncharacteristically aggressive competitive posturing from tweedy IBM Research in anticipation of the puffy-vested Google team's big announcement.

The announcement had been foreshadowed by some drama fueled by that competition. Someone at NASA, a collaborator on Google's experiment, had accidentally posted the research paper to the NASA Technical Reports server a week before Pichai's blog post, where it wouldn't have been noticed if Google's own search engine hadn't ingested it and added it to Google Scholar, where it was fed into the email alerts of anyone monitoring the topic of quantum computing! IBM used the advance peek to formulate a rebuttal, seeking to undermine Google's claims, including members of IBM's quantum computing efforts.

IBM had a secret weapon in its effort to undermine Google's claims: Summit, a supercomputer at the US Department of Energy's Oak Ridge National Laboratory. IBM had installed Summit in 2018, and its 150 petaFLOPS, or quadrillions of operations per second, ranked it as the most powerful computer in the world at the time. IBM, leveraging their comprehensive knowledge of Summit's architecture, hypothesized that using

petabytes of hard drive storage as working memory, it would be able to complete the same experiment in two and a half days instead of hundreds of thousands of years. The rebuttal, posted to IBM's own blog a week ahead of Google's announcement, didn't do much other than make it even harder to make sense of the whole thing.

Regarding investment in the private sector, Yaffe-Bellany cited a *Nature* article by Elizabeth Gibney, who found that between 2012 and 2019, "venture capitalists have invested more than $450 million into start-ups exploring the technology." On the global stage, quantum computing was characterized as a strategic race at the superpower scale, with "China…spending $400 million on a national quantum lab" and "the Trump administration…promising to spend $1.2 billion on quantum research, including computers." The US investment sounded like an attempt to catch up, because China had "filed almost twice as many quantum patents as the United States in recent years."

This familiar context of a race to be first to market with a new technology, along with the sheer scale of global investment, may have played a role in creating a sense that quantum computing was imminent. The idea that this was the "hello world" moment for quantum computing, a phrase used by

Pichai and repeated in Elizabeth Gibney's article in *Nature* about the experiment, certainly helped create this impression. The true nature of the experiment seemed as difficult to explain as the quantum behaviors that underpinned the workings of the machine. In fact, the experiment was designed to take advantage of that "quantumness" in a way that would be particularly challenging for a nonquantum classical computer to emulate.

The most cogent explanation of just what Google's announcement meant came from physicist Scott Aaronson, the director of the Quantum Information Center at the University of Texas in Austin. Aaronson had been involved in the design of the experiment carried out by Google's research team, and had a reputation as a science communicator and prolific blogger. Writing an op-ed in the *New York Times* a week after the initial announcement, Aaronson set the context for understanding appropriately, writing, "the calculation doesn't need to be useful: much like the Wright Flyer in 1903, or Enrico Fermi's nuclear chain reaction in 1942, it only needs to prove a point." Drawing parallels to both the invention of flight and the atomic bomb ultimately seemed to further confuse the search for clarity in the press.

The challenges faced by the journalists covering the event and their readers contributed to a fundamental misapprehension about the "supremacy" moment—that the technology was on the cusp of emergence, and that in some reasonable amount of time it could be expected to be useful for something, and further, that it was an evolutionary step forward for the information technologies with which we were already familiar. This is the understanding of how information technology works since the late 20th century: a new chip, a new application, faster wireless networking, better functionality, or some other measure of progress through innovation makes its way into the media, into our consciousness, and a year or three later, there it is in the store or on the ecommerce site, or in the App Store. Generations of technology iteratively improve on their predecessors, and Apple Newtons are supplanted by Palm Pilots and then RIMM Blackberries, which are in turn kicked to the curb by the iPhone.

It's a comfortable narrative because it's familiar, but that doesn't make it appropriate framing for understanding what quantum computing was in 2019. It's not like there wasn't sufficient backstory for the Google accomplishment to signal an arrival, a coming out. *Time* magazine had run a cover story in 2014, five years earlier, heralding the launch of D-Wave's "Infinity Machine," the first commercially available quantum

computer. The company, based in British Columbia, Canada, claimed it would soon be capable of cracking all computer encryption schemes. It was true: the math that made this feat theoretically possible did exist (and we will explore it in Chapter 7). However, given the limitations of all known quantum devices, the claims in the article were seen as highly dubious by the quantum physics and quantum computing communities. Even at press time, D-Wave's claims of operating on the quantum level had been under fire by scientists, and by 2019, no one had managed to produce any type of performance advantage using the D-Wave machine. However, given the time that elapsed between the *Time* story on D-Wave and this Google accomplishment, it seemed reasonable to believe the technology had matured considerably.

But no matter what impression the collective coverage of the experiment conveyed, the reality of quantum computing was far less familiar and far less mature. What the Google team, the IBM team, and similar teams at various tech giants and startups as well as researchers in academic labs around the world were working toward was actually a complete paradigm shift, in literal terms. The machines they were building were attempting to rewrite the foundational information theory that ruled every transistor on every chip, every instruction in every piece of software, every interaction with a computer since computers

were invented in the middle of the 20th century. As Aaronson wrote, quantum computing was an attempt to rewrite the "rules that Charles Babbage understood in the 1830s and that Alan Turing codified in the 1930s," rules adhered to by "every computer on the planet—from a 1960s mainframe to your iPhone." If successful, their machines would create solutions to classes of problems that were practically impossible to solve using the computers, no matter how powerful we made them. IBM's Dario Gil had his own take on Pichai's playful "googol" explanation, in an interview with the *New York Times* that was packaged with IBM's rebuttal. He said that if one were to build a system with 100 perfect qubits, "you would need to devote every atom of planet Earth to store bits to describe that state of that quantum computer."

In 2019, however, no one even had built a single perfect qubit, nor does one exist at the time of this writing. The Google experiment was contrived, pointless from the perspective of the kind of useful things we expect from computers. Tasked with anything we'd recognize as useful, the best quantum computer would lose in a head-to-head battle with a 10-year-old laptop. It wouldn't even be close.

So why the excitement? Why all the investment? The context the media missed or glossed over was that the incredible

progress delivered by classical computing has, in the past decade, been showing signs of coming to an end, and that struck fear into the captains of tech and the heads of state alike. A growing sense of panic was gripping the tech world, causing it to cast about for a new engine of growth. Despite quantum computing being nowhere near ready for its mainstream debut, the industry was looking further and further afield for a solution to a looming crisis. Moore's law, the massively powerful driving force behind decades of technological advancement and economic production, was showing signs of falling apart. Computer industry insiders were increasingly concerned about finding ways to maintain their progress, and quantum was emerging as a potential savior.

In fact, Moore's law is not a law at all. In 1965, Gordon Moore, the director of research and development at Fairchild Semiconductor, wrote an article in which he observed that the number of components on a chip had doubled each year since 1958, leading to roughly doubling what a chip could do as well. He expected this trend to continue for at least another 10 years. A decade passed, and Moore reassessed his prediction at a conference, where he said he believed the trend would continue, revising the period of doubling to every two years. Carver Mead, a famed professor at Caltech, coined the term "Moore's law," and it stuck. It stuck in the language, largely

because it was demonstrated by the entire chip industry for the next 30 years, as illustrated in Figure 1-2.

Moore's law explains the progression from the Intel 8088 chip in 1980, with its 29,000 transistors, to Intel's Itanium chip in 2000, which had 220 *million* transistors. In many ways, global trade, economic growth, and geopolitics have all been calibrated in tight coordination with this reliable, remarkable progress. The progress has been so remarkable that it seems to defy probability, with predictions of its failure having been made again and again, only for a new chip to deliver yet another leap forward. However, in the early 2000s, the concerns began to accumulate. Some attributed slowing productivity in that era to the slowdown of microprocessor advancements.

As these concerns grew, the search for alternative technological approaches has gotten more attention. Exotic ideas like designing chips to operate more like our brains, computing using DNA, or even fungi have been able to mobilize more funding for research than they may have without the concerns about Moore's law. Quantum computing is a very compelling idea from this perspective, as its mathematical ideas provide a neat solution to a fundamental bottleneck of classical computing that is present in certain types of problems that get exponentially more challenging as they increase in scale.

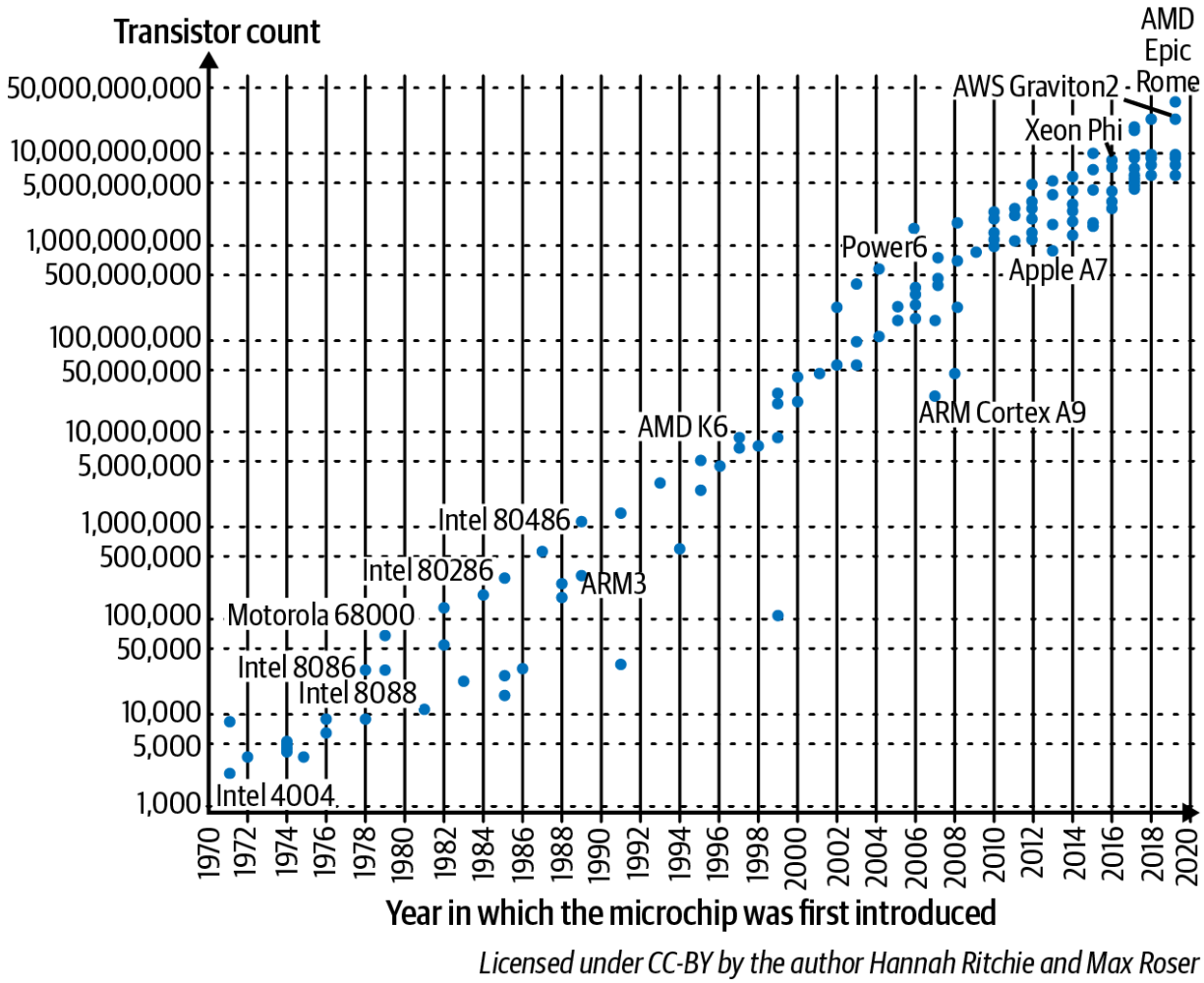


Figure 1-2. A logarithmic graph showing the timeline of how transistor counts in microchips are almost doubling every two years from 1970 to 2020 ("Moore's Law: The Number of Transistors on Microchips Doubles Every Two Years," OurWorldinData.org)

In fact, the Google experiment was explicitly designed to demonstrate this core characteristic of quantum computing. Its chip, known as Sycamore, had 53 operational qubits, on which a random set of quantum operations was executed, the

equivalent of a very small computer program, and then their output was read out. A collection of quantum operations is known as a “circuit” to those in the field. Because of the way Google’s Sycamore chip exploited the characteristics of quantum mechanics, representing all the possible states of the circuit and therefore calculating the same output, it would take 2^{53} classical bits, or 9 *quadrillion* bits. The largest number of transistors on a general-purpose microprocessor as of 2023 was Apple’s M-1 chip, with 114 billion. A quadrillion is a million billions.

This was the reasonable basis for Google’s claim that it had proven the “supremacy” of quantum computers over classical. Notwithstanding IBM’s hypothesized solution on Summit’s 250 petabytes of storage, the fact remained that it took Sycamore a few seconds to complete the task, still several orders of magnitude faster than the classical supercomputer.

The Google experiment established a credible claim that supremacy was provably real, no matter how regrettable the use of that particular word. It’s worth noting that none of the coverage of Google’s experiment seemed to notice the use of a word with prejudicial associations. Since then, many people have expressed their dislike of the word “supremacy” to describe the milestone, and John Preskill, the term’s originator,

has voiced his own regrets on numerous occasions. A preferred term has emerged, “quantum advantage,” but since Google used “supremacy” originally, it’s less confusing to stick to that term when discussing the experiment.

This oversight regarding word choice aside, the news and analysis around the experiment left much to be desired. Overreliance on frames of reference from decades of emerging technologies, tech rivalries, venture capital investment trends, and global economic dynamics caused most pieces to miss the real news in the announcement. Appreciating the breakthrough requires a deeper dive into exactly what quantum computing is, what makes it so radically different from classical computing, what its relationship is to the origins of quantum mechanics, and how adding quantum to information theory is opening the door to an entirely new way of computing.

Quantum computation may provide a way to sustain the trajectory of Moore’s law, not by sustaining the relentless doubling of transistors but by replacing classical bits with quantum ones. It is hoped the quantum gambit will allow us to sidestep the plateau with an entirely different technology that may break classical computation bottlenecks and open up miraculous new applications in bioscience, climate science, material science, communications, and financial services,

among others. These machines' ability to simulate natural systems with far greater fidelity than any classical system was hoped to lead from the demonstration of a contrived "supremacy" ultimately to remarkable breakthroughs in our understanding of our universe itself. But first, the field was going to have to move past its Wright Flyer moment and prove there was a path to a productive, useful technology. This was going to require an unprecedented effort to bridge the worlds of theoretical physics with engineering and technology development as well as effort that had begun decades before Google's results in 2019.

Chapter 2. Don’t Watson My Quantum

> *Once a new technology rolls over you, if you’re not part of the steamroller, you’re part of the road.*
>
> —Stuart Brand

IBM’s peevish “well, actually” response to Google’s supremacy experiment went beyond mere objections to the experiment’s design and Google’s resulting claims. The company felt it had been crucially involved in the roots of the field and an active investment that went back at least twice as long as this bunch of Silicon Valley startup kids. Given its legacy of contributions right down to coining the phrase “quantum information theory,” one could understand the feeling of having called dibs on the emerging technology. There was another emotional response that Google’s move to grab the quantum spotlight elicited, a deep gut reaction culturally unique to the IBM of the 21st century, best described as “oh no, not again.”

While IBM itself has a history that stretches back over 100 years as a manufacturer of recordkeeping and measuring systems like meat scales and punch card readers founded in 1911, its

identity as a computer company began with its IBM Research division. Founded in 1945, contemporaneously with the birth of classical computing, IBM's original research teams focused on the fundamental science that underpins the technologies derived from information theory. As the effort grew from its origins in a lab at Columbia University to purpose-built buildings of its own, the research division leveraged chemistry, physics, and other disciplines to create intellectual property and earn Nobels and Turing prizes. In 1961 IBM built the Thomas J. Watson Research Center in Yorktown Heights, New York, designed by the midcentury modern architect Eero Saarinen (see Figure 2-1), and it serves as the Research headquarters to this day. In many ways, it is the last of the great industrial R&D powerhouses at General Electric, Westinghouse, Bell Labs, and others that drove the transition from the late industrial era to the early Information Age. The Research division filed more patents than any other firm every year from 1992, until finally losing the top spot to Samsung in 2019, after an unparalleled run of 27 years.



Figure 2-1. IBM Thomas J. Watson Research Center, Yorktown Heights, NY, ca. early 1960s

Fueled by Research's intellectual property output, the IBM corporation had defined the early part of the information technology industry with its "mainframe" computers entering the market in 1952. The mainframe product family converged under the System/360 architecture and brand in 1964, and by 1973, IBM's $11 billion in revenue represented one-third of the entire US technology market. IBM powered banks and governments and every other large-scale computing application

so ubiquitously that “no one lost their job buying IBM” entered the business lexicon as the Department of Justice (DOJ) investigation into alleged anticompetitive monopolistic practices, opened in 1969, gathered steam.

In the 1980s, a new product for a new market was presented as a way to take the DOJ heat off IBM by demonstrating IBM’s largesse to the little fish swimming around in the same market. The product was the personal computer or “PC,” a device that had a tiny fraction of the power of the machines marketed to industry and governments and that was a fraction of the cost. The IBM PC had been developed as a kind of “skunkworks” project, not on the formal R&D roadmap but tolerated as an oddball experiment. Skepticism about the size of the market for the PC created a constant battle for resources to get it made, but by the late ’70s, it was nearly ready for launch.

A question remained about what the operating system for the PC would be, and here an executive decision was made that would change the course of history not just for IBM but for the entire world. Part of the pressure from the DOJ around the mainframe arose from the fact that IBM controlled both the hardware and the software, and no competitor could gain even the tiniest of a toehold with a hardware peripheral or a software product without risking being sidelined by IBM. The

PC was already leveraging off-the-shelf hardware components built to open specifications to create a more open system, and so why not license the operating system as well? Enter Microsoft and exit the IBM technological hegemony, and with it a vital piece of IBM's own perception of its strength and capabilities.

Though it would take decades, the mainframe and the very notion of massively centralized computing power would slip from dominance to a footnote in the IT industry, and IBM slowly slipped down the power rankings. The dominance of the cloud since the 2000s has largely been the story of Google and Amazon Web Services, with Microsoft as a fast follower. IBM's cloud offering is not seen as defining the market in any way, which is particularly galling for IBM. The concept of a virtual machine created for a single user within a large multiuser system was popularized in the late 1990s by internet service providers hosting multiple customers' websites on one server and on the personal computer by VMWare, founded in 1998. IBM had first shipped code to run a virtual machine on a mainframe host in 1967 but had failed to capture any of the value created by those virtualized products.

As a unique intellectual wellspring, IBM Research was seen by many in the company as the likely source of innovations that

could recapture market leadership. However, the one model of generating commercial value from intellectual property that IBM Research knew was outbound licensing, the practice of allowing others to use patented knowledge to create commercial products for a fee. In the early 2000s, IBM's income from IP licensing peaked at nearly $2 billion per year, a powerful stimulant that ensured the mountain of patents would continue to grow. Internally, however, commercializing its own IP proved an elusive goal. Research teams were tasked to work more closely with product teams, in the hopes they could create novel IPs just over the horizon on existing product roadmaps, with little success. Experiments with various incubation and acceleration techniques were similarly fruitless. Its two greatest market triumphs in the late '90s and early 2000s—the WebSphere application server and its Power line of servers—owed their market success in large part to others' innovations, Sun's Java programming language and the open source Linux operating system, respectively.

Hope for IBM Research to deliver a win accumulated around artificial intelligence (AI) projects. Advances delivered by the division in the mid-1980s in the use of advanced statistical analysis had been successfully commercialized by the company in products like dictation software and automated voice response phone systems, and IBM was seen as a market leader

in that niche. The statistical and probabilistic approaches to AI continued to evolve, resulting in IBM's Deep Blue software defeating world champion player Garry Kasparov in a chess match in 1997. Chess had been an intellectual benchmark for AI systems since their inception in the 1950s, so this was a historic achievement.

Building on what was learned from Deep Blue, IBM Research then focused on more general models for simulating intelligence, and delivered a ground-shaking *Jeopardy* win in 2011 with a system IBM named Watson, after the company's founder, Thomas J. Watson, who also lent his name to IBM Research's headquarters. Under CEO Ginni Rometty's leadership, IBM Watson became a "bet the farm" play in the 2010s, with enormous marketing campaigns and the creation of an entirely new division of the company, underscoring how big a deal this was seen to be.

There was certainly reason to be enthusiastic. AI had been the dream from the dawn of the industry, when early devices were often referred to as "electric brains." With Deep Blue and Watson's triumphs over human minds in chess and general trivia, it seemed that the time was right to deliver significant commercial value from the technology. For a number of years, it seemed that IBM might have a new winning horse in the race.

Then, in 2015, Google entered the chat. A London-based company named DeepMind, acquired by Google in 2014, formed the nucleus of a major effort in AI in a division called Google Brain. DeepMind's AlphaGo software beat the world Go master at the game, a feat many, many times more challenging than beating humans at chess. AlphaGo used a very different approach than Watson, based on a "neural network," a software approximation of the way we believe our own brains learn to recognize and act on patterns. Neural networks weren't a new idea by any means, rooted in our first information theory models for understanding human reasoning as far back as the late 19th century. The idea of an artificial neuron originated in a paper from 1948, and various attempts to build more sophisticated versions paralleled the development of computers from the very beginning. However, in many ways the theory had outstripped the available computational power by the mid-1980s, and attention had moved on to other approaches. What DeepMind had realized was that it could now muster enough computational power to train a neural net powerful enough to tackle the challenge of Go. Played on a 19 × 19 grid, a game of Go has 10^{360} possible moves, an unimaginably large number, which dwarfs the 10^{123} possible moves on an 8 × 8 chess board. Using approximately 50 Nvidia GPUs that were originally designed for computer graphics but had been repurposed for

the math that drives neural networks, AlphaGo was trained over a period of weeks on a diet of millions of Go matches to recognize patterns that would suggest a probable path to a win.

The triumph of AlphaGo over the world champion human player by itself did little to upset IBM's dominant position in the emerging AI market. Buzz in the mainstream press isn't what sells enterprise software, after all. Perceived value and inherent trust are required to compel a CIO to sign on the dotted line of a software contract, and IBM felt it could again rely on the lore of "no one lost their job buying IBM."

However, in the internet and open source era, something had happened to the decision-making power dynamics of the IT market, something that IBM simply didn't have the corporate DNA to understand. Instead of a strictly top-down decision-making process driven by the C-suite, technology investments had become increasingly driven by software developers themselves. Faced with a challenge at work, devs would tend to find potential solutions in technologies downloadable from the internet, try them out informally, and develop solutions that were deployed without centralized approval or even knowledge. The "world wide web" itself had spread that way, with the first web server software packages like HTTPd, which

was made available as a source code download from the University of Illinois in 1993.

Grassroots technology adoption was embedded in Google's corporate DNA: its origin story in a dorm room at Stanford University was a canonical example of the phenomenon. And so it was natural for Google, on the heels of the AlphaGo win, to release free, open source tools for developers to build their own neural nets. Called TensorFlow, the software developer kit (SDK) was released in 2015 and almost instantly began to take the air out of IBM's Watson balloon, reaching a stunning 89% market share of AI projects by 2018. Suddenly IBM was in catch-up mode, again.

All this was the backdrop to IBM's reaction to Google's 2019 quantum computing supremacy announcement. The supremacy experiment seemed like AlphaGo toppling Watson all over again, a demonstration of capability in an area where IBM took pride in its first-to-market technology that might be followed by a giant sucking sound as the developer community flocked to Google's platform and left IBM high and dry.

There was reason to feel territorial about the fledgling quantum computing market as well. Not only had IBM launched its cloud service—amusingly named the IBM Quantum Experience—in

2016 and then in 2017 launched its commercial offering, the IBM Q Network, and open sourced its SDK, Qiskit, but its contributions reached all the way back to the very beginning of the field. In fact, two years before Google's announcement of the supremacy experiment, IBM Research had staged its own quantum computing triumph at an event called the Think Summit, held in December 2017.

At the event, the pioneers of the field presented their work on algorithms and theoretical proofs of the performance advantages quantum computers would bring over classical technologies. Its primary function, however, was to build on the progress IBM had made in recent years, building on the offering on the public cloud, the SDK, and the commercial offering by announcing a system with 53 qubits, several times larger in scale than anything that anyone had built to that point. The system publicly launched in October of 2019, the same month Google announced its supremacy experiment. It came like a bolt out of the blue, taking the IBM quantum team completely by surprise. IBM felt it had a lock on the technology. After all, it had been involved with the very earliest stages of developing the notion that quantum phenomena might provide some advantage for computation.

In fact, the Think Summit was consciously an echo of an event IBM had helped organize decades earlier, in 1981, a conference known as "The Physics of Computation". The small event brought together a small group of people who had individually been poking around at the edges of an idea that quantum physics may be useful to incorporate into information theory. The idea had been slowly emerging over the previous decade, and this conference had the explicit goal of determining whether it had any merit. The gathering is universally recognized as the launch pad for the field of quantum computing. It was this history that gave IBM its sense of clear leadership in the field—leadership that Google now was threatening just as it had in AI. The fear experienced around the supremacy announcement was that the firm's efforts in quantum would suffer the same fate as Watson.

Competitive threats were far from the only risk facing anyone vying for the quantum computing crown. The technology was not merely nascent, but it was still mostly theoretical, with dependencies on many science breakthroughs still to come. At the final session of the 2017 IBM summit, Dario Gil, director of IBM Research and champion of IBM's investment in quantum computing, moderated a panel of experts. A pioneer in the field, UK-based physicist Richard Jozsa was on the panel, due to his collaboration with David Deutsch to create the first quantum

algorithm in the ’90s. Gil asked the panel what they would like to see done with this new marvel, the 53-qubit system to be launched in the coming year. Jozsa said he’d like to see the fundamental rules of quantum mechanics tested on the machine, because he “wasn’t convinced we had the physics right.”

Chapter 3. A Quantum Mechanical Kludge

> *Those who are not shocked when they first come across quantum theory cannot possibly have understood it.*
>
> —Niels Bohr

That an event celebrating the progress being made toward computers harnessing quantum physics would culminate in a challenge to the axioms of the field itself might seem odd, but the controversy was not new. In fact, from its inception as a collection of answers to the most vexing problems in physics at the turn of the 20th century, quantum mechanics has struck many scientists as plainly impossible and incorrect or at least incomplete. This may have something to do with the fact that the answers were struck upon out of a desperate attempt to make our explanations of the universe work.

Humans are very odd animals. Not content with securing requisite calories and water for continued existence, fending off or evading predators, and spending the rest of our time procreating and lazing around, at some point in our hidden collective past, we decided that a good thing to do with our

unusually large brains was to ask the question, “why?” Our first answers were more fiction than fact, contributions to the corpus we call “ancient myth.” However, knowledge about how the world worked that proved handy, often in the form of advice about how to grow things or raise animals, tended to be passed along, orally at first, from generation to generation, slowly accumulating more value. As recordkeeping improved in the Fertile Crescent and North Africa about 5,000 years ago, the accumulation of understanding really got going, spurring on written language and also innovative inquiries into the physical world such as mathematics and astronomy from the African continent, South Asia, and Asia. Our mastery of the world slowly grew. The stories from our myths were recapitulated in the technologies our growing scientific knowledge allowed us to engineer. By the 18th century, “science,” from the Latin root *scire*, or “know,” very closely resembled our contemporary understanding of the word.

Physics as a subdomain of science is a thread woven into this human story, also thousands of years old. A landmark in the recorded history of natural science was the insistence of Thales of Miletus, a pre-Aristotelian philosopher, that all natural phenomena must have a natural cause. Mathematics developed from simple recordkeeping—crucial for managing resources in any agrarian society—proved surprisingly helpful in making

sense of astronomical observations. The observable heavenly bodies' motion could be articulated in mathematical formulas that predicted when and where the object would appear in our night sky with startling accuracy. These formulas, when considered as parts of a larger system, enabled scientists like Galileo and Giordano Bruno to construct entire working hypotheses for how our globe interacted with the rest of the planets in our system, with the sun at the center. Those with a vested interest in the mythological story of how our universe operates were unhappy, but the scientists had the benefit of experimental validation of their ideas. Eventually, those experiments created fields of engineering and masterful technologies that would take us to visit the solar system and the heavenly bodies we observed within it. The origins of the wood engraving seen in Figure 3-1, used by French astronomer Camille Flammarion in his 1888 book *L'atmosphère: météorologie populaire*, remain unclear, but it appears to show a medieval pilgrim piercing the boundary of a heliocentric cosmology to find the various systems of nature beyond the veil.

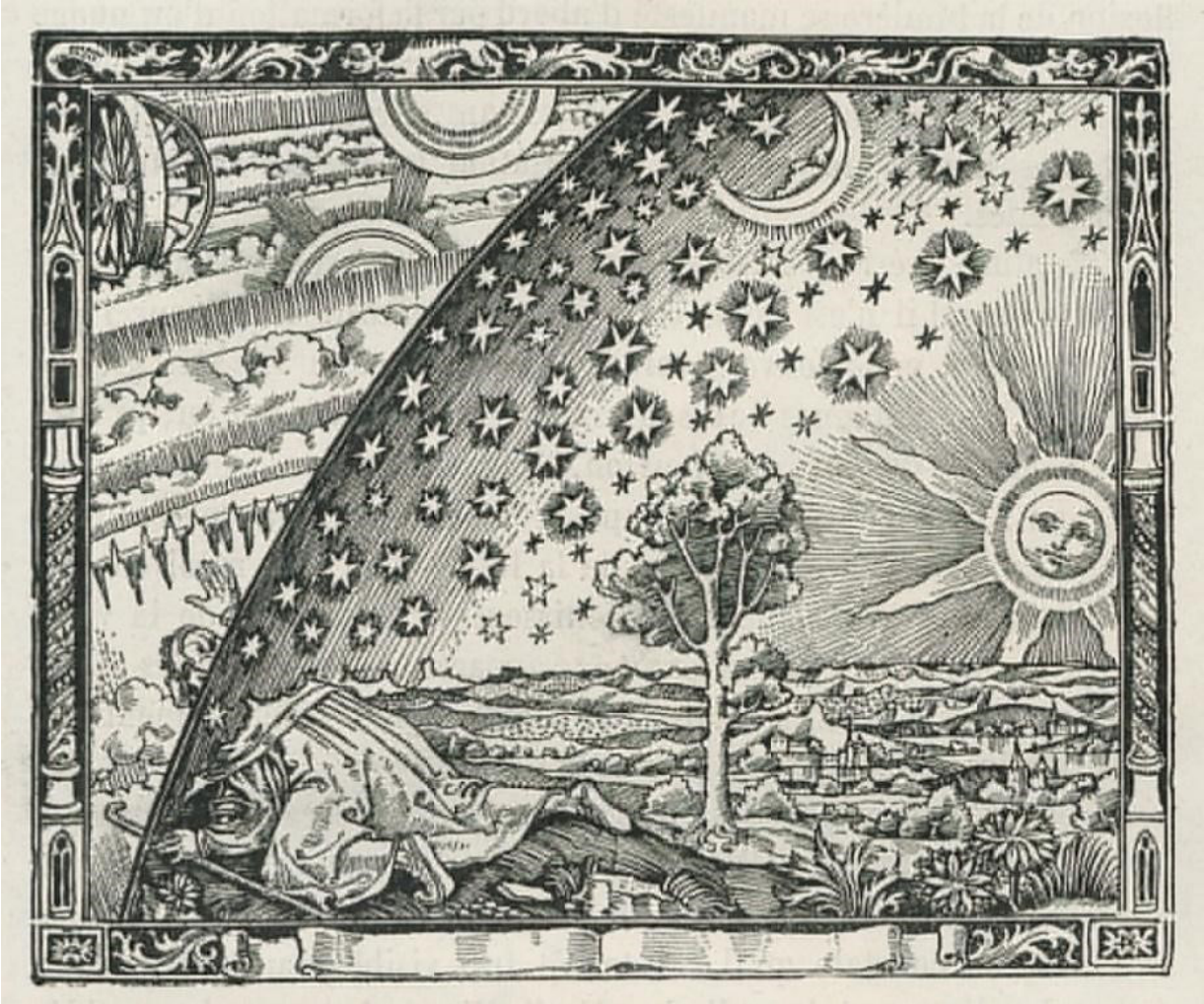


Figure 3-1. Flammarion engraving, ca. 1888, depicting a man in pilgrim's robes piercing the pre-Galilean firmament to see various "spheres within spheres" of the natural world (via Wikipedia)

Given the impressive progress and scope of the physicists' ability to create logical, mathematical models that precisely captured and predicted the behavior of nature, the growing confidence in the power of the field comes as no surprise. Rationalism as a philosophical belief formally emerged in continental Europe in the 17th century, rooted in the idea that, given a set of foundational principles such as the axioms of

geometry, one could work outward to deductively reproduce all other knowledge. German Enlightenment-era polymath Gottfried Leibniz and Dutch philosopher Baruch Spinoza in particular believed that a reliance on empiricism, or mere measurement of the world, was a necessary crutch for our limited human minds. God was thought of as an invisible watch maker and the universe his magnificent watch works.

The Enlightenment as a Western European cultural era was the embodiment of rationalist thinking and, at its most optimistic, held a growing optimism in humankind's eventual completion of the project to create a complete and predictive model of the universe that would all just make sense. The generic pattern employed by natural philosophers such as Leibniz and Sir Isaac Newton was to either devise an experiment to demonstrate a behavior of nature or to work out the mathematical equations that could explain and predict that behavior. This division of tasks persists in the field of physics today, with scientists identifying as either experimentalists or theorists.

