PRINCETON UNIVERSITY

Ph410
Physics of Quantum Computation[1]

Kirk T. McDonald

(This version created September 21, 2022)

kirkmcd@princeton.edu

http://kirkmcd.princeton.edu/examples/ph410problems.pdf

Reference: M.A. Nielsen and I.L. Chuang, *Quantum Computation and Quantum Information* (Cambridge UP, 2000)

See also the courses by Meglicki, Mermin and Preskill:

http://beige.ucs.indiana.edu/B679/

http://people.ccmr.cornell.edu/~mermin/qcomp/CS483.html

http://www.theory.caltech.edu/~preskill/

Copies of many papers relevant to this course are in the directory
http://kirkmcd.princeton.edu/examples/QM/
in the format author_journal_vol_page_year.pdf

An electronic archive of unpublished papers on Quantum Physics (and many other topics) is at http://www.arxiv.org/

An electronic archive of published papers on Quantum Information is at
http://www.vjquantuminfo.org/quantuminfo/

[1] A classic "quantum" computation: π = 3.1415 92653 58979... = "How I need a drink, alcoholic of course, after the heavy lectures regarding quantum mechanics..." PC users, please change "alcoholic" to "cranberry" or "pineapple". Skeptics can substitute "weirdness" for "mechanics".

Problems

Solutions

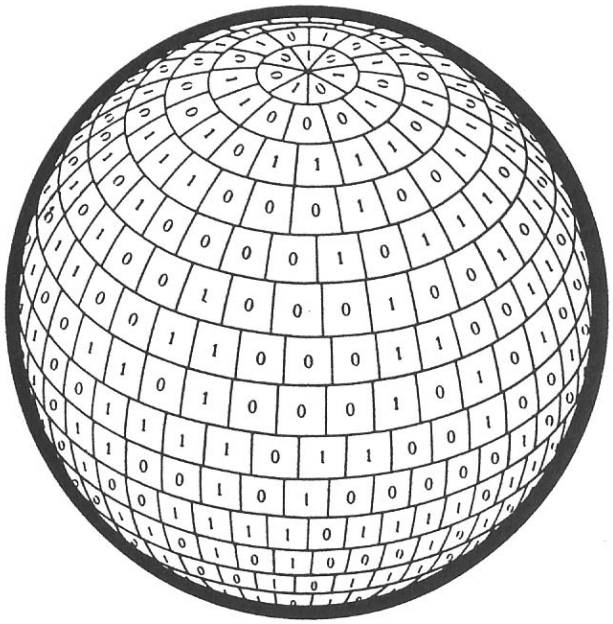
1. **The Ultimate Laptop**

   A laptop computer weighs about 1 kg and occupies a volume of about 1 l.

   Without knowing exactly what a quantum computer is, deduce limitations on the speed (N operations per second) and memory size (M in bits) of a laptop quantum computer, based on the uncertainty principle and thermodynamics/statistical mechanics.[2] How does the capability of the laptop depend on its temperature T? Compare the operation of the laptop at room temperature to the case where all of the rest energy of the laptop is available.

   It is instructive to relate the memory size of the laptop to its entropy.[3]

   There might be some computational advantages to making the laptop as small as possible. The ultimate compact laptop would be a 1-kg black hole. Given that the entropy of a black hole is roughly kA/L_P^2, where k is Boltzmann's constant, A is the surface area, and L_P is the Planck length, what is the memory size of the black-hole laptop?



"It from Bit" – John Archibald Wheeler

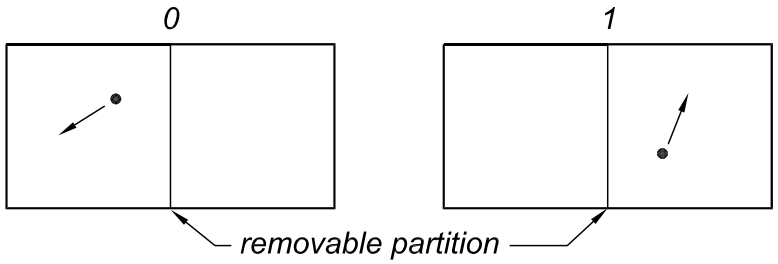
[2] If you feel the urge to "peek" at the literature, note that this problem is based on
http://kirkmcd.princeton.edu/examples/QM/bekenstein_prl_46_623_81.pdf
See also, http://kirkmcd.princeton.edu/examples/QM/lloyd_nature_406_1047_00.pdf

[3] For a high-level "reminder" about entropy (much more than needed here!), see
http://kirkmcd.princeton.edu/examples/QM/wehrl_rmp_50_221_78.pdf

2. Maxwell's Demon

One of the earliest conceptual "supercomputers" was Maxwell's Demon[4], who uses intelligence in sorting molecules to appear to evade the Second Law of Thermodynamics.

To place the demon in a computational context, consider a computer "memory" that consists of a set of boxes (bits), each of volume V and each containing a single molecule. A (re)movable partition divides the volume into "left" and "right" halves. If the molecule is in the left half of a box this represents the 0 state, while if the molecule is in the right half of a box we have the 1 state.



The boxes are all at temperature T, as maintained by an external heat bath. By averaging over the motion of each molecule, we can speak of the pressure P in each box according to the ideal gas law, $P = kT/V$, where k is Boltzmann's constant.

(a) A Model for Classical Erasure of a Bit

A memory bit can be erased (forced to the 0 state) *without knowledge as to the value of that bit* by the following sequence of operations:

- Remove the partition, permitting a free expansion of the gas from volume v to $2V$.
- Isothermally compress the volume of the box from $2V$ back to V by means of a piston that moves from the far right of the box to its midplane. The molecule is now in the left half of the box, no matter in which half it originally was.
- Reinsert the partition (at the right edge of the compressed volume).
- Withdraw the piston, restoring the box to its original shape, with the molecule in the left half of the box and nothing in the right half = the 0 state.

Deduce the total entropy change of the system of memory + thermal bath for the combined processes of free expansion followed by isothermal compression.

Exercise (a) illustrates Landauer's Principle[5] that in a computer which operates at temperature T there is a minimum entropy cost of $k \ln 2$ to perform the "logically irreversible" step of erasure of a bit in memory, while in principle all other types of operations could be performed (reversibly) at zero entropy and zero energy cost.[6]

---

[4] J.C. Maxwell, Letter to P.G. Tait (1867), http://kirkmcd.princeton.edu/examples/statmech/maxwell_demon_quotes.pdf
The Theory of Heat (1871), p. 308, http://kirkmcd.princeton.edu/examples/statmech/maxwell_heat_71.pdf.

[5] http://kirkmcd.princeton.edu/examples/QM/landauer_ibmjrd_5_183_61.pdf

[6] The principle enunciated by Landauer himself is that erasure has an energy cost of at least $kT \ln 2$. Verify that the present example illustrates this claim.

An important extrapolation from Landauer's Principle was made by Bennett[7] who noted that if a computer has a large enough memory such that no erasing need be done during a computation, then the computation could be performed reversibly, and the computer restored to its initial state at the end of the computation by undoing (reversing) the program once the answer was obtained.

The notion that computation could be performed by a reversible process was initially considered to be counterintuitive – and impractical. However, this idea was of great conceptual importance because it opened the door to quantum computation, based on quantum processes which are intrinsically reversible (except for measurement; see prob. 5).

A second important distinction between classical and quantum computation (*i.e.*, physics), besides the irreversibility of quantum measurement, is that an arbitrary (unknown) quantum state cannot be copied exactly (prob. 6).

(b) **Classical Copying of a Known Bit**

In Bennett's reversible computer there must be a mechanism for preserving the result of a computation, before the computer is reversibly restored to its initial state. Use the model of memory bits as boxes with a molecule in the left or right half to describe a (very simple) process whereby a bit, whose value is known, can be copied at zero energy cost and zero entropy change onto a bit whose initial state is 0.

A question left open by the previous discussion is whether the state of a classical bit can be determined without an energy cost or entropy change.