Sometimes, however, the experiment and theory wouldn't line up. These discrepancies have long been irresistible bait for natural philosophers of old and modern physicists alike, since they flag areas of our understanding that are incomplete or incorrect, though fixing those inconsistencies can be very

challenging. Observations of Mercury's orbit made in the middle of the 19th century weren't theoretically understood until Einstein's general theory of relativity 50 years later. Another example of curious experimental results took even longer to explain, and resulted in perhaps the largest disruption to our scientific understanding of nature.

In 1801 Thomas Young, a British polymath known as "the last man who knew everything" by his contemporaries, sought to improve the understanding of the interference phenomenon, in which light can interact with itself to create puzzling results. The prevailing theory of light as "corpuscular," or made of particles, had been described by Sir Isaac Newton in his book *Opticks*, published in 1704. However, if light particles collided, they would bounce off each other like marbles, which could not explain some of the observed behaviors of light, and this contradiction motivated Young to explore further.

Starting with a shallow vessel of water divided in two by a wall with two small openings, he demonstrated how waves passing through those openings will create two sets of ripples fanning out from the other side of the barrier. Where the waves overlap with one another, an interference can be observed. If the waves are aligned, they will create a wave with either a higher peak or deeper trough. If unaligned, they will diminish one another,

even to the point of rendering that point of water perfectly level if they are evenly and oppositely matched. These are called constructive and destructive interference, occurring when the waves are said to be in superposition with one another.

Young then moved on to light, creating an opaque barrier with two slits for light to pass through, just as the waves in water had passed through the gaps in the tank wall. When observing the light illuminating a surface on the other side of the barrier, Young could see clear lines of interference. The bands of light and dark showed that the light passing through each opening had fanned out just as the waves of water had, overlapping with one another, and when superimposed in that way, they had either amplified or diminished the intensity of the resulting light. This double-slit experiment showed light behaving in precisely the way waves do in water, posing a serious challenge to Newton's corpuscular theory of light (see Figure 3-2).

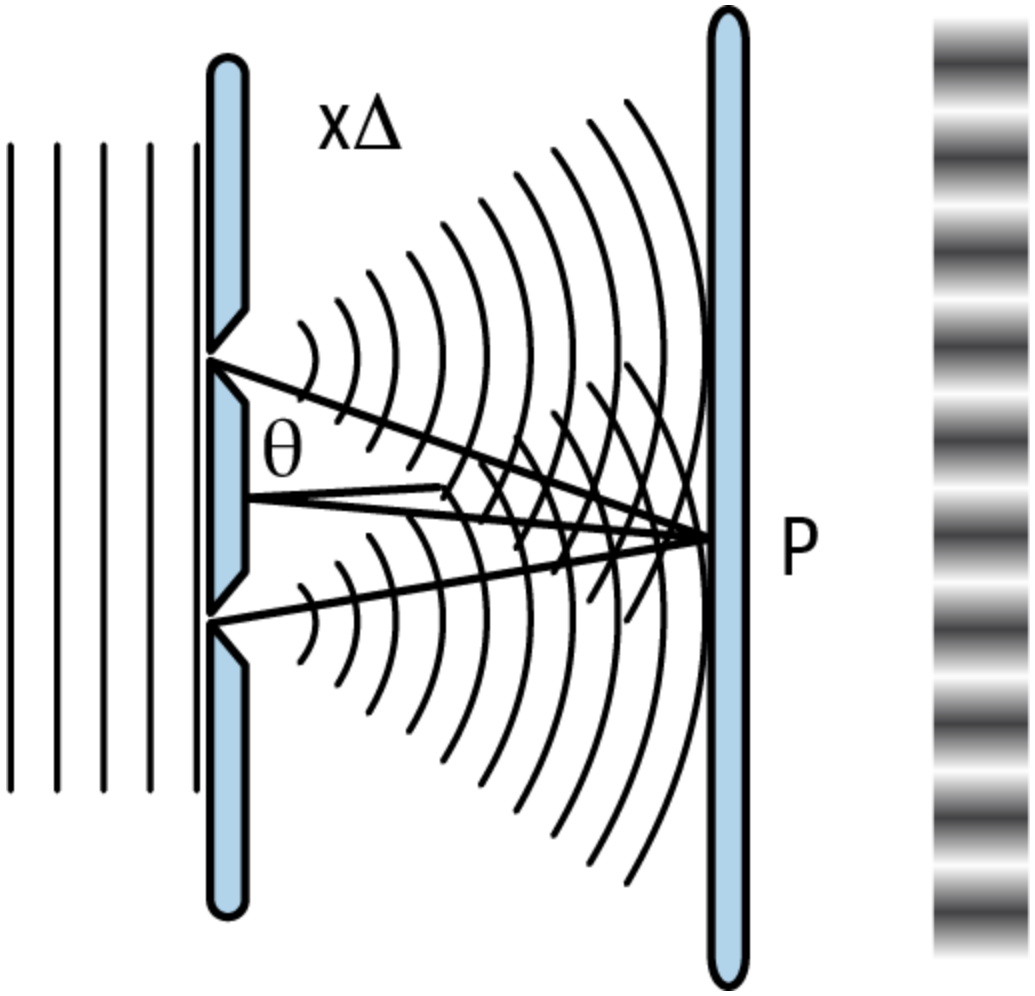


Figure 3-2. The double-slit experiment, using two parallel openings in a barrier, a light source, and a flat plane, demonstrates how particles of light, or photons, can behave as waves and interfere with one another as ripples in water would, creating an interference pattern on the plane.

A replacement theory would take another 60 years, until James Clerk Maxwell, a Scottish physicist working on a comprehensive theoretical foundation of electromagnetic behavior, realized that he could explain Young's experimental results and many other behaviors of light using equations developed for electromagnetic phenomena. The publication of Maxwell's "A Dynamical Theory of the Electromagnetic Field" in 1864 is an

ideal example of scientific progress in physics, unifying disparate theories and experimental results, producing a foundation large and stable enough for a huge scope of exploration, leading to additional breakthroughs and even paving the way for future technologies such as radio, television, and telecommunications. The rationalist dream would have been for work such as Maxwell's to perpetuate throughout physics, unifying theory and explaining experimental results, culminating in a complete model of our universe that can predict any natural system's behavior.

However, the universe had other ideas. Scientific inquiry along a particular thread initially plucked by French civil engineer and physicist Augustin-Jean Fresnel in the first years of the 19th century, well in advance of Maxwell's work, would eventually lead to serious challenges for the rationalist way of thinking. Aside from inventing the lens used in lighthouses to increase their range and visibility, Fresnel's contributions to optics included quantification of the interference of Young's double-slit experiment and other aspects of the wave-like behaviors of light. Presenting his work to the French Academy of Sciences in 1819 eliminated support for Newton's corpuscles once and for all. Light was now understood to exist as a wave and, as such, fit into the broader wave function theory that culminated in Maxwell's classical wave theory. Light was a fully integrated

phenomenon in the broader electromagnetic spectrum of radiation. However, not all the details of this comprehensive theory actually made sense.

In 1862, German physicist Gustav Kirchhoff, working on the interactions of electromagnetic radiation and thermal energy, created an idealized, abstracted object he called a "black body," which was nonreflective, opaque, and of uniform density. Its simplified form could be a stand-in for planets or other real-world objects that could be observed to emit heat when exposed to light or other electromagnetic radiation. What Kirchhoff was seeking was a mathematical model that would predict the amount of heat emitted in proportion to the energy of the radiation to which it is exposed. Experimentalists such as Scottish meteorologist and physicist Balfour Stewart picked away at the problem, trying to make enough observations for the theorists to stitch together into a generalized mathematically provable theory. By the early years of the 20th century, a theory named after British physicists Lord Rayleigh and Sir James Jeans, the Rayleigh-Jeans law, provided an accurate model for black-body radiation, but only at longer wavelengths. Extrapolating to the shorter wavelength radiation, the calculated energy emitted becomes infinite, an obvious flaw in the theory that became known as the "ultraviolet catastrophe,"[1] illustrated in Figure 3-3.

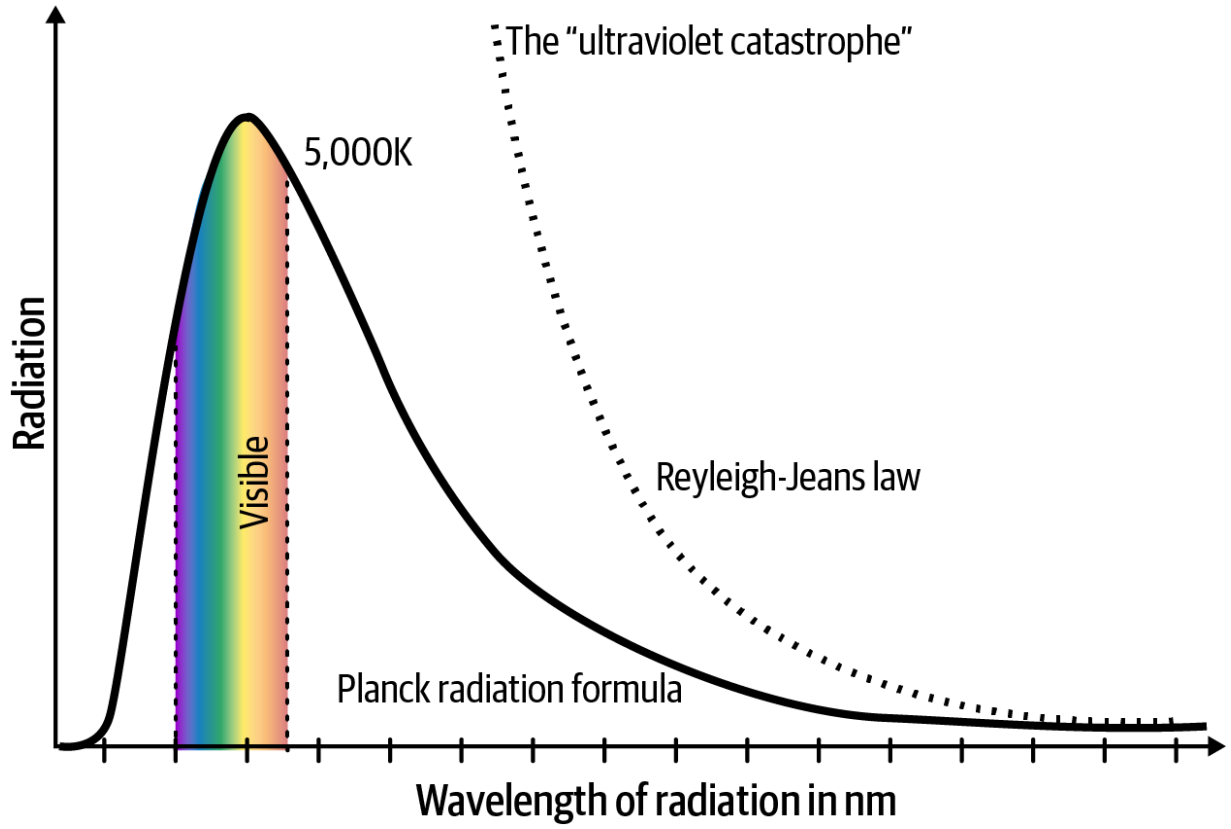


Figure 3-3. A plot of the incorrect calculations provided by the Rayleigh-Jeans law, which result in radiation energy going infinite, a physical impossibility

There was a solution on the horizon for this problem of the Rayleigh-Jeans law's incorrect result. However, it would open a door that would fundamentally undermine physics as understood in the 19th century and create knotty problems for the rationalist view of the universe that remain in many ways unresolved to this day. The originator of this creatively destructive solution was German physicist Max Planck, who would be awarded the Nobel Prize in 1918 for his breakthrough. Leveraging a mathematical framework developed by Austrian physicist Ludwig Boltzmann for his

work in statistical mechanics, he posited that energy existed in discrete quantities rather than a continuous field. This was no small leap—though Planck himself said he didn't think deeply about the theoretical implications as a threat to classical physics, calling his postulate about energy quantization "an act of desperation…I was ready to sacrifice every one of my previous convictions about physical laws."

Planck's desperate act consisted of applying a kind of brute force hack to the theoretical model of black-body radiation to make it work. The recasting of energy as something that comes in discrete "quanta," or quantities, had no foundation in the principles of electromagnetic radiation, and in fact, it was not done in response to Rayleigh-Jeans, as it is sometimes framed in textbooks. It was merely Planck's "kludge," or hasty improvisation, to make the model work, not Planck flipping the table on the classical world. In the same letter he went on: "purely a formal assumption…I really did not give it much thought except that no matter what the cost, I must bring about a positive result."

With this one impulsive reach for a solution to a prickly problem, Planck had inadvertently taken a crucial step toward a profound revolution in physics. Natural phenomena like electromagnetic radiation were presumed to exist as a

continuous spectrum, like the rainbow of colors extracted from sunlight by a prism. At a human scale, even at a microscopic scale directly observable by technological means, the continuity was readily apparent. Below that scale, extrapolations of the visible continuum fell apart. Now Planck's solution suggested that the radiation physicists were predicting theoretically and measuring experimentally was less like a volume knob on an old stereo, smoothly transitioning from 0 to 10 (or 11, as the case may be), and more like a channel changer on a television, snapping between channels with only nothingness in between.

The next shock wave came for physicists in 1905, when Albert Einstein tackled the difficulty that classical physics had with explaining the photoelectric effect. Experimentalists had observed that when metal is exposed to ultraviolet light, electrons are emitted. This is the phenomenon we rely on to make smoke detectors, digital cameras, solar panels, and all manner of other technologies work. However, the wave theory of light was unable to explain some of the specific behaviors the experiments observed. Einstein used Planck's "quanta" of energy and applied it to light, proposing that it exists in discrete quantities he called photons. This seemed to contradict the work Fresnel had done to establish the wave nature of light, even while the classical wave theory from Maxwell could predict and explain so much of light's observable behaviors.

Despite the paradox, Einstein's explanation produced results that matched the experimental results, the idea slowly began to take hold, and Einstein would go on to win a Nobel prize for the work in 1921.

Whether the Nobel for photons legitimized the early work of Planck and Einstein or their ideas had simply had enough time to incubate, the rest of the decade was a very busy one for quanta. In 1924 the French physicist de Broglie proposed in his doctoral thesis that the wave-particle duality was not restricted to massless photons, but that electrons—and by extension, all matter—could exhibit both wave and particle behaviors. This theory suggested there was no neat solution to the questions around subatomic behaviors coming around the corner. In fact, the more physicists investigated, the more they found confounding and mystifying behavior.

Mystifying behavior was further validated in 1927 by Werner Heisenberg, a German physicist who had been working at the University of Göttingen, along with Planck, until he was summoned by Niels Bohr to Copenhagen to help work with a problem. Bohr, a giant in the field already, had created formulas that could accurately predict properties of chemicals such as the frequency of light they would emit when heated before experimental results were achieved. However, the internal logic

of Bohr's theory mystified him, and on the advice of Wolfgang Pauli, he had called on Heisenberg to help. The issue was that, in order to work, the formulas had to assume that electrons orbited their nuclei only at very precisely defined orbits and not at some arbitrary distance from the center. What's more, when excited, electrons didn't seem to ascend linearly from a lower orbit to a higher one, the way you'd expect a satellite orbiting the earth to smoothly move into larger and larger orbits by firing their jets. Instead, the electrons seemed to leap from one allowable orbit to another. The accuracy of the predictions strongly implied those assumptions were correct, but Bohr had no idea why.

In June of 1925, suffering from hay fever, Heisenberg traveled to Helgoland, an island off the German coast with no trees—hence no pollen—to seek relief. There, at 3 a.m. one morning, he completed work on a hypothesis he later described as leaving him shaken and disturbed. Like Planck, he'd found himself desperate for ideas that would provide him with a way to make the mathematics work, and so he'd decided to take it as given that the strange leaps predicted in Bohr's equations reflected the actual situation at the subatomic scale, and had constructed tables of numbers that represented the allowed positions for each electron. In one position in the table was the electron's current orbit, in another, the orbit it was leaping to,

its destination. The table or matrix mathematics was clumsy, but it worked. The initial paper written from the Helgoland insight provided not only a working model to explain Bohr's strange leaps but also the raw material Heisenberg, Born, and Jordan Pasqual needed to create the foundational model for quantum mechanics. Bohr's orbital leaps are not arbitrary but seem to be part of the universe's internal design.

The following year, 1926, Erwin Schrödinger, an Austrian working at the University of Zurich, delivered a generalized theory of de Broglie's wave-particle duality that also succeeded in calculating Bohr's results. Schrödinger sought to create a mathematical way to describe and predict the behavior of a particle when it acted more like a wave. The resulting equation seemed a very elegant solution, calculating the function of a wave Schrödinger calls ψ, or psi, after the Greek letter, and it reproduces the same results as Heisenberg, without the awkward matrices. Heisenberg saw a critical flaw in the wave model for the primacy of Schrödinger's wave function in the observable behavior of waves. Traveling over a distance, a wave will arrive at a distant point a diffused version of itself, the ψ having spread out evenly over the course of its journey. An electron arriving at a location, however, arrives as a single point, a phosphor on a television screen, not a smear. Heisenberg found Schrödinger's work "repulsive" and was

greatly motivated to reassert his own explanation for the physical world.

One of the oddities of the matrix mathematics of Heisenberg's model was noncommutativity, in other words, the order of calculations matters. In most types of math, the terms on one side of an equation can be read right to left or left to right; it doesn't matter because "4 × 5" yields the same value as "5 × 4." However, in matrices, the order does matter. Multiplying the position of a quantum system by its velocity does not give the same answer as multiplying its velocity by its position. Because of the relationship between noncommutative pairs like position and velocity or momentum, an uncertainty exists between them, and Heisenberg realized that knowing the value of one made the value of the other attribute impossible. In 1927, he published his work in a paper called "On the Perceptual Content of Quantum Theoretical Kinematics and Mechanics," in which he made assertion regarding the uncertainty of the value of one of a noncommutative pair when one measures the other attribute of the pair.

Heisenberg's uncertainty principle provided a theoretical basis that proved the wave-like behavior of particles put limits on how precisely we could measure their properties. In fact, the more precisely we measured one attribute of certain related

pairs, like position and momentum or time and energy, the more uncertain the other attribute became. That is, if you tried to fix a particle's position in space, its momentum, or the product of the particle's velocity multiplied by its mass, became less and less certain. In this way, the uncertainty principle integrated Schrödinger's wave function as a behavior exhibited within Heisenberg's framework of quantum mechanics. Unmeasured, an electron or a photon's attributes were undetermined, a probabilistic wave, Schrödinger's ψ. Once observed, the wave function collapsed and the particle was a particle again. Heisenberg's uncertainty principle described a dead end for the path of rationalist scientific progress in the sense that it proved that it was impossible to precisely define all the characteristics of a given subatomic particle and therefore to fully understand the larger system of which it was a part.

Heisenberg, by formulating probability distributions that described a particle, implied that a particle had no fixed attributes but rather that they were determined probabilistically when measured. Worse, Schrödinger's equation implied nonlocal influences on those probabilistic outcomes, meaning a quantum particle's attributes might be influenced by some other particle at an arbitrary distance. This new physics was extinguishing the rationalist mission to trace the phenomena of the physical world and everything in it back

to its root causes, its Godhead, presumably a set of equations from which one could derive all the contents of the universe and all of the interactions and ultimate fates. Humans are fond of the assumption that there is an objective reality that may be extremely difficult to accurately discern but ultimately does have a solution, an ultimate answer, a rational reason for everything that has happened and will happen. As quantum physics continued to take shape, it seemed to pose a threat to the entire notion of objective truth.

These disturbing implications became the backdrop for a major physics conference held in 1927, in Brussels. Named for the Belgian industrialist and philanthropist who provided support for the event, the Solvay Conference had begun in 1911 and in 1927 was being held for the fifth time. Given the recent results from de Broglie, Heisenberg, Schrödinger, and others, the implications of the emerging theory for the behavior of quanta, or quantum mechanical physics, was understandably the central topic of debate.

A consensus was forming among a majority of the Solvay participants, who were aligning with the formal convergence of the recent work in quantum mechanics led by Niels Bohr, Heisenberg, Max Born, and others. Einstein, despite his role in discovering that light also existed as discrete quanta of energy,

was convinced that quantum mechanics was incomplete or incorrect in some way. In a letter to Max Born the previous year, he'd expressed his concerns with the disturbingly nondeterministic universe suggested by emerging physics, writing that "the theory says a lot, but does not really bring us any closer to the secret of the 'Old One.' I, at any rate, am convinced that He is not playing at dice." Einstein's concern connected back to the rationalist dream of a universe that would make sense to *someone*, even if it remained beyond the abilities of any individual to understand. With perfect knowledge—that of a hypothetical supreme being, for example—every aspect of the universe should follow from some other aspect, an unending chain of cause-and-effect certainty. What quantum mechanics suggested was that no such perfect knowledge was possible, that events in the universe were, in the end, entirely random and impossible to predict.

The participants at Solvay, seen in Figure 3-4, squared off into two factions. A group of the quantum skeptics, led by Einstein, faced off with Bohr and the others, who were stitching together the work of the previous 20 years into a coherent formalism.

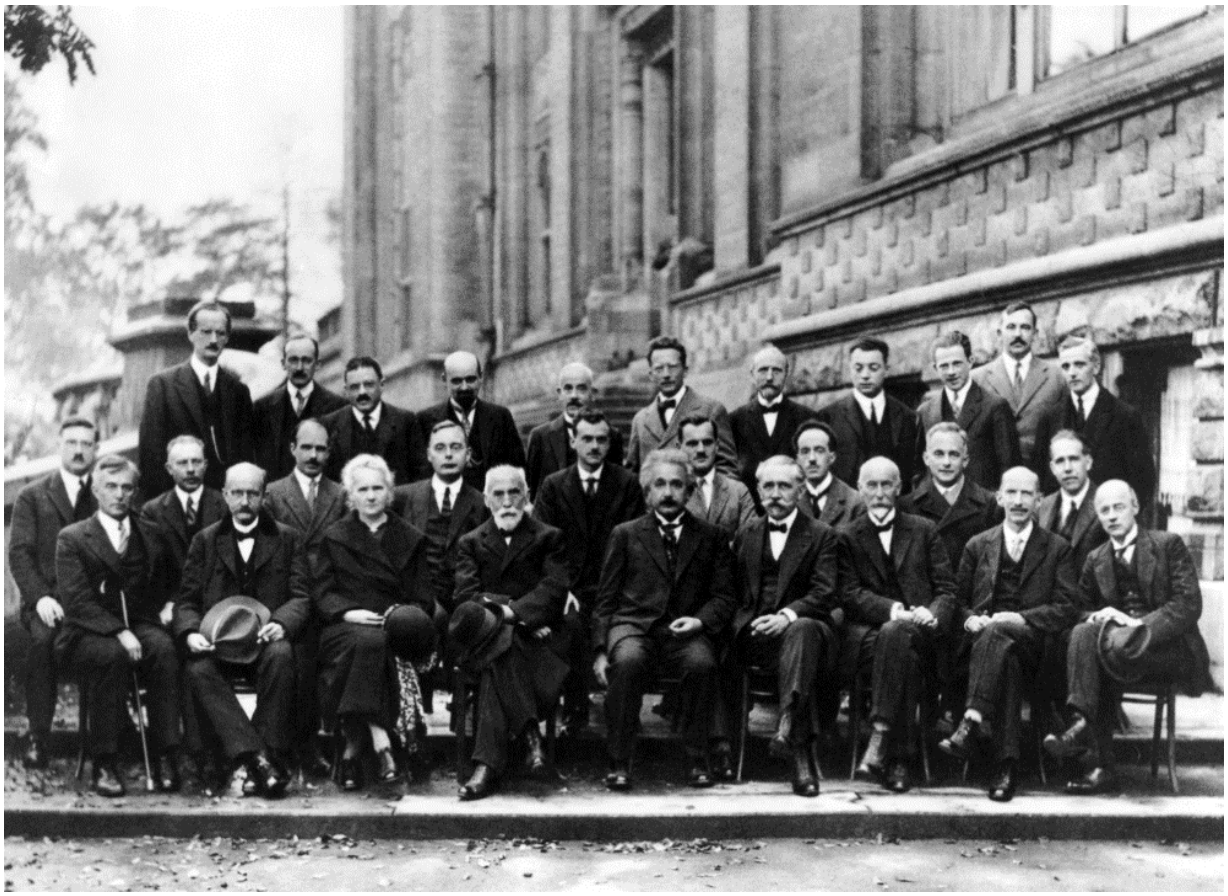


Figure 3-4. The participants of the Solvay Conference, 1927

This set of axioms included a belief that the wave-particle duality applied to quantum particles such as electrons and photons, which could behave as waves in certain contexts. While behaving as a wave, a particle can be in superposition of states, a concept that defies human intuition and most analogies used to describe it. If a particle has an attribute that has two possible states, let's call them "up" and "down," in superposition, that particle's state will be some probabilistic combination of both states. It is not accurate to say it is both up and down simultaneously nor that it is some deterministic mix,

for example, 30% up and 70% down. In fact, the probabilistic parts, the “amounts” of up and down, are unknowable. Once the particle is measured, its wave function will “collapse,” resolving into a single state, up or down, no longer in superposition.

As Heisenberg had previously discovered, certain pairs of attributes of quantum particles such as position and momentum can’t both be known at the same time, and the more accurately you measure one, the less precisely you know the other. Finally, Bohr’s contribution to all this quantum weirdness was the concept of complementarity, which held that certain phenomena such as the wave-particle duality, are determined by the experimental setup. Put in terms of Young’s experiment, if light is shone through a single slit, the photons will behave as particles, and if it is shone through two slits, they behave as waves. If one attempts to measure which slit each individual electron goes through, the wave behavior will disappear and photons will behave as particles again. In terms of our previous explanation of superposition, when in superposition, the photons have a probability density to pass through one slit or another, and when measured, their wave function collapses and they behave like particles again.

After Solvay, the two camps continued to develop their respective positions, with the quantum mechanics theory coalescing at Bohr's institute at the University of Copenhagen, coined the "Copenhagen interpretation" by Heisenberg in 1935. Einstein for his part published work with Boris Podolsky and Nathan Rosen that cast doubts upon quantum mechanics' completeness. The implications of the Copenhagen interpretation included the basis of the thought experiment that came to be known as the Einstein-Podolsky-Rosen (EPR) paradox. The EPR paradox posited a pair of particles prepared in such a way that, if one were to measure an attribute such as momentum of one particle, the momentum of the other would also be known instantaneously, even if the particles were separated by an arbitrary distance. This, the EPR paradox contended, would violate the theory of relativity, which held that nothing could travel faster than the speed of light. Since the measurement of one entangled particle would determine the collapsed state of the other particle regardless of the distance separating them, the implication was that the information was conveyed from one to the other faster than light itself could travel. In another letter to Max Born, Einstein laid out his objections to quantum mechanics, writing that "physics should represent a reality in time and space, free from spooky action at a distance." The crux of the problem for Einstein was

“nonlocality,” or the action of one particle on another particle at any distance from the first. The paper posited the existence of “hidden variables” not known to quantum theory that would provide a rational resolution to the paradox.

Einstein wasn’t the only pioneer of quantum theory who felt it must be incomplete. Erwin Schrödinger, whose wave function was so crucial to the Copenhagen interpretation, created his own thought experiment, also in 1935, in an attempt to prove the absurdities that quantum mechanics implied were possible. Writing in response to the EPR paradox, Schrödinger’s paper, “The Present Situation in Quantum Mechanics”, imagined a cat in a steel box along with a small amount of radioactive material and a device that detects radiation, a Geiger counter. If the Geiger counter detects a particle emitted by the radioactive material, it will trigger a mechanism that breaks open a vial of poison, and the cat will die. Since the emitting of radiation is modeled using the wave function, the Copenhagen interpretation holds that the wave will collapse when “measured.” Schrödinger’s cat would not “collapse” into a state of alive or dead until an observer opened the box and determined the wave function collapse. Before the box was opened, the cat would be both alive and dead in some probabilistic combination.

Various responses from Bohr and others suggested fallacies and alternative perspectives on both the EPR paradox and Schrödinger's cat, and the Copenhagen interpretation survived and even thrived. The long, gradual erosion of the primacy of classical physics was now complete. The Copenhagen interpretation remains to this day the dominant understanding of the workings of our universe. Classicism and rationalism had run their courses and been ultimately defeated by uncertainty and probabilities.

However, it's important to note that the Copenhagen interpretation, while historically significant and still widely taught, is just one competing view of quantum mechanics, and there are others with their own supporters and detractors. The debates from Solvay were just the beginning, with the Copenhagen interpretation setting the stage for a broad set of theoretical and experimental explorations yet to come. After the watershed year for quantum physics in 1935, its esoteric pursuits had to take a back seat to more pragmatic concerns. The world was at war, again, and scientists were being called on to contribute more directly and profoundly than in any prior conflict.

- Not to be confused with the album of the same name by the American band Trotsky Icepick, released in 1991.

Chapter 4. Midcentury Classical

> *In natural science, Nature has given us a world and we're just to discover its laws. In computers, we can stuff laws into it and create a world.*
>
> —Alan Kay

While the debate over the completeness and interpretation of quantum mechanics was captivating the attention of the physics community, the rest of the world was preoccupied with a very different topic in 1935. The previous two years had seen Hitler declare himself chancellor and then president of Germany, a self-styled "Führer," and in 1935 he declared the country would begin re-arming itself, violating the Treaty of Versailles. Also in 1935 Mussolini invaded Ethiopia in defiance of the League of Nations, revealing its powerlessness to stop aggressions, and the world braced itself for war. In the same year, Robert Watson-Watt, a Scottish physicist, demonstrated a method that used radio waves to detect a British bomber at a significant distance, proving the feasibility of a system called "radio detection and ranging," or radar. Daydreaming about the behavior of subatomic particles might be fascinating, but it did nothing for national security. Physicists turned from the

theoretical to the practical and began to contribute to their nations' readiness for war.

As the world descended into conflict, scientists continued to work on practical applications of the theoretical advances of the prior decades. Radar was improved through the collective work of allied scientists and engineers delivering improvements like the cavity magnetron, which used higher frequency microwaves to detect aircraft, greatly extending the range over radio waves. Wireless communication brought greater agility to the battlefield, but it also exposed a nation's plans and strategies to a new vulnerability. Radio transmissions were routinely intercepted, so the warring nations began to increase the sophistication of their encryption methods through the use of complex mechanical devices like the German Enigma machine.

Even capturing an Enigma device, as the British Royal Navy did in 1941, was not enough to break the German code, as the elaborate scrambling scheme had so many variations based on the seed number chosen before encryption that the best mathematicians and puzzle solvers in the UK could not defeat it in a timely fashion. The UK launched two separate secret projects to overcome the German code, both at Bletchley Park. One effort sought to build machines using mechanical relays,

called bombes, that would use the captured knowledge of the Enigma machine to accelerate the breaking of an encrypted message. Critical to the development of the bombes was a young mathematician named Alan Turing. While the bombe was not a computer per se, since it was only designed to carry out a narrow set of computations and was not programmable, the other project at Bletchley Park was. Known as “Colossus” and developed by engineer Tommy Flowers, it was the first electronic, programmable, digital computer, and it was used to crack an even more complex German code, the Lorenz cipher.

Flowers, son of a bricklayer, took evening classes at the University of London to earn a degree in electrical engineering, then joined the telecommunications branch of the Royal Post Office. When Turing asked his director in 1941 for help with the bombe, he was introduced to Flowers. The bombe used mechanical relays, and Turing was exploring how vacuum tubes—“valves,” as they were known in the UK—might improve his machine. Mechanical relays are electromagnetic switches, where an electrical current is used to physically open and close a circuit—a light switch with a magnet taking the place of a human finger to flick it up and down. Vacuum tubes had no moving parts, instead using current in one lead in the tube to induce current to flow between the other two leads. In effect, tubes moved the wall switch’s function *inside* the light bulb.

Flowers had designed and built telephone networks entirely using electronic valves instead of electromechanical relays, and he provided some advice. Turing was impressed by Flowers and encouraged him to develop his ideas further. Using his own money to develop the prototype, Flowers eventually convinced the government his system could work. By December 1943, the Colossus Mark I, with its 1,800 vacuum tubes, was operational at Bletchley Park. A restored Colossus Mark I is seen in Figure 4-1.

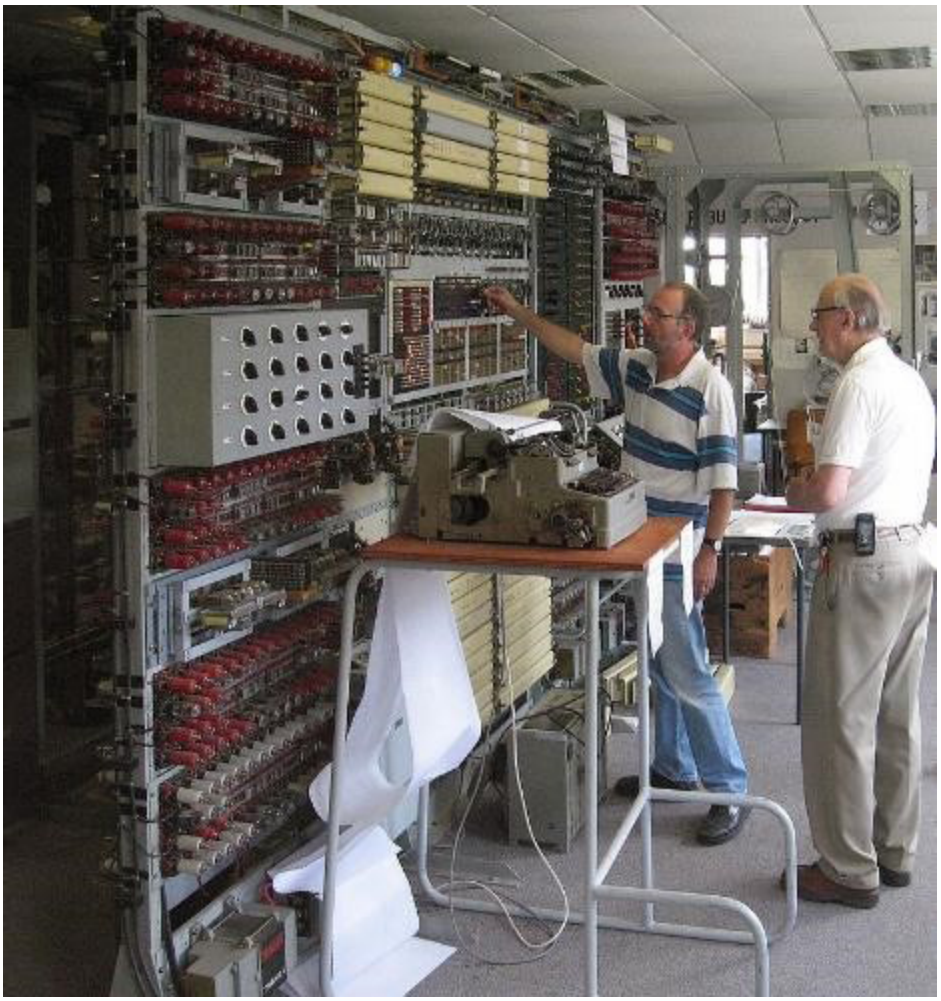


Figure 4-1. A rebuilt Colossus breaks a cipher at Bletchley Park in 2006 (MaltaGC, used under CC BY-SA 3.0, via Wikimedia Commons)

The Colossus system was programmable using a variety of switches, patch cables, with data entered via a paper tape input, all of which in combination could provide the inputs the machine needed to decrypt the Lorenz cipher message. Most notably, the input on the tape consisted of a series of holes that represented the encrypted message in binary form. While it was not a general-purpose computer, the Colossus could be said to be the realization of Charles Babbage's analytical engine.

Conceived of in the 1830s, the engine (plans for which are seen in Figure 4-2) was a mechanical computer designed to take punched cards as input for computations performed by a central “mill,” with results delivered into a memory store. The analytical engine was never built, but in the sense that it was designed as a general-purpose computer, it could be said to have been more sophisticated than the Colossus. Babbage had initially designed and prototyped a simpler machine, the difference engine, which was more calculator than computer and not general purpose.

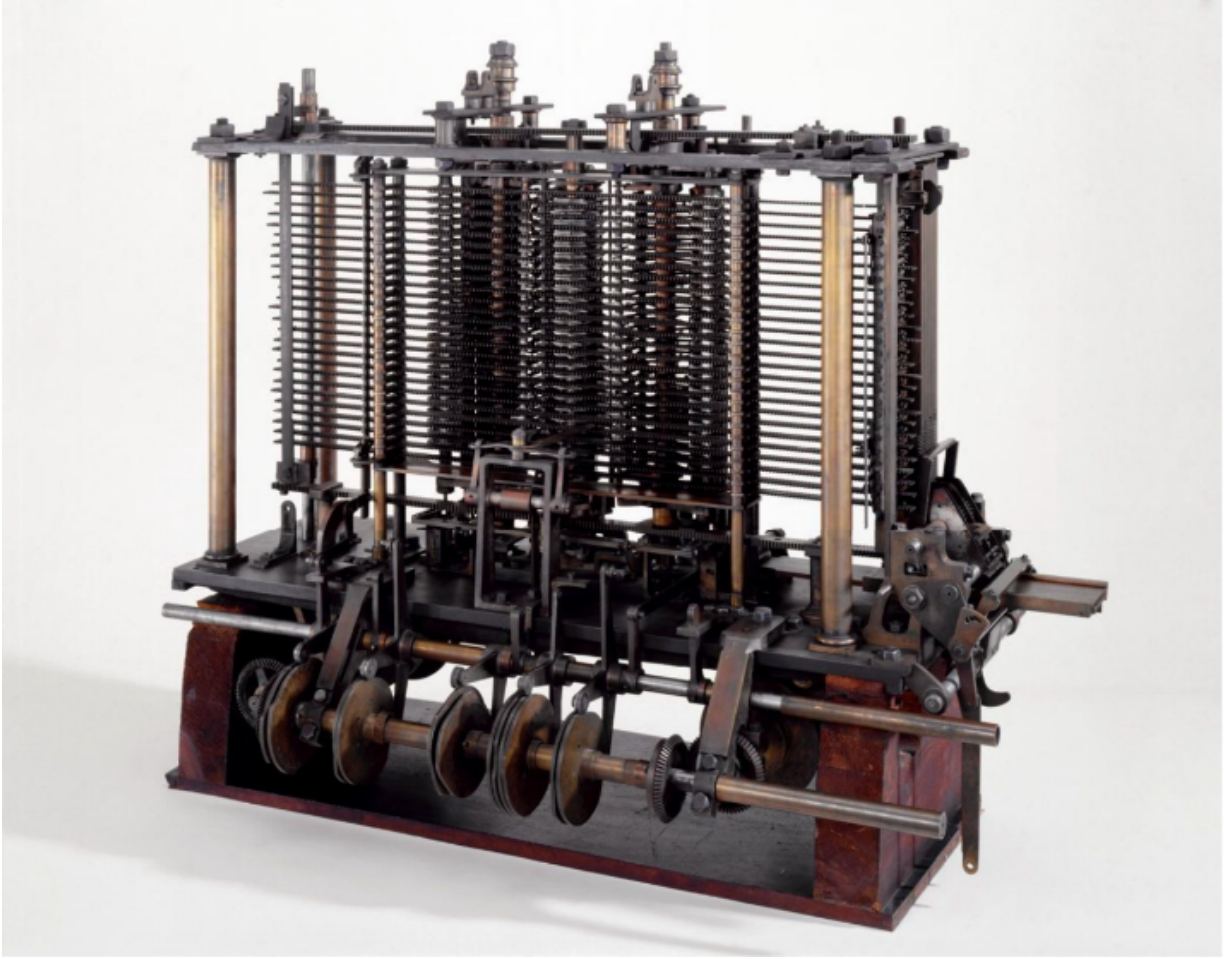


Figure 4-2. Trial model of Babbage's analytical engine, on display at the Science Museum in London, UK

His collaborator from 1840 onward, Ada Lovelace, was the first to recognize the potential of a general-purpose computer for more than just math, writing that it

> *might act upon other things besides number, were objects found whose mutual fundamental relations could be expressed by those of the abstract science of operations, and which should be also susceptible of adaptations to the action of the operating notation and mechanism of the engine…Supposing, for instance, that the fundamental relations of pitched sounds in the science of harmony and of musical composition were susceptible of such expression and adaptations, the engine might compose elaborate and scientific pieces of music of any degree of complexity or extent.*

Babbage and Lovelace form the canonical example of the inseparable dependence of computer engineers on computer scientists, and vice versa. The collaboration between experimentalists and theorists is sometimes called “codesign,” a term that attempts to capture the symbiotic and interdependent nature of the work done by both camps. In the middle of the 20th century, codesign was a vital ingredient in the astonishing pace of progress in computation. Before Alan Turing was helping design the bombes at Bletchley Park, he’d been a gifted mathematician, who in 1937 had published what has been called “easily the most influential math paper in history,” where he laid out the foundation for the entire field of computation.

Titled "On Computable Numbers, with an Application to the Entscheidungsproblem" the paper was Turing's attempt to address a problem posed by the mathematician David Hilbert, the Entscheidungsproblem," or "decision problem." Hilbert had asked in 1928 if there was a standard method or process that can be applied to a mathematical proposition, which gives a decision as to the truth or falsehood of the proposition. Hilbert was pursuing a question first raised by Gottfried Leibniz, who had built a calculating engine in the 17th century, 200 years before Babbage's analytical engine. The machine had caused Leibniz to contemplate the notion of calculable and incalculable problems. Hilbert's reformulation into the "decision problem," also known as the "halting problem," set the stage for a revolution.

Turing approached the problem through the theoretical work of another mathematician and logician, Kurt Gödel, whose incompleteness theory in 1930 laid out a proof that any set of logical axioms will have limits to what they can prove. Extending that line of thinking to Hilbert's question, Turing's 1937 paper described a hypothetical machine that read instructions and data from a tape and carried out operations based on the input. Having described the idea of a digital computer as a culmination of a centuries-long conversation among theorists, Turing found himself just a few years later

involved in the efforts to practically realize the dream of a universal computing machine.

Before the war efforts, Turing's studies brought him to Princeton university in the US, where in 1936 he began his PhD with Alonzo Church as his advisor. Church had also been working on the problem of computability, and in fact had published a paper very much along the lines of Turing's work. Their parallel efforts are recognized today as the Church-Turing thesis.

Flowers and Turing were not the only ones working on programmable computers. In fact, under the Nazis, the Germans had not only developed the Enigma and the Lorenz code machines, limited-purpose computers in their own right, but had also built, in 1941, the world's first automatic, digital, programmable computer. Built by Konrad Zuse, who, like Flowers, was an engineer, the Z3 system used a punched tape to load instructions and was able to perform floating point arithmetic. Meant to speed the calculations required for aircraft design, the system failed to capture the imagination of the Nazis, and when it was destroyed in allied bombing raids in 1943, thankfully, there was no effort to replace it.

Meanwhile, in the US, similar scientific and engineering efforts were afoot. At MIT, Norbert Wiener—an American-born child of Jewish immigrants, a child prodigy, a philosopher, and mathematician who had studied with David Hilbert—worked on systems to control antiaircraft guns automatically, using analog electronic means to track and predict the path of enemy planes. The US army set up the Signal Intelligence Service as a sort of counterpart to Bletchley Park, with a focus on breaking the codes used by the Japanese in the Pacific theater. And at Bell Labs, a researcher by the name of Claude Shannon was working on theoretical understanding of information.

Working on analog computing devices during his graduate studies in electrical engineering at MIT in the ’30s, Shannon had drawn on the work of George Boole to create switching circuits that allowed the simplification of the electromechanical relays on the machine. This, like the experience of Tommy Flowers, had applications in the design and operation of telephone networks. In fact, Shannon had, as a child, used a barbed wire fence as a telegraph wire to communicate with a friend. Shannon’s work after earning his PhD at MIT was more theoretical in nature. He joined Bell Labs in 1941 to work on fire control systems. along the lines of Wiener’s cybernetics work, and on cryptographic systems.

His involvement in code breaking brought him in contact with Alan Turing, who in 1943 had been sent to the US to share techniques from Bletchley Park with the Navy. During a visit to Bell Labs, Turing and Shannon met in the cafeteria, and Turing walked Shannon through his paper “On Computable Numbers,” which had a huge impact on Shannon. Much of his work during the war at Bell Labs culminated in his 1948 paper “A Mathematical Theory of Communication,” where among his many contributions he coined the word “bit” to mean a binary digit with a value of 0 or 1. His theoretical work was to lay the foundation for information theory and much of digital communications networks for decades to come.

The work of Tommy Flowers to build Colossus had its doppelgänger in the US as well. The US Army, wanting a faster way to calculate the large number of differential equations required for artillery firing tables (used to aim artillery barrages), funded an effort at the University of Pennsylvania to build a computer. John Mauchly and J. Presper Eckert began building ENIAC, or Electronic Numerical Integrator and Computer, starting in 1943. Inspired by the design of Vannevar Bush’s differential analyzer, for which Shannon had created his logical circuits, the ENIAC was, like Colossus, a programmable, general-purpose computer that used vacuum tubes for gates.

With the problems of breaking Nazi and Imperial Japanese codes largely solved, the champions of computing began to look to other challenges that their increasingly capable machines could tackle. Even before ENIAC was fully operational, the idea of using it to ease the calculations needed for the Manhattan Project was raised. The project to build the atomic bomb was enormously complex, requiring the work of a small army of human “computers,” mostly women, to carry out the mathematics. Frustrated by the slow pace of progress, Richard Feynman, then managing a group of computers at Los Alamos for the Manhattan Project, came up with an innovative approach to breaking down a large problem into smaller ones that could be solved in parallel, a technique that closely resembles modern approaches to parallel digital computing.

Another key figure at Los Alamos, John von Neumann, was well aware of both the ENIAC effort and Feynman’s and others’ efforts to make human computers more efficient. Von Neumann had been at Princeton’s Institute of Advanced Study (IAS) since shortly after its creation in the early ’30s. Von Neumann began lining up support from Princeton and secured funding from the Office of Naval Research and the Atomic Energy Commission before starting to build his machine in 1946.

With the war over, von Neumann's efforts were helped by the availability of surplus components such as vacuum tubes, which, like ENIAC or Colossus, the IAS machine used for logical gates. Memory was implemented using a novel approach where the cathode ray tubes were scanned continuously and what eventually would be called a pixel's state was either read out or changed. The IAS team took advantage of the fact that the glow of an activated spot on the tube would persist between scans of the electron gun and that the gun itself could "sense" when a spot was illuminated or not. Forty tubes were used to store an array of 1,024 "words," each 40 bits long, for a total 40,960 bits of memory. The IAS machine (see Figure 4-3) was also innovative in the way it loaded its program from punched tape into the same memory space as the data, creating an instance of Turing's tape abstraction in those tubes.

The team was challenged by the quality of the surplus components, especially the 6J6 vacuum tubes, about 2,000 of which the team used in the machine. Faulty tubes were easy to detect with startup diagnostics and could be replaced. The real issue was the tubes operating out of spec, introducing subtle errors over time to the machine's operations. Ingenuity and engineered redundancy led to solutions to the team's problems, all of which have left a mark on contemporary computers.

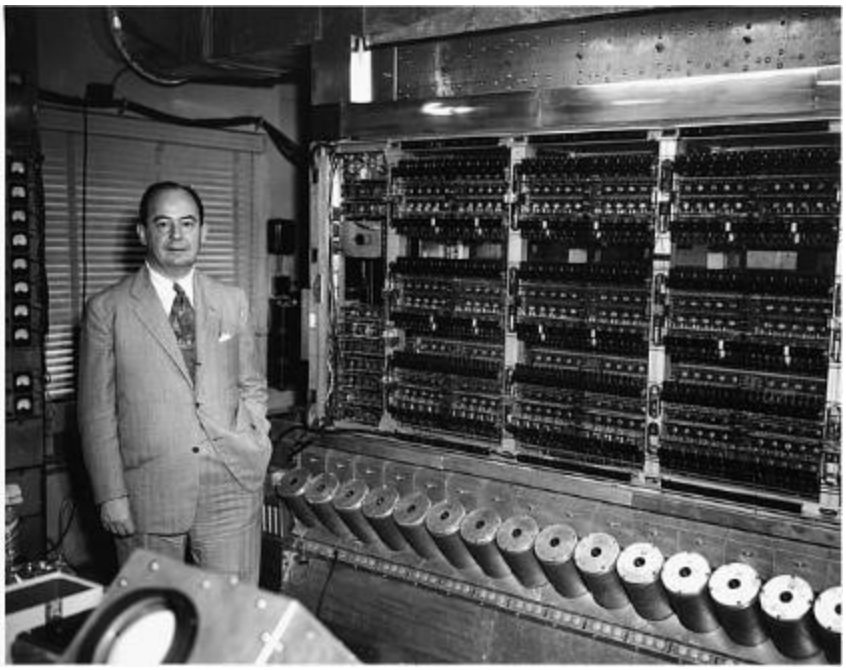


Figure 4-3. John von Neumann with the system built at the Princeton Institute for Advanced Study ("John von Neumann and the Electronic Computer at the Institute for Advanced Study, Princeton, New Jersey," ResearchGate)

What was potentially more challenging for von Neumann was to maintain support from the sources of funding in the new postwar context. His involvement in the Manhattan Project allowed him to create solid arguments for the use of his computer to continue research into nuclear weapons, a topic that seemed germane given the newly simmering rivalry with Soviet Russia. For those who wanted a postwar focus on peace and prosperity, von Neumann could emphasize the role of computation in the development of nuclear reactors for energy generation, beckoning to the promise of the "atomic age." More broadly, von Neumann seemed to have an instinct for the

ability of digital computers to aid in scientific inquiry, especially through the use of simulation or modeling.

In an effort to demonstrate the potential for computers to simulate natural processes other than the compression and blast of a nuclear weapon, von Neumann brought a leading meteorologist, Jule Charney, to IAS, where he built the first computer models to simulate weather patterns. Another researcher invited to use the IAS machine was Nils Barricelli, an Italian-Norwegian scientist who pioneered the field of artificial life, his programs representing mutations and evolutionary pressure among life forms simulated in the vacuum and cathode ray tubes of the IAS machine.

The researchers and engineers working around the IAS machine created the Monte Carlo method, an immensely powerful statistical approach that creates an approximate model of a complex system using random sampling. Invented by Stan Ulam, the Monte Carlo method remains tremendously useful to this day in physics, engineering, finance, and artificial intelligence.

Though the IAS machine was decommissioned in 1958, just 13 years after the beginning of the project, its impact on the development of classical computing was profound. Von

Neumann seemed to have a keen sense for the most powerful theoretical and experimental ideas from the scientists, mathematicians, and engineers contributing to the war effort, which he drew on to powerful results. Most importantly, he demonstrated a very open attitude toward the creations of the team at the IAS, happily running off copies of the machine's specification to anyone who asked. As a result of being shared freely in this manner, there were 15 documented "clones" of the IAS machine built at other institutions around the world.

In contrast, Eckert and Mauchly, who led the ENIAC and UNIVAC projects at the University of Pennsylvania, were quite territorial about their work. They complained that von Neumann and others were "snooping around" their work, and in 1947, they formed the Eckert-Mauchly Computer Corporation and began filing patents to protect what they viewed as their intellectual property. The legal fight over these foundational patents took decades to play out, first being granted to Eckert and Mauchly in 1964 and finally being overturned in the early '70s. In the meantime, the IAS clone devices, with derivative and humorous names like "MANIAC" at Los Alamos or "Johnniac" at the RAND Corporation, had a profound and lasting influence on computing. In what might be the first example of the power of an "open source" strategy in computing, von Neumann's influence is apparent to this day in

classical computing architectures in everything from a web server to the smartphone in your pocket, while the ENIAC architecture is lost to the ages.

However, if the computational demands that provided the motivation and funding for the Bletchley Park Bombes, Colossus, ENIAC, and the IAS system had remained their sole motivation, the momentum gained in the 1940s may have dissipated in the postwar era. Instead, it seems that every researcher who got their hands on the machines or even read about the work that was being carried out on them was able to imagine more uses for the technology, and dreams of a digital future continued to gain steam. In fact, immediately after World War II came one of the critical defining moments in the history of classical computing: the invention of the transistor.

The devices built during the war had transitioned from mechanical relays to vacuum tubes, which had brought huge gains in speed and scale. However, like a light bulb, tubes were prone to physical failure as their components heated and cooled repeatedly through their duty cycle. Thankfully, as interest in machines that used dynamic circuits to solve Boolean logical problems surged during the war, physicists were hard at work building out our working knowledge of the universe in ways that could improve upon vacuum tubes. In

particular, John Bardeen and Walter Brattain, two researchers at Bell Labs in New Jersey, were exploring the characteristics of materials known as “semiconductors.” The term was invented in the World War II era to describe a category of materials first identified at the end of the 19th century, which had a mix of characteristics of conductors like copper and insulators like glass. Of particular interest to researchers like Bardeen and Brattain was the ways in which semiconductors’ behavior could be manipulated and controlled. Their experiments resulted in a demonstration of a germanium strip with two electrodes, which, when a small current was applied to one end, would produce a much-amplified current from the other end.

Others at Bell Labs such as William Shockley built upon this experiment and, in 1947, created what they called the junction transistor. Consisting of three layers of semiconducting material sandwiched together, a junction transistor was capable of powerfully amplifying electrical signals, which would lend its name to the new, cheap, and compact “transistor radio.” More importantly, a junction transistor could act as a digital switch: when current was applied to the middle section of the junction, current could flow from one end to the other. Turn off the current to the middle, and the current could not flow.

The typical vacuum tube used in computing devices could last a few thousand hours in operation before its filaments burnt, its vacuum failed, or some other mechanical flaw ended its useful existence. With no moving or delicate components, the “solid state” transistor far exceeded the lifespan of tubes, with some transistors from the ’50s and ’60s in operation to this day. A huge barrier to building complex, useful computational devices that didn’t require the continuous application of specialized skills to maintain had been removed.

Bell Labs’ breakthrough ushered in the electronic age due to the simplicity, low cost, and durability of the transistor. The earliest pioneers of the vacuum tube era of computing—such as Remington-Rand, who had bought the UNIVAC from Mauchly and Eckert; IBM, with its model 701 “Defense Calculator”; or Ferranti and English Electric in the UK—all began transitioning to solid state machines. However, there were still challenges to building robust universal computers capable of productive work at scale. Transistor-based machines had to be assembled by hand, with workers soldering each tiny component to circuit boards. Solder joints are fallible, and a single cracked joint could cause the machine to immediately stop. Even the most skilled workers with the best materials and the most controlled environment will have a nonzero probability to create a joint that will fail. Multiply that probability by a sufficiently large

number of components soldered to a board, and the overall chance of failure will rise. The largest transistor-based mainframes by the late ’50s, such as the IBM 7090, had 50,000 transistors.

This challenge was the dominant theme of the 1950s, with huge amounts of energy and R&D budget focused on potential solutions, none of which made much of a dent. Approaches like MERA, or Modular Electronics for Radar Applications, funded by the US Air Force, which tried to establish a LEGO-brick system of electronic components that snapped together, typified the attempts to solve the problem.

Then, at the very end of the decade, researchers at Texas Instruments and Fairchild Electronics realized at almost the exact same time that the solid state of transistors meant they could be “printed” on a monolithic sheet of semiconducting material. By 1962 the defense industry was using integrated circuits for the guidance system of the Minuteman missile, and in 1964 IBM launched the System/360, with a combination of printed resistors and soldered transistors. The electronic age gave way to the computer age, and integrated circuits paved the way to the microprocessor.

In the early, heady days of the computer age, it seemed as though these machines were capable of anything, with each generation yielding ever more powerful devices. Gordon Moore's observation that compute power was doubling every year was originally made in 1965, and though the frequency had stretched to every two or so years by 2020, it's been an astonishingly constant rate of innovation.

These miracles of science and engineering, however, were imagined, designed, and built in the context of Shannon and Turing's classical information theory. While extremely powerful, the reliance on Boolean logic creates a weakness, with certain types of problems demanding exponentially more computing power to solve than even our most powerful machines can deliver. This shortcoming may be solved by the introduction of quantum mechanical concepts into information theory, if we can invent the machines to do it.

Chapter 5. Gedankenexperiment

> *Ordinary classical information, such as one finds in a book, can be copied at will and is not disturbed by reading it. Quantum Information is more like the information in a dream. Recalling a dream and describing it to someone else inevitably changes your memory of it. But unlike dreams, quantum information obeys well-known laws.*
>
> —Charles H. Bennett[1]

Turing's information theory had been turned into a realized technology by teams working at Princeton and the University of Pennsylvania in order to aid in the calculations needed to make quantum theory into a terrible weapon. After the war, the focus for those building computers was for applications of a less destructive nature, and so fields of information science and physics headed off, for the most part, in different directions. Quantum physics was there to help create the semiconductor-based integrated circuits and microprocessors, and create the lasers in CD-ROMs and fiber optics, but it had no role in algorithms and applications. That separation would not last long.

It is not usually possible when tracing the development of a scientific theory to find a tidy origin story for a breakthrough. However, this is not the case with quantum computing. The work of Rolf Landauer, a physicist who had joined IBM's research division in 1952, can be seen as a clear point of origin for the field. Landauer, a Harvard-trained physicist whose family had left Germany to escape Nazism, was interested in the physics of computation, especially the thermodynamic limits of computational operations. This was motivated at first by practical reasons, to understand how to better manage waste heat in computer systems. His interests would come to influence others in the field, inspiring new work or cross-pollinating with other scientists ideas until culminating in an integration of quantum mechanics into information theory to create a new quantum information science.

Landauer's own inspiration arose in part from another physicist, one we encountered in Chapter 3, James Clerk Maxwell, the creator of the classical wave theory of electromagnetism. Asked for feedback on a manuscript in 1867, Maxwell created an imaginary scenario in which two chambers are filled with gas and a passage connects one to the other. If the gas in one chamber were hotter than the other, the second law of thermodynamics dictates that the hotter atoms would tend to migrate toward the cooler chamber and vice versa, until

the two chambers had equalized. However, Maxwell imagines a "finite being," seen in Figure 5-1, who can close the passage between the two chambers at will, and who also possesses perfect knowledge of the energy, position, and momentum of each molecule of gas in each chamber. When a molecule approaches the passage, if its energy is higher than the average of all the molecules, the agent allows the molecule to pass into the other chamber. If its energy is lower, the passage is blocked and the molecule stays in the chamber. Over time, one chamber will become hotter and the other cooler, decreasing entropy and violating the second law of thermodynamics.

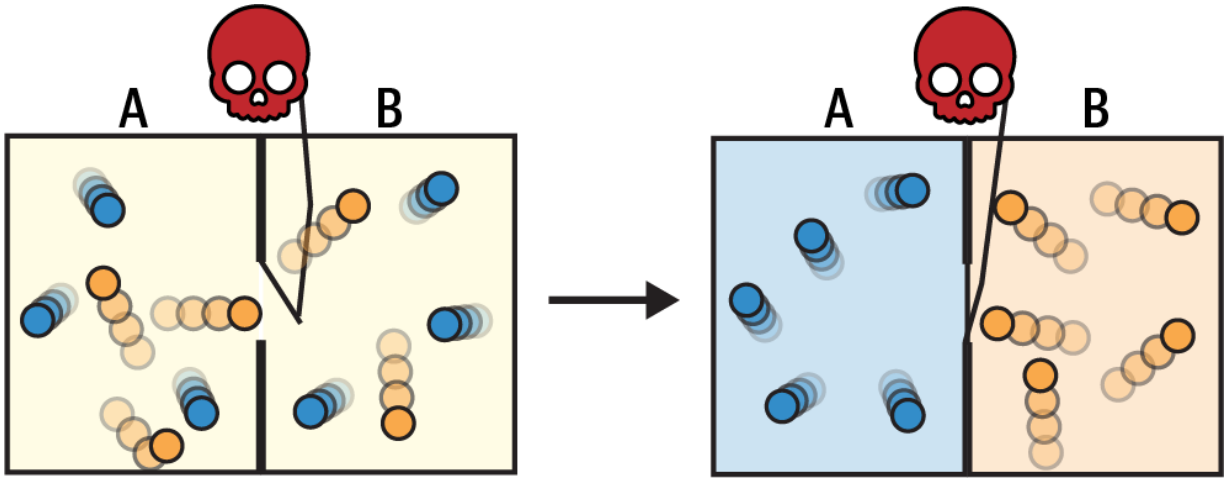


Figure 5-1. Maxwell's demon (adapted from Htkym, used under CC BY-SA 3.0)

Lord Kelvin later referred to the being as "Maxwell's demon," a name that stuck despite the fact that the very religious Maxwell wouldn't think of including a demon in his thought experiment. These "what ifs" are a powerful tool in physics and are often

referred to by the German word *gedankenexperiment*, at least since its use by Hans Christian Ørsted in his 1811 work "First Introduction to General Physics."

Maxwell's demon seems to be the type of idea that provokes irritation and a desire to explain why it is wrong in nearly everyone who encountered it. To many, it seemed self-evident that the demon's closing and opening of the connection between the two chambers must require work, as would acquiring information about the molecules to make the decision about that connection. However, none of the irritated responses generated solid theoretical rationales, until Landauer's paper "Irreversibility and Heat Generation in the Computing Process", published in 1961. His argument was that some of the logical operations carried out by computing machines were irreversible—once executed, there was no way to reverse them. This logical irreversibility, Landauer reasoned, was tied to a physical irreversibility, an increase in entropy of the system measurable by an increase in heat. Erasing a bit, the smallest unit of information, would incur a minimum energy expenditure, even in a perfectly efficient computing device. In a biographical memoir about him, Charles Bennett wrote that Landauer believed "computer designers and even theorists of computing ought never to lose sight of the fact that every bit must be embodied and every logic operation accomplished by

some real, physical apparatus." Information is intrinsically physical, as much a part of nature as mass or velocity.

Before Landauer, the physics of information was not rigorously studied, and the field operated with a rough approximation of heat generated by computation tasks. His work proved stimulating to the scientific community by widening the scope of physical science to include information, but also through his stipulation that only an irreversible change of information consumed energy. This implied that there may be a way to compute without consuming any energy at all, if one's operations were all reversible.

Aside from the practical implications of better managing energy consumption and heat output, reversibility hinted at deeper implications for computing. When we process information, whether in our heads on in classical computers, the operations are quite often irreversible. Logical operations in a computer are called "gates" because conceptually they can be thought of as something the information—represented in classical computers as electrical current—can be controlled by. The gate evaluates the information and, depending on its value and the nature of logical operation, allows it to pass through or does not. Some of these gate operations include the erasure of information and are irreversible. The "AND" gate performs a

Boolean logical operation, which takes two inputs and, if both inputs have the logical value 1, will output 1. If either input has the value 0, the output will be 0. Evaluating the output of an AND gate that produces 0, one has no way to know which of the inputs was 1 and which was 0, and therefore no way to precisely reverse the operation.

However, if all the operations performed to arrive at a result were reversible, the computation could be run backward to its original inputs. Some logical gates were reversible, such as the NOT gate, which simply inverts the value of a bit. If the output of a NOT gate is 0, it is certain that the input was 1, and so the operation can be reversed. In the early sixties, however, a set of reversible gates that were universal—capable of carrying out any possible computation—did not exist. The idea of reversibility proved to be an irresistible idea to Charles Bennett. Bennett was a PhD student at Harvard who happened to attend a talk by Landauer. Bennett had been studying chemistry, but the idea of reversible computation had a profound impact on him, so he switched to studying physics and ended up at IBM Research with Landauer.

A few others were also noodling on this question. Konrad Zuse, who had actually built the first working computer in 1938, the Z1, wrote briefly in his 1969 book *Calculating Space* on the

significance of reversibility but didn't take his thinking much further. Another was the largely self-taught polymath Ed Fredkin, who at MIT was playing with the idea of a computer made of a frictionless billiards table. If the information could be encoded in the position of the balls, with logical operations on their information formalized as a series of collisions and bounces and trajectories, Fredkin believed a reversible computer could be created. However, "frictionless" is the kind of assumption that tends to limit thought experiments to thought alone—there certainly isn't any easy way to eliminate friction on a billiards table! Fredkin's thinking did result in creating the "controlled swap" logical operation, also known as a Fredkin gate, in 1972. Interestingly, sources for the original date of the billiard ball computer and the Fredkin, or controlled swap (CSWAP), gate are hard to come by, possibly because Fredkin didn't follow a standard academic path in his career and had less pressure to document his scientific work. In any event, Fredkin did publish a paper in 1981 with Tommaso Toffoli called "Conservative Logic" in which they laid out both the billiard ball computer and CSWAP gates. This was a logical operation involving three bits that was both universal and preserved the number of 1's and 0's from the input through to the output. Preserving the balance of the input values in the output allowed for reversing the operation, and its universality

meant that it could be used for any conceivable logical operation or computation. Fredkin gates were the first theoretical proof that reversible logical computation was possible.

The very next year, Bennett published his paper "Logical Reversibility of Computation", which made the case that a universal computer could not only be made logically reversible but could also be made "thermodynamically reversible," meaning that, since no information was destroyed via irreversible operation, no energy would be lost. In the decade since Landauer's principle was published, Zuse, Fredkin, Bennett, and others had completed the theoretical foundation for logical and physical reversible computation. All that was lacking was any idea of how one could build a device that could realize the theory. The means of computation known in the early 1970s seemed very far from the kind of efficiency that reversibility would require.

However, a few years before Bennett published his reversibility paper, he had been talking with a friend of his, Stephen Wiesner. The son of Jerome Wiesner, science advisor to the Kennedy and Johnson administrations who later would become president of MIT, Wiesner was a PhD student at Columbia during the tumultuous convergence of the civil rights

movement, the rise of feminism, the war in Vietnam, and the rest of the emerging baby boomer cultural wave. During student strikes in 1968, Wiesner had hit upon an odd idea. Using what he had learned about quantum mechanical principles, he dreamt up “quantum money” that could not be counterfeited. Wiesner’s quantum money encoded the serial number of a bank note in the polarity of a stream of photons, which, being quantum in nature, could exist in both wave and particle form. In transit, they behaved as waves and carried the information, but if someone tried to read the serial number by measuring their polarity, that coherent wave state would “collapse,” destroying the information and making it impossible to create a copy of the bill.

Wiesner also came up with an idea he called conjugate coding, which took advantage of some other aspects of weird quantum behaviors to convey information and can make a credible claim to being the first proposal that quantum information’s idiosyncrasies could actually be useful. In addition to leveraging the impossibility of measuring quantum information without changing it in an obvious way, Wiesner also pointed out that two quantum bits could be used to carry three bits of classical information. These two insights would, in the coming years and decades, come to be known as the “no-cloning theorem” and “hyper dense encoding,” but in 1970 they were so

far ahead of their time as to sound insane. Wiesner wrote up his ideas in a paper that, according to colleagues, was submitted to the *IEEE Transactions on Information Theory*, and was rejected.

When Wiesner spent time with Bennett in 1970, he told him about his ideas, with Bennett feverishly taking notes, following along with Wiesner's logic. At the top of the page of notes pictured in Figure 5-2, Bennett wrote "quantum information theory," the first known use of the phrase.