In a computer, the way we show that we know the state of a bit is by making a copy of it. To know the state of the bit, *i.e.*, in which half of a memory box the molecule resides, we must make some kind of measurement. In principle, this can be done very slowly and gently, by placing the box on a balance, or using the mechanical device sketched on the following page,[8] such that the energy cost is arbitrarily low, in exchange for the measurement process being tedious, and the apparatus somewhat bulky. Thus, we accept the assertion of Bennett and Landauer that measurement and copying of a classical bit are, in principle, cost-free operations.[9]

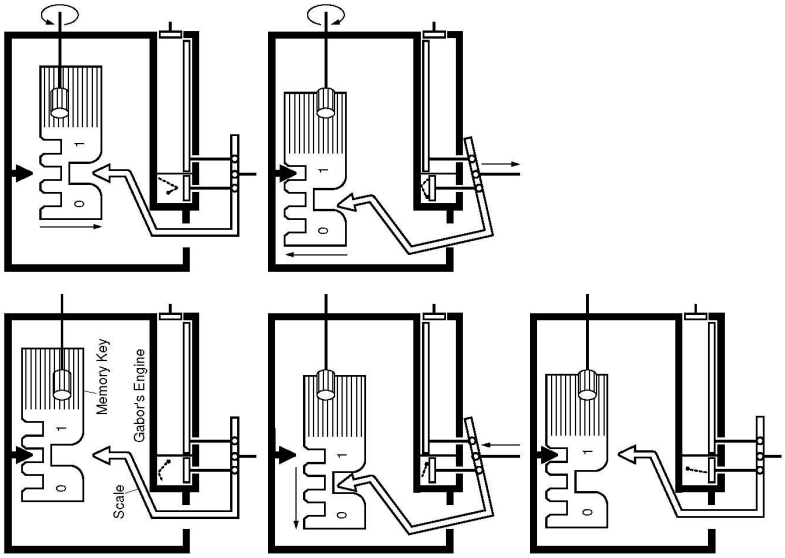
We can now contemplate another procedure for resetting a classical bit to 0. First, measure its state (making a copy in the process), and then subtract the copy from the original. (We leave it as an optional exercise for you to concoct a procedure using the molecule in a box to implement the subtraction.) This appears to provide a scheme for erasure at zero energy/entropy cost, in contrast to the procedure you considered in part (a). However, at the end of the new procedure, the copy of the original bit remains, using up memory space. So, to complete the erasure operation, we should also reset the copy bit. This could be done at no energy/entropy cost by making yet another copy,

[7] A thoughtful review by Bennett is at http://kirkmcd.princeton.edu/examples/QM/bennett_ibmjrd_32_16_88.pdf Bennett's original paper is at http://kirkmcd.princeton.edu/examples/QM/bennett_ibmjrd_17_525_73.pdf

[8] http://kirkmcd.princeton.edu/examples/QM/zurek_quant-ph-9807007.pdf

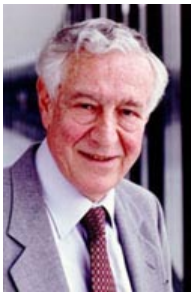
[9] However, there is a kind of hidden entropy cost in the measurement process; namely the cost of preparing in the 0 state the bits of memory where the results of the measurement can be stored.

and subtracting it from the first copy. To halt this silly cycle of resets, we must sooner or later revert to the procedure of part (a), which did not involve a measurement of the bit before resetting it. So, we must sooner or later pay the energy/entropy cost to erase a classical bit.



Recalling Maxwell's demon, we see that his task of sorting molecules into the left half of a partitioned box is equivalent to erasing a computer memory. The demon can perform his task with the aid of auxiliary equipment, which measures and stores information about the molecules. To finish his task cleanly, the demon must not only coax all the molecules into the left half of the box, but he must return his auxiliary equipment to its original state (so that he could use it to sort a new set of molecules...poor demon). At some time during his task, the demon must perform a cleanup (erasure) operation equivalent to that of part (a), in which the entropy of the molecules/computer decreases, but with an opposite and equal (or greater) increase in the entropy of the environment.

The demon obeys the Second Law of Thermodynamics[10] – and performs his task millions of times each second in your palm computer.



The "moral" of this problem is Landauer's dictum:

Information is physical

.

[10]For further reading, see chap. 5 of *Feynman Lectures on Computation* (Addison-Wesley, 1996),
http://kirkmcd.princeton.edu/examples/QM/feynman_computation.pdf
Also, http://kirkmcd.princeton.edu/examples/QM/bennett_ijtp_21_905_82.pdf
http://kirkmcd.princeton.edu/examples/QM/bennett_ibmjrd_32_16_88.pdf
http://kirkmcd.princeton.edu/examples/QM/bub_shpmp_32_569_01.pdf

When the gas is so far condensed that it assumes the liquid or solid form, then, as the molecules have no free path, they have no regular vibrations, and no bright lines are commonly observed in incandescent liquids or solids. Mr. Huggins, however, has observed bright lines in the spectrum of incandescent erbia and lime, which appear to be due to the solid matter, and not to its vapour.

LIMITATION OF THE SECOND LAW OF THERMODYNAMICS.

Before I conclude, I wish to direct attention to an aspect of the molecular theory which deserves consideration.

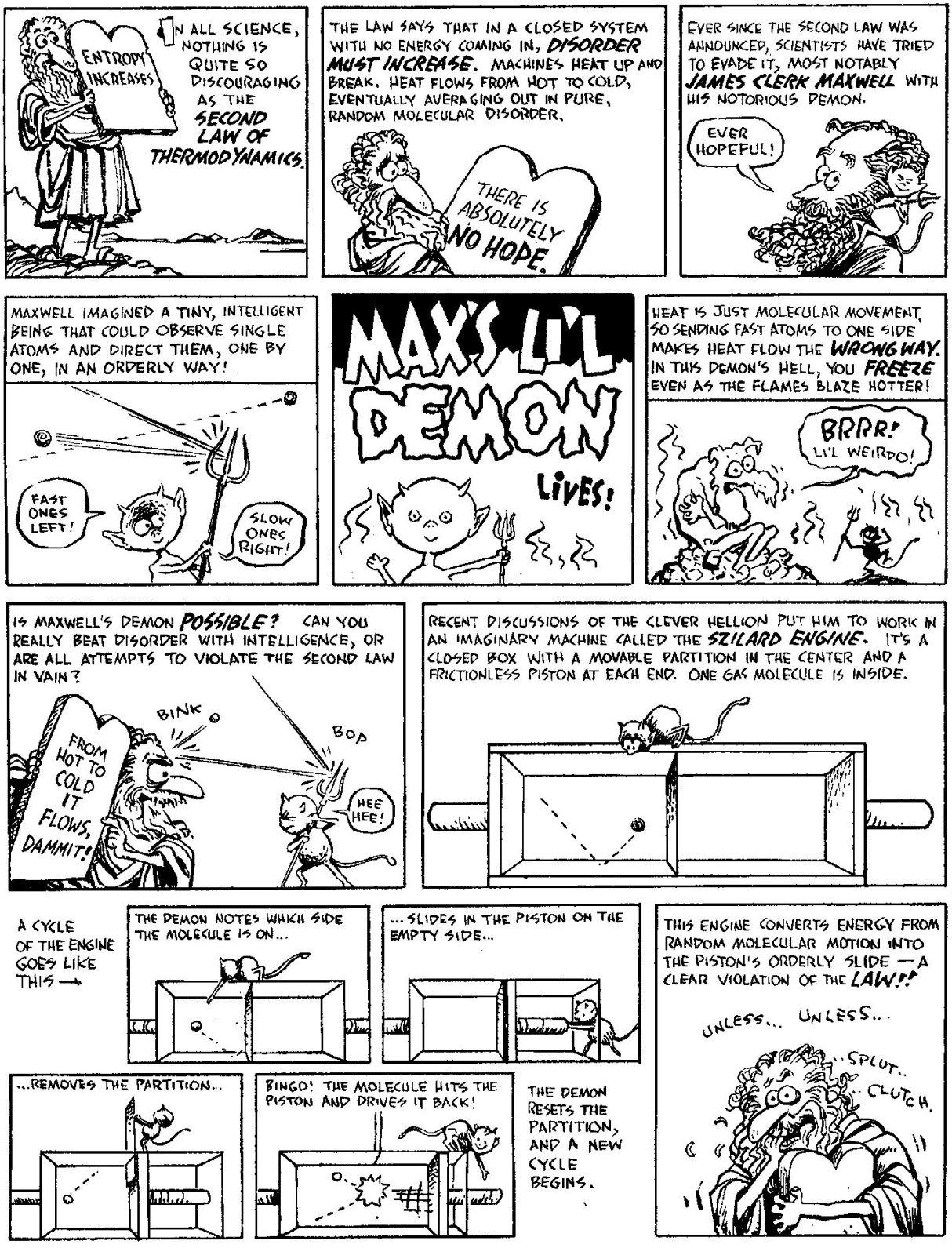
One of the best established facts in thermodynamics is that it is impossible in a system enclosed in an envelope which permits neither change of volume nor passage of heat, and in which both the temperature and the pressure are everywhere the same, to produce any inequality of temperature or of pressure without the expenditure of work. This is the second law of thermodynamics, and it is undoubtedly true as long as we can deal with bodies only in mass, and have no power of perceiving or handling the separate molecules of which they are made up. But if we conceive a being whose faculties are so sharpened that he can follow every molecule in its course, such a being, whose attributes are still as essentially finite as our own, would be able to do what is at present impossible to us. For we have seen that the molecules in a vessel full of air at uniform temperature are moving with velocities by no means uniform, though the mean velocity of any great number of them, arbitrarily selected, is almost exactly uniform. Now let us suppose that such a vessel is divided into two portions, A and B, by a division in which there is a small hole, and that a being, who can see the individual molecules, opens and closes this hole, so as to allow only the swifter molecules to pass from A to B, and only the slower ones to pass from B to A. He will thus, without expenditure of work, raise the temperature of B and lower that of A, in contradiction to the second law of thermodynamics.

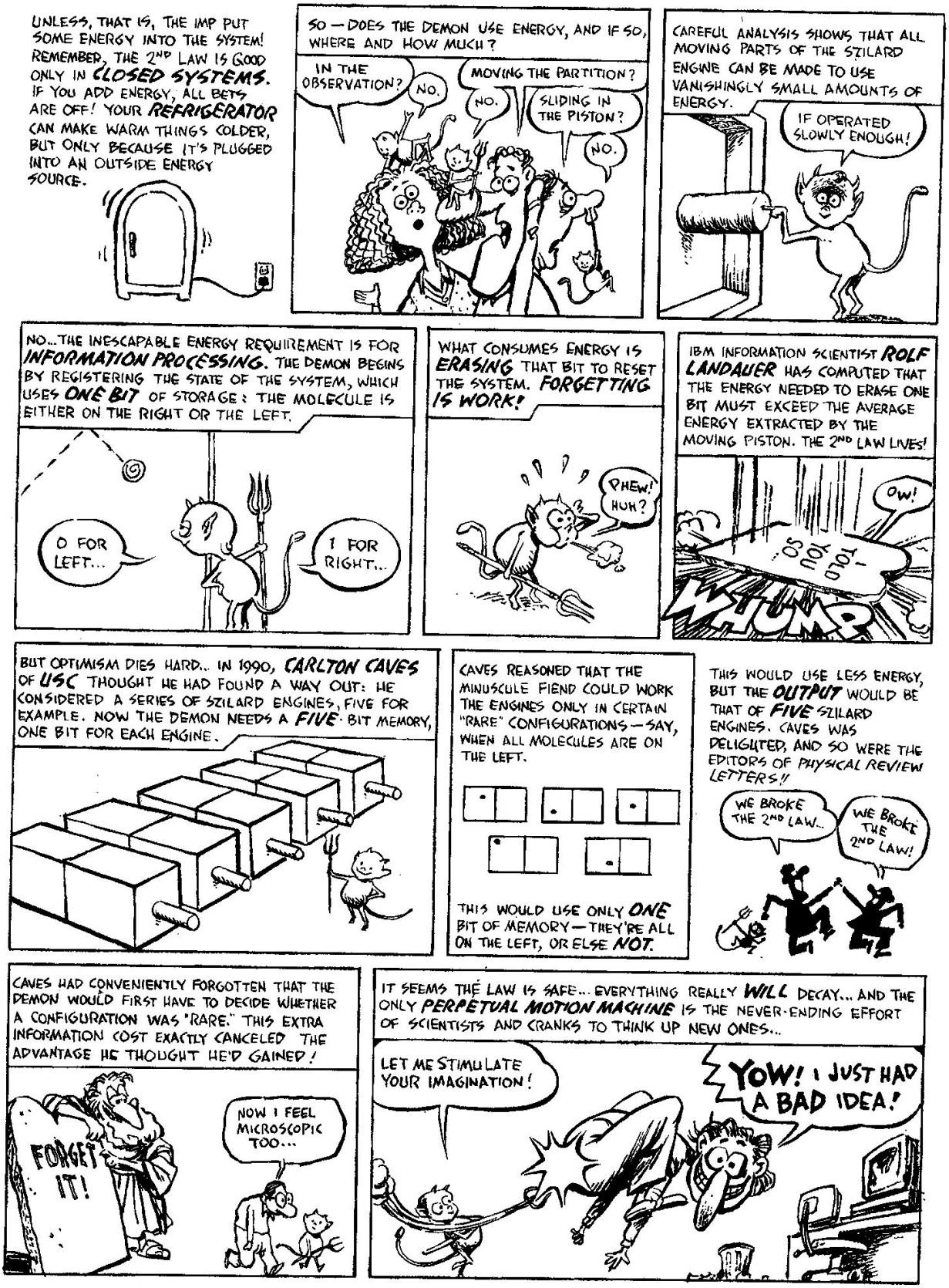
This is only one of the instances in which conclusions which we have drawn from our experience of bodies consisting of an immense number of molecules may be found not to be applicable to the more delicate observations and experiments which we may suppose made by one who can perceive and handle the individual molecules which we deal with only in large masses.

In dealing with masses of matter, while we do not perceive the individual molecules, we are compelled to adopt what I have described as the statistical method of calculation, and to abandon the strict dynamical method, in which we follow every motion by the calculus.

It would be interesting to enquire how far those ideas about the nature and methods of science which have been derived from examples of scientific investigation in which the dynamical method is followed are applicable to our actual knowledge of concrete things, which, as we have seen, is of an essentially statistical nature, because no one has yet discovered any practical method of tracing the path of a molecule, or of identifying it at different times.

I do not think, however, that the perfect identity which we observe between different portions of the same kind of matter can be explained on the statistical principle of the stability of the averages of large numbers of quantities each of which may differ from the mean. For if of the molecules of some substance such as hydrogen, some were of sensibly greater mass than others, we have the means of producing a separation between molecules of different masses, and in this way we should be able to produce two kinds of hydrogen, one of which would be somewhat denser than the other. As this cannot be done, we must admit that the equality which we assert to exist between the molecules of hydrogen applies to each individual molecule, and not merely to the average of groups of millions of molecules.