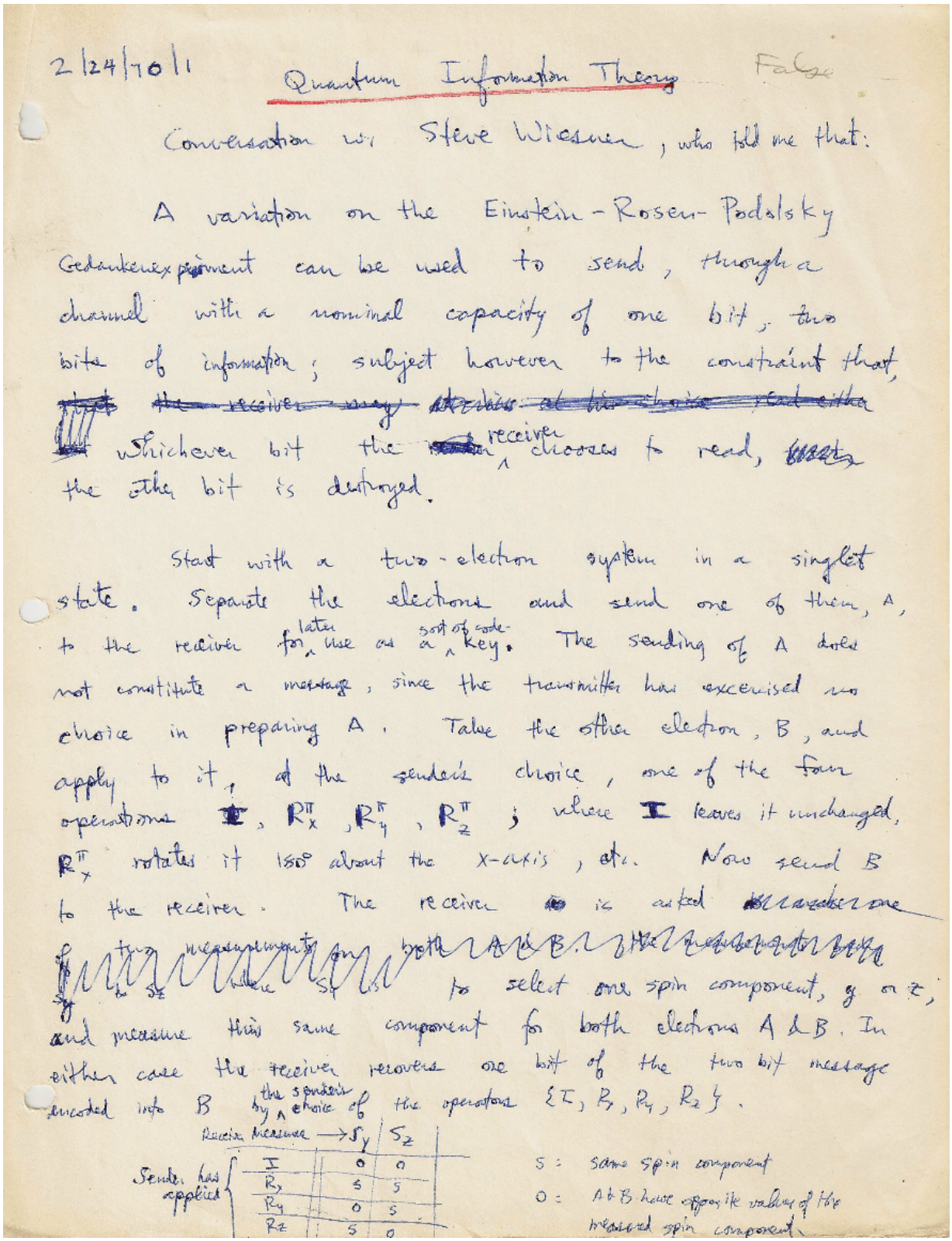


Figure 5-2. A page from Charles Bennett's notebook ("Stephen Wiesner", *My Quantum Thoughts: Or Sattath's Blog*, August 14, 2021)

Connecting principles known from quantum mechanics into information theory was the catalyst that the theory of reversible computing needed. What physicists call the arrow of time, meaning its tendency to move in one direction, toward the future, is a function of entropy, or lost energy. At the quantum mechanical scale are a class of operations used to describe the time evolution of closed systems that are unitary, meaning they do not result in an increase in entropy. They preserve the state of the quantum system and are therefore reversible. What Wiesner harnessed in his *gedankenexperiment* were examples of those unitary operations, and for Bennett, a connection from reversible computing to quantum mechanics was forged: quantum information theory.

Slowly, other unorthodox thinkers began pulling at this thread of quantum information, joining the efforts of Bennett and Fredkin. By the close of the decade, seminal work was being produced by pioneers such as Paul Benioff, Tommaso Toffoli, and Yuri Manin. Benioff, a researcher at Argonne National Lab, wrote a paper titled "The Computer as a Physical System" in 1980 describing how a theoretical model of a computer could be thought of as a quantum mechanical system. Another paper in 1980 by Yuri Manin similarly had talked about the theoretical possibility of quantum computing, though it contained scant details. Toffoli's "Reversible Computing" did not touch on

quantum mechanics but it did introduce a new reversible logical gate, the Toffoli gate, advancing the understanding of the theory.

Fredkin had also drawn his friend Richard Feynman to the topic of reversible computing. Feynman, a professor at Caltech in Pasadena, California, since 1950, was widely renowned for his contributions to quantum mechanics, and by 1980 he was a revered figure in the field. Through his books, public talks, and leadership of the investigation into the explosion of NASA's Challenger space shuttle, he would go on to become one of the best-known physicists of all time, rivaling Einstein. His popularity was aided by his curious mind and his love for inquiry and oblique intellectual strategies above all else, and he was famous for creative thinking.

Feynman's interest in computers traced back to his role on the Manhattan Project at Los Alamos, where some of the very first computations were performed by ENIAC for the design of the device. He'd led a team of physicists performing calculations the old-fashioned way, with brain power, devising a method for breaking down and running his teams' tasks in parallel, techniques that influence computer systems' architectures to this day. Fredkin had gotten Feynman interested in reversibility's potential to eliminate energy dissipation, which

posed engineering challenges when building computers, but he remained skeptical of Fredkin's enthusiasm for quantum information theory.

Finally, Fredkin and Bennett conspired to organize a conference on the subject, the previously mentioned Physics of Computation conference, to be jointly sponsored by MIT and IBM in May 1981. Cochaired by Fredkin and Rolf Landauer, the conference, held at an MIT property called Endicott House, would be an opportunity to get the little community of unorthodox thinkers together to thoroughly hash out the merits of quantum information theory. The group picture, seen in Figure 5-3, does not feature Bennett, because he was taking the picture, a role he'd repeat at similar gatherings of quantum information enthusiasts for years to come.



Figure 5-3. Group photo of the participants of the Physics of Computation Conference, 1981

Over the weekend, they presented their thoughts and debated various aspects of the idea of quantum information theory. To conclude the conference, Feynman delivered a keynote in the form of extemporaneous reflection on his perspective on the proceedings. In that talk, he expressed his belief that simulation systems from nature would be easier done with computers built on quantum phenomena like nature itself, and discussed the challenge of building such computers, remarking, “it looks like a darn hard problem to solve, by golly.” A certified genius calling a problem hard to solve is irresistible bait to a certain kind of mind, and the fledgling field of quantum information science was now truly a legitimate topic to work on. The community formed at Endicott fanned out across the world to spread the gospel of quantum computing. The idea that information was a part of the physical world and that information theory should incorporate quantum mechanical ideas was increasingly taken as rote.

Writing a review of the physics of computation published in 1982, Bennett used the ideas that began with Landauer to finally solve the mystery of Maxwell’s demon. While measuring the energy of the molecules and even opening and closing the passageway between the chambers could be accomplished without expending energy, the demon was not violating the second law of thermodynamics. It could be assumed that the

demon had finite storage for the information about the molecules' energy and so at some point would have to erase some of that information to make room for new inputs. As Landauer had shown, permanent destruction is the indelible link between information and the physical world. Erasing a bit will always expend a minimum amount of energy and so continues to increase the entropy of the system as a whole, as the second law dictates.

Bennett and the rest of IBM Research contributed far more than resolving this old *gedankenexperiment*. As the field spread across academic institutions, national labs, and research institutions, there was no organization more invested than IBM Research. Landauer was already a larger-than-life figure in the company by 1981, and Bennett would end up being elevated as well to IBM Fellow in 1995. Contributions, both theoretical and experimental, continued to be made by researchers at T. J. Watson and other IBM labs around the world, culminating in the public and commercial offerings starting in 2015 and reaching a kind of crescendo at the Think Q Summit in 2017.

The arc from Endicott in 1981 to T. J. Watson in 2017 would, for any classical technology, be the end of the origin story. The technologies we've seen emerge in the digital era would have transitioned from fundamental research questions to mere

engineering challenges, finding ways to leverage the fresh scientific knowledge into ever more capable systems. Companies would be working to find a product-market fit with increasing investment from the private and public sectors alike. The competitive struggle among the earliest stakeholders would play out, and IBM, Google, or another company would emerge victorious, or power would end up being shared by an uneasy coalition. The ensuing years would see its maturation, increasingly productive output, unintended consequences, and adaptive societal behaviors, as we have seen from the microprocessor, the internet, mobile devices, and so many others.

Quantum computing is not like all those familiar technologies. It was not unreasonable to ask if quantum computing even qualified as a technology in 2017, in the sense that no one had definitive proof that a physical machine could be built that would be able to solve a practical problem better than classical computers. That proof remains elusive to this day. Quantum computing is so outside the bounds of what we are used to from digital technologies because of its reliance on the inherent weirdness of quantum physics, a field that still experiences its own controversies and skepticism.

- Charles H. Bennett, “Information Is Quantum: What Weird Physical Phenomena Discovered a Century Ago Have Taught Us about Information and Information Processing”, presented at Imaginado o Futuro Rio de Janeiro, May 9, 2018

Chapter 6. Universes of Information

> *In the development of ideas, primitive pre-ideas often lead continuously to modern scientific concepts...It is, for instance, possible to trace the development of the idea of an infectious disease from a primitive belief in demons, through the idea of a disease miasma, to the theory of the pathogenic agent.*
>
> —David Deutsch

By the time the motley crew assembled at Endicott House in 1981 to debate whether quantum mechanics might be useful for computation, classical computing was settling into a form that would remain familiar to us to this day. The mainframe and terminal were giving way to the personal computer with monitor, mouse, and keyboard, and the Osborne 1, the first portable computer, had just been launched, pointing toward our mobile computing present (admittedly, at 25 pounds, it wasn't *very* mobile). From the perspective in the mid-2020s, the computers and networks of the '80s seem feeble, with bandwidth, memory, and storage measured in kilobytes, not gigabytes. However, that perspective is due to the extremely rapid rate of progress the technologies were enjoying. Thanks to

Moore's law, the 68,000 transistors in Motorola's eponymous 68K chip launched at the beginning of the 1980s was eclipsed by the Intel 80486's 1.5 million by the end of the decade.

With that barreling engine of progress at full tilt, it seems peculiar that anyone would be looking for alternative approaches to computation, as were the Endicott House conference-goers. For most, of course, the main motivator was pure curiosity—a very tricky puzzle to work out, regardless of the utility of any possible outcome. However, Feynman's explicit justification for the effort to define and establish quantum information science was to address the deficiencies of classical information science when it came to simulating natural systems. Classical computers, for all their increasing power, did have vulnerabilities that were already obvious to those who had been involved from the very start.

Feynman's central role at Endicott was not merely due to his stature in the field. At Los Alamos, his experience leading a team of human computers working on the math to model aspects of the Manhattan Project put him at the very point of focus for the earliest digital computing projects at the University of Pennsylvania and Princeton's Institute for Advanced Study. These machines were intended to perform the Manhattan Project calculations faster than any team of humans,

and though neither was operational before the Trinity test shot, both were seen as huge leaps in our ability to simulate natural systems. Given his experiences with manual calculations and then classical computerized calculations, in conjunction with his continued work on quantum electrodynamics and other fundamental contributions to the understanding of nature, it was unsurprising that he was able to see the promise of using quantum computers to simulate quantum systems so readily.

The issue that Feynman and the others in those early days had seen with classical computers was the limitations of Boolean logic when attempting to simulate highly correlated systems, like those described by quantum mechanics. In these systems, each element is interacting strongly with many other elements in some way. The simulation must represent those inter-relationships on a 1:1 basis, and so each additional correlated element adds exponentially to the connections necessary to represent it. Classical methods exist to work around this challenge. In an atom, for example, the electrons' spin interacts with that of all the other electrons present. If there are two electrons in orbit, you have two interactions to model. If you add an electron, you have four. Add another electron, you get eight. In addition, each electron interacts with the nucleus, and there are other interactions that will affect our model, all of

which dramatically increase the computational burden of the simulation.

It's the classic parable of the grains on the squares of the chessboard, which was first told in the 13th century about the game's creator in India. When asked by the king what reward he wanted for his marvelous invention, he asked for one grain of wheat to be placed on the first square of the chessboard, two on the next square, four on the next. and so on, doubling the number each time. The king laughingly agrees, saying it's a meager reward, and then his advisors tell him that by the 64th square, they will need to find a total of 18,446,744,073,709,551,615 grains of wheat. That's exponential growth—2 raised to the exponent representing the 64th doubling, or 2^{63}.

Nature, however, "computes" the answer to these highly correlated natural systems so efficiently that we, as observers, are unable to perceive the computation time. An atom bomb just explodes without waiting for its force to be calculated by equation. A wooden table is an unimaginably vast collection of subatomic particles arrayed in such a way that it just sits there, being a wood table, in real time. The universe must possess some knowledge or mechanism that allows the efficient calculation of these highly correlated quantum mechanical

systems, and so a computer built on quantum mechanical principles might afford those advantages as well. Most radically stated, information itself was being recognized as a natural phenomenon, intrinsic to the inner workings of the universe. A natural resource utilized by our abstract tools, instead of a human-made technology.

The origin of “information” as a distinct human technology is impossible to trace, given its self-referential nature—the tool is, after all, used to describe itself first and foremost. We can imagine, however, groups of humans recognizing various observations as generally important to share, repeat, preserve, and at some point these observations becoming a rote category —knowledge, or wisdom, or lore. Over time, knowledge would have inevitably split into subcategories. Examples include optimal agricultural and animal husbandry practices for a region’s climate, or the generally accepted means for a community to make collective decisions, a favorite story told around the fire pit, or a recipe for a well-loved dish.

As communities grew into societies and specialization required a means of tracking the production and consumption of collective resources like grain, means of accounting and mathematics emerged. Observations were recorded and preserved using marks in clay or papyrus, or through other

means. Accumulated observations and records represented the collected and preserved knowledge about a particular subject. Certain people began to devote their lives to understanding and extending the knowledge of particular subjects, and the information encoded in all those marks and pictures took on a life of its own, seemingly independent of its human origins.

Eventually Boole created his algebra, Turing created his abstract machine, von Neumann built his machine, countless other women and men made their individual contributions, and we had a universal logical machine equipped with a set of tools that could solve any problem expressed by the information theory understood in the middle of the 20th century.

At its heart, computation as we understand it is classical information science transformed into practice through a set of gates or transformations that can manipulate the states of bits to create any Boolean logical operation. A NOT gate turns a 1 into 0 and vice versa. An AND gate takes two inputs, and if both are 1, the output is 1. If either is a 0, the output is 0. Other gates such as XOR (“exclusive or”) and NAND (“not and”) are also defined, and in a classical computer, a set of these gates are implemented so that any logical “circuit” or operation can be carried out.

A classical algorithm such as a sorting routine or a Monte Carlo simulation is ultimately translated into a set of gates to be applied to a collection of bits, creating and executing a number of circuits, which, in totality, carry out the desired action. The algorithm describes the logical steps required to get to the desired outcome. For example, given an unordered list of integers, compare the first with the second in the list, and if the second is smaller, swap their positions. Move to the second item in the list and the third, and repeat the comparison and swap if necessary. Continue down to the bottom of the list, then return to the first number in the list, and repeat. If you get to the end of the list and you haven't swapped the order of any two pairs, stop. You have performed a bubble sort, and the list now starts with the smallest integer at the beginning of the list and each item larger to the last, the largest number at the end of the list.

A bubble sort or any other algorithm can be translated into logical operations in a classical computer, a blazing fast sequence of circuits composed of gates dynamically constructed and executed until the final result is found. Every application we use classical computers for, from *World of Warcraft* to a simulation of a nuclear reaction, is composed of algorithms instantiated as circuits made of logical gate operations. An algorithm can be thought of as a practical expression of an information technology.

But like an abacus, all of classical information is symbolic—a human abstraction meant to represent some actual thing. The beads on the rods of the abacus are merely representing the bushels of grain, though they can be used to perform arithmetic on them to forecast how far into the winter your stores will last. Bits are abacus beads instantiated in etched layers of semiconducting materials. They are a purely human invention, operating by rules of logic we defined, performing calculations in the way we invented. If there are limitations to classical information science, we must ask if those are a byproduct of the limitations of our own minds. Despite the efficacy of logic and mathematics, the question still remains, are they “real” or merely useful inventions of our evolved mammalian brains, the dreams of the universe, not the universe itself?

Quantum bits are realized through the manipulation of quantum mechanical systems that either exist in nature or are engineered to reflect a system from nature, and in that way they are not merely symbolic representations of information that may be following our own arbitrary rules. Qubits contain systems that exist in nature managed and manipulated to achieve desired effects. Quantum information does come from the laws that nature is giving us, but just as for classical systems, human minds have had to divine and distill those laws into a workable technology. The same way we had to break

down a pile of grains and figure out minimal universal laws to count them, we also had to learn how to break down piles of electrons and find out minimal laws to describe them universally, and that is what we call quantum logic. Our own experience with quantum behaviors was necessary to lead us to quantum logic, but we validate that logic against the behavior of the quantum world, unbiased by our history of experiences, thus offering a new space for doing calculations. The universe is made of information, and quantum information science imagines a method for plugging into the computer that is the universe. Put another way, by exploring the physics of computation, Landauer, Bennett, Benioff, Manin, Fredkin, Feynman, and the rest of the attendees of the conference in 1981 had realized that quantum mechanics could be used to reshape the technology of information science, to reform it out of the clay of the universe, and so surpass the limitations of the information science of our own invention. These dreams of the universe would emerge from the stuff of the universe itself.

After the Endicott conference, the focus of the work shifted from “why” to “how.” In order to grasp the overwhelming otherness of a quantum computer, we’re going to have to build more than a vague grasp of how they work; we’re going to have to look at the lowest-level building blocks. Don’t worry too much: we won’t be doing any real mathematics, just skimming

the conceptual framework of how the whole thing holds together.

Building a quantum computer requires the implementation of operations that mimic the set of gates that power a classical computer, but instead of being implemented as variations in current in microscopic electrical components, these gates need to operate at the quantum scale. Thankfully the foundations of quantum mechanical theory defined in the early part of the 20th century provided certain mathematical discoveries, which lend themselves to logical operations. In particular, the work of a physicist named Wolfgang Pauli offered a promising starting point. In 1927 Pauli had published a paper that defined a set of matrix operations acting on the spin of certain types of quantum particles such as electrons, photons, and protons. Pauli's matrices had been recognized as a powerful tool by Heisenberg, Born, and Pasqual as they had worked on the uncertainty principle, and over 60 years later, they were seen as a useful tool again, this time in the development of a theory of quantum information.

These matrices, known now as Pauli-X, Pauli-Y, and Pauli-Z, describe "rotations" of a particle's "spin" around three axes. The use of scare quotes is entirely intentional, as the term "spin" is used conceptually only—in no way is a photon or electron

spinning like a planet or top. The complete inadequacy of physical phenomena or objects to act as metaphors for the quantum scale results in its reputation as being nonintuitive. Nothing we have ever experienced at our human scale, or even at the smallest and largest reaches we can observe with the help of optical or other sensory technological aids, has any analog in quantum mechanics. Even the visible evidence of quantum phenomena, such as that exhibited by the double-slit experiment, defies comprehension. Everything the physics describes at that scale is utterly alien, and yet it constitutes everything that exists.

In purely abstract terms, these spin ½ particles' state could be represented as a vector on a three-dimensional graph and simplified further to a familiar x- and y-axis two-dimensional plot. The vector could be in two distinct states, with x equal to 0 and y equal to 1, or with x equaling 1 and y at 0, as illustrated in Figure 6-1.

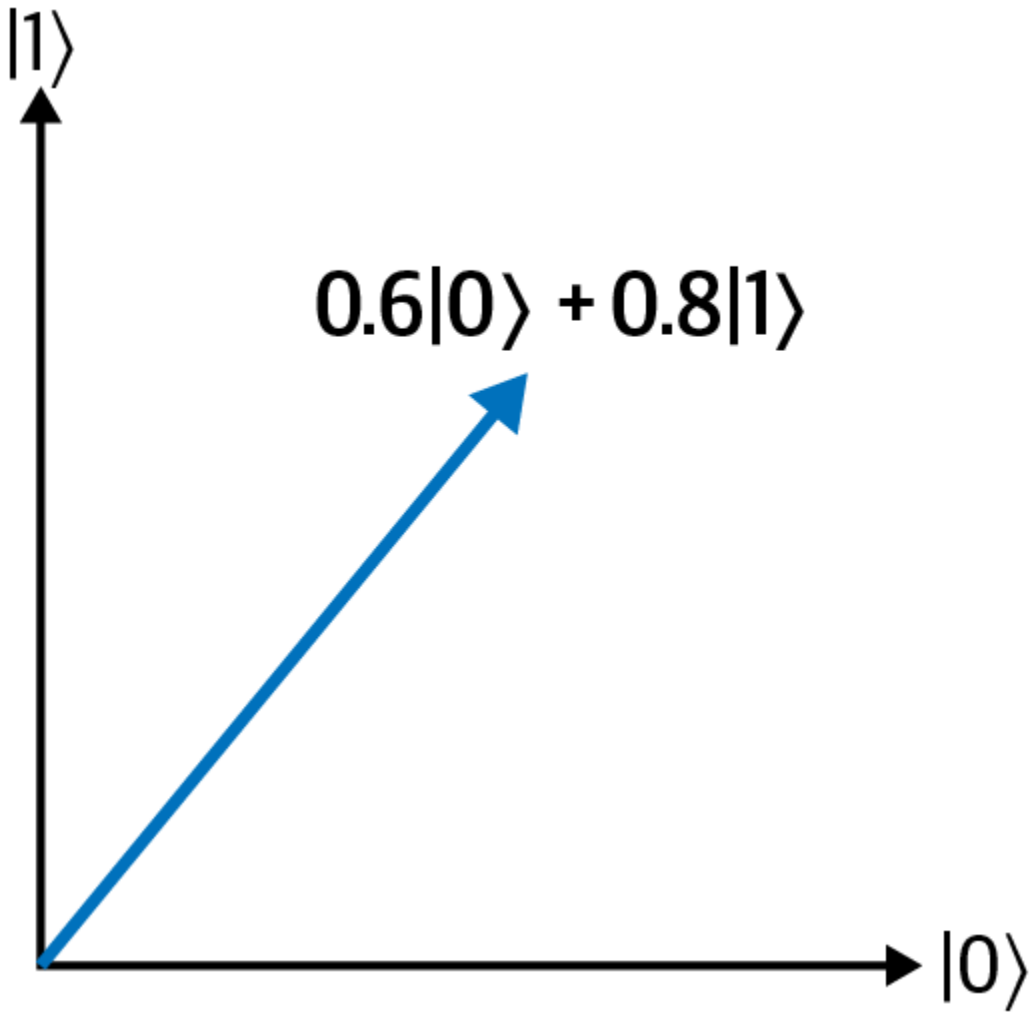


Figure 6-1. A two-dimensional plot of the vector representing a qubit's state

A vector can be represented as a matrix, with the two states of the particle represented as $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Pauli's operations multiply the particle's state matrix against other matrices he formalized due to the transformations they carried out on the state. If a Pauli-X matrix multiplication is applied to a particle's state, the x and y values will flip to the opposite. If x is 0, it becomes 1, a 1 becomes 0, and the same happens to the y

value. In the context of the conventions emerging for quantum computing in the 1980s, the Pauli-X was identified for its ability to act as a Boolean NOT operation.

You've noticed that the graph in Figure 6-1 doesn't use the large square bracket matrices to represent the computational basis states of a qubit, instead using $|0\rangle$ and $|1\rangle$ to represent the qubit's states. These symbols are named after Paul Dirac, a British physicist whose PhD thesis in 1926 was the first on the topic of quantum mechanics. Dirac notation is used in quantum information science to represent these vectors in the state space, using a symbol called "ket," or $|x\rangle$. In the spin ½ systems Pauli was working with, the [1,0] vector is designated as $|0\rangle$, while the [0,1] vector is $|1\rangle$. These two states are known as the "basis states." For computation, these map to our 0 and 1 states of classical bits. To represent a generic qubit of unknown state, we can write $|\Psi\rangle$, Schrödinger's psi, used by physicists to denote an arbitrary quantum state.

This set of logical gates carried out by quantum mechanical operations may appear to merely represent classical computing with extra steps. However, with the classical basics out of the way, we're ready for our first uniquely quantum operation: the Hadamard gate. Recall that the double-slit experiment demonstrated that photons can be thought of as particles that

act as waves under certain conditions, or waves that sometimes act as particles. Calculating the attributes of a particle behaving as a wave, according to Werner Heisenberg's uncertainty principle and the work of Erwin Schrödinger, the particle's state can be an indeterminate combination of possible states. This "superposition" is one of those quantum concepts that doesn't translate well into simple analogies because it has no near comparison in our scale. It is not the same as a flipping coin, or as being "all possible values at once," both of which are popular attempts to explain superposition that tend to make physicists very cranky. Nor is the cat actually alive and dead at the same time.

However, it is undeniable that superposition is a combination of probabilities of multiple states, and that when the state is measured, it "collapses" into a single state, with the probabilities of the alternatives irrevocably lost. Intuitively, this bizarre aspect of quantum information seems like it may be part of the mysterious "powers" the universe possesses as it computes its own existence so efficiently we can't observe it as it occurs. And so, we should have a gate operation for that.

Here, a French mathematician named Jacques Hadamard provides the equations for our use. The Hadamard transform, first published in 1893, can be used to manipulate our vector so

that its length remains 1, but its composition is now a combination of the $|0\rangle$ state and the $|1\rangle$ state, with some probability. In terms of the vector depicted in Figure 6-1, we have rotated it 90 degrees around the y-axis followed by a 180-degree rotation around the x-axis. Mathematically, if we call the vector length of $|0\rangle$ alpha, or α, then the absolute value of α squared, or $|\alpha|^2$, is the probability of $|0\rangle$. Similarly, the probability of $|1\rangle$ can be calculated as the absolute value of its vector length squared, or $|\beta|^2$. Combined, the probabilities of α and β always equal 1, because, well, the probability of the universe existing is 100%, so far as we know. In other words, the amplitude, or length of the state vector, remains 1, no matter what gate operations are applied.

While in superposition, the values of α and β are available for computation, but as soon as we measure the qubit, they collapse and are irretrievably lost, and the qubit is in one of its two basis states. The extra computational "space" during superposition is made even larger due to the complex nature of α and β. They are each represented by the equation a + b*i*, where a and b are real numbers, but b is the coefficient of the imaginary number *i*, where i^2 is equal to –1. You don't have to fully understand the math to recognize that $\alpha|0\rangle+\beta|1\rangle$ is *quite* a bit more sophisticated as a state space than that of a classical bit, that is, a 0 or a 1. If your algorithm requires elements of

linear algebra like matrix multiplication or very high dimensionality, qubits have got you covered. This concept of multidimensional space, with each qubit providing another dimension, maps to a concept called Hilbert space, named after David Hilbert, the same German mathematician who had posed the *Entscheidungsproblem* that had inspired Turing. Hilbert had hit upon the idea of infinite dimensional Euclidean spaces while working on differential and integral equations.

Crucial to quantum operators is the preservation of the total of the probabilities of the basis states summing to one. This total probability is represented by the length of the state vector of a qubit, which is also equal to one. Recall the fascination with reversible computation that helped motivate the investigation of quantum information by Charles Bennett, Ed Fredkin, and others. Many quantum gate operations such as the Pauli matrices or the Hadamard gate preserve the amplitude of the state vector as 1, and are said to be unitary, so that if the entries in the matrices are all real, the operations can be applied twice to a qubit, which will return to its original value, or “identity” matrix. In terms of the vector on a 2D plane, you can think of this as applying the same rotation twice, returning to the original value. The computations are reversible.

Additionally, unitary gates mean that the evolution of a closed quantum system, that is, one not being measured, is deterministic and therefore can be used for logical computation. Interestingly, the collapse of a superposition is considered probabilistic, with an uncertain outcome rather than deterministic. Despite the implications of this upsetting Einstein and others, it is another truly unique attribute of quantum computing, part of the native "language" of the universe, and a clue to the immense potential of quantum computing.

Even adding the Hadamard gate, our work so far can seem to be a mere projection of Boolean logic onto the quantum state space—a reconstruction of the "artificial" information science of classical computing. In order to support logical operations needed for anything humanly recognizable as an algorithm, this set of building blocks is necessary. It does not, however, limit the complexity of the quantum state space. The two-dimensional projection is the simplest representation of a qubit. The other Pauli matrices and another set of gates known as phase gates operate on other axes. Pauli-Y gates operate on the y-axis, and Pauli-Z and phase gates operate on the z-axis. Considering all three axes enables one to visualize a three-dimensional projection, represented by the Bloch sphere, as seen in Figure 6-2.

This three-dimensional projection reflects full informational complexity of a single qubit in quantum computing. The two-dimensional projection contains the classical 1 and 0 values as well as the superposition of those basis states, created by a Hadamard gate. Pauli-X, Y, and Z gates flip those basis states along the named axis. Phase shift, or S gates, rotate the state around an axis.

The collection of the gates we've examined so far are referred to as Clifford gates, and include the CNOT, or controlled NOT gate, which acts on two qubits to entangle their states. Clifford gates can be used to create a Hilbert space with a theoretical maximum number of dimensions limited by the number of qubits in the system. However, they do not represent capabilities unique to a quantum computer. In fact, Clifford gate operations are possible to simulate on classical computers with a relatively high efficiency, and therefore they don't represent a path to a substantial performance advantage from hypothetical quantum computers.

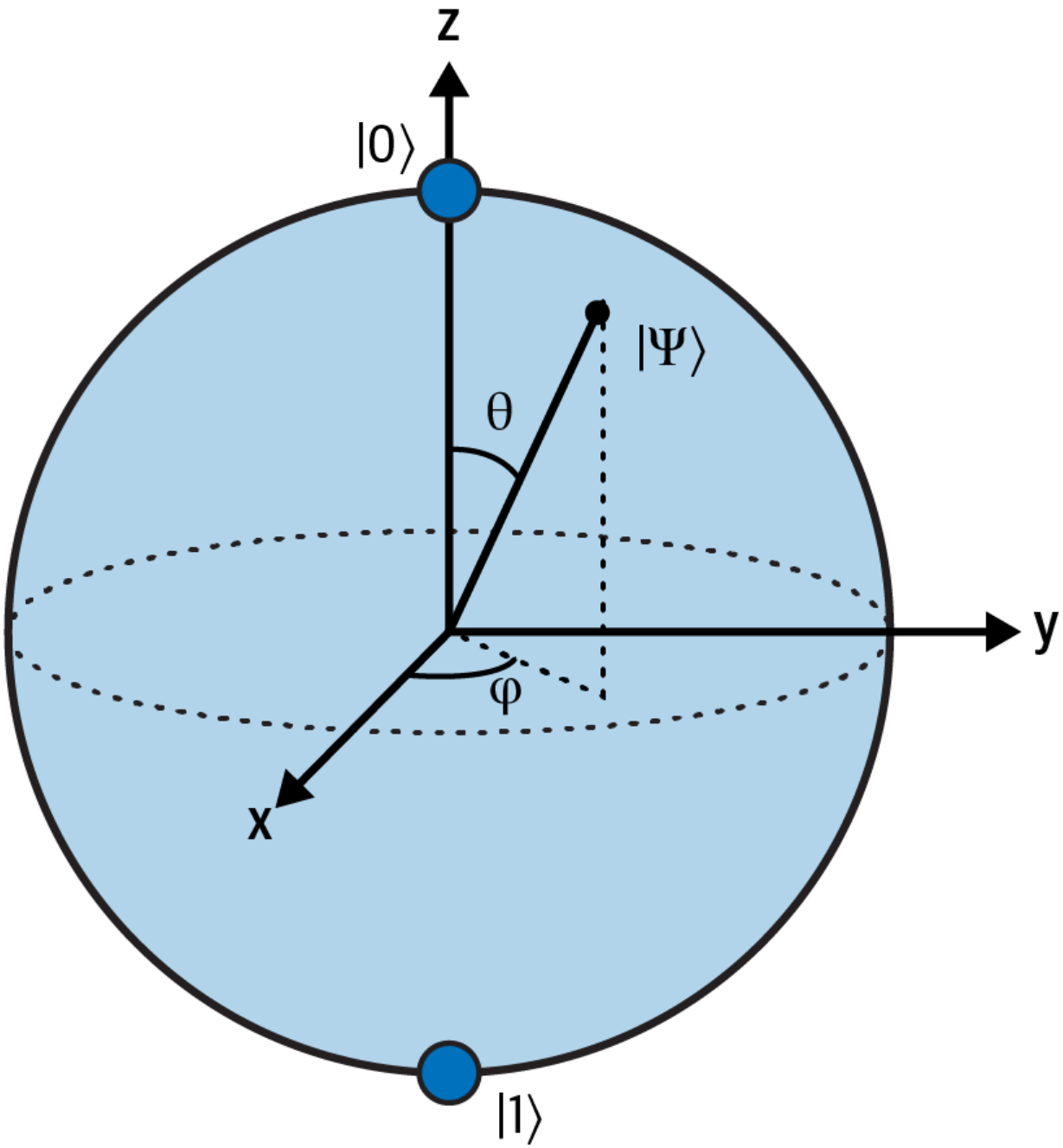


Figure 6-2. A Bloch sphere representation of a qubit state

Tommaso Toffoli's pre-Endicott paper, "Reversible Computing", mentioned in Chapter 5, helped crystallize the field. Toffoli is an Italian physicist and computer scientist who not only set the

stage for Endicott, but he attended the conference, attracted by the promise of unitary quantum operations' reversibility. In his paper, Toffoli described a gate with three inputs, where if the first two bits are set to 1, the value of the third bit will flip. This gate's operations are reversible and can be used to create a universal set of logical gates, able to solve any computable problem.