ENTROPY INCREASES
IN ALL SCIENCE, NOTHING IS QUITE SO DISCOURAGING AS THE SECOND LAW OF THERMODYNAMICS.
THE LAW SAYS THAT IN A CLOSED SYSTEM WITH NO ENERGY COMING IN, DISORDER MUST INCREASE. MACHINES HEAT UP AND BREAK. HEAT FLOWS FROM HOT TO COLD, EVENTUALLY AVERAGING OUT IN PURE, RANDOM MOLECULAR DISORDER.
THERE IS ABSOLUTELY NO HOPE.
EVER SINCE THE SECOND LAW WAS ANNOUNCED, SCIENTISTS HAVE TRIED TO EVADE IT, MOST NOTABLY JAMES CLERK MAXWELL WITH HIS NOTORIOUS DEMON.
EVER HOPEFUL!
MAXWELL IMAGINED A TINY, INTELLIGENT BEING THAT COULD OBSERVE SINGLE ATOMS AND DIRECT THEM, ONE BY ONE, IN AN ORDERLY WAY!
FAST ONES LEFT!
SLOW ONES RIGHT!
MAX'S LI'L DEMON LIVES!
HEAT IS JUST MOLECULAR MOVEMENT, SO SENDING FAST ATOMS TO ONE SIDE MAKES HEAT FLOW THE WRONG WAY. IN THIS DEMON'S HELL, YOU FREEZE EVEN AS THE FLAMES BLAZE HOTTER!
BRRR! LI'L WEIRDO!
IS MAXWELL'S DEMON POSSIBLE? CAN YOU REALLY BEAT DISORDER WITH INTELLIGENCE, OR ARE ALL ATTEMPTS TO VIOLATE THE SECOND LAW IN VAIN?
FROM HOT TO COLD IT FLOWS, DAMMIT!
BINK
BOP
HEE HEE!
RECENT DISCUSSIONS OF THE CLEVER HELLION PUT HIM TO WORK IN AN IMAGINARY MACHINE CALLED THE SZILARD ENGINE. IT'S A CLOSED BOX WITH A MOVABLE PARTITION IN THE CENTER AND A FRICTIONLESS PISTON AT EACH END. ONE GAS MOLECULE IS INSIDE.
A CYCLE OF THE ENGINE GOES LIKE THIS →
THE DEMON NOTES WHICH SIDE THE MOLECULE IS ON...
...SLIDES IN THE PISTON ON THE EMPTY SIDE...
THIS ENGINE CONVERTS ENERGY FROM RANDOM MOLECULAR MOTION INTO THE PISTON'S ORDERLY SLIDE — A CLEAR VIOLATION OF THE LAW!!
UNLESS... UNLESS...
...REMOVES THE PARTITION...
BINGO! THE MOLECULE HITS THE PISTON AND DRIVES IT BACK!
THE DEMON RESETS THE PARTITION, AND A NEW CYCLE BEGINS.
SPLUT..
CLUTCH.


UNLESS, THAT IS, THE IMP PUT SOME ENERGY INTO THE SYSTEM! REMEMBER, THE 2ND LAW IS GOOD ONLY IN CLOSED SYSTEMS. IF YOU ADD ENERGY, ALL BETS ARE OFF! YOUR REFRIGERATOR CAN MAKE WARM THINGS COLDER, BUT ONLY BECAUSE IT'S PLUGGED INTO AN OUTSIDE ENERGY SOURCE.
SO — DOES THE DEMON USE ENERGY, AND IF SO, WHERE AND HOW MUCH?
IN THE OBSERVATION?
NO.
MOVING THE PARTITION?
NO.
SLIDING IN THE PISTON?
NO.
CAREFUL ANALYSIS SHOWS THAT ALL MOVING PARTS OF THE SZILARD ENGINE CAN BE MADE TO USE VANISHINGLY SMALL AMOUNTS OF ENERGY.
IF OPERATED SLOWLY ENOUGH!
NO...THE INESCAPABLE ENERGY REQUIREMENT IS FOR INFORMATION PROCESSING. THE DEMON BEGINS BY REGISTERING THE STATE OF THE SYSTEM, WHICH USES ONE BIT OF STORAGE: THE MOLECULE IS EITHER ON THE RIGHT OR THE LEFT.
0 FOR LEFT...
1 FOR RIGHT...
WHAT CONSUMES ENERGY IS ERASING THAT BIT TO RESET THE SYSTEM. FORGETTING IS WORK!
PHEW! HUH?
IBM INFORMATION SCIENTIST ROLF LANDAUER HAS COMPUTED THAT THE ENERGY NEEDED TO ERASE ONE BIT MUST EXCEED THE AVERAGE ENERGY EXTRACTED BY THE MOVING PISTON. THE 2ND LAW LIVES!
OW!
SO - YOU FOOL
WHUMP
BUT OPTIMISM DIES HARD.. IN 1990, CARLTON CAVES OF USC THOUGHT HE HAD FOUND A WAY OUT: HE CONSIDERED A SERIES OF SZILARD ENGINES, FIVE FOR EXAMPLE. NOW THE DEMON NEEDS A FIVE-BIT MEMORY, ONE BIT FOR EACH ENGINE.
CAVES REASONED THAT THE MINUSCULE FIEND COULD WORK THE ENGINES ONLY IN CERTAIN "RARE" CONFIGURATIONS — SAY, WHEN ALL MOLECULES ARE ON THE LEFT.
THIS WOULD USE ONLY ONE BIT OF MEMORY — THEY'RE ALL ON THE LEFT, OR ELSE NOT.
THIS WOULD USE LESS ENERGY, BUT THE OUTPUT WOULD BE THAT OF FIVE SZILARD ENGINES. CAVES WAS DELIGHTED, AND SO WERE THE EDITORS OF PHYSICAL REVIEW LETTERS!!
WE BROKE THE 2ND LAW...
WE BROKE THE 2ND LAW!
CAVES HAD CONVENIENTLY FORGOTTEN THAT THE DEMON WOULD FIRST HAVE TO DECIDE WHETHER A CONFIGURATION WAS "RARE." THIS EXTRA INFORMATION COST EXACTLY CANCELED THE ADVANTAGE HE THOUGHT HE'D GAINED!
FORGET IT!
NOW I FEEL MICROSCOPIC TOO...
IT SEEMS THE LAW IS SAFE... EVERYTHING REALLY WILL DECAY... AND THE ONLY PERPETUAL MOTION MACHINE IS THE NEVER-ENDING EFFORT OF SCIENTISTS AND CRANKS TO THINK UP NEW ONES...
LET ME STIMULATE YOUR IMAGINATION!
YOW! I JUST HAD A BAD IDEA!

3. States, Bits and Unitary Operations

States and Bits. States of a system, or subsystem, involved in a computation will be denoted in various equivalent ways. A Greek symbol such as ψ will sometimes be used. When considering a quantum system, we will often embed the symbol in a ket, following Dirac:[11] $|\psi\rangle$. Our enthusiasm for this notation is such that we will often use it for classical states as well.

We will deal almost exclusively with subsystems that have a finite number of independent (basis or eigen)states, in which case it is also useful to consider the state as a vector. Thus, for a two-state system we write the state as a column vector:

$$\psi = |\psi\rangle = \begin{pmatrix} \psi_0 \\ \psi_1 \end{pmatrix}. \tag{1}$$

If a classical system $|\psi\rangle$ has n states, it can be in only one of those states, say state $|j\rangle$. Then the vector components ψ_j of state $|\psi\rangle$,

$$|\psi\rangle = \sum_j \psi_j |j\rangle, \tag{2}$$

have values $\psi_j = 0$ or 1.

A classical 2-state system could be used as a bit (or Cbit for classical bit) of a computer memory, and we define,

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{3}$$

An obvious mathematical generalization of a classical state is a vector in which the components ψ_j are complex numbers ($\psi_j = a_j + ib_j$ where a_j and b_j are real numbers, and $i = \sqrt{-1} = e^{i\pi/2}$). An ever-amazing physical fact is that quantum systems can be well described by such vectors.

The physical meaning (as first explained by M. Born in 1926[12]) of the complex vector components of a quantum system is that the absolute square, $\left|\psi_j\right|^2$, of a vector component ψ_j is equal to the probability that the system will be found in state j IF a measurement is made to determine the state of the system.

A measurement is a process that involves some kind of interaction of the quantum system with a more classical system such that the quantum system emerges with the classical property of being in only one of its possible basis states. See prob. 5 for more discussion of measurement.