Implementing Toffoli gates, also known as controlled-controlled NOT, or T gates, in quantum information science achieves the same goal as in a classical system. Pairing Toffoli gates with the Clifford set of gates achieves universality, allowing a quantum system to solve any computable problem. What's more, its operations can perform tasks that are not efficiently simulated on a classical computer, and therefore it holds the promise of significant quantum advantages to computational performance.

Drawing on work by a disparate set of mathematicians, theoretical physicists, and computer scientists stretching out a century or more, the raw ingredients of a science of quantum information were gathered together. As a whole, these gates were needed to fully reformulate information theory through the principles of quantum mechanics. In the decade after Endicott, building on the earlier work of Paul Benioff, Manin, Toffoli, and others, a British physicist named David Deutsch

worked toward this formalized model of universal quantum computing. Deutsch recalls a discussion with Charles Bennett in the early ’80s in which he complained that working on a theory of computational complexity seemed to him a waste of time since it would always be dependent on the capabilities of a given computer—there was no idealized form of a universal computer. Bennett offered his contrary view that “the thing is, there *is* a fundamental computer. The fundamental computer is physics itself.” In 1985, Deutsch published “Quantum Theory, the Church-Turing Principle and the Universal Quantum Computer”, which extended the prior work to a formal theory of quantum computing.

Chapter 7. Shor and Everything After

This Is No Clockwork Universe

If the eternal dance of molecules
Is too entangled for us mortal fools
To follow, on what grounds should we complain?
Who promised us that Nature's arcane rules
Would make sense to a merely human brain?

—Peter Shor

Contemplating the years following Endicott, it can be difficult to understand what kept people engaged in the work in a field that had yet to prove if it would be good for anything. Deutsch's quantum update to the Church-Turing thesis, as important an achievement as it was, did not provide any justification for the effort it would take to build a quantum computer. Classical computing is universal, meaning that it is theoretically possible to solve any problem that can be properly expressed. Quantum computers, thanks to Deutsch's work, had theoretical proof that they were capable of exactly the same feat, no more and no less. In order to justify the investment required to build such an exotic machine, a promise of performance exceeding the capabilities of classical computers would be necessary.

It seems as though this question must have been preoccupying David Deutsch as well, because in 1992, he and his collaborator, Richard Jozsa, proposed an algorithm that provided the first provable quantum advantage over a classical computer. Provided a function hidden in a hypothetical black box, the algorithm would determine if the function's output would be "constant" or "balanced." A constant output function would take *n* bits as input and return the same value no matter what the input was, and a balanced output would produce 1 for exactly half of the input values and 0 for the other half of the inputs. That is, if your input is 4 bits of data, the constant function would always output 1 no matter what your input value was, while a balanced function would produce 1s for half of the possible values of the input, and 0s for the other half. The challenge is to properly identify the nature of the black box function in as few iterations as possible. Taking a classical approach, the worst-case solution would take just over half of the possible input values. In other words, take our 4 bits, which provide 16 possible input values. Suppose our black box function is balanced, and that it will output 0 for values 1–8, and 1 for values 9–16. If we start with our input set to 1, our output will be 0. If we increment the input by 1 and try again, we'll get another 0. At the 9th iteration, the black box will output 1, and we'll have proven it is balanced. If the black box

was constant, the 9th iteration will also be 0, and its nature will be revealed. Our classical algorithm will, in the worst case, with our n of 4, take $2^{(n-1)} + 1$, or 9 iterations to determine the nature of the algorithm, constant or balanced.

Deutsch-Jozsa uses quantum information to perform a nearly magical act: it invokes the black box function twice, and can determine whether it is constant or balanced without fail. It does so by encoding the input values into quantum bits and applying a quantum operation that puts them in superposition. It is then provided as input to a quantum version of the black box function, and after the second run, we have our answer. This was a demonstration of what Deutsch called "quantum parallelism," because it allowed for multiple input values to be tested simultaneously, reaching an answer very efficiently. How efficiently? The classical version of this algorithm required in the worst case $2^{(n-1)} + 1$ iterations to find an answer, where n is the number of bits provided as input. In other words, if our n is 4, as we saw previously, the worst case will take 9 iterations. If our n is 8, we may need a maximum of $2^{(8-1)} + 1$ inputs to get our answer, or 129. In both cases, in every case, Deutsch-Jozsa does it in 2.

You may be wondering, what good is an algorithm that can perform this task? What is the point of this task at all? In fact,

the point of this task is that it was designed to be particularly difficult for classical information theory and particularly yielding to quantum information theory. It has no other value than as an opportunity for the quantum theorists to show off. But what a result! What Deutsch-Jozsa proved was that a quantum algorithm could solve certain problems exponentially faster than classical approaches. As our value of *n* increases, the worst-case number of classical iterations increases exponentially, so that doubling the size of the input set to 16 bits increases our worst case number of iterations to 32,769, and doubling again to 32 bits increases to over 2 billion! Deutsch-Jozsa still gets an answer in one go, provided we have enough quantum bits to represent our set of inputs.

This was the result that the theorists working in the field were waiting for—a verifiable speed-up produced by a quantum algorithm—a justification for investing more time and effort into exploring this new frontier. Leading theorists from the University of California–Berkeley, Ethan Bernstein and Umesh Vazirani, leveraged Deutsch-Jozsa to create their own algorithm, Bernstein-Vazirani, which, instead of determining whether a function was balanced or constant, tried to learn a string encoded in a function. Their goal was to provide a more generally applicable framework for evaluating the complexity of quantum algorithms.

It's worth pointing out that what these quantum algorithms did was prove there was an advantage for at least one kind of problem, but in no way did it suggest there would be advantage across the board versus classical computers. This was an early confusion that drove some of the hype and skepticism around quantum computers, traces of which can be found to this day. No one has ever suggested that quantum computers would be good at spreadsheets or video games or other algorithmic tasks that have driven the design and optimization of classical computers decade after decade. Instead, what excited information theorists was the possibility that quantum computers could carve out new territory in the field of computational complexity, and be able to solve certain types of problems that were too challenging for classical computers.

The study of computational complexity has its roots in the same work that spawned computers, and began to be formalized with the Church-Turing theorem and subsequent work to refine it. Essentially, in complexity theory exists a "polynomial space," referred to as PSPACE, which contains all possible problems. Within that space, as illustrated in Figure 7-1 and summarized in Table 7-1, there are different classes of problems characterized by the amount of time it takes to arrive at a solution via computation. The class of problems that can be solved in "polynomial time" are called P, and are considered to

be solvable efficiently on classical computers. They are referred to as polynomial because the time to arrive at a solution scales linearly with the difficulty of the problem. In other words, a P problem that is solved in x time would, if you were to double its complexity, still be solvable, now in $2x$ time.

Other classes of problems are not as yielding and are referred to as NP, for “nondeterministic polynomial” time scaling. While NP problems are solvable, no efficient algorithm is known for them, and in the worst case the time they take to solve may scale exponentially, so instead of $2x$ for a problem that is doubled in complexity as we saw with the P example previously, an NP problem may take x^2 time. What the work of the first pioneers in quantum algorithms suggested was the existence of a new complexity class, a class that contained not just P but some elements of the NP class. These NP problems that potentially scaled exponentially as their complexity increased may be solvable in polynomial time.

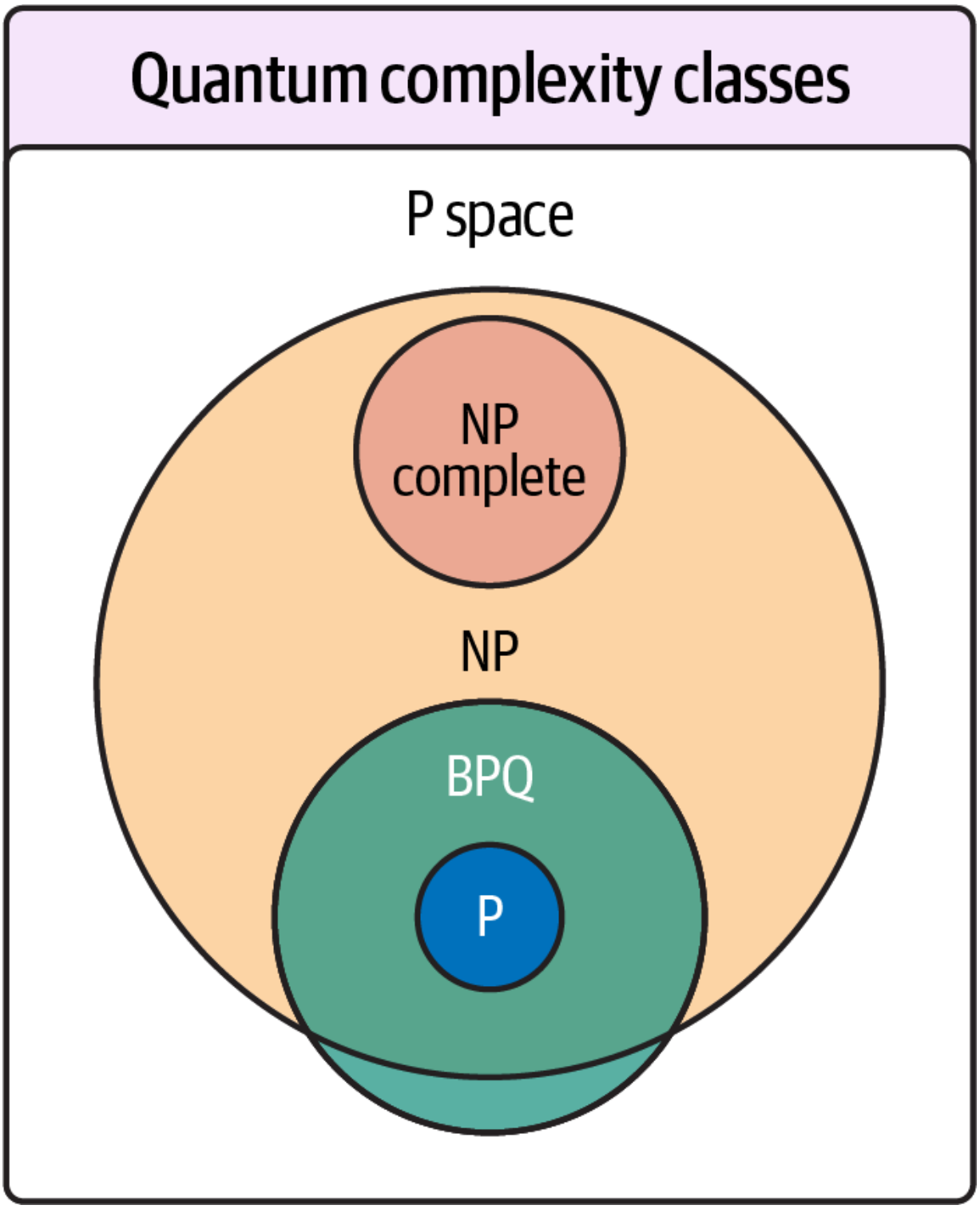


Figure 7-1. Computational complexity classes, including the quantum classes (adapted from Himanshu Sahu and Hari Prabhat Gupta, "Quantum Computing Toolkit from Nuts and Bolts to Sack of Tools," arXiv, March 6, 2023)

Table 7-1. Classes of problems in complexity theory

Polynomial space problem (PSPACE)	**Space requirements are polynomial bound (e.g., NxN chess).**
Nondeterministic polynomial time (NP)	Solutions can be guessed and verified in polynomial time (e.g., graph isomorphism).
Nondeterministic polynomial time complete (NP-complete)	This is a class of problem for which no efficient solution is available (e.g., traveling salesperson).
Bounded-error quantum polynomial time (BQP)	Problems can be efficiently solved by a quantum computer (e.g., factoring discrete log).
Polynomial time solvable (P)	This means the problem can be efficiently solved by both quantum and classical computer (e.g., graph connectivity).

This new quantum complexity class became known as bounded-error quantum polynomial time (BQP). The excitement that Deutsch-Jozsa and Bernstein-Vazirani

generated boiled down to the fact that quantum computers might be exponentially faster than classical computers for these tasks.

To excite interest in those whose research did not immerse them in computational complexity theory, however, it was going to take a proof with some practical purpose. The next step on that journey was taken just two years later, by Dan Simon, a researcher working at the University of Montreal. Simon's algorithm didn't do anything useful either, at least not directly. It did, however, demonstrate some very interesting attributes in terms of the task it tackled. Simon expanded on a concept that Bernstein-Vazirani had extracted from Deutsch-Jozsa, called a quantum oracle. This was an abstraction of Deutsch-Jozsa's black box. In this case, Simon used it to speed up his solution to something called the "hidden subgroup problem" (HSP). Thanks to quantum parallelism, Simon was also able to achieve an exponential speed-up, reducing the time from 2^n to simply n time. HSP is a technique that can be used to solve problems where you are trying to find an optimal path, called a graph problem. An example is the "traveling salesman problem," where the challenge is to find the most efficient route that encompasses all the salesman's stops. More generally, HSP can be used to find the period, or hidden repeating structure, of a

signal. This last fact proved highly suggestive to another researcher, Peter Shor.

Shor was a researcher at AT&T Bell Labs, where, after seeing a lecture from a visiting Charles Bennett, he developed an interest in quantum information science. He was drawn to Simon's algorithm in particular, realizing that it could be adapted to calculating discrete logs very efficiently. Discrete logarithms, or logs, are the exponent required to apply to one number to get another. In other words, given real numbers a and b, the logb(a) equals x such that bx = a. This is easier to demonstrate than to explain. In base 10, the powers of 10 are 0.01, 0.1, 1, 10, 100, etc. For any number in this list, you can easily compute the log. For example, the $\log_{10}(100)$ is 2, and the $\log_{10}(0.01)$ is –2. In other words, 100 is 10 to the power of 2, and 0.01 is 10 to the power of –2. This example for base 10 is simple. However, in classical computing, calculating discrete logs of other integers is considered intractable—not in the sense that it is unsolvable, but that producing a solution will take so long as to not be useful. Shor's adaptation of Simon's algorithm, on the other hand, yielded an exponential speed up in its solution, leveraging the oracle, quantum parallelism, and a variation of Simon's HSP solution. Put simply, it could produce an answer in a practical amount of time.

This initial breakthrough was sufficiently impressive that Shor was invited to give a talk at Bell Labs by Len Adleman, a very highly regarded mathematician and cryptographer. Adleman introduced Shor by saying in effect that “there are two kinds of geniuses. There are the da Vincis, whose inventions are beautiful but ultimately don’t work, and the Babbages, whose inventions are beautiful and eventually do work. I don’t know if Peter is a da Vinci or a Babbage.”[1] This compliment made such an impression on Shor that it is a permanent part of his memory of the extraordinary year he had in 1994.

Members of the audience must have been impressed as well, because the news of Shor’s algorithm made its way to Umesh Vazirani in Berkeley seemingly overnight. Vazirani actually recognized something even more powerful hidden in Peter’s initial algorithm, and called him on the weekend to ask if he’d looked at factoring. In fact, Peter had. As he tells the story, after the talk on a Tuesday, he realized his work could be adapted to efficient factoring, which he was able to describe to Vazirani over the phone just a few days later, despite suffering from a cold at the time.

In April 1994 Peter published his work, and his factoring algorithm—what would be known as Shor’s algorithm—hit like a lightning bolt. Prime numbers are those that, as factors, have

only themselves and 1—no other number divides evenly into a prime. It is, of course, trivial to multiply two large prime numbers, n and p, together, to arrive at their product, q. However, it is very inefficient, given q, to reverse engineer the prime factors of q, namely, n and p. The classical algorithms known for this task have to "brute force" a solution by trying to divide any candidate answer by all numbers smaller than it to prove its prime status. The larger the potential answer you are testing, the more numbers you need to test it with, and the larger the number you're trying to factor, of course, the more numbers might be one of its prime factors. The time taken to solve this classically expands rapidly into intractable territory—while Shor's algorithm can find those primes exponentially faster.

Shor accomplished his magic with a very original approach to the problem of factoring and unique quantum mechanical phenomena. First, inspired by Simon's oracle, the algorithm prepares a state that is projected onto each of the system's basis states—our measured, probabilistic combination of $|0\rangle$ and $|1\rangle$ from before. Each of these basis states represents a possible answer, and collectively they represent the set of all possible answers. Then, he very cleverly recast the factoring problem as one of finding periodicity. Using a classical algorithm to find the period at which the factors of the input number repeat, Shor

then uses that period value to drive a quantum version of a function called a Fourier transform across the quantum bits in superposition, utilizing what Deutsch called quantum parallelism to achieve his exponential speed-up to produce all the number's factors. A simple classical computation after the quantum steps determines if the results are prime numbers. The classical pre- and postprocessing added some overhead to the quantum computation, but not enough to dull its impressive performance advantage.

The real impact of Shor's algorithm wasn't just that it did something much more efficiently than classical computing. It had to do with what exactly it was doing. Cryptography is a set of techniques for obscuring information one wishes to keep private, and in the digital age, we have relied on one technique in particular, namely, using the prime factors of very large numbers as our keys to unlock the secret information. The classical factoring problem is of sufficient complexity that it forms the basis for many modern cryptographic systems, most notably Rivest-Shamir-Adleman (RSA)—yes, the same Adleman who hosted Shor for his first public talk on his discrete log algorithm. RSA and other cryptographic algorithms like Diffie-Hillman, are public-key or asymmetric key encryption schemes. These cryptographic keys rely on the "one-way" nature of prime factoring, making encryption very efficient for anyone

possessing the public key but highly inefficient without the private key. The private key contained knowledge of the primes required to decrypt the message. Modern cryptography on classical computers relies on the fact that there are no shortcuts to factoring a large number to its primes, and Shor seemed to have just invented such a shortcut.

And so, overnight quantum computers went from an odd thought experiment that might not do anything for anyone to seeming to bring the movie *Sneakers* to life. In the film, released in 1992, a technology is created that can decrypt all secret data, instantaneously. Shor's algorithm would not deliver instantaneous results, but it could reduce the code-breaking time drastically. Some models estimate that a 2,048-bit RSA code, which would take classical computers trillions of years to break, would yield to a quantum computer with about 4,000 quantum bits in about 10 seconds. In another cosmic coincidence, when the filmmakers wanted to make the script for *Sneakers* seem more believable, they hired a mathematician with significant cryptographic experience to act as a consultant: Len Adleman.

Shor's algorithm was no Hollywood fantasy, and the intelligence agencies of the world sat up and took notice. Quantum computing was finally graduating from an eccentric's hobby to

a legitimate technology with stupendous potential power. The catch was, of course, that qualifier on the estimate that it would take a computer with 4,000 qubits. In 1994 there were no systems with that description, and there was no well-defined path to get there. All this work, up to the time of Shor's publication of the factoring algorithm, had been theoretical. No operating qubit had been created, and no one knew if it was even possible to build an operational quantum computer. Attention shifted from the theory to potential realization. Given the sensitive information that Shor's algorithm seemed to threaten to unlock, the field found itself suddenly very well funded, and a first quantum bubble began, fueled by public sources of funding.

Personal communication with Peter Shor.

Chapter 8. Experimentalists Just Wanna Have Fun

> *No one doubts that building a useful quantum computer (one capable of addressing challenging computational problems) is difficult, but is it actually impossible? Do fundamental physical principles pose a truly serious obstacle?*
>
> —John Preskill

The foundations of quantum computing were built by theorists' equations, from Wiesner, Deutsch, Shor, and the rest. Now that it had been established that universal quantum computing was theoretically possible, and that it could do certain things *much* faster than classical computers, attention began to shift to the experimentalists working in their labs with lasers, liquid helium, and other hazardous materials. The fact that Shor had attracted the attention of certain government agencies, who now wanted to know who could build a machine to run the algorithm and how much it would cost, was only to their advantage. The public funding sluices opened up, and anyone with a reasonable idea about how one might build a qubit

found themselves flush with budget to pay students and postdocs and buy all kinds of equipment and components.

The theorists had begun working toward realizing physical quantum bits by describing them in the abstract, with quantum mechanical terms. All the experimentalists needed to do was to identify quantum systems that could fulfill those characteristics, and away they'd go. Specifically, the system had to have two or more distinct states measurable in attributes such as spin, polarization, or energy level that could be reliably induced, maintained, manipulated, and read out. These two-level systems had a nice new portmanteau, "qubit," for quantum bit, coined by Ben Schumacher in 1993 in a paper that eventually was published in 1996. Since the field of study had a slick new name and plenty of funding, inventing qubits ought to be easy.

In Wiesner's paper, he'd used the polarization of photons as the attribute used to store and represent information, and in 1989, Charles Bennett and John Smolin at IBM Research had created a working quantum communication demonstration using polarized photons supplied by lasers, as seen in Figure 8-1. Peter Shor had in fact first encountered the field via a talk that Bennett gave at Bell Labs about that experiment. Photons were attractive because polarization is easy to measure and easy to

maintain over a long period of time. They are stable at room temperature, so far as quantum systems go, and not sensitive to disturbances from their surrounding environment, or noise. However, the reason photons are not susceptible to noise is that they are very resistant to interactions in general and are therefore difficult to manipulate. Operations critical to performing quantum computation, like entanglement, can only be accomplished probabilistically. In other words, the experimentalist can set up the conditions that may induce entanglement between photons, but the particles may or may not cooperate. This made progress with photonic systems slow after the Bennett-Smolin experiment.

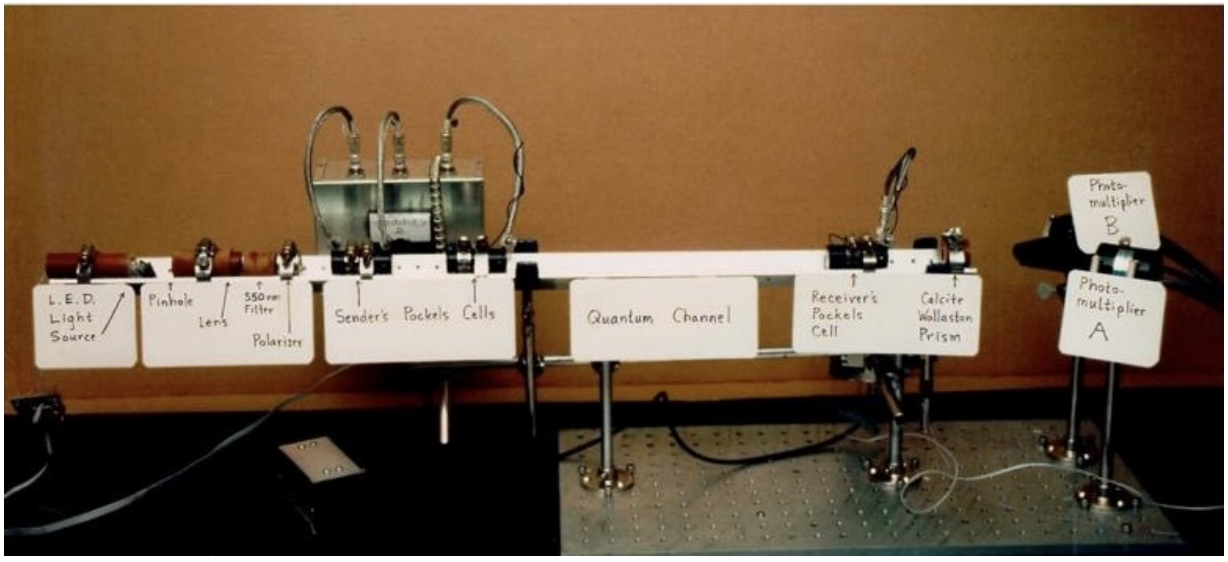


Figure 8-1. Bennett and Smolin's experimental implementation of BB84 protocol for quantum communication (Charles H. Bennett, "Physics of Computing: Quantum Computing and the Thermodynamics of Computing", Quantware, January 31, 2006)

After Endicott, experimental physicists had started evaluating the various themes of their research in collaboration with theorists, looking for potential qubit candidates that would be easier to make work than photons. Early proposals to create novel systems tended to be more hypothetical than practical using techniques that didn't exist yet combined in fairly far-fetched schemes, which, if you squinted your eyes and tried very hard to believe, *might* function as computers. The first example of this occurred in 1988, when two Japanese researchers, Yoshihisa Yamamoto and K. Igaeta, proposed using atoms for two-qubit operations (e.g., entanglement) and photons to transmit the qubit states around the system. On paper, the scheme suggested it would support a universal set of quantum operations including the CNOT gate that produces entanglement, but implementation details were thin.

A set of formal criteria for suitable qubit implementations began to crystallize thanks to the efforts of another IBM researcher, David DiVincenzo. Informally circulating in the community trying to conjure up experimental ideas that could be useful, he ended up publishing them in 2000, by which point they were as follows:

1. The ability to scale physically to many qubits, with the qubits uniform and with predictable behavior

2. The ability to initialize the state of the qubits to a well-defined starting point, such as $|000\ldots\rangle$
3. A set of quantum operations, or gates, that could support universal logical computation
4. The ability to remain coherent for much longer than it took to apply a gate—in other words, low noise and high stability
5. The ability to measure the state of the qubits accurately

The DiVincenzo criteria, as they came to be known, were emblematic of an important shift among the physics community toward thinking in terms of systems engineering. That is, DiVincenzo's criteria were about the characteristics of the device being designed rather than the typical focus on a standalone experiment. Physics labs tend to produce instruments that, with months, years, sometimes even decades of preparation, can perform a set of operations that yield data for the scientists to analyze ex post facto, a process that may again take months or years to perform. The resiliency, reliability, manageability, and other aspects of the instrument's behavior as a system are typically not priorities.

For example, the first criterium addresses scalability, or the ability to make not just one or two qubits but many orders of magnitude more. Classical computers, after all, had only been made possible by the ability of silicon microprocessors to scale

into the hundreds of millions of transistors. Initial calculations of the number of qubits required to factor a number on the scale of those used in contemporary cryptography using Shor's algorithm estimated that it could take tens of millions of qubits. Any serious attempt to build qubits would need to address that requirement.

The third criterium that the qubits would need to implement a universal set of gates represents the overlapping of the theorists' world with that of the experimentalists. A Turing-complete quantum computer, capable of computing all that is computable, would require physical devices that could carry out all the linear algebra needed to implement the required logical operations. In other words, Deutsch and Shor may have been able to conjure the mathematics needed to reimplement a Turing machine using quantum mechanics, but only experimentalists could design and build machines that would bring that to reality.

The remaining three criteria—that the qubits must be able to be initialized into a well-defined initial state, to withstand noise and retain coherence longer than it takes to carry out operations, and to have their resulting data read out reliably—all speak to the stable operation of a hypothetical machine. Here was the central challenge faced by those seeking to build

these devices: engineering a stable system that exposed and made stable the behaviors of the subatomic realm was far more challenging than anything ever accomplished to date. The scale of quantum mechanics defies our sense of scale, with dimensions starting at billionths of a meter, or nanometer—in numerical terms, 0.000000001 meters, 10^{-9}—and descending from there. Classical microprocessors in the late ’90s were being fabricated with circuits at the scale of about 150 to 200 nanometers.

The seeming impossibility formed part of an irresistible draw for ambitious experimentalists, and the late ’90s saw a burst of creative productivity. A number of more feasible proposals were floated by researchers in Europe, the Americas, and Asia, ideas cross-pollinating around the world. Various other domains within the experimental community that had produced exotic lab instruments began to be reevaluated in the light of the requirements laid out by DiVincenzo. In fact, the concepts for qubit modalities that would become dominant in the 2010s can all trace their roots to the period between Shor’s algorithm in 1995 and the end of the decade. Qubit designs based on cutting-edge experimental physics devices such as ion traps, neutral atom arrays, superconducting circuits, and even more speculative ideas like quantum dots and electron spins were first defined in papers from this era, though truly

successful implementations would take another decade to be achieved.

At a high level, the explorations began to fall into two camps: those seeking to design systems that brought quantum phenomena up to a larger scale than was typical in nature; and those hoping to extend our engineered control down to the "native" quantum scale. One lab experiment seeking to manipulate systems at a quantum scale that was relatively mature by the 1990s was nuclear magnetic resonance (NMR). The "nuclear" in NMR refers to the nucleus of an atom, which, when bathed in a strong magnetic field and then pulsed with a weaker radio frequency signal, produces an electromagnetic pulse of its own, characteristic to its own composition. In other words, electromagnetic signals can be used to coax an atom into revealing the subatomic, or quantum, state of its nucleus. This effect was first observed by Isidor Rabi in his lab at Columbia University, where his experimentation established an enormous amount of knowledge about the atom.

Rabi was an incredibly gifted scientist. Born in the Austro-Hungarian empire in 1898 and raised as an immigrant in the Lower East Side of New York City, he was an experimentalist par excellence. At Columbia University, where he took a position in 1929, he set about extending the result of a prior

experiment carried out by Otto Stern and Walther Gerlach in Frankfurt in 1922. The Stern-Gerlach experiment stands as one of the first great validations of quantum mechanical theory. The Stern-Gerlach setup had been an adaptation and extension of Young's double-slit experiment from 1801, but instead of shining a light at gaps in an opaque screen, they aimed a beam of silver atoms through a magnetic field that could be varied in strength and direction. The atoms split into discrete spots on the detection area, forming an interference pattern similar to the photons in Young's experiment. This established that the magnetic "moment" of the atoms existed in discrete quantities, and therefore their angular momentum as they travelled from the source to the destination was also "quantized."

Rabi's extension of Stern-Gerlach maintained the beam of atoms or molecules, achieved by heating a substance in the experiment's oven, and the variably tuned magnetic field, used to orient the resulting atoms' or molecules' magnetic moment in a specific way. After the initial magnetic field, the particle stream was subjected to an oscillating magnetic field that was perpendicular to the direction of the beam and whose frequency could also be varied. When the frequency of the oscillating field matched the natural frequency of the magnetic moments of the particles, resonance would occur, which would cause a change in the orientation of the particles' magnetic

moment. Finally, the beam hit a detector, which would measure the intensity of the beam and reveal the magnetic properties of the atoms.

The resulting technique was called molecular beam magnetic resonance, and with this experiment, Rabi had invented the field that would be known as NMR, earning a Nobel Prize for his work in 1944. Rabi's initial experimental results were further built upon by Felix Bloch and Edward Mills Purcell to use on a variety of liquid and solid material in 1946, work which earned them the Nobel Prize in 1952, and the basis for modern NMR devices was fully realized.

In the decades since its discovery, physicists had found many applications of NMR, using its ability to probe the inner state of materials for analytical chemistry and biochemistry experiments, a technique called spectroscopy, eventually leading to the commercialization of the technology as "nuclear magnetic resonance imaging" in the healthcare industry. The term "nuclear" evoked destructive weapons in the general public, which was avoided by shortening the name to magnetic resonance imaging (MRI). Industrial development of MRI devices in the 1970s and 1980s meant that mature, stable, and relatively inexpensive devices were available "off the shelf" for

physicists to experiment with, and so it was no surprise that it was considered as a platform for creating qubits.

As early as 1993, just such a proposal was made by the theorist Seth Lloyd of MIT, followed by an expanded proposal by David DiVincenzo in 1994. Missing crucial implementation details, the proposals languished for a few years, until 1997, when two separate experimental groups, one at MIT and the other at Stanford and UC Santa Barbara, published papers nearly simultaneously that provided the missing details, namely, how to implement gates using NMR devices.

Following the ideas laid out by the papers, a group at Oxford led by Jonathan A. Jones and Michele Mosca created two working NMR qubits and used quantum gates to run the Deutsch-Jozsa algorithm on them. The threshold between theory and experimental proof had been crossed, and history had been made. And as if to further underline the event for posterity, a group of researchers from UC Berkeley, Stanford, and MIT replicated their experiment, and so did Isaac L. Chuang, who had by then moved to IBM Research's Almaden Lab, as shown in Figure 8-2.



Figure 8-2. Isaac Chuang in the IBM Almaden lab working on computation using NME qubits

These NMR qubits were not physical devices or even singular entities. They were created out of collections of molecules, called ensembles, that were manipulated with the magnetic and radio frequency fields in the NMR machine. The way the researchers implemented qubits in NMR devices was not dissimilar to the way the technology is used to scan our bodies. The strong magnetic field causes the protons in the hydrogen atoms to align, then the radio signals pulse and cause the protons to become unaligned, and when they stop, the sensors detect the energy released by the protons reassuming their alignment. The different types of tissue in our bodies emit energy differently during realignment, which can be transformed into imagery. In NMR qubits, the basic NMR function could read out the spin of the molecules in bulk, and additional, custom tailored radio signals were used to modify their spin, implementing a universal set of logical gates. Different atoms within a molecule could be addressed

separately, realizing multiple qubits out of the components of a molecule. It took at least 10^8 molecules to generate a strong enough signal to detect the state of the individual qubits. Control of these ensembles was, right from the beginning, the main challenge facing the NMR qubits.