An important distinction between a classical and a quantum state is that prior to a measurement, a quantum state can be said to be in a superposition of several of its possible basis states, while a classical system can only be in one basis state at time. This greater flexibility of quantum states compared to classical states is the core reason why we might expect quantum computation, involving manipulation of quantum states, to offer greater opportunities than classical computation.

[11] http://kirkmcd.princeton.edu/examples/QM/dirac_pcps_35_416_39.pdf
[12] http://kirkmcd.princeton.edu/examples/QM/born_zp_37_863_26_english.pdf

A physical restriction is that the total probability must be unity for finding a quantum system in some state. So, we always assume the normalization condition,

$$|\psi|^2 = \sum_j \left|\psi_j\right|^2 = 1, \tag{4}$$

for our quantum states (2). If we think of the state ψ as a (column) vector $\boldsymbol{\psi}$ as in eq. (1), then the normalization condition (4) could be written as,

$$\boldsymbol{\psi}^\star \cdot \boldsymbol{\psi} = 1, \tag{5}$$

where the $\star$ implies complex conjugation. Anticipating the use of matrices along with the state vectors, we can think of the vector $\boldsymbol{\psi}^\star$ in eq. (5) as a row vector,

$$\boldsymbol{\psi}^\star = (\psi_0^\star, \psi_1^\star, \ \ldots, \ \psi_n^\star). \tag{6}$$

And, in the notation of Dirac, we introduce the bra of ψ as,

$$\boldsymbol{\psi}^\star = \langle\psi|. \tag{7}$$

In sum, the various ways of writing the normalization condition for a quantum state are,

$$|\psi|^2 = \langle\psi|\psi\rangle = \boldsymbol{\psi}^\star \cdot \boldsymbol{\psi} = \sum_j \left|\psi_j\right|^2 = 1. \tag{8}$$

Given two states $|\psi\rangle$ and $|\phi\rangle$ we can define a vector dot product (inner product or scalar product) as,

$$\langle\psi|\phi\rangle = \boldsymbol{\psi}^\star \cdot \boldsymbol{\phi} = \sum_j \psi_j^\star \phi_j. \tag{9}$$

When $\langle\psi|\phi\rangle = 0$ we say that states $|\psi\rangle$ and $|\phi\rangle$ are orthogonal.

The simplest quantum system of interest for computation is a 2-state system, which could function as a quantum bit (or Qbit). The zero and one states can be written as in eq. (3), and a general Qbit can be written as in eq. (1), together with the normalization condition that,

$$|\psi_0|^2 + |\psi_1|^2 = 1. \tag{10}$$

The general Qbit could also be written as,

$$|\psi\rangle = \psi_0|0\rangle + \psi_1|1\rangle, \tag{11}$$

which is a superposition of the Qbits $|0\rangle$ and $|1\rangle$.

The "meaning" of a Qbit state (11) to a (classical) observer is ascertained by a measurement, the results of which are that it is found to be in state $|0\rangle$ with probability $|\psi_0|^2$ and in state $|1\rangle$ with probability $|\psi_1|^2$. Hence, there is no change to this meaning if the state were multiplied by an arbitrary phase factor $e^{i\phi}$. Thus, there is not a unique identification of a Qbit state with its meaning as determined by measurement. For example, the Qbit $|0\rangle$ can be represented by the equivalent forms,

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} -i \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} (1+i)/\sqrt{2} \\ 0 \end{pmatrix}, \qquad etc. \tag{12}$$

Note that the complex coefficients ψ_0 and ψ_1 that characterize the Qbit state (11) cannot be determined by a single measurement. Only if a large number of copies of a Qbit were available could its state be determined to good accuracy, via a large set of measurements.[13]

Multiple Bit States.

We digress slightly to record our notation for states involving multiple bits, whether classical or quantum.

The simplest states of a multiple bit system are those that can be expressed as a direct product (= tensor product) of single bit states. For example, in a system of 3 bits $|x\rangle$, $|y\rangle$ and $|z\rangle$ we can denote the direct product state $|\psi\rangle$ using the direct-product symbol $\otimes$ or not, as we find more convenient:

$$|\psi\rangle = |x\rangle \otimes |y\rangle \otimes |z\rangle = |x\rangle|y\rangle|z\rangle = |xyz\rangle. \tag{13}$$

The last form of eq. (13) is the most compact notation, and will be the one most used. We can also express the state $|\psi\rangle$ as a column vector of length 2^n for an n-bit state. For example, a system of 3 bits could be written as,

$$|\psi\rangle = \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \otimes \begin{pmatrix} y_0 \\ y_1 \end{pmatrix} \otimes \begin{pmatrix} z_0 \\ z_1 \end{pmatrix} = \begin{pmatrix} x_0y_0z_0 \\ x_0y_0z_1 \\ x_0y_1z_0 \\ x_0y_1z_1 \\ x_1y_0z_0 \\ x_1y_0z_1 \\ x_1y_1z_0 \\ x_1y_1z_1 \end{pmatrix}. \tag{14}$$

Note that the vector for a classical state of n bits has exactly one nonzero element (whose value is, of course, 1), in contrast to a quantum state vector in which all elements can be nonzero (and complex, subject to the normalization condition (8)). For example, a system of 3 Cbits, 1, 0 and 1 is written as,

$$|1\rangle|0\rangle|1\rangle = |101\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}. \tag{15}$$

[13] This lack of "completeness" to the operational meaning of a single Qbit state is a source of unease for many people. However, our task is to master the computational and physical richness of quantum states, rather than to get bogged down over metaphysical questions of "meaning".

While the Qbit (11) will be observed (by a suitable apparatus) to have only the value $|0\rangle$ or $|1\rangle$, the claim is that it does not have either of these values prior to being observed, and that there is an intrinsic probabilistic character to the observation of a single Qbit. This claim is counterintuitive to "classical" thinking, but it appears to be well supported by experimental evidence.

In more ordinary language, this is the integer 5.

We will often abbreviate these n-bit, direct-product basis states as $|j\rangle_n$, meaning,

$$|j\rangle_n = \prod_{l=0}^{n-1} \otimes |j_l\rangle = \prod_{l=0}^{n-1} |j_l\rangle, \tag{16}$$

where $j_l = 0$ or 1. An identity that will be useful on occasion is,

$$\sum_{j=0}^{2^n-1} |j\rangle_n = \prod_{l=0}^{n-1} \sum_{j_l=0}^{1} |j_l\rangle = \prod_{l=0}^{n-1} (|0\rangle + |1\rangle). \tag{17}$$

In contrast to classical states, quantum states can be constructed by adding together (superposing) other quantum states.[14] Multiple bit states that are the sums of direct product states are said to be entangled. For example, the state,

$$|\psi\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}} \tag{18}$$

is entangled, and is not meaningful in a classical system. The existence of entangled states in quantum systems proves to be one of the richest features of such systems.[15]

Unitary Bit Operations. In both classical and quantum computation we consider operations on bits. The most basic of these operations takes one bit into another (during some time interval not specified here). If symbol O represents such an operation, which takes initial state ψ_i into final state ψ_f, we write,

$$\psi_f = \mathsf{O}\psi_i, \qquad \text{or} \qquad |\psi_f\rangle = \mathsf{O}|\psi_i\rangle. \tag{19}$$

If we think of the states as n-dimensional vectors, then the operator O can be regarded as an $n \times n$ matrix whose elements O_{jk} are real in case of classical operations and complex in case of quantum operations. Thus, we can write the effect of operation O on column vectors as,

$$\psi_{f,j} = \sum_{j,k} \mathsf{O}_{jk} \psi_{i,k}. \tag{20}$$

For the corresponding row vectors, whose elements are the complex conjugates of the elements of the column vectors, the effect of operation O can be written as,

$$\psi^\star_{f,j} = \sum_{j,k} (\mathsf{O}_{jk} \psi_{i,k})^\star = \sum_{j,k} \psi^\star_{i,k} \mathsf{O}^\star_{jk} = \sum_{j,k} \psi^\star_{i,k} \mathsf{O}^{T\star}_{kj} = \sum_{j,k} \psi^\star_{i,k} \mathsf{O}^\dagger_{kj}, \tag{21}$$

[14] To restore the normalization condition (8), the sum of n states should be rescaled by a factor. If the component states are orthogonal, that factor is simply $1/\sqrt{n}$.

[15] Indeed, many of the features of quantum systems that so bothered such people as Einstein and Schrödinger can be traced to the existence of entangled states. The wonderful word entanglement was introduced to physics by Schrödinger in his "cat" paper of 1935, which in my opinion was the greatest contribution of that paper. The understanding the intimate relation of entanglement to measurement (see prob. 5), is due to von Neumannn (1932).
http://kirkmcd.princeton.edu/examples/QM/einstein_pr_47_777_35.pdf
http://kirkmcd.princeton.edu/examples/QM/schroedinger_cat.pdf

where O^T is the transpose of O ($\mathsf{O}^T_{jk} = \mathsf{O}_{kj}$), and adjoint $\mathsf{O}^\dagger$ is the transpose of the complex conjugate of O ($\mathsf{O}^\dagger = \mathsf{O}^{T\star}$). Hence, in Dirac's notation the bra of ψ_f is,

$$\langle\psi_f| = \langle\psi_i|\mathsf{O}^\dagger. \tag{22}$$

If operator O is to represent a physical operation, then the final state must also obey the normalization condition (8). Thus,

$$1 = \langle\psi_f|\psi_f\rangle = \langle\psi_i|\mathsf{O}^\dagger\mathsf{O}|\psi_i\rangle = \langle\psi_i|\psi_i\rangle = 1. \tag{23}$$

We conclude that,

$$\mathsf{O}^\dagger\mathsf{O} = \mathbf{I}, \tag{24}$$

where $\mathbf{I}$ is the identity operator. When written as a matrix, operator $\mathbf{I}$ is, of course, the unit matrix. For operations on a single bit,

$$\mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{25}$$

Since the inverse O^{-1} of operator O obeys,

$$\mathsf{O}^{-1}\mathsf{O} = \mathbf{I} \tag{26}$$

by definition, we see that an operator O must be unitary if it represents a physical transformation on a classical or quantum state, meaning,

$$\mathsf{O}^{-1} = \mathsf{O}^\dagger = \mathsf{O}^{T\star}. \tag{27}$$

An alternative notation for bit operators in terms of Dirac's bras and kets is sometimes useful. A general 2×2 unitary matrix,

$$\mathsf{U} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{28}$$

that acts on states in the $[|0\rangle, |1\rangle]$ basis can also be written as,

$$\mathsf{U} = a|0\rangle\langle 0| + b|0\rangle\langle 1| + c|1\rangle\langle 0| + d|1\rangle\langle 1|. \tag{29}$$

Thus,

$$\begin{aligned} \mathsf{U}|\psi\rangle &= \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} \psi_0 \\ \psi_1 \end{pmatrix} = \begin{pmatrix} a\psi_0 + b\psi_1 \\ c\psi_0 + d\psi_1 \end{pmatrix} = (a\psi_0 + b\psi_1)|0\rangle + (c\psi_0 + d\psi_1)|1\rangle \\ &= (a|0\rangle\langle 0| + b|0\rangle\langle 1| + c|1\rangle\langle 0| + d|1\rangle\langle 1|)(\psi_0|0\rangle + \psi_1|1\rangle). \end{aligned} \tag{30}$$

An introduction to quantum computation that emphasizes this notation is given in
http://kirkmcd.princeton.edu/examples/QM/knill_quant-ph-0207171.pdf

(a) **2×2 Classical Unitary Operators**

Deduce the form of all possible 2×2 unitary matrices than can act on a single classical bit. Show that there is only one nontrivial such operator, the NOT gate:

$$\mathsf{NOT} = \mathsf{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{31}$$

We denote the NOT gate (operator) by the symbol X because it simply flips a bit. How many distinct unitary matrices are there for a system of n classical bits?

(b) **Square Root of NOT**

As an example of a quantum operation on a Qbit that has no classical analog, construct the $\sqrt{\mathsf{NOT}}$ operator (which, of course, will be a unitary 2×2 matrix with complex elements).

(c) **Arbitrary 2×2 Unitary Matrix**

An arbitrary 2×2 unitary matrix U can be written as,

$$\mathsf{U} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \tag{32}$$

where a, b, c and d are complex numbers such that $\mathsf{U}\mathsf{U}^\dagger = \mathbf{I}$. The decomposition (32) is somewhat trivial. Express the general unitary matrix U as the sum of four unitary matrices, times complex coefficients, of which two are the classical unitary matrices $\mathbf{I}$ and X that were found in part (a). Denote the "partner" of $\mathbf{I}$ by Z and the "partner" of X by Y such that,

$$\mathsf{X}\mathsf{Y} = i\mathsf{Z}, \qquad \mathsf{Y}\mathsf{Z} = i\mathsf{X}, \qquad \mathsf{Z}\mathsf{X} = i\mathsf{Y}. \tag{33}$$

You have, of course, rediscovered the so-called Pauli spin matrices,[16]

$$\boldsymbol{\sigma}_x = \mathsf{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_y = \mathsf{Y} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_z = \mathsf{Z} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{34}$$

As usual, we define the Pauli "vector" $\boldsymbol{\sigma}$ as the triplet of matrices,

$$\boldsymbol{\sigma} = (\boldsymbol{\sigma}_x, \boldsymbol{\sigma}_y, \boldsymbol{\sigma}_z). \tag{35}$$

Show that for ordinary vectors $\mathbf{a}$ and $\mathbf{b}$,

$$(\mathbf{a} \cdot \boldsymbol{\sigma})(\mathbf{b} \cdot \boldsymbol{\sigma}) = (\mathbf{a} \cdot \mathbf{b})\, \mathbf{I} + i\, \boldsymbol{\sigma} \cdot \mathbf{a} \times \mathbf{b}. \tag{36}$$

With this, show that a general 2×2 unitary matrix can be written as,

$$\mathsf{U} = e^{i\delta} \left(\cos\frac{\theta}{2}\, \mathbf{I} + i \sin\frac{\theta}{2}\, \hat{\mathbf{u}} \cdot \boldsymbol{\sigma} \right) = e^{i\delta} e^{i\frac{\theta}{2}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}}, \tag{37}$$

[16] http://kirkmcd.princeton.edu/examples/QM/pauli_zp_43_601_27.pdf The Pauli spin matrices (and the unit matrix $\mathbf{I}$) are not only unitary, they are also hermitian, meaning that they are identical to their adjoints: $\boldsymbol{\sigma}_j^\dagger = \boldsymbol{\sigma}_j$.

where δ and θ are real numbers and $\hat{\mathbf{u}}$ is a real unit vector.[17]

What is the determinant of the matrix representation of U? The subset of 2×2 unitary matrices with unit determinant is called the special unitary group SU(2). What is the version of eq. (37) that describes 2×2 special unitary operators?

Are the $\sqrt{\mathsf{NOT}}$ operators found in part (b) special unitary operators?

You may wish to convince yourself of a factoid related to eq. (37), namely that if A *is a square matrix of any order such that* $\mathsf{A}^2 = \mathbf{I}$*, then* $e^{i\theta\mathsf{A}} = \cos\theta\,\mathbf{I} + i\sin\theta\,\mathsf{A}$*, provided that* θ *is a real number. It follows that* A *can also be written in the exponential form.*

$$\mathsf{A} = e^{i\pi/2}e^{-i\frac{\pi}{2}\mathsf{A}} = e^{-i\pi/2}e^{i\frac{\pi}{2}\mathsf{A}}. \tag{38}$$

There are several unitary operators of interest, such as the Pauli matrices, that are their own inverse. If we call such an operator V, then its exponential representation of V can be written in multiple ways,

$$\mathsf{V} = e^{i\delta}e^{i\frac{\theta}{2}\hat{\mathbf{v}}\cdot\boldsymbol{\sigma}} = \mathsf{V}^{-1} = e^{-i\delta}e^{-i\frac{\theta}{2}\hat{\mathbf{v}}\cdot\boldsymbol{\sigma}}. \tag{39}$$

(d) Why isn't the operator Z of eq. (34) a valid 2×2 classical unitary operator?