Still, the experimentalists forged on, improving their techniques of control and measurement, increasing the number of qubits from two to three, and moving from Deutsch-Jozsa to a more complex search algorithm named after a researcher named Lov Grover. In addition to the challenge of controlling such massive numbers of molecules in a liquid all in perfect resonance, the NMR work faced theoretical challenges as well. A researcher named Samuel L. Braunstein published work with his colleagues that proved that none of the experiments on NMR qubits to date had included any quantum entanglement. Without entanglement, it was very unlikely that a quantum computer made up of NMR qubits would ever out-perform a classical computer, since everything but quantum behaviors could be efficiently simulated.

Research continued, especially at IBM Research's Almaden Lab, where Isaac Chuang and colleagues continued to scale up their control systems, taking on ever more ambitious algorithmic execution. In 2000, Chuang and team ran the first part of Shor's

algorithm, and saw a path to be able to run the entire thing. They just needed seven qubits. This was a real test of the scalability of NMR qubits, because Chuang and his postdoc Lieven Vandersypen had to design and have synthesized a molecule that would be stable and have seven atoms suitable to act as qubits. The candidate molecule was hard enough to specify because the atoms had to be sufficiently separated in their energy states for the NMR signals to act on them each in isolation. The group did find a suitable molecule, and then work began on synthesizing it. A full year later, the sample was ready in a glass tube that could be fitted into the NMR machine. They prepared the state of the seven-qubit system, building on work done at Los Alamos National Laboratory the prior year. The group then managed to run a full implementation of Shor, factoring 15 to its primes, 3 and 5. Chuang moved back to MIT shortly thereafter and, sensing that further advances in scale would be nearly impossible to realize, abandoned NMR qubits for good.

It would be easy to say NMR was a dead end, the daguerreotype of qubits, abandoned as soon as more effective methods came along, but in fact the effort was quite valuable. First and foremost, successfully running Shor's algorithm was a moment that immediately earned a spot in the history of the field. Sure, it was only 15, but that record stood for a decade, finally being

defeated in 2012, when another team using photons factored… 21. More importantly it proved that it was possible, striking a blow against the skeptics and naysayers.

Even the criticisms leveled against the specific modality of qubits had value. What Braunstein's work pointed out was that it was not sufficient to make a computer out of a quantum system. Rather, that system needed to support truly quantum phenomena and use them for the purposes of computation. The tricks employed by Deutsch, Jozsa, Simon, Shor, and the rest of the earliest pioneers of quantum algorithms were leaning heavily on superposition and entanglement to achieve their speed-ups, and without them, all you were doing was building an extra-complicated, less capable classical system. The need to verify the use of quantum entanglement in its operations has become a standard part of the evaluation of new devices.

Perhaps most importantly, the work on achieving control of NMR qubits impacted the development of control technologies for other qubits. The equipment used was being built for industrial users in a healthcare setting, and as such, it had a robust embedded language to control the various parameters of the device. The quantum researchers leveraged that control language and built a sophisticated suite of software tools for initializing qubits, applying logical operations and measuring

qubits states once calculations were complete. And these tools were shared and maintained by the various groups using similar hardware for their experiments. This was not the usual scenario for experimental physics, where if software was written, it was typically hacked together for one researcher's very specific purpose, not documented, not usable by anyone else, and abandoned when the experiment was complete. This is, in fact, still very often the case in modern physics labs.

However, when the brief reign of NMR qubits ended, at least one notable researcher, Isaac Chuang, pivoted to an alternative qubit modality, and his postdocs and students started looking at what components of the NMR software toolkit could come along with them. That new modality being set up in Chuang's lab was qubits implemented using isolated charged atoms suspended in electromagnetic fields—known as trapped ions.

Chapter 9. Nature's Perfect Qubits

> *Atoms on a small scale behave like nothing on a large scale, for they satisfy the laws of quantum mechanics. So, as we go down and fiddle around with the atoms down there, we are working with different laws, and we can expect to do different things…We can use, not just circuits, but some system involving the quantized energy levels, or the interactions of quantized spins.*
>
> —Richard Feynman

The argument for using individual atoms for qubits is simple: every atom is identical and is perfectly consistent and predictable according to the rules of the physical world. This is quite unlike anything humans design and fabricate from materials with defects and minute variations. While NMR qubits used molecules as the basis for computation, others sought means to dive deeper into the quantum scale, isolating and manipulating individual atoms in systems that could support computation. In fact, the story of physics is in large part the efforts of experimentalists to extend our perception and awareness down to the subatomic scale, a challenge that is difficult to imagine for those not in the field. The methods

employed by these scientists to observe what is beyond the limits of our direct sensory observation are, by necessity, wildly creative.

Subatomic experimentation began in earnest in 1897, when Joseph John Thomson, a British physicist and Cavendish Professor of Physics at Cambridge University, conducted a number of experiments that established the existence of the electron. Thomson's experiments used a cathode-ray tube, a device that consisted of a glass chamber where a vacuum was formed and an electrical current was applied to each end. Applying a current, a beam would flow from the cathode, or negative electrode, to the anode, or positive electrode (see Figure 9-1). Devices of this type had been used by scientists for generations in their experiments.

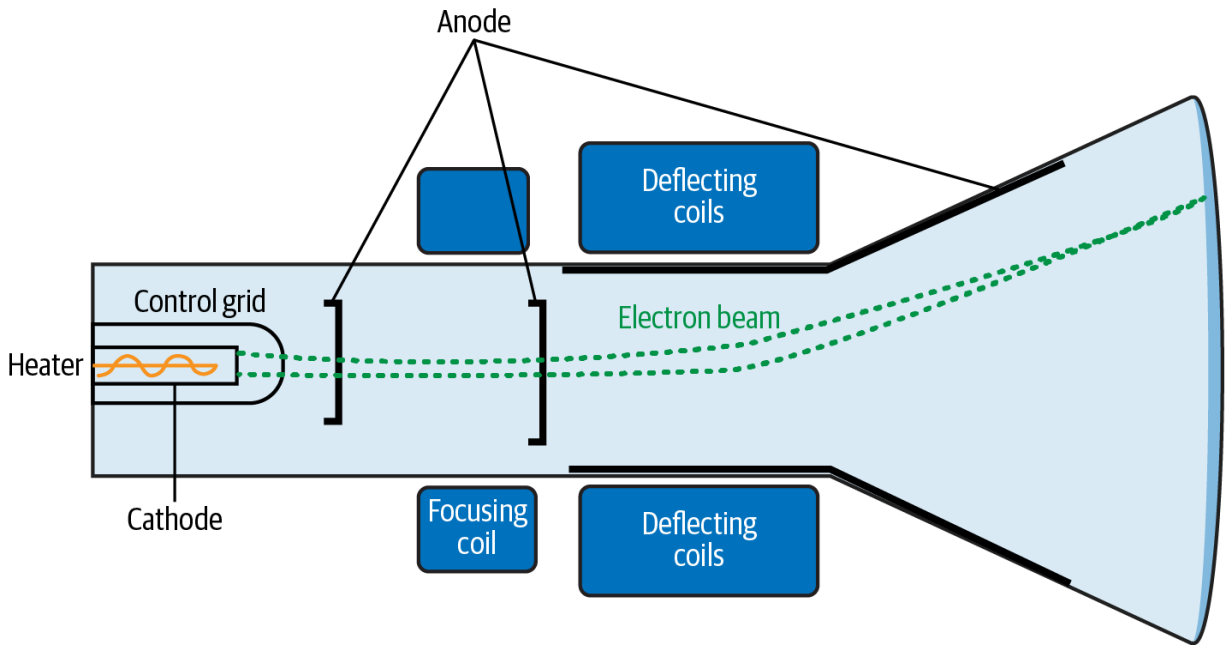


Figure 9-1. Cathode-ray tube similar to that used by J. J. Thomson to experimentally establish the existence of the electron

If this instrument looks familiar, it may be because you watched television between 1927, when Philo Farnsworth used a CRT designed in 1897 by Ferdinand Braun to create the first prototype, and 2015, when the last television using a CRT was built. Before television, CRTs were used in oscilloscope and radar displays and, as we saw in Chapter 4, were used as rudimentary digital memory devices by von Neumann and his team at Princeton.

Before Braun came along, the original tube design had been modified by William Crookes in 1890 with the capability to produce negative and positive electrical charges midway along the tube, and these charges would deflect the ray from its

normally straight path. Thomson was the first to measure the amount of the deflection compared to the amount of current. The amount of deflection from a straight path compared to the strength of the electrical charge provided a measure of the mass of the particles in the stream. Before Thomson's experiments, the atom was thought to be the smallest unit of matter and was believed to be indivisible and indestructible, a concept that dated back to ancient Greek philosophers like Democritus. In the early 19th century, the work of John Dalton and others formalized an atomic theory that described atoms as solid, indivisible spheres that made up the elements. This view was not without controversy, and questions remain to this day about Dalton's methods and whose work he may have used without attribution, but by Thomson's time, the model was generally accepted in the scientific community. According to this model, the smallest element, hydrogen, was indivisible and therefore the smallest mass in existence. However, the deflections produced by Thomson's experiments suggested that the ray was made up of particles thousands of times lighter than the smallest element.

The design of Thomson's experiment, and its implications for atomic theory, provides insight into the methods used to extend our understanding of the world beyond the reach of our senses. Unable to directly observe what was traveling through the tube,

Thomson found a way to manipulate where the ray hit the fluorescing material on one end of the beam by varying the charge affecting the beam's path. Having proved the beam was composed of something much smaller than any atom, Thomson named these particles "corpuscles," borrowed in part from Newton's corpuscular theory of light, and proposed that atoms were like a pudding with the negatively charged corpuscles embedded in them. His work was tied back to experiments by George Johnstone Stony, who in 1891 coined the term "electron" for what he believed was the fundamental unit of electricity. This was the first experimental evidence of a subatomic particle, significantly altering the understanding of atomic structure, which previously was thought to be indivisible.

This experiment of Thomson's was not only groundbreaking because of the implied existence of electrons as a smaller component of the atom, but also because it paved a path toward mass spectrometry and, eventually, ion traps. By the early 1900s, Thomson had built upon his cathode-ray experiments, modifying the instrument to bombard gases inside the vacuum tube with electrons and then discharge the electrons via the anode, leaving behind positively charged atoms of the gas. These positive ions were then attracted toward the cathode and formed the basis of Thomson's analysis. Turning again to the magnetic and electric fields from the former experiment,

Thomson could deflect these ions onto a detection screen. By analyzing the deflection patterns, he could infer properties like the charge-to-mass ratio of the ions.

Following the initial experiment, many advancements in technique and equipment were made by various experimentalists, becoming a common technique for identifying and quantifying molecules, based on their mass, referred to as “mass spectrometry.” By the 1950s, however, limitations in the precision of measurements produced by the best experimental setups inspired enhancements that would bring these ion beam devices closer to their eventual intersection with the field of quantum computing.

Wolfgang Paul, a German physicist not to be confused with Austrian physicist Wolfgang Pauli, had created a nonmagnetic mass filter in an attempt to achieve greater precision than was possible with the magnetic deflection methods of spectrometry to that time. The device was developed with his student, Helmut Steinwedel, in 1953, and consisted of four parallel metal rods, called a quadrupole. As seen in Figure 9-2, each diagonally opposed pair of rods is electrically connected, and an oscillating voltage in the radio frequency range, combined with a direct current voltage, is applied to the rods. One pair of rods has a positive voltage, or potential, while the other pair has a

negative potential. Ions are introduced into the quadrupole mass filter at one end and travel along the axis of the quadrupole. As they pass through the quadrupole, they experience the dynamic field produced by the rods, and only ions of a certain mass-to-charge ratio have a stable trajectory through the filter. Ions with a different mass-to-charge ratio experience unstable trajectories and collide with the rods, getting neutralized and removed. By adjusting the voltages dynamically, the mass-to-charge ratio of the ions that can pass through the filter can be precisely controlled. This approach to mass filtering was a significant advancement in the field of spectrometry, allowing for more precise, efficient, and flexible ion separation and analysis.

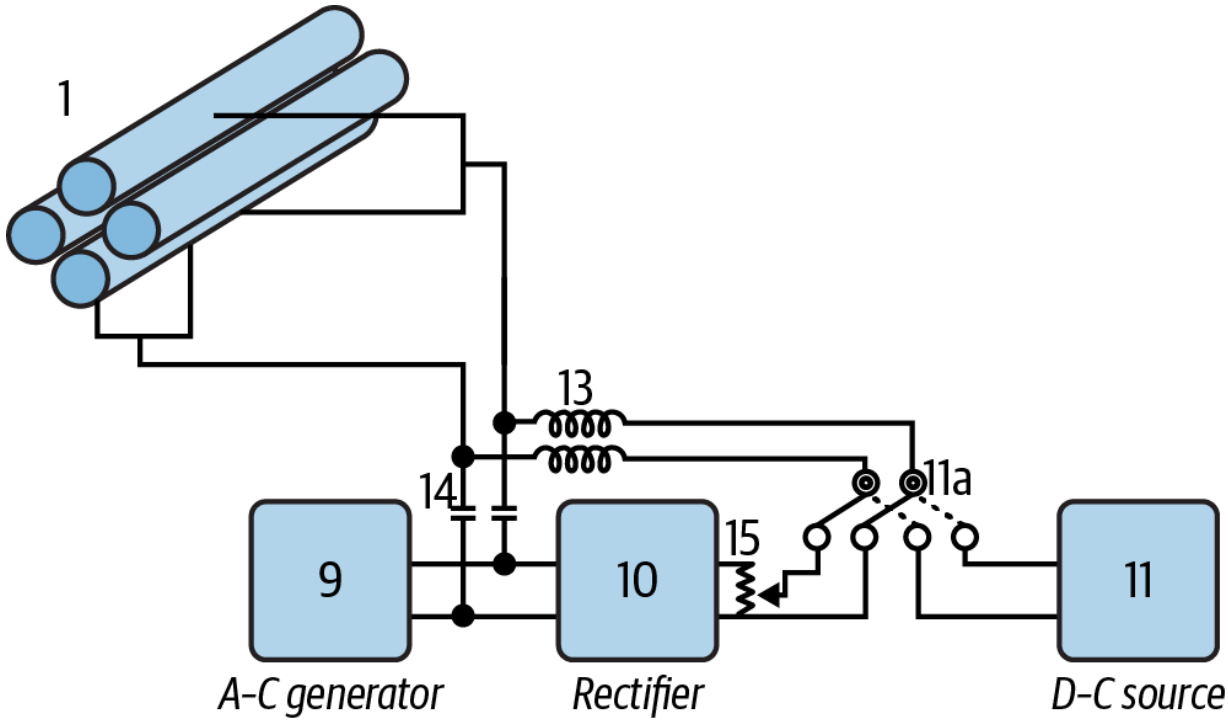


Figure 9-2. A diagram of the quadrupole mass filter (Paul et al. “Apparatus for Separating Charged Particles of Different Specific Charges,” patent number: 2939952, United States Patent and Trademark Office, June 1, 1960, Wikipedia)

However, Paul’s innovation did not end with the quadrupole mass filter. Continuing to explore the potential of the dynamic field, Paul realized the combination of radio frequencies with the static charge could be manipulated to confine ions in three-dimensional space. Ions would enter one end of the quadrupole and, after filtering, could be suspended in space, stable for long periods of time in a kind of trap. At the time, there was a growing interest in precision measurements in atomic physics, particularly for fundamental constants and atomic weights, and the ability of Paul’s trap to confine and study individual ions for extended periods was essential for such measurements. Paul

would go on to be awarded a Nobel Prize in 1989 for his contributions to the field.

At this point, the story of ion trapping intersects with Isidor Rabi. In addition to the work discussed in Chapter 8, measuring nuclear magnetic resonance, Rabi also played an integral role in the scientific discoveries upon which all modalities of qubits depend. After some years working with his molecular beam magnetic resonance (MBMR) setup, Rabi suggested that his device might be used as a clock. The idea of using quantum mechanics to provide timekeeping was not Rabi's originally, having been proposed by Scottish physicist James Clerk Maxwell in his *Treatise on Electricity and Magnetism*, writing in 1873, "a more universal unit of time might be found by taking the periodic time of vibration of the particular kind of light whose wave length is the unit of length." Like all great theorists, Maxwell needed a brilliant experimental physicist to come along before his idea could be realized. After Rabi's suggestion in 1945, a project was initiated at the National Bureau of Standards (NBS), led by Harold Lyons, to use MBMR for timekeeping. In 1949, Lyons's efforts created a clock based on the absorption of microwaves by ammonia molecules, the first of its kind. Called the NBS-1, the clock used the incredibly unwavering vibration of the molecules in the way Maxwell had

suggested, and was accurate to within 1 second every 3,000 years.

By 1955, advances in MBMR allowed the creation of clocks based on individual atoms at both the NBS and the National Physical Laboratory in the UK. These cesium clocks, named NBS-2 and Caesium I, respectively, achieved much greater accuracy than the ammonia-based NBS-1, gaining or losing 1 second only every 20,000 years. These devices represented quantum physics–based technologies with significant practical applications in everything from telecommunications to space exploration. Most importantly, it enabled a revolution in navigation: the Global Positioning System would not be possible without the accuracy of atomic clocks. In the physics community, atomic clocks brought about a new competitive drive: who could deliver another order of magnitude in precision?

One of the groups contributing to the drive toward greater precision was that of Norman Ramsey. Ramsey had been one of Rabi's collaborators on the MBMR experiments at Columbia before World War II, and had taken a position at Harvard after the war, where he continued to work with molecular beam experiments. In the early days of atomic timekeeping, the Ramsey group created what was known as the "separated

oscillatory fields method," in 1949. This addition of another oscillating field separate from the original field contributed substantial additional accuracy to MBMR measurements. By the mid-1960s, a PhD student named Dave Wineland was completing his degree under Ramsey's mentorship.

Under Ramsey's direction, Wineland worked on precise measurements of the magnetic interactions between the nucleus and the electrons in orbit around it. The measurements were done using masers, which Ramsey had just recently invented, a predecessor of lasers that emitted focused microwaves rather than visible light. This time with Ramsey would ignite in Wineland a complete fascination with high-resolution spectroscopy and precision that would carry through the rest of Wineland's career.

After earning his PhD in 1970, Wineland went to the University of Washington to join Hans Dehmelt's group, who were working on trapping ions in very similar ways to Wolfgang Paul, with whom Dehmelt would eventually share a Nobel Prize in 1989. Wineland was initially drawn to Dehmelt's work by the advantages provided for the hyperfine measurements Wineland had been carrying out with the maser at Harvard. However, once in Washington, his work instead focused on magnetic traps, known as Penning traps after Frans Michel

Penning, whose work at Philips Research in the 1930s inspired their design.

After five years with Dehmelt, he joined the Time and Frequency Division at National Bureau of Standards, where he was tasked with helping make the NBS-6 clock operational (see Figure 9-3). This was a next-generation cesium ion beam clock, with several advancements creating a new level of precision. The ions produced in the oven at one end traveled through the vacuum chamber, passing through a microwave cavity to the other end, where a detector counted the atoms in each of two energy states after interacting with the microwaves. Despite its increased accuracy, the ions in NBS-6 operated at room temperature, and therefore were limited to traveling close to the speed of sound, at an average of 300 meters per second, which made them susceptible to time dilation, per Einstein's special theory of relativity.

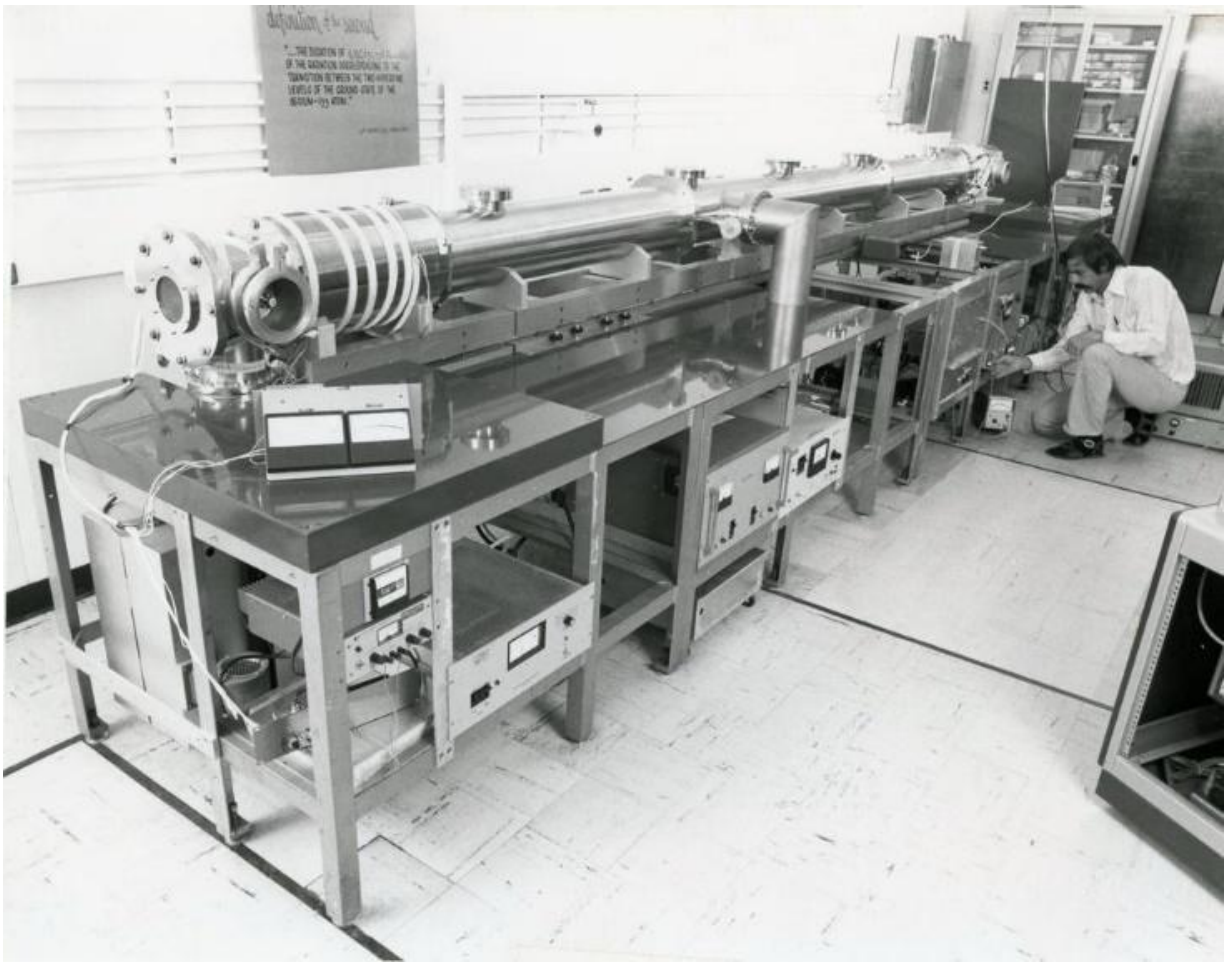


Figure 9-3. David Wineland with the NBS-6 atomic clock ("Context of a Discovery: Dave Wineland," NIST, August 31, 2021)

While in Washington, Wineland had written a paper with Dehmelt proposing a method to cool trapped ions using laser light. Atomic physicists had known for quite some time that higher levels of precision in their clocks would require cooling, but requiring a heat source that could produce a stream of charged atoms traveling at 300 meters per second did not lend itself to any normal method for cooling. With NBS-6 now online,

Wineland began to work on the problem of laser cooling, returning to the trapped ion regime from his prior work.

Despite NBS's lack of focus on basic research, Wineland's group leader, Helmut Hellwig, believed the laser cooling might be valuable for the future of their timekeeping devices, so Wineland and some colleagues began work in the fall of 1977, using magnesium ions in a Penning trap. As is often the case in the sciences, another, unrelated effort was working in parallel toward a similar goal. The group, led by Peter Toschek at the University of Heidelberg, had been joined by Dehmelt, who was on sabbatical from the University of Washington. Their work used a Paul trap with a barium ion, but both groups had the same goal: to use lasers to cool their ions to as low a temperature as possible.

The cooling effect relied on the fact that every atom has a very specific natural resonance frequency. A laser with a frequency that matches that of a particular atom will, when aimed at that atom, cause it to either emit or absorb a photon, changing its energy state. The absorption of a photon from the laser also imparts a change in the motion of the atom, increasing momentum in the opposite direction of the laser. Emitting a photon occurred in a random direction and so had no net effect on the atom's momentum. The subtle trick dreamt up by

Wineland and Dehmelt was to tune the laser just below the atom's natural frequency. In the same way that the Doppler effect makes sounds coming toward you shift up in frequency, when the atom moved toward the laser, the laser's frequency would shift up to the atom's resonant frequency, a photon would be absorbed, and the atom would be kicked in the opposite direction and emit a photon, diminishing the overall energy of the atom. Over the course of many such frequency shifts, the atom would be cooled.

In the fall of 1978, both groups had achieved the laser cooling effect on their trapped ions, with their respective papers reaching the scientific journal's offices within one day of each other. Laser cooling immediately became a major area of research, with some atomic physicists exploring its applications on ions and others using it on neutral atoms, with both Paul and Penning traps being used experimentally, based on the desired goals. At NBS, the Time and Frequency Division produced its first laser-cooled atomic clock in 1984. Using a beryllium ion in a Penning trap, Wineland and his colleagues achieved comparable precision to the best cesium beam clock, and they were just getting started. Thus, the story of Rabi's molecular beams and Paul's and Penning's traps finally intersected in Boulder in the 1980s, setting the stage for the experimental realization of qubits made of atoms.

One group using lasers on neutral atoms was led by Carl Wieman at University of Colorado Boulder. In Wieman's lab, gases made up of noncharged atoms were held in a vacuum chamber, where laser cooling was hoped to produce temperatures low enough to create Bose-Einstein condensates. This state of matter occurs when atoms are supercooled and condensed to their lowest energy state. At that point, their quantum state becomes indistinguishable from one another, providing a macroscopic stage on which one may study quantum mechanical effects. Predicted by the theoretical work of Satyendra Nath Bose and Albert Einstein in the early 20th century, Bose-Einstein condensates required conditions of near absolute zero temperatures that had not yet been achieved, but in the '80s, physicists believed they would soon have experimental results. One member of Wieman's group was a PhD student named Christopher Monroe, who, upon completing his degree, would join Dave Wineland's team.

The National Bureau of Standards had been renamed the National Institute of Standards and Technology (NIST) in 1988, and Wineland had founded the Ion Storage Group to continue the work begun by the laser-cooled ion trapping experiments of the late '70s. Once at NIST, Monroe began to help with efforts to entangle ions within the same trap. Each entangled ion would multiply the accuracy of the device when used as a clock by

augmenting the frequency of the system. Ten entangled ions would create ten times the “run rate” of the clock, adding yet another order of magnitude to the resulting precision of its timekeeping. By 1990 they had also demonstrated putting trapped ions into superposition of both their ground (or low-energy) and excited (or high-energy) states. This was accomplished using very carefully tuned and timed laser pulses after cooling the ion as much as possible.

What had begun with streams of atoms aimed at detectors in Rabi’s lab was now fully supplanted by the trapping model, where small collections of atoms could be held suspended in three-dimensional space. This incredible feat required careful preparation of the fields used to create the capture zone and not a small amount of luck to find the ion once it was in there. Monroe and the other staff members could sometimes hunt for weeks before finding the tiny flash of emitted light that signified the presence of the ion. Once located, however, the system was remarkably stable, with the ion’s state capable of being preserved for hours or even days at a time. It made for an incredibly powerful new platform for physics experiments.

In 1994, the International Conference on Atomic Physics was held at the University of Colorado Boulder, and a guest lecture was delivered by Artur Ekert, a physicist from the UK who had

done his PhD at Oxford under David Deutsch. Ekert's thesis had described a method for using Bell's inequality as a mechanism for quantum encryption, which he arrived at independently. In the pre-internet days, research papers were distributed very inefficiently, so Ekert only learned of prior quantum encryption work by Bennett and others once the similarities occurred to Deutsch, and he mentioned them to him. Ekert became a kind of quantum evangelist, giving talks, hoping to inspire and recruit new enthusiasts to the field, and especially hoping to rope experimentalists in so that they might build something that worked for theorists to use.

At NIST, Ekert gave a talk on the theory of quantum computing, explaining its potential power and also that no one knew how to build a quantum computer, or even if such a thing were possible. In the audience were two theorists from Europe, Peter Zoller and Ignacio Cirac. Zoller had taken a professorship at Boulder earlier that year, and Cirac, who had done part of his PhD at the University of Innsbruck with Zoller, was a visiting professor from University of Castilla-La Mancha in Spain. Inspired by the challenge, Zoller and Cirac immediately set about working on a scheme to create a physical qubit. Working in quantum optics at Boulder and at the Joint Institute for Laboratory Astrophysics (JILA), a facility where academic researchers collaborated with NIST scientists, it was natural for

them to think of ion traps and lasers. In fact, through JILA, they were familiar with Wineland and the work his group was doing.

In a few months, they had a draft paper, "Quantum Computation with Cold Trapped Ions", which they submitted to the journal *Physical Review Letters*, and shared with Wineland. The journal referees initially rejected their work, saying it was impossible to build such a device. Wineland and Monroe, however, were galvanized into action. Their work on improving the control of the ions for timekeeping devices was so well aligned with Cirac and Zoller's proposal that it seemed almost trivial for them to implement. In what may be the most efficient collaboration between theorists and experimentalists, Wineland and Monroe demonstrated a gate operation on an ion in a Paul trap just a few months after the paper was published, in 1995. Trapped ion qubits had arrived.

As much as it was a huge milestone, in practical terms, the initial result out of NIST was modest. While Zoller and Cirac had laid out proofs that trapped ions could support a universal set of gates capable of supporting quantum algorithms, the Wineland group's experiment had simply applied a two-qubit gate, a controlled NOT (CNOT), to a single ion. The internal hyperfine states of the atom represented one qubit, and

external motional states represented the other. However, it was a solid basis for continuing work, and by 1997, Wineland's group had demonstrated a CNOT applied to two ions in single file inside a trap. In multiple ion devices, the hyperfine states, or spin states, of the ions are used as the two-level systems of an individual qubit. Two qubit gates are applied by interacting with each atom's external motion, called motional state, with the two ions' motional states interacting as a single "phonon," or quantized vibrational state.

Other groups, inspired by both the theory paper from Cirac and Zoller and the results out of NIST, turned to building experimental qubits out of ion traps. Back at the University of Innsbruck, Cirac and Zoller found experimental partners in Rainer Blatt's group, while Hartmut Haffner led a group at the Max Planck Institute for Quantum Optics in Munich, Germany. Back in the US, a group at Caltech under the leadership of H. Jeff Kimble was also conducting experiments with trapped ions, though their work focused mostly on the extension of Feynman's quantum electrodynamics (QED) to the context of the microwave cavity in the trap. This work, known as cavity QED, was to play an important role in the development of other types of qubits, especially those made from superconducting circuits.

The theorists weren't done with their contributions either. While the CNOT gate operation proposed by Zoller and Cirac had been made to work at Innsbruck in 2003, it was very challenging to do reliably. It required the two ions to be held in their joint motional ground state during the operation—in other words, sharing the lowest possible energy of their collective motion—and also required a complex sequence of one- and two-qubit actions performed by the device's lasers. This required extreme precision with the device and made the entire enterprise quite sensitive to noise, thermal or otherwise. Any disturbance or imprecision would cause the qubit to undergo "decoherence," a transition from quantum to classical behavior, and any computation would be lost. The problems would be exacerbated in larger systems, where maintaining a large number of ions in their joint ground state would increase exponentially in difficulty.

Two physicists at Aarhus University in Denmark, Klaus Mølmer and Anders Sorensen, suggested another approach that was insensitive to the vibrational state, and published a paper in 1999 titled "Multi-Particle Entanglement of Hot Trapped Ions", describing their technique for operating on all ions in a one-dimensional trap at once. An Australian researcher, Gerard Milburn, suggested some enhancements, and by 2000, a solid theoretical model for what came to be called Mølmer-Sorensen

gates had been formed. Wineland's group performed an experimental validation that same year at NIST, operating on four trapped ions with reasonable fidelity and relative ease compared to Cirac and Zoller's method. Mølmer-Sorensen gates completely supplanted Cirac-Zoller gates, and they dominate the operations performed on ion trap quantum computers to the present day.

Very soon after, Isaac Chuang, having abandoned liquid state NMR qubits when he left IBM, was setting up his new lab at MIT with trapped ion qubits in 2001. Chuang had earned his PhD at Stanford with Yoshihisa Yamamoto as his advisor, and was a coauthor on the paper laying out the proposal for an atom- and photon-based quantum computer back in 1988. In that sense, the pivot to ion traps was almost inevitable, as both atoms and photons were utilized in this scheme as well, and had the advantage of having a physical manifestation. Chuang became involved in improving trap design itself, contributing to efforts to realize a planar ion trap, with the quadrupole and its electrodes laid out on a two-dimensional "chip" of sorts.

This technique, often called "surface electrode traps," was an important innovation, allowing the use of CMOS, or complementary metal oxide semiconductor, technologies for fabrication, the same as classical computer chips. Forward-

looking researchers, anticipating the engineering challenges of creating systems with enough qubits to do useful work, wanted to create a fabricated platform that not only would accommodate the trap but, eventually, the optics as well. They imagined this architecture paving the way to systems with hundreds, thousands, and even millions of trapped ion qubits.

After the disappointment of NMR's limitations, it would have made for an excellent *denouement* for trapped ions to have proven to be unambiguously the best technology for the job. However, that did not seem to be the case. Despite the surface electrode trap, advances in scale came slowly, and the size of the systems that could be built in the 2000s stalled out around five qubits. Enthusiasm began to wane for trapped ion qubits, and a notable paper that implemented all possible two-qubit gate operations, ostensibly to lay the foundation for a larger-scale device, instead was seen as proof that all the interesting experiments had been completed and there wasn't value in continued efforts.

This conservatism may have been a by-product of the world of atomic timekeeping, with its emphasis on precision. After all, the devices being built at NIST were providing the accurate, reliable timing on which critical systems around the world relied. While ion trappers were proud of the long coherence

times and high gate fidelities of the devices they were building, no one had any compelling strategy to get to systems with the scale and execution times required to outperform classical computers. The gate times were slow, for one thing, taking microseconds for each operation. At those speeds, a trapped ion computer's operations would be orders of magnitude slower than the gates on classical systems. The theoretical advantages of quantum algorithms had mostly to do with making intractable classical problems solvable, so if the system took an impractical amount of time to solve a problem, it offered no practical solution at all.

More daunting was a method to achieve the scale required for practical usefulness. The ions in a Paul trap line up single file in a single dimension. Penning traps could be two-dimensional but had significant hurdles to overcome before they could be used for qubits, and no one was sure if that would even prove to be possible. This one-dimensional trap had a hard upper limit of about 30 ions, with larger numbers of ions creating instability and degrading performance. The only way to get into the range of thousands of qubits seemed to be to couple traps one to another using photons. This was feasible but again faced significant headwinds and major challenges. Some were going to require advances in engineering, while others required novel science. Others simply were not yet identified.

By 2010, the funding agencies who had been enticed and terrified by the potential for Shor's algorithm to crack all codes were taking stock 15 years in. The qubit counts were stalling out under 10, several orders of magnitude fewer than would be needed for Shor's algorithm. Some influential voices in the researcher community were casting doubt on the whole enterprise. Even if it proved possible, the road ahead was going to be long. The clock community continued to improve trap-based devices to gain ever more orders of magnitude of precision. Experimental physicists continued to use the technique for the study of atomic physics and quantum mechanics. But the vision of using nature's perfect qubits in large-scale engineered computing devices seemed elusive. Perhaps another approach was needed, one that expanded the scale of the quantum world to something closer to our own. Instead of atoms emitted from an oven and caught in midair, these qubits were constructed out of superconducting circuits held at colder temperatures than occur anywhere in nature.

Chapter 10. Man-Made Atoms

> *Mercury has passed into a new state, which on account of its extraordinary electrical properties may be called the superconductive state.*
>
> —Heike Kamerlingh Onnes

The story of superconducting qubits begins with the phenomenon of superconductivity itself. When electrons travel through a material, they encounter resistance, a sort of friction against their forward movement that causes some energy to dissipate in the form of heat and other by-products. Resistance is why an electrical device will feel warm to the touch when plugged into a power supply. The phenomenon was discovered in 1911 by Dutch physicist Heike Kamerlingh Onnes, but finding an explanation for the behavior would take decades of effort. In the late 1950s, a trio of physicists at the University of Illinois Urbana-Champaign collaborated on research into the conditions under which resistance disappears and a material behaves as a superconductor. John Bardeen, Leon Cooper, and their grad student Robert Schrieffer were working with materials that, when cooled to very low temperatures using liquid helium, would cause electrons to be attracted to one

another and form pairs. These are now called Cooper pairs, and once they have formed, all the pairs of electrons move in correlation with one another, with absolutely no resistance. A current flowing through a superconducting circuit will go on forever with no energy dissipation, even after the source of the current is disconnected, so long as the conditions for superconductivity remain. This trio's understanding of supercooling was formalized in 1957 as the Bardeen-Cooper-Schrieffer (BCS) theory, for which the three men earned a Nobel Prize in 1972.

BCS naturally set off a superconducting craze in the physics community. There were fantastic ideas about levitating trains, lossless energy transmission, the hunt for materials that might be superconductors without supercooling, and the less fanciful desire to understand the further implications of these new insights into quantum mechanics. In 1962, a 22-year-old PhD student at the Mond Laboratory of Physics at Cambridge named Brian Josephson predicted that under certain conditions, Cooper pairs of electrons would form across an insulating gap in a circuit fabricated from superconducting material. This prediction was widely regarded as untrue and foolish. The young graduate student found himself debating John Bardeen himself, in a public forum. Experimental verification came, however, the very next year, thanks to two other British

physicists, Philip W. Anderson and John M. Rowell. The device they built consisted of a superconducting circuit punctuated by an insulator, as seen in Figure 10-1, which should have prevented electrons from traveling through the circuit but didn’t. The electrons in this device, now known as a Josephson junction, formed Cooper pairs on either side of the insulator, effectively tunneling through the nonconductive material. The theory and proof of electron tunneling earned Josephson vindication in the form of a shared Nobel Prize in 1973.

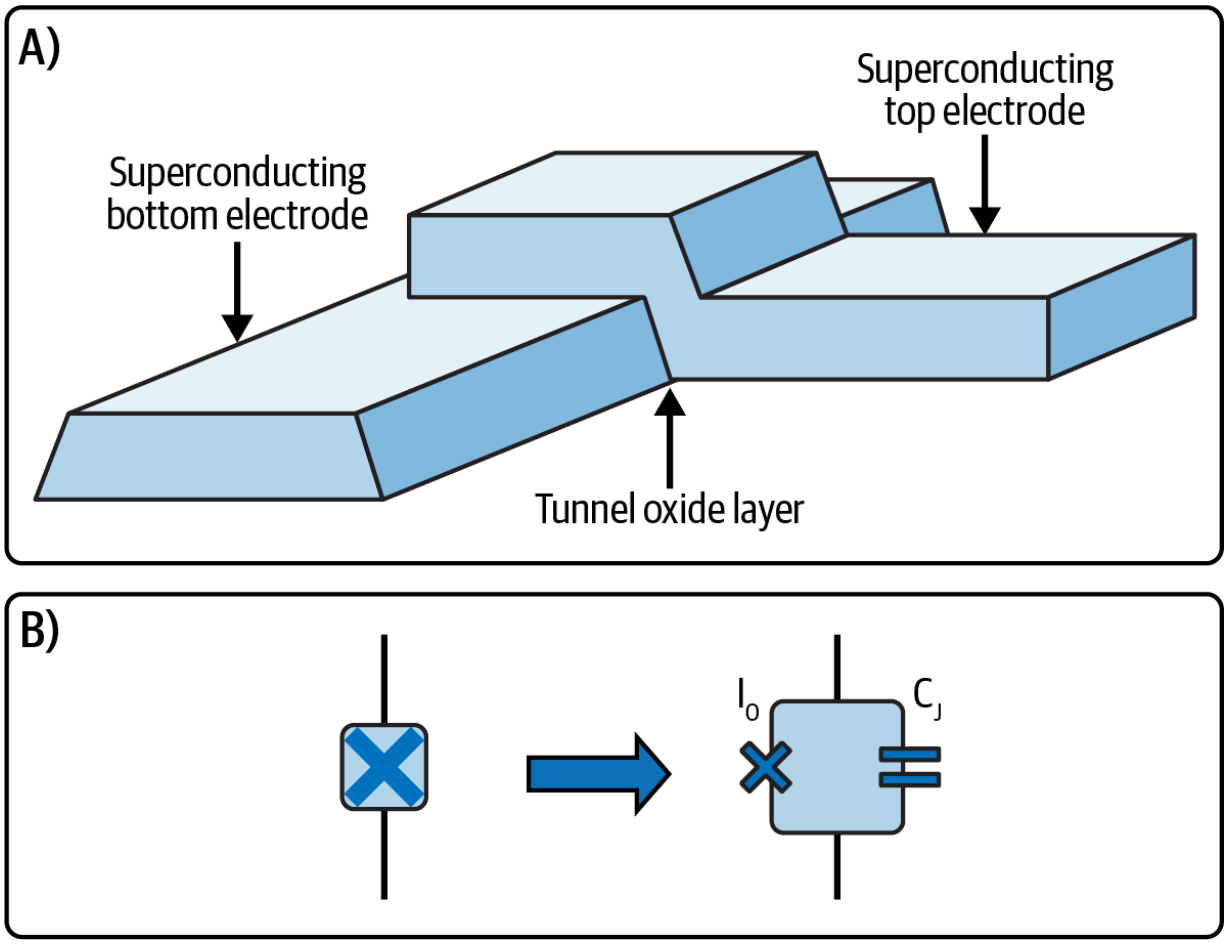


Figure 10-1. a) is a typical design for a Josephson junction made with two superconducting films separated by a nonconductive oxide layer; b) is the schematic representation of a Josephson junction

Though the Josephson junction was initially just an experiment designed to validate the hypothesis of electron tunneling, it inspired enthusiasm as a device unto itself. It was easy to fabricate—add an insulating barrier of something like aluminum oxide a few nanometers thick to a simple circuit made of material that was a superconductor at very low temperatures. Most importantly, though, it exhibited quantum coherence at a macroscopic level. The electrons in Cooper pairs

are held together by a force that counteracts the normal repelling charge one would expect of two like-charged particles. This attractive force extends further out from one electron than the charge force—so far, in fact, that it can bridge the insulating gap in the Josephson junction. Across the insulating gap, a Cooper pair maintains its connection, and current flows through the broken circuit. The wave function that describes that Cooper pair applies to *all* of the Cooper pairs in the superconducting circuit, and so the phase coherence, the result of the force exerted from one electron to the other across the insulating gap, is present in the entire circuit.

When studying the impossibly small scale of quantum mechanics, any opportunity to increase to merely incredibly small is very welcome, and so a wave of experimentation with Josephson junctions followed. In addition to adding a powerful new tool to the experimentalist's lab, there was also interest in possible practical applications of the scientific breakthrough. At Ford Research Labs, researchers exploited the electromagnetic characteristics of the Josephson junction to build what became called a SQUID, for superconducting quantum interference device. The electron tunneling is extremely sensitive to the magnetic flux created by the current in the superconducting circuit, and so, properly instrumented, a junction can become a

very sensitive magnetometer, able to detect minute changes in magnetic fields.

Two academic researchers were particularly interested in SQUIDs: John Clarke, who earned his PhD at Cambridge before joining the physics department at UC Berkeley in 1969; and Tony Leggett, who studied at Oxford before spending the late ’60s at UIUC with Bardeen and others exploring superconductivity and its effects. Clarke was drawn to the Josephson junction as a tool, seeming to have an intuition that it was more than just an experimental demonstration of theory and might be put to other uses.

Leggett, in addition to the substantial contributions to the theory of superconductivity and superfluidity, for which he would earn a Nobel Prize in 2003, began encouraging the physics community to explore the possibility of macroscopic quantum mechanical phenomena in SQUIDs and other Josephson tunneling-based devices. While at Sussex, Leggett’s initial work had focused on nuclear magnetic resonance (NMR), but by the early ’80s, Leggett had taken the John D. and Catherine T. MacArthur Chair at UIUC and continued work on SQUIDs, studying “the quantum superposition of states that are by some reasonable criterion macroscopically distinct.” Leggett’s ideas were tantalizing to physicists because what they

implied was that it might be possible to engineer artificial atoms that would provide the opportunity to study behaviors normally hidden from view at the subatomic scale.

While neither Clarke's nor Leggett's work directly referenced quantum computing at this stage, both had indirect, outsized influence on the development of the field through their work on Josephson junctions and SQUIDs. An undergraduate at the University of California Berkeley named John Martinis took a condensed matter course from Clarke in his fourth year and was captivated by the challenge of understanding the noise characteristics of Josephson devices. He stuck around and ended up with Clarke as his PhD advisor. Martinis had a background tinkering with electronics that he brought to his study of physics, and was attracted by the fact that the frequencies that were used to interact with SQUIDs were in the microwave range, a well-understood domain. This observation would play a critical role in the future of the technology for computation.

At conferences, Martinis encountered Leggett's study of macroscopic quantum phenomena and his interest in Josephson junctions deepened. During his PhD he was joined at UC Berkeley by another SQUID enthusiast, Michel Devoret, doing a postdoc with Clarke after earning his PhD from the

Commissariat à l'Energie Atomique (CEA) in Saclay, France. At CEA, his advisor, French physicist Anatole Abragam had deep experience with nuclear magnetic resonance devices, which equipped Devoret exceptionally well for the experimental work in Clarke's group, especially due to his experience with dilution refrigerators. Martinis and Devoret seemed to complement each other well, and accelerated the progress in SQUID experimentation enormously. Martinis hit upon the idea of adding a second resonant Josephson junction for readout, achieving some degree of isolation from noise in the device. Most critically, Devoret and Martinis supplanted the qualitative analysis typical of the work done to that point with more rigorous, quantitative data collection and analysis.

In 1985, Devoret, Martinis, and Clarke published "Measurements of Macroscopic Quantum Tunneling out of the Zero-Voltage State of a Current-Biased Josephson Junction," which presented analysis of the microwave signals collected from a Josephson junction that validated the quantum mechanical behavior present in the device. Since the researchers lacked the theoretical foundation for quantum computing, there was no way at this stage that they could recognize the suitability of the system to support computation. Martinis recalls a conference held around the time of the Clarke, Martinis, and Devoret paper, where, in addition to the

topic of macroscopic quantum tunneling in the superconducting experiments at UC Santa Barbara, the last speaker was Richard Feynman. Feynman spoke on the topic of quantum computing, which strongly implied that some conference organizers saw the possibility of an intersection of the two topics in the future.

Leggett's work at UIUC continued to explore macroscopic quantum tunneling and coherence, and to popularize the notion that Josephson junction-based systems could provide a window into the dynamics usually relegated to the subatomic scale. One researcher with NEC's Fundamental Research Group in Japan, Yasunobu Nakamura, who had been working with single-electron transistors, read Leggett's work and was intrigued. Shifting his work to Josephson junctions, he and his supervisor, Jaw-Shen Tsai, began building superconducting devices. Within a year, they gained access to a dilution refrigerator and began cooling their aluminum devices down to the millikelvin range.

Nakamura, Tsai, and their collaborator Yuri Pushkin rapidly assimilated learnings from various groups who'd worked with Josephson junctions like Clarke, Martinis, and Devoret, as well as more recent enthusiasts such as Hans Mooij at Delft in the Netherlands and Daniel Esteve at CEA. Nakamura also drew on

work he'd encountered in single-electron transistors from a variety of Russian researchers, one of whom, Yuli Nazarov, had written a paper on quantum dots, an application of single-electron transistors, which contained formulas that described the dynamics in the device being built at NEC. Traveling to Europe for the first time in 1996, Nakamura met many of these colleagues, in particular, hitting it off with Devoret, who passed the word through the community that Nakamura was doing interesting work. This led to meetings with Delsing at Chalmers, Esteve at CEA, Mooij at Delft, and others. When Nakamura met with Nazarov, however, the Russian had trouble understanding where Nakamura had found his formulas, even when Nakamura explained they had come from his own published work.

After this fruitful trip, Nakamura returned to Japan and shortly thereafter demonstrated superposition of the two charge states in their device. One Josephson junction on the device was designed to facilitate Cooper pair tunneling, creating what is referred to as a "quasiparticle," or particle-like behavior out of the interaction between the two electrons in the pair. On the other side of the device, another junction with a high voltage bias was designed for electron tunneling. The junctions were linked such that, when a Cooper pair tunnels into the reservoir between them, two electrons can escape across the other

junction. Without the Cooper pair present, the electrons are blocked. Monitoring the current across the electron tunneling junction therefore serves as a very sensitive probe for the presence of the quasiparticle. Careful measurements showed that the two charge states, the Cooper pairs present and not present, could be put in superposition. This, finally, was the two-level system that would be required to create a qubit using a Josephson junction.

Published in 1997 in *Physical Review Letters*, Nakamura shared those results with his colleagues in Europe, and sent an email to Tony Leggett, asking if he agreed it was evidence for the macroscopic quantum coherence that he had proposed. Leggett responded that he didn't consider a single Cooper pair "macroscopic," as he had been talking about large numbers of electrons acting in a quantum mechanical way at a larger scale. The debate over what exactly "macroscopic quantum coherence" meant remained a sideshow to the practical progress on superconducting qubits for many years.

Strictly speaking, the proposal that Nakamura, Pushkin, and Tsai had published would not produce a qubit at this stage. In fact, the 1997 paper didn't mention quantum computing at all. The motivating force was simply curiosity: could they demonstrate superposition in their system? Others who they

had been sharing results with were thinking about quantum computing, however, and a variety of papers in 1997 and 1998 from researchers at CEA, Karlsruhe, and Delft all pointed the way to developing proper qubits from these superconducting devices.

In 1998, using an expensive high-frequency pulse generator borrowed from an optical communications group at NEC—which was capable of pulses of just 100 picoseconds, or trillionths of a second—and despite skepticism from their peers, Nakamura, Pushkin, and Tsai were able to demonstrate coherent control of a two-level system. Their device had a coherence time of only one or two nanoseconds, but it was capable of quantum coherence at a macroscopic scale: a qubit made out of an artificial atom. Twelve years after the Berkeley result, the superconducting qubit was born.

The paper marks the beginning of a period of extremely rapid progress and wildly creative innovation among experimentalists working toward superconducting qubits. A design component originating in Esteve's lab at CEA, an iteration of which formed the basis of the Nakamura, Pushkin, and Tsai device, was refined and earned the name "Cooper pair box" around the same time as the Nakamura paper was published. This box optimized the superconducting "island"

separated from a “reservoir” by a Josephson junction so that precise control of the charge state of the whole system became more easily achievable and formed the basis of much of the subsequent innovation.

By 2001, Nakamura, Pushkin, and Tsai had refined their device and published another paper measuring a Rabi oscillation—a periodic transition between two energy states in resonance with a driving electromagnetic field. Around the same time, Mooij and others at Delft created what they called a flux qubit, where the two levels were represented by the electromagnetic state of the system rather than the charge, which presented a different set of strengths and weaknesses. Esteve’s group at CEA came up with something they called a “quantronium” qubit with an enlarged superconducting island in an attempt to make the device less susceptible to charge noise. The neologism formed by tacking “-onium” on the end of a physics word followed a tradition in the field that began in the middle of the 20th century, when work on antimatter began referring to “positronium” to refer to the bound state of an electron and its antiparticle, a positron. With quantronium, Esteve resurrected the convention and started a trend in the superconducting community to name designs in a manner that sometimes sound better suited for comic books.

In 2002, Devoret left CEA and took a role at Yale University, where he began collaborating with Rob Schoelkopf and Steve Girvin. Yale rapidly became a major center of gravity for superconducting qubits, beginning with a major theoretical breakthrough just two years after Devoret arrived. Quantum electrodynamics (QED) was Richard Feynman's contribution to the Standard Model of particle physics, providing a complete model to describe the interactions between photons and charged particles like electrons. QED was the work for which Feynman was awarded a Nobel Prize in 1965. Working with trapped ions, Serge Haroche had updated the theory for that specific context, producing work in 1996 known as "cavity QED." Now, the Yale Quantum Institute team, including Alexandre Blais, Andreas Walfraff, and Ren-Shou Huang working under the supervision of Schoelkopf and Girvin, worked on a further update to QED, which they called circuit QED.

The paper, published in 2004, laid out a blueprint for understanding and improving on the superconducting qubit. The framework presented a one-dimensional "cavity" in the form of a superconducting circuit to stand in the place of an atom in cavity QED. The dynamics of this artificial atom were thoroughly explored as a means of realizing qubits, including addressing issues of noise and coherence and advocating a

focus on the strong coupling between the Cooper pairs in the circuit and the microwave photons of the control signals as a means to address noise. The authors also described in detail the design of these devices and the means for applying gates and reading out qubit state.

Following on from that theoretical baseline, the researchers at Yale and their collaborators went on to produce the two dominant species of superconducting qubits, creating a reference design and theory for charge qubits in 2007 and for flux qubits in 2009. The comic book naming trend continued, as the flux qubit earned the name “fluxonium,” and the charge qubit was dubbed the “transmon.” Transmon was supposedly a portmanteau for “transmission line shunted plasma oscillation qubits,” but that seems like the first example of quantum computing retcon. Whatever its true origin, the naming of the transmon contributed a new suffix to naming trends, and has spawned many further variations such as unimon, coaxmon, Xmon, and others.

The transmon qubit is designed to reduce sensitivity to charge noise, a common problem in earlier qubit designs like the Cooper pair box. The key innovation in the transmon is its use of a large shunt capacitor, which increases the qubit’s coherence time while maintaining the nonlinear characteristics

essential for qubit operation. This stabilizing feature improved the overall performance of the transmon, making them competitive with flux qubits. Flux qubits, due to their use of electromagnetic flux as the means of storing and manipulating state, have reduced sensitivity to charge noise, giving them a natural advantage over charge qubits like the transmon, in addition to their faster operations thanks to higher energy gaps. However, the transmon's simpler design provided an easier path to overall improvements in system performance.

Experimentalists continued developing innovations such as three-dimensional resonator cavities that led to improved coherence times, all electronic control of qubits, and innovation around error correction, but the major contribution would remain the transmon and the fluxonium qubits. Transmons in particular simplified the design and fabrication of the devices sufficiently to make building qubits much more accessible, effectively leveling the research landscape and becoming the dominant design in labs around the world.

In parallel to the work on gate-based superconducting qubits, another way to use quantum tunneling had been developed in 1999 by two Japanese researchers, Kadowaki and Nishimori, called quantum annealing. Annealing is the process in metallurgy where controlled cooling is used to achieve desired

effects, for example, making sheet metal less prone to cracking while being shaped. In 1983, computer scientists used simulated annealing to solve the “traveling salesman problem,” and by the late ’90s, it was a standard heuristic tool used for similar applications. The idea of the quantum researchers was to use the Josephson tunneling to arrive at a solution more efficiently than classical methods. The advantage of quantum annealing is that gates are not needed, which lowers the level of engineering sophistication required of the overall system architecture. However, because annealing is heuristic and not algorithmic, meaning that at its core it is based on guesses that are not verifiable, there is no guarantee of arriving at the best final answer. However, the idea was attractive enough to launch the world’s first quantum computing startup, D-Wave Systems, founded in Vancouver, Canada, within the same year as the paper.

As work on gate-based superconducting devices progressed, so too did the work in D-Wave’s R&D labs, and by 2011 the company made the first quantum system commercially available, the D-Wave One, with 128 superconducting annealing qubits. John Martinis, having spent time at NIST and now at University of California Santa Barbara (UCSB), met the D-Wave team at a conference in Vancouver just before the commercial launch, in 2010. Martinis was drawn in by the engineering feat

of shipping a quantum technological product, and along with other researchers including Matthias Troyer, he began working on characterizing the performance of the unique machine.

In 2013, D-Wave Two was launched, equipped with 512 annealing qubits, and a system was purchased by the NASA AMES lab, which had entered into a collaboration with Google to explore the potential applications of quantum annealing. The same year, Martinis, Troyer, and others published two papers showing that the initial apparent performance gains achieved on the D-Wave system could be matched by clever classical simulation code primarily written by Troyer. This is the first example of what has become known as "dequantization," an awkward term for challenges to claims of quantum advantage with clever approaches using classical computing, such as the effort by IBM in response to Google's supremacy experiment.

While doubts were beginning to be cast on the performance of the annealing qubits, progress on gate-based qubits was continuing. Martinis's group at UCSB had a system with five gate-based qubits and were planning an upgrade to nine qubits, with hopes they could implement an error correction scheme called "surface codes" on the chip. This technical roadmap could feasibly deliver on the potential promised by the theorists like Shor and Deutsch a decade earlier, and between the papers

and the working system, Martinis attracted the attention of the head of Google's quantum team, Hartmut Neven. After delivering a talk to the Google researchers, Neven asked Martinis if he and his group would consider bringing their project to Google, and Martinis agreed.

Meanwhile, the "Yale gang" in the northeast of the US was having its own impact on the superconducting qubit landscape. Chad Rigetti, Jay Gambetta, and Jerry Chow all emerged from the transmon efforts at Yale to find their own ways to IBM Research, where foundational work had steadily continued since the '80s. The company was now ready to make a serious investment in moving toward commercializing the nascent technology, and the transmon seemed the most suitable vehicle. Gambetta focused IBM's efforts, enlisting the expertise of the chip fabricators and the equipment at T. J. Watson and convincing the company to cease investment in other qubit modalities. Chow led the chip design team, and Gambetta assumed control of the company's efforts to incubate a new business based on quantum computing.

Another tech giant had decided to enter the fray along with Google and IBM. Microsoft Research, led by mathematician Michael Freedman, had begun investing in a superconducting qubit effort in 2012, hiring Matthias Troyer as their lead

theorist. Instead of duplicating the approach used at UCSB or Yale, Microsoft's team had landed on a novel approach to the modality, announcing their intent to build something called topological qubits. Topological quantum states were pointed out to have significant advantages for fault tolerant computation by Alexei Kitaev, a famed researcher at Caltech. When a quantum system is in a topological state, its information is preserved through noise that would destroy the state of a nontopological system. Making a qubit that could induce and maintain a topological state as part of the operation of its two-level system seemed like a shortcut to a fault-tolerant quantum computer, and in that context, Microsoft's technical direction made sense.

However, to say that topological qubits were elusive would be a massive understatement. Theoretically, a variation on the Josephson junction architecture should be able to conjure up a very exotic quasiparticle called the Majorana Fermion, but no one had to that point been able to achieve that result in a lab. Undaunted, Microsoft went ahead, hiring Leo Kouwenhoven, a Dutch researcher at Delft who had published a promising glimmer of something resembling a Majorana in 2012, and Charlie Marcus, an American working at the Niels Bohr Institute in Copenhagen and widely seen as the best fabricator of nanoscale materials. The hunt for a Majorana was on, and in Marcus and Kouwenhoven's labs as the researchers fabbed one

test device after another, the quasiparticle proved as elusive as the physicist who lent it his name, Ettore Majorana, who had disappeared without a trace in 1938.

IBM's choice to focus on transmons quickly began to pay off, catching them up to others' superconducting efforts with its own five-qubit system, which instilled the IBM Research team with confidence in the roadmap they were working toward. The simplicity of IBM's design made transmons easier to fabricate, and helped IBM Research bring to bear the decades of nonsuperconducting chip fabrication know-how. IBM Research believed it could iterate on its design faster than its competitors, increasing quality and scale at a pace that could sustain progress toward commercialization. In May of 2016, IBM Research launched the IBM Quantum Experience, a website that provided access to a five-qubit machine.

Here for the very first time, anyone with an internet connection could use a visual interface, the Composer (seen in Figure 10-2), to drag and drop quantum gate operations on representations of the physical qubits in the device. The sequence of gates, strung along "wires" that represented the evolution of each qubit through time and resembled a musical staff, was referred to as a circuit. A quantum circuit could be constructed to represent a simple algorithm and then be submitted via API call

to run on the real device. It was intoxicating to imagine the symbolic representation of a simple algorithm being translated into a series of microwave pulses, then transmitted down wires leading into the dilution refrigerator at T. J. Watson, down past four Kelvin into the millikelvin zone, the “coldest place in the universe,” as experimentalists like to point out, to ultimately poke and prod at the artificial atom coursing through the superconducting circuit, returning the signal that measured its quantum state.

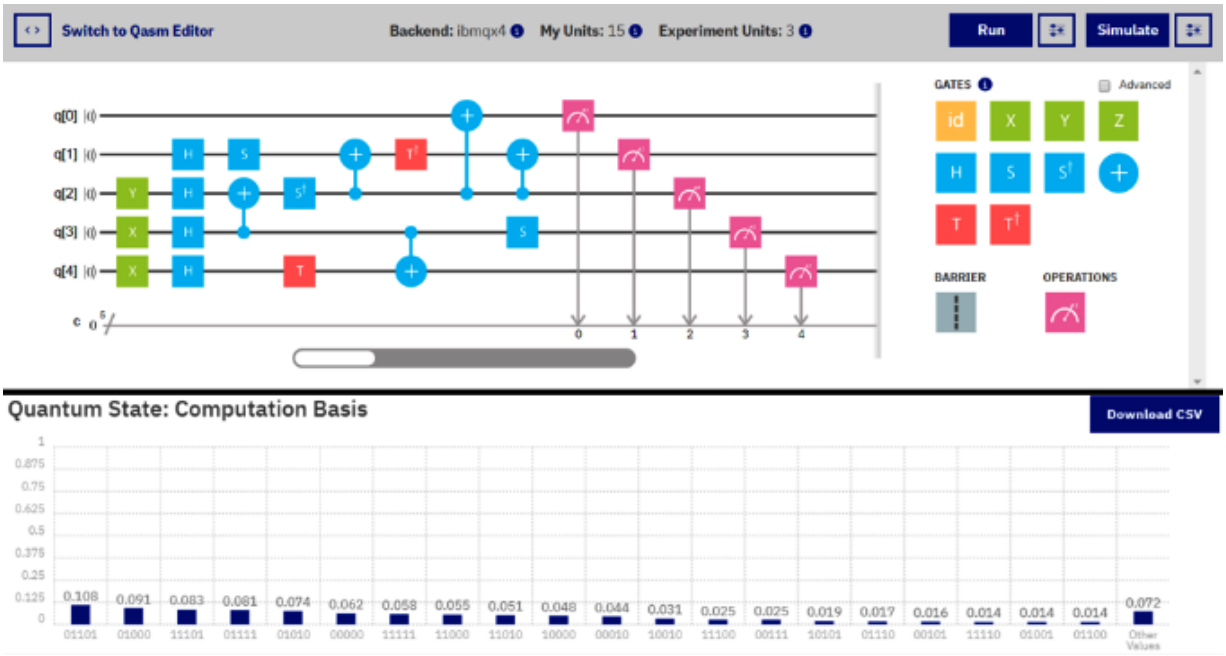


Figure 10-2. IBM’s Composer interface for constructing quantum circuits

It was also unlike anything any software developer working in 2016 had ever seen, more like the hand-crafted programs run on von Neumann’s computer in the ’40s than even the most

low-level programming of the modern era. If there ever was going to be a true quantum computing industry, it would need an entirely new kind of software developer possessing the requisite skills too. Encouraged by the positive reception of the Quantum Experience, the next year IBM launched a commercial offering called the IBM Q Network. For an annual fee, the largest IBM customers could gain access to larger experimental devices not available to the general public and, more importantly, access to the IBM research scientists who could hold their hands through their first exploration into the unfamiliar territory.

Around the same time as the Q Network launch, IBM released an open source software developer kit (SDK), called Qiskit. The decision to release the SDK as open source was made by Gambetta after he and his lead software architect hastily coded up the first version over a weekend, anticipating the decision would get bogged down in IBM's sometimes bureaucratic processes. With the SDK, developers everywhere in the world could write code in Python, a very familiar high-level language with tens of millions of users, running quantum circuits via the Quantum Experience APIs and getting the results right on their laptops. IBM Research had radically democratized access to these exotic devices, making the transmon the de facto standard

of what now seemed to be the next big information technology revolution.

Chapter 11. Qambrian Explosion

> *But to me, quantum computing is not an impossible dream; it is a possible dream. It is a dream that can be held without flouting the laws of physics as currently understood. It is a dream that can stimulate an enormously productive collaboration of experimenters and theorists seeking deep insights into the nature of decoherence. It is a dream that can be pursued by responsible scientists determined to explore, without prejudice, the potential of a fascinating and powerful new idea. It is a dream that could change the world. So let us dream.*
>
> —John Preskill

Making a real quantum computer available for anyone to try out on the internet in May of 2016 will forever cement IBM's place in all future histories of the field. However, if the company hoped it would enable it to exhibit the dominance of this new technology it had enjoyed during the mainframe era, it would rapidly prove not to be the case. For many, the event served as a starter's pistol, creating more awareness than anything that had happened before, and waking the private sector to the technology's potential. Quantum computing research had been

experiencing diminishing interest from the public sector funding sources like DARPA. Since the original jolt of Shor's algorithm in 1995, the realization was dawning on the intelligence community that "breaking encryption" was going to take thousands and thousands of high-quality qubits, while all current efforts were topping out under 10, and low-quality ones at that. Alternative approaches like D-Wave's annealers possessed more "qubits," but there was real debate regarding the technology's claim to quantum behavior or any possible advantage over classical computation, and the excitement about its launch was also dimming by the mid-2010s.

However, if the likes of IBM, Google, and Microsoft were investing in internal R&D efforts, why shouldn't the private sector venture-capital class take its shot? Perhaps sensing that Google and IBM's efforts in the field would make it easier to convince venture capitalists to invest, Chad Rigetti had left IBM to found the second quantum startup in existence, the eponymous "Rigetti," in 2013. In March of 2016, weeks before the IBM Quantum Experience launch, Rigetti raised $24M in an A round led by Andreesen Horowitz, with the intent to build superconducting qubits of its own.

Over in the ion trapping community, Chris Monroe was energized. He had left NIST in 2000, joining the University of

Michigan to develop a highly capable group working in ion trapping, aiming to build more expertise and momentum for the technology. From there, he had moved to University of Maryland and the Joint Quantum Institute (JQI), where University of Maryland collaborated with NIST in 2007. At the JQI, his group had demonstrated entanglement of two ions in separate traps using photons, a critical challenge for reaching commercial scale. He was ready to take the entrepreneurial plunge.

Monroe cofounded IonQ in 2015 with Jung Sang Kim with a $2 million seed investment from New Enterprise Associates, a VC firm from the valley. Kim was a researcher from Bell Labs with whom he'd been collaborating on the photonic interconnects. Kim had been working on optical switches for telecommunications networks at Bell Labs, a project that became less important as multiplexing on single fibers proved more efficient and effective. IonQ raised $20 million in 2016 in a round led by Google Ventures and NEA. Not only was IBM not without competitors in superconducting qubits, but now the modality itself had a rival for the race to commercially viable quantum computing business. IonQ wasn't even the only ion trap project for long. The pioneers of ion trapping at Innsbruck University, including Zoller, Cirac, and Rainer Blatt, founded a spin-off startup called Alpine Quantum Technologies, and

longtime US defense contractor Honeywell began its own ion trap qubit R&D effort as well.