(e) **2-Bit Classical Unitary Operations**

How many 2-bit classical unitary operators are there?

While these operators can be expressed as 4×4 matrices, it is also instructive to catalog them using 2×2 matrices. For this, we regard a pair of Cbits x and y as elements of a 2-dimensional vector, $\begin{pmatrix} x \\ y \end{pmatrix}$.

Show that the number of 2-bit classical unitary operators is the same as the number of operators described by the transformations,

$$\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} x' \\ y' \end{pmatrix} = \mathsf{M}\begin{pmatrix} x \\ y \end{pmatrix} \oplus \begin{pmatrix} a \\ b \end{pmatrix}, \tag{40}$$

where M is an invertible 2×2 matrix (but not necessarily unitary), and a and b are constant Cbits. The symbol $\oplus$ means addition modulo 2, and the additions resulting from the matrix multiplication are also modulo 2.

Since the matrices M are invertible, eq. (40) describes reversible transformations. Explain how/why these transformations are also unitary (and hence are representations of all of the 2-bit classical unitary operators).

An important conclusion that can be drawn from the identification of 2-bit classical unitary operators with the form (40) is that they are all linear functions of the input bits x and y.

It turns out that nonlinear, classical, unitary bit operators require 3 or more Qbits. Hence, the classical AND *and* OR *operations will require 3 Qbits for a quantum implementation. See prob. 9.*

[17] Note that if make the replacements $\theta \to -\theta$ and $\hat{\mathbf{u}} \to -\hat{\mathbf{u}}$ we obtain another valid representation of U, since the physical operation of a rotation by angle θ about an axis $\hat{\mathbf{u}}$ is identical to a rotation by $-\theta$ about the axis $-\hat{\mathbf{u}}$.

4. **Rotation Matrices**[18]

The form (37) of a general 2×2 unitary matrix suggests that these matrices have something to do with rotations. Certainly, a matrix that describes the rotation of a vector is a unitary transformation.

A general single-Qbit state $|\psi\rangle = \psi_0|0\rangle + \psi_1|1\rangle$, where $|\psi_0|^2 + |\psi_1|^2 = 1$, can also be written as,

$$|\psi\rangle = e^{i\gamma}\left(\cos\alpha|0\rangle + e^{i\beta}\sin\alpha|1\rangle\right). \tag{41}$$

As noted above, the overall phase γ has no meaning (to a measurement of $|\psi\rangle$). [19] So, it is tempting to interpret parameters α and β as angles describing the orientation in a spherical coordinate system (r, θ, ϕ) of a unit 3-vector that is associated with the state $|\psi\rangle$. The state $|0\rangle$ might then correspond to the unit 3-vector $\hat{\mathbf{z}}$ that points up along the z-axis, while $|1\rangle \leftrightarrow -\hat{\mathbf{z}}$.

However, this doesn't work! The suggestion is that the Qbit 0 corresponds to angles $\alpha = 0$, $\beta = 0$ and Qbit 1 to angles $\alpha = \pi$, $\beta = 0$. With this hypothesis, eq. (41) gives a satisfactory representation of the Qbit 0 as $|0\rangle$, but it implies that Qbit 1 would be $-|0\rangle = -0$.

We fix up things be writing,

$$|\psi\rangle = e^{i\gamma}\left[\cos\frac{\alpha}{2}|0\rangle + e^{i\beta}\sin\frac{\alpha}{2}|1\rangle\right], \tag{42}$$

and identifying angles α and β with the polar and azimuthal angles of a unit 3-vector in an abstract 3-space (sometimes called the Bloch sphere). That is, we associate the Qbit $|\psi\rangle$ with the unit 3-vector whose components are, $\psi_x = \sin\alpha\cos\beta$, $\psi_y = \sin\alpha\sin\beta$ and $\psi_z = \cos\alpha$. Now, the associations.

$$0 \leftrightarrow (\alpha = 0, \beta = 0) \leftrightarrow |0\rangle, \qquad 1 \leftrightarrow (\alpha = \pi, \beta = 0) \leftrightarrow |1\rangle, \tag{43}$$

given by eq. (42) are satisfactory.

(a) **Pauli Spin Matrices and Rotations**

Verify that the Pauli spin matrices $\boldsymbol{\sigma}_x$, $\boldsymbol{\sigma}_y$ and $\boldsymbol{\sigma}_z$ transform the general single-Qbit state (42) to states whose "angles" α' and β' are the result of rotations $\mathbf{R}_x(180°)$, $\mathbf{R}_y(180°)$ and $\mathbf{R}_z(180°)$ by 180° about the x, y, and z axes, respectively, of the 3-vector on the Bloch sphere that is associated with the Qbit $|0\rangle$. Your argument need not be based on the formal rotation matrices discussed in part (b); it suffices to consider the effect of rotations on the angles α and β, in comparison to the effect of the spin matrices on the state (42).

(b) **Rotation Matrices**

A general rotation in 3-space is characterized by 3 angles. We follow Euler in naming these angles as in the figure on the next page.[20] The rotation takes the axis (x, y, z) into the axes (x', y', z') in 3 steps:

[18] The "Three R's" of quantum computation are Reading, (w)Riting and Rotating.

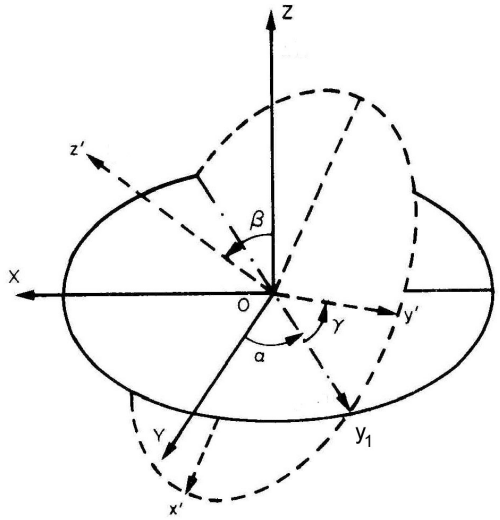
[19] Mathematically, one can imagine a Qbit (2-spinor) as a kind of pole + flag, where the overall phase describes the azimuthal angle of the flag. See, for example,
http://kirkmcd.princeton.edu/examples/QED/steane_1312.3824.pdf

[20] From sec. 58 of Landau and Lifshitz, *Quantum Mechanics*.

i. A rotation by angle α about the z-axis, which brings the y-axis to the y_1 axis.

ii. A rotation by angle β about the y_1-axis, which brings the z-axis to the z'-axis.

iii. A rotation by angle γ about the z'-axis, which brings the y_1-axis to the y'-axis (and the x-axis to the x'-axis).