John Martinis announcing the supremacy experiment on stage at a conference in 2019 was clearly Google's attempt to create the perception that the company had pulled away from the pack, but the reality was quite different. In addition to sharpening the rivalry with IBM, Google's move may have also played a role in announcements of cloud service offerings from Microsoft and Amazon Web Services in 2020. Both vendors planned to launch services that would provide access to multiple devices running different modalities of qubits. This seemed to be an explicit recognition that the market was not converging on a single vendor's qubits any time soon, and also that Microsoft's hunt for Majorana particles, vital to its topological qubit design, was still not producing any promising results. In the absence of its own hardware, Microsoft would provide their customers with access to vendors whose efforts had proved more successful.

Within the year, both vendors had followed through with their plans. Microsoft's Azure cloud division launched Azure Quantum with access to IonQ's trapped ion device and the superconducting system from Rigetti. Amazon Web Services' Braket launched with IonQ and Rigetti as well as D-Wave's

annealing system. Now customers wondering about quantum computing had a number of choices available to them: standalone websites that offered access to a single vendor's device like IBM's; or multidevice services in the cloud with unified interfaces like Amazon Braket and Azure Quantum. That both services focused on superconducting and trapped ion systems from IonQ and Rigetti reinforced the market perception that the race to commercial viability lay between those two modalities, and soon the superior technology from the best vendor would deliver value. However, by the early 2020s, no one's qubits, including superconducting or trapped ions, were capable of commercial, practical applications, only interesting experiments.

In fact, by 2023, it seemed as though perhaps all the investment and excitement in superconducting qubits may have been premature. Martinis had left Google in 2020, shortly after the supremacy experiment, and the company's chip R&D efforts seemed to have stalled, or at least had withdrawn from the public stage. Google's cloud access for developers shut down in 2022, and while its group of world-class researchers produced a steady flow of top-notch papers, there wasn't much talk about bringing a product to market. Rigetti had been joined by several other superconducting startups, such as Oxford Quantum Circuits, IQM, and SeeQC, but no system had demonstrated any

performance advantage over any classical computer for any task, and Rigetti's stock price had sunk nearly to the penny stock range. Rigetti himself departed the company in 2023.

IBM's resolve remained strong in 2023, at least in terms of product marketing. The company had updated its product roadmap a few times since 2020, with the 2023 plan promising a $200 million "hybrid quantum/classical supercomputer" with 100,000 qubits by 2030. IBM's public plans involved a data center–sized collection of dilution refrigerators with as many qubits in each as could be fit, and classical links allowing parallel quantum computing at scale. Questions remained about IBM's program, however, namely how its chip topology can deliver greater connectivity between qubits while maintaining control over the system's noise. Connectivity is crucial because if two qubits are not physically connected on the chip, one can only entangle them through a series of operations called SWAP gates, which literally swap the state from one qubit to another. If this game of telephone between two qubits goes on too long, the computation collapses into noise, the short coherence times on superconducting devices insufficient to complete any computation. Additionally, classical technology used to link quantum devices will impose significant computational overheads, which make it difficult to imagine realizing any true quantum advantage.

Meanwhile, Microsoft's topological qubit program was in shambles. In 2022, Leo Kouwenhoven had to retract a paper his group had written that claimed to find signatures of Majorana particles in a device; the group had mischaracterized the results. The mistake may have been unintentional but was likely a reflection of the mounting pressure on the researcher to deliver progress for his corporate source of funding. Charles Marcus also had a published result challenged, though in his case, the independent investigation found nothing amiss; it was merely a result of being in the "blast zone." Both men left Microsoft, and the company's topological qubit program showed no signs of recovering from their departure.

The uncertain fates of privately funded superconducting qubits over the decade since they were launched is a cautionary tale about the nature of the challenges still faced by the technology. Josephson junctions may have provided a macroscopic stage on which quantum effects could be observed, but controlling those phenomena is by no means a solved problem. At this stage, the best minds in the world can fabricate a superconducting chip with little insulating gaps at just the right spots, hook that chip up to microwave cabling in the bottom of a dilution refrigerator, cool the whole thing to eleven thousandths of a degree Kelvin above absolute zero, and then use room temperature electronics to talk to that chip—and every aspect

of that is a minor miracle. However, the road to fault-tolerant quantum computing at large scale is blocked by knowledge we simply do not have today. Without large numbers of qubits able to precisely execute long-running programs, we have very interesting lab instruments but not industrial-scale computers. Every so often, a claim will be made that the science of superconducting qubits is “solved,” and all that is left is incremental improvements in the systems engineering, but that is simply not true. The artificial atom may eventually have its day, or it may, like the NMR qubit, terminate its forward progress, a fascinating bit of science with no practical use, abandoned and eventually forgotten.

Things in the ion trap camp were not faring much better. IonQ had been joined by another commercial ion trap device vendor, Quantinuum, which spun out of Honeywell’s quantum computing R&D program in 2021. Both had very capable systems with about 30 qubits each, and both faced the same limitations of their approach—those of scale and speed. It was certainly true that trapped ions had much, much longer coherence times than superconducting qubits. However, gate operations were so time-consuming that the trapped ion’s effective “clock speed,” or rate at which it could carry out operations, was a couple of orders of magnitude slower than superconducting systems—the equivalent of a classical

computer running at a few kilohertz. Given that modern classical computers run in the billions or even trillions of cycles per second, that performance gap seems important.

More troubling is the question of scaling up ion trap–based quantum computers. One dimensional trap, where the ions are held in a single line, can only contain around 30 qubits before the thermal energy generated by gate operations overwhelms the capacity of the lasers to maintain the required temperature. A number of schemes for scaling up are being explored. Some traps consist of two one-dimensional traps in an “X” formation, and the qubits from one trap can interact with those from the other. The lasers used to cool and apply gate operations are also held up as a possible means to optically link traps. However, this entanglement operation is painfully slow, as the photons only interact with the ions probabilistically, failing as often as it succeeds. Alternative trap designs are being experimented with as well. Oxford Ionics has a Paul trap implemented on a chip that also has an embedded antenna, which allows the device to apply gates through microwaves, similar to superconducting devices, which promise faster operation and larger-scale chips. Universal Quantum plans to solve the problem of linking qubits across traps by having the individual ions actually *leap* from one trap to another. Others are returning to research Penning traps, hoping to build two-dimensional arrays of micro traps,

each ion in its own magnetic field—another theoretical solution to scaling.

Like their colleagues working with superconducting devices, the ion trappers have achieved incredible feats of engineering and creativity with their devices. Charged atoms are produced and deposited into a trap with automated efficiency and precision. Coherence times started off impressive and have continued to improve. Gate fidelities are the best of all existing qubit types. The devices are miraculous achievements, and ion trap–based clocks have achieved a stunning precision of 9.4×10^{-19}, meaning that if one could run for 33 billion years, twice the age of the known universe, it would still be within one second of accuracy. However, they are still not practical quantum computers capable of carrying out tasks faster than your average laptop circa 2023. Actually, most things a laptop can do, an ion trap device can't even begin to replicate. In 2023 Chris Monroe left IonQ, and Tony Uttley, who had led the effort at Honeywell and been Quantinuum's CEO at the spin-out, also left. IonQ's official statement about Monroe's departure was that the fundamental science was done, and Monroe wanted to get back to academia since the company only had engineering challenges ahead of it. A charitable interpretation would be that it is difficult for a world-class physicist to lead engineering

efforts, but perhaps it was that the challenges of commercial qubits at this point were more of the entrepreneurial variety.

The narrative of the two-horse race between superconducting Josephson junctions and trapped ions as qubit modalities was weakening as well. Both approaches and all the variations being explored of each are so far from being solved problems that the field is effectively wide open. A number of alternative qubit modalities have emerged, some with roots going as far back as ions and superconducting technologies, some with more recent origins. Several are currently being researched, and even though most are still very early stage, others are already approaching the maturity and capability of systems that are based on either atoms or superconducting circuits.

It is sometimes said that both trapped ion and superconducting qubits are using electrons for computation. In superconducting qubits, of course, the current in the circuit is made of electrons, which form the artificial atom that provides the basis for the qubit. Trapped ion qubits, while based on the entire atom, use lasers to excite the electrons of that atom to different states to represent the $|0\rangle$ and $|1\rangle$ states. Both these qubits' vulnerability to noise is in large part due to their reliance on electrons, because their state is very easily disrupted by heat, electromagnetic interference, and even cosmic rays.

There is a subatomic particle available in abundance, much less sensitive to noise from heat or other sources, that also can serve as a quantum two-level system suitable for computation: the photon. Theoretical evidence for the photon's existence was provided by no less than Albert Einstein himself, in a 1905 paper explaining how the photovoltaic effect works. Initially, Einstein used the term "quanta" to describe the discrete units in which light occurred, a usage that helped pave the way to the general adoption of both the phrase and the theory of quantum mechanics. An American physical chemist, Gilbert Lewis, popularized the term "photon" in 1926, adapted from the Ancient Greek *photos*, for light. Of course, photons had been inextricably linked to the developments leading up to quantum mechanics, all the way back to Young's double-slit experiment, which, though Young didn't know it at the time, used photons to demonstrate superposition.

The roots of quantum information science are similarly associated with photons. Stephen Wiesner's breakthrough thought experiment from the late '60s, published over a decade later, used polarized photons to represent the serial numbers of his money that would be impossible to counterfeit. Around the same time, in 1972, quantum entanglement was first demonstrated using photons. Charles Bennett and John Smolin's experiment demonstrating cryptographic key exchange was

inspired by Wiesner's work and used polarization in much the way he had suggested.

Despite these milestones and ongoing use as a workhorse in physics labs, where optical tables, lasers, and pieces of glass are habitually used in a broad set of experiments, photonic quantum computing did not develop at the same pace as trapped ion or superconducting qubits. Progress, however, has been made, and several companies are working on photonic qubits, grappling with their own set of hurdles to get to quantum advantage.

One notable example is PsiQuantum, founded in 2015 by Jeremy O'Brien, Terry Rudolph, Mark Thompson, and Peter Shadbolt, physicists working at several UK universities. O'Brien had experimentally implemented a CNOT gate on photons in 2003, and Rudolph contributed a unique approach to photonic-based computation the firm refers to as "fusion-based quantum computing." PsiQuantum had raised almost $800 million by 2023, but its progress is difficult to gauge as the company has famously chosen to remain in "stealth mode" until it can demonstrate fault-tolerant computing with a large number of qubits.

Other photonic companies have been less secretive, such as Xanadu Quantum Technologies, founded in Toronto in 2016, out of the Creative Destruction Labs' quantum track. Xanadu's approach initially used squeezed states of photons to perform Gaussian bosonic sampling, which promised better performance for a certain subset of problems. In 2022 the Xanadu team demonstrated a device it called Borealis to perform computations it claimed were 50 million times faster than the classical alternative. Like Google's supremacy experiment, the problem was contrived as a demonstration and had little practical value, and the Borealis device was not a universal computer. As a non-gate-based device, Borealis was something of a revival of approaches like quantum annealing. While analog quantum annealing devices like D-Wave's had not shown any practical advantage in the 15 years since founding, others were still considering an analog approach.

While the clock builders at NIST had been focused on charged atoms in their ion traps, there existed a larger segment of the atomic physics community who had been working with neutral atoms. Before arriving at NIST, Chris Monroe had been building neutral atom devices at the University of Colorado in Boulder before joining NIST, and it was at Boulder where some of the most important work on neutral atoms were performed. The motivator for these experimental devices was a state of matter

theorized by work by Satyendra Nath Bose and Albert Einstein in the 1920s. At extremely low temperatures, atoms would share the same quantum state, effectively demonstrating quantum behavior at macroscopic scale. In 1995, the Bose-Einstein condensate (BEC) was realized in the lab at Boulder by Eric Cornell and Carl Wieman, earning them a Nobel Prize in 2001.

While the pursuit of Bose-Einstein condensates and the Nobel Prize everyone assumed would go with it was on, neutral atoms attracted more physicists and funding than trapping ions. Once BECs had been created, that community began to think of other applications for their expertise, and as quantum computation became more imaginable, their attention naturally fell on that topic. Lacking a charge to manipulate via electromagnetic fields, neutral atoms must be held in an optical lattice created by lasers. Photons are also used to cool and then to excite the atoms, creating what's called a "Rydberg state" so that the outer electron orbitals can be made to interact with one another.

The Rydberg state of an atom is difficult to manipulate, as it is highly excited and therefore prone to noise. However, Mikhail Lukin, working at Harvard, was successful in building a stable array of 256 atoms in the summer of 2021. The atoms in this device are sometimes referred to as qubits, but the term is

misleading, because Lukin's device did not support gate operations. Instead, the atom's energy and position could be adjusted and tuned to model the time evolution of a specific Hamiltonian equation.

A Hamiltonian, named after Sir William Rowen Hamilton, who created it in 1833, is a crucial equation for understanding isolated dynamical systems. Loosely speaking, it is used to describe the total energy of a system, both kinetic and potential, and is used with Schrödinger's equation to determine how the quantum state of a system changes over time. Energy levels are used to predict the outcomes of interactions and transitions within the system, and so an efficient method for modeling a Hamiltonian is essential to the challenge of quantum simulation. In this way, what Lukin's group had built at Harvard was a realization of Feynman's original vision in 1981, an efficient means of simulating quantum systems, built using quantum mechanics.

Progress on the machine at Harvard had been promising enough for Lukin to found a company, QuEra, along with colleagues Vladan Vuletić, Markus Greiner, Dirk Englund, Nathan Gemelke, and John Pena in 2018. And in 2021, QuEra made its first commercial device, Aquila, based on the prototype in Lukin's lab, available on the cloud via Amazon's

Braket service. Other companies are also pursuing neutral atoms, such as Pasqal, in France, and Atom Computing and Infleqtion, both based in Colorado. Infleqtion was formerly known as Cold Quanta, and actually predates the others by more than a decade, founded in 2007, focusing initially on BEC before pivoting to computation more recently.

It turns out there are many ways, other than through photons or neutral atoms, to make qubits in addition to trapped ions or superconducting circuits. Companies like Quantum Brilliance are using defects in diamonds as qubits. The defects consist of a missing carbon atom in the diamond crystal lattice, replaced by a nitrogen atom, and are referred to as “nitrogen vacancy” (NV) qubits. EeroQ, a Chicago-based company, is trapping single electrons on the surface of superfluid liquid helium and using those as qubits. In many ways, however, the race to qubits can be reduced to a question of electrons versus photons. Even in devices that use entire atoms in radio frequency or optical traps, it is the electrons’ energy state or spin that is used to encode information.

Using electrons feels familiar, of course, since electrons flowing through circuits are the basis for classical computing. In some ways, the dream is to replicate the scale accomplished by the titans of classical, leveraging CMOS technologies to build

quantum computers of the future. The most direct attempts to use CMOS are called spin qubits, where electrons are trapped in tiny wells or dots on the surface of an otherwise very familiar semiconductor chip. Spin qubits were first proposed by David DiVincenzo and Daniel Loss in 1997, and in April 2020, QuTech at Delft University made a system with two-spin qubits available to the public. Spin qubits use the intrinsic one-half spin in the degrees of freedom of individual electrons, with the benefit of being fabricated using our familiar CMOS technologies. Like photonic qubits, they promise massive scalability and more practical deployment, if they can overcome certain critical challenges to implementation. Several startups have been founded to work on silicon-based spin qubits, including Silicon Quantum Computing in Australia, Equal1 in Ireland, and Quantum Motion in the UK. In addition to startups, Intel has made a serious investment in spin qubits, leveraging its unparalleled experience and resources. In 2023 the company announced Tunnel Falls, a 12-qubit silicon chip, available for researchers.

With CMOS-based spin qubits a possibility, one might ask why startups like QuEra, PsiQuantum, and IonQ and even giants like IBM are bothering to pursue their exotic qubits modalities. Like each of their approaches, spin qubits on silicon face multiple challenges that require fundamental scientific breakthroughs.

In basic research, there really is no way to gauge timelines or even feasibility—a scientific black swan with no way to measure the odds of success or failure. More exotic ideas about how to build qubits continue to crop up around the world, all of which face numerous unsolved scientific challenges before they can be usable. This evolutionary diversity only serves to emphasize that in the 2020s, 40 years past the nominal start of the race, there simply is no low-risk path to making devices large enough, stable enough, and performant enough to win over all the others—or to do anything useful to industry, for that matter.

What we're facing is a gaping hole in our knowledge and no proof that void even has a viable solution to fill it. In typical industrial research and development, the focus is more on engineering challenges that are incremental in nature—issues of miniaturization or finding the right material or making fabrication more accurate or cheaper. Each team working on qubits is making its own (highly) educated guess about the solvability of its modality's unique challenges. Ion trappers believe in the fundamental stability of their qubits and believe gate speed and scalability won't prove insurmountable. The superconducting community believes its fabrication skills will continue to improve and its system's speed is unmatched, and the community doesn't worry about the short coherence times,

the connectivity challenges, or its reliance on dilution refrigerators and Helium-3, which is an isotope of Helium that is only produced as a by-product of radioactive decay. The photonic proponents point out that photons make the most stable qubits and maintain that they can overcome the difficulties of interacting with them or losing them to absorption, and anyway, scaling any of those other types of qubits will require using photons as “flying qubits.”

Each variation on a modality of qubit, whether from a VC-backed startup or a corporate-backed R&D team, arcs out into this wash of unknown unknowns, a flare describing a path up into the darkness, its apex hopefully finding commercial usefulness. The truth is, no existing qubit modality has a clearly defined path to being stable enough at a large enough scale to be anything more than very elaborate academic experimental devices. With all the investment in a technology that has not yet proven itself to be valuable, it’s no wonder that some are reminded of the dot-com bubble and other instances where the hype has led the public astray.

Chapter 12. The Heroic Era of Qubits

> *This scheme, like all other schemes for quantum computation, relies on speculative technology, does not in its current form take into account all possible sources of noise, unreliability and manufacturing error, and probably will not work.*
>
> —Rolf Landauer

Researchers at the Shanghai Artificial Intelligence Laboratory announced in June 2024 that they had surpassed the results achieved in Google's quantum supremacy experiment using classical computers. The supremacy experiment is, of course, the one described in Chapter 1, the one held up as proof that quantum computers can provide superior performance, and a widely cited justification for continued investment in the field. In 2019, Google's Sycamore chip took 10 minutes to complete the task of generating randomized quantum circuits, and in 2024, the team from China used a cluster of Nvidia GPUs to complete the same task in 14.22 seconds. A far cry from the 10 thousand years that Google had said it would take a classical computer. Even more impressive, the Shanghai AI lab's solution used less power than Google's device, consuming just 0.29 kWh

of electricity, compared to the 2.39 kWh used by Sycamore. The claim of supremacy seems utterly extinguished. Could this be the mass extinction that culls the enormously diverse efforts to build quantum computers?

The quest to build a quantum computer that can create economic value has always been fraught with risks, and many skeptics have doubted that it is even possible from the very beginning. Scott Aaronson, in one of the classes he teaches at the University of Texas Austin, managed to collect 11 of the most pointed attacks on quantum computing, all of which he provides counterarguments for, with varying degrees of persuasiveness. He summarizes his opinion in much the same was as others such as Richard Jozsa or Tony Leggett have, writing, “it’s entirely conceivable that quantum computing is impossible for some fundamental reason…That would be much more interesting than if quantum computing were possible, because it changes our understanding of physics.”

From a scientist’s perspective, a failed experiment is more valuable than a success because it provides concrete proof that your hypothesis is wrong. A success could merely mean that you haven’t been clever enough to find the weakness in your incorrect hypothesis. This attitude is usually at odds with the process by which projects are evaluated by investors of all

types. This is why science is systematically underfunded in a capitalist society. Who wants to put a dollar into an experiment that the experimentalist is *hoping* will fail? After Shor provided proof that quantum computing could massively outperform classical in at least one task, the field was transformed. What had been a niche topic of interest to a very small community, numbering in the dozens globally, began to attract a lot more attention. For researchers, it started to look like a rich vein of the raw material for scientific glory. For the intelligence community, a worrisome new potential threat to evaluate. Funding flowed from the public sector, even if some agencies, like the NSA, did not have preexisting programs set up for the purpose. In a sense, both the funding agencies' and the scientists' goals were aligned: they were both hoping to uncover a fatal flaw in the idea. As we've seen, the slow progress from 1995 to the early 2000s convinced the intelligence community the threat was not imminent, and they moved on to other topics, constricting the flow of funds to the established teams. To continue the work, alternate sources of money needed to be found.

Tech startups are, at this point in the digital age, a very familiar phenomenon. The attributes of the earliest stages of building a business based around a new technology are part of the cultural fabric of the 21st century, fodder for popular

entertainment and memes. Some of the quantum computing community, possessing sufficient self-confidence, decided it was time to start pitching their projects for private sector investment. It was, in some sense, a simple pitch: give us money, and we'll make a computer that can do something impossible. There were challenges, duly disclosed, and quite a bit of uncertainty around the real timeline, but the potential returns were deemed large enough to justify the risks. The physicists who worked to secure funding didn't have special insight into how to blaze a trail from small-scale noisy devices to something that could create significant economic value. Instead, they were gripped by the vision of becoming the singular person to realize, against all odds, the dream of quantum computing. A number of venture funds and corporate innovation leaders found the vision convincingly enticing, and in the early 2000s the quantum computing industry was born. If there were errors committed in this phase of the development of the technology, they were not caused by willful omission or distortion of information about the risks. Instead, what seems to have happened was a kind of muddying of the waters. The startup-style pitches created the impression that quantum was somewhere just over the horizon, a disruptive technology, a challenging technology, but, in the end, an *incremental* innovation.

If this book has managed to convey one fact thus far, I hope it is this: quantum computing is not in the same category as any emerging technology in modern memory. It isn't, in fact, just a new category; it's an actual new paradigm of computing, honest to goodness, not like the other million times you heard that phrase used hyperbolically. We've seen that the ways in which you build a qubit are not the ways the pioneers of classical computing built bits. We've seen that the math that is the native language of qubits is not native to classical bits, and so the algorithms look *very* different. The nature of a paradigm shift is that it is discontinuous, and nothing that exists before it will accurately predict the world that exists after the shift occurs. This is in stark contrast to the incremental progress classical computers have enjoyed since their last true paradigm shift, the invention of the microprocessor. The engineers hand-soldering transistors to boards to build guidance computers for Apollo spacecraft could not have imagined chips infiltrating our kitchen appliances, children's toys, cars, and every other thing. Put another way, from a contemporaneous perspective, the launch of the iPhone was impressive, but many companies had tried their own versions of that device already, and we all had a sense of what the mobile internet would bring us, at least in broad strokes.

That cognitive framing is also deeply embedded in the funding models dominant in the private sector for funding new and emerging technology development. The typical venture fund has a horizon, past which the venture capitalists or VCs managing the fund need to start exiting their investments so that they can show returns for investors who are the sources of the capital in the fund. Corporate CFOs similarly have time scales on which they need to show returns on technology R&D spending in the form of new lines of business growing revenues. The time scales vary somewhat from shorter to longer, but very rarely do they exceed 10 years, which is comfortable when incrementally improving on existing technology but can be problematic for more fundamental scientific research. This has led to the decline of the great industrial R&D labs of the latter half of the 20th century like Xerox Parc or Bell Labs, the engines of so many massive leaps forward in computation and communication. It is also the force that tends to concentrate venture funding in technologies and applications that are recognizably adjacent to already successful ones. So these scientific leaders seeking to build practical qubits had little choice but to find compelling arguments for the CFOs and the VCs to make a bet on their grand projects.

Not that these were dumb bets, not at all. A new paradigm for computing has only happened once before in modern history, and the earliest investors in the successful applications of that paradigm made huge returns. The issue with timing is that to get it right through rational planning, one must have a way to model out when the relevant events will occur. This is why smart money hates investing in fundamental science. Scientific breakthroughs are almost always achieved by human curiosity and undirected inquiry—the only two strategies proven to be effective when we "don't know what we don't know." American-Lebanese essayist Nassim Nicholas Taleb wrote about the difficulties posed by the lack of priors on which to predict future outcomes in 2001, calling them "black swans." In Europe, it was believed that all swans were white until 1697, when Willem de Vlamingh saw a black swan while exploring Western Australia. Black swan events are challenging for investment due to not only uncertain timelines but to the more fundamental inability to measure the chances of either a positive or negative outcome. Before de Vlamingh saw the black swans there, Europeans believed that all swans were white, and the probability that a black swan existed was zero. After the discovery, the probability of a black swan's existence had increased infinitely. This is a difficult set of circumstances to construct a rational investment case around.

The other dynamic that has become familiar to observers of early stage technology investment is the market bubble. The Netscape initial public offering (IPO) on Nasdaq in 1995 was seen as a watershed moment for the commercialization of the internet and the World Wide Web. Spun out of the National Center for Supercomputing Applications at the University of Illinois Urbana-Champaign in April 1994, Netscape was initially funded by famed venture capitalist Jim Clarke, with $4.1M from his firm, Kleiner Perkins. By the end of the day of its IPO, just 16 months later, the firm was worth almost $3B. The dream of the meteoric rise to wealth, fame, and power as the result of modern technology and venture capital was born. Startups sprung up and were funded faster than anyone could track. Many had clever ideas based on realistic plans, but many did not have the first clue how to build what they were proposing or, if it were built, how it would ever make money. Startups found funding through the “strength of the founding team,” which amounted to little more than a positive vibe from the interaction in a Sand Hill Road boardroom and confidence they could fake it till they made it. The investment strategy that seemed to prevail was to invest in ten random startups, of which nine would fail and one would produce a hundredfold return. The entrepreneurial spirit among those startup founders was admirable but perhaps relied too much on

enthusiasm and self-confidence and not enough on fundamentals of microeconomics.

Eventually, the optics of business models that valued “eyeballs” but offered no path to profitability caught up with the dot-com boom. There’s only so much faking one can do, and time runs out if one is not eventually making it. Financial markets’ appetites for public offerings waned, interest rate hikes ended the flow of nearly free venture capital, and in March 2000, the house of cards came tumbling down. In the five years between Netscape’s IPO and the dot-com crash, the Nasdaq had gained 800%. The quantum computing industry can’t compare to the scope and scale of the dot-com bubble, but there are enough similarities between the two to raise concerns.

And so, 25 years after the founding of the first privately funded quantum computing startup, D-Wave, with new uncertainty about the validity of the milestones achieved so far, it is natural for the investors in the field to be feeling some uncertainty. Even before the Shanghai result toppled Google’s claim of supremacy, there have been a variety of worrisome signals. Quantum algorithms which have been purported to have an advantage over classical approaches have had those claims refuted by researchers like Ewin Tang, a postdoctoral fellow at University of California Berkeley. Techniques that sought

shortcuts to commercial applications by combining quantum computing with classical in hybrid algorithms have consistently fallen short. Most importantly, perhaps, the hardware and algorithm innovation driven by the field of AI has outperformed expectations, diminishing the immediacy of the perceived threat to Moore's law. That the threat to chill interest in quantum is coming from AI is certainly ironic, because the use of the term "winter" to describe when a technology falls out of favor was coined by AI pioneers like Marvin Minsky in the early '80s about their own experiences.

Maybe it's simply time to shift back to a more purely scientific footing. National and regional governments began investing in quantum computing again in earnest around the time of the Google experiment, and by 2024 more than $40B has been committed worldwide to fund quantum technology research. With the classical microprocessor industry now hyperconcentrated in just a handful of companies around the world, there is a general desire not to see a repeat of that pattern with quantum technologies. Public sources of funding have a much broader idea of the benefits they are seeking to gain through their money, and so fundamental research can progress without unrealistic timelines for commercial milestones weighing it down. The challenge may be to shift from a commercial focus back to a focus on the public good,

without creating the perception that somehow the mission overall has been a failure.

It is undeniable that efforts of the quantum community have been producing impressive steps forward. We've evolved far beyond the tiny handfuls of short-lived qubits to near maximum yields for this stage in the evolution of each modality. Superconducting devices from IBM have broken records, exceeding 1,000 qubits on a single chip, and others, like Finland's IQM, are not far behind. Ion traps are at the 30-plus qubit limit for one dimension, with stable, long-lived coherent chains in the machines from IonQ and Quantinuum. Neutral atom arrays, the first cousin to ion traps, are developing rapidly in devices from QuEra, Atom Computing, Pasqal, and others. Photonic and spin qubit boosters continue to ask for patience and promise stunning accomplishments at some point in the future. The gate fidelities and coherence times on all devices continue to increase, which is enabling novel and valuable science, even on noisy, error-prone qubits. The machines Feynman dreamt of, able to efficiently simulate the quantum systems found in nature, are here.

And then there's fault tolerance. Long-running, precise computation will require it, and it's been a vital area of research from the beginning. Shor's scheme for error

correction was just the starting point, but conceptually all attempts to make qubits more resilient follow a similar pattern. Groups of physical qubits are operated in a coordinated fashion where errors can be either detected or avoided, creating a single, logical qubit out of the whole ensemble. As you might imagine, this compounds the challenges of controlling a single qubit, already a high engineering bar. There have been demonstrations of single logical qubit error correction schemes, especially within the ion trap community, but there's not much you can do with a single qubit. The technical roadmaps discussed by hardware vendors working toward fault-tolerant quantum computing suggested it could be five to ten more years before we see logical qubits in larger numbers.

However, more recent work from IBM and QuEra has shaken those assumptions and holds the promise of larger-scale capabilities in the near future. Near the end of 2023, the scientists at Harvard and MIT who founded QuEra published a paper detailing work they had carried out in their lab. Using an optical array of 280 atoms, they built a kind of disembodied von Neumann machine, with areas designated for storage and computation, and created 48 logical qubits out of coordinated ensembles of atoms. Not to be outdone, in March 2024, IBM published a paper that described theory work defining a biplanar chip architecture where 288 physical qubits are

combined using a scheme called quantum low-density parity codes to create 12 logical qubits. It is expected that both companies will bring devices to market in the next year or two that will make logical qubits available for general use.

What's particularly exciting is that the generally accepted limit for efficient simulation of fault-tolerant qubits by classical computers is about 50. Since the Hilbert space of a quantum computer, containing all possible states, increases exponentially, 50 qubits represent a quadrillion-dimensional space—for all intents and purposes, impossible to simulate classically. That the experiment QuEra is working on bringing to market had 48 logical qubits puts them at the very brink of that limit. The chip described in IBM's paper only has 12 qubits, but the company has been working very hard to cluster machines together, working with cryo vendor BlueFors on the Kide system, a hexagonal fridge with five chambers. Filling that cluster with new chips would give IBM a system with 60 logical qubits, well beyond the classical simulation limit. Systems of this scale suggest the true era of quantum advantage is very close, since anything useful carried on those devices would be impossible to achieve any other way.

The overarching challenge that unites the field is to find the balance required to stay the course. Enterprises that will

eventually be using quantum computers to achieve new operational capabilities must calibrate their investments to maintain a focus that is sufficient and sustainable. This means sufficient to ready them for rapid adoption when the time is right, at a level that is sustainable over an unknowable span of time. Investors, too, must be rigorous with their own expected horizons on returns. Either settle in for the long haul, using funds that can roll from one 10-year fund to another if need be, or target shorter-term, lower-yield exits through supply chain plays, professional service models, acqui-hire, and the like. The public sector also has a vital role to play. It is far too early to step aside for the private sector to take the lead on this highly strategic technology. Innovative private-public partnership models need to be experimented with on a continuous basis at the regional, national, and global scale, until such time as the arc of development surges past any nurturing framework to take its place as a spectacular source of the creation of value.

Most importantly, the scientists and engineers dreaming of our quantum future need to practice radical transparency on all the unknown and known risks for any aspect of the overall endeavor. It may be difficult to answer questions about applications and timelines with a perfectly honest “we don’t know,” but introducing or tolerating imprecision or confusion in nonscientific audiences is inadvisable. Gaps in

understanding, inaccurate analogies with classical emerging technology patterns, or other errors may keep the funding train on the track in the near term but eventually will cause disillusionment, cynicism, and loss of trust. The crucial mission is to make the case for the transformative power of creating tools that operate at the quantum limits. Quantum physics remains mystifying even as it continues to be one of the most successful scientific theories in history. The computing and communication technologies we are building derive their power directly from the most mystifying aspects of that theory, and therefore are highly resistant to the human imagination. In this way, a central challenge of quantum information science is one of communication. We desperately need the metaphorical and narrative tools to begin to grasp what it is we are building, to both maintain our focus on the task at hand and to prepare for the future these tools will enable us to create.

About the Author

Sebastian Hassinger is a veteran of emerging technology markets stretching back to early commercialization of the internet. He built Apple's first customer support website, IBM's first internet systems management technologies, and mobile computing products, and he holds over a dozen patents. As an entrepreneur, he cofounded two ISPs (the one in Canada was too early, the one in Texas too late) and an iOS development studio in Ho Chi Minh City. He has coauthored technical books on HTML, CGIs (remember those?), VRML (bet you don't remember that!), and other topics. Since 2017 he has been working on the early development of quantum technologies, first for IBM and now for AWS. He created and cohosted the podcast *Forwards & Backwards: A History of Quantum Computing* while at IBM and now cohosts *The New Quantum Era*, a podcast that features interviews with the scientists and engineers at the very bleeding edge of this new technology.

Colophon

The cover illustration is original art by Susan Thompson. The cover fonts are Gilroy and Guardian Sans. The text fonts are Minion Pro and Scala Pro; the heading and sidebar font is Benton Sans.