The 2×2 unitary matrix that corresponds to this rotation is,

$$\begin{aligned}
\mathsf{R}(\alpha,\beta,\gamma) &= \begin{pmatrix} \cos\frac{\beta}{2}e^{i(\alpha+\gamma)/2} & \sin\frac{\beta}{2}e^{i(-\alpha+\gamma)/2} \\ -\sin\frac{\beta}{2}e^{i(\alpha-\gamma)/2} & \cos\frac{\beta}{2}e^{-i(\alpha+\gamma)/2} \end{pmatrix} \\
&= \begin{pmatrix} e^{i\gamma/2} & 0 \\ 0 & e^{-i\gamma/2} \end{pmatrix} \begin{pmatrix} \cos\frac{\beta}{2} & \sin\frac{\beta}{2} \\ -\sin\frac{\beta}{2} & \cos\frac{\beta}{2} \end{pmatrix} \begin{pmatrix} e^{i\alpha/2} & 0 \\ 0 & e^{-i\alpha/2} \end{pmatrix} \\
&= \mathsf{R}_{z'}(\gamma)\mathsf{R}_{y_1}(\beta)\mathsf{R}_z(\alpha), \qquad (44)
\end{aligned}$$

where the decomposition into the product of 3 rotation matrices[21] follows from the particular rules,

$$\mathsf{R}_x(\phi) = \begin{pmatrix} \cos\frac{\phi}{2} & i\sin\frac{\phi}{2} \\ i\sin\frac{\phi}{2} & \cos\frac{\phi}{2} \end{pmatrix}, \qquad (45)$$

$$\mathsf{R}_y(\phi) = \begin{pmatrix} \cos\frac{\phi}{2} & \sin\frac{\phi}{2} \\ -\sin\frac{\phi}{2} & \cos\frac{\phi}{2} \end{pmatrix}, \qquad (46)$$

$$\mathsf{R}_z(\phi) = \begin{pmatrix} e^{i\phi/2} & 0 \\ 0 & e^{-i\phi/2} \end{pmatrix}. \qquad (47)$$

Convince yourself that the combined rotation (44) could also be achieved if first a rotation is made by angle γ about the z axis, then a rotation is made by angle β about the original y axis, and finally a rotation is made by angle α about the original z axis.

[21] The order of operations is that the rightmost rotation in eq. (44) is to be performed first.

There is unfortunately little consistency among various authors as to the conventions used to describe rotations. I will now adopt the notation of Barenco et al.,[22] *who appear to write eq. (44) simply as,*

$$\mathsf{R}(\alpha,\beta,\gamma) = \mathsf{R}_z(\gamma)\mathsf{R}_y(\beta)\mathsf{R}_z(\alpha). \tag{48}$$

Occasionally (for example in part (c)) we will need to remember that in eq. (48) the axes of the second and third rotations are the results of the previous rotation(s). Also, Nielsen and Chuang consider rotations to be by the negative of the angles that I do. Thus, the operator that I call $\mathsf{R}_x(\phi)$ *is called* $\mathsf{R}_x(-\phi)$ *by them.*

Rather than rotating the coordinate axes, we may wish to rotate vectors in Bloch space by an angle ϕ about a given axis $\hat{\mathbf{u}}$, while leaving the coordinate axes fixed. The operator,

$$\mathsf{R}_u(-\phi) = e^{-i\frac{\phi}{2}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}}, \tag{49}$$

performs this type of rotation.

Note that according to eqs. (45)-(47),

$$\boldsymbol{\sigma}_x = -i\mathsf{R}_x(180^\circ), \qquad \boldsymbol{\sigma}_y = -i\mathsf{R}_y(180^\circ), \qquad \boldsymbol{\sigma}_z = -i\mathsf{R}_z(180^\circ), \tag{50}$$

and also,

$$\boldsymbol{\sigma}_x = i\mathsf{R}_x(-180^\circ), \qquad \boldsymbol{\sigma}_y = i\mathsf{R}_y(-180^\circ), \qquad \boldsymbol{\sigma}_z = i\mathsf{R}_z(-180^\circ), \tag{51}$$

so that the Pauli spin matrices are equivalent to the formal matrices for 180° rotations only up to a phase factor i.

Show that a more systematic relation between the Pauli spin matrices and the rotation matrices is that eqs. (45)-(47) can be written as,

$$\mathsf{R}_u(\phi) = e^{i\frac{\phi}{2}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}}, \tag{52}$$

which describes a rotation of the coordinate axes in Bloch space by angle ϕ about the $\hat{\mathbf{u}}$ axis (in a right-handed convention).

Equations (51) and (52) can be combined to write,[23]

$$\boldsymbol{\sigma}_j = ie^{-i\frac{\pi}{2}\boldsymbol{\sigma}_j} = e^{i\frac{\pi}{2}}e^{-i\frac{\pi}{2}\boldsymbol{\sigma}_j}, \tag{53}$$

which permits us to define an arbitrary power of a Pauli matrix as[24]

$$\boldsymbol{\sigma}_j^\alpha = (e^{i\frac{\pi}{2}}e^{-i\frac{\pi}{2}\boldsymbol{\sigma}_j})^\alpha = e^{i\frac{\pi\alpha}{2}}e^{-i\frac{\pi\alpha}{2}\boldsymbol{\sigma}_j}. \tag{54}$$

Show that the phase gate with phase angle $\phi = \pi\alpha$ can be written as,

$$\begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi\alpha} \end{pmatrix} = \boldsymbol{\sigma}_z^\alpha = \mathsf{Z}^\alpha. \tag{55}$$

[22] http://kirkmcd.princeton.edu/examples/QM/barenco_pra_52_3457_95.pdf

[23] This also follows from eq. (38).

[24] Since the Pauli matrices are their own inverses, it is tempting to write $\boldsymbol{\sigma}_j^\alpha = e^{-i\frac{\pi\alpha}{2}}e^{i\frac{\pi\alpha}{2}\boldsymbol{\sigma}_j} = \boldsymbol{\sigma}_j^{-\alpha}$. However, this relation is not true in general, because fractional powers of the Pauli matrices correspond to rotations by angles that are a fraction of π, so the direction of the rotation matters.

(c) **More Square Roots of NOT**

Use the facts about rotation matrices presented in part (b) to construct additional representations of the NOT and $\sqrt{\mathsf{NOT}}$ operators that act on Qbit $|\psi\rangle = \psi_0|0\rangle + \psi_1|1\rangle$, supposing that the overall phase of $|\psi\rangle$ is irrelevant. Recall that the basic intent of the NOT operator is to flip the bits 0 and 1.

(d) **Rotation Matrices and the General Form (37)**

From eq. (44) we see that the determinant of $\mathsf{R}(\alpha,\beta,\gamma)$ is unity, so that there is a one-to-one correspondence between rotation matrices and 2×2 special unitary operators. Recalling the general form (37), we also infer that a general 2×2 unitary operator U can be written as,

$$\mathsf{U} = e^{i\delta}\mathsf{R}(\alpha,\beta,\gamma) = e^{i\delta}\mathsf{R}_z(\gamma)\mathsf{R}_y(\beta)\mathsf{R}_z(\alpha). \tag{56}$$

What is the relation between parameters α, β and γ of eq. (56) and the parameters θ and $\hat{\mathbf{u}}$ of eq. (37)?

(e) **Double NOT**

Among the many identities involving rotation matrices, demonstrate that,

$$\boldsymbol{\sigma}_x\mathsf{U}\boldsymbol{\sigma}_x = \mathsf{X}e^{i\delta}\mathsf{R}(\alpha,\beta,\gamma)\mathsf{X} = e^{i\delta}\mathsf{R}(-\alpha,-\beta,-\gamma), \tag{57}$$

which will be used later in the course.

(f) **Basis Change**

The result of a rotation R in 3-space can be thought of as a change of basis from the orthonormal triad $(\hat{\mathbf{x}},\hat{\mathbf{y}},\hat{\mathbf{z}})$ to a new orthonormal triad $(\hat{\mathbf{x}}',\hat{\mathbf{y}}',\hat{\mathbf{z}}') = \mathsf{R}(\hat{\mathbf{x}},\hat{\mathbf{y}},\hat{\mathbf{z}})$. Similarly, a 2×2 unitary matrix,

$$\mathsf{U} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{58}$$

can be thought of as causing a change from the basis of orthonormal states $(|0\rangle,|1\rangle)$ to a new orthonormal basis,

$$|\psi\rangle = \mathsf{U}|0\rangle = a|0\rangle + c|1\rangle, \qquad |\phi\rangle = \mathsf{U}|1\rangle = b|0\rangle + d|1\rangle. \tag{59}$$

Verify that the states $|\psi\rangle$ and $|\phi\rangle$ are indeed orthonormal.

(g) **Hadamard Transformation**

It will be of interest on occasion to switch from the basis $[|0\rangle,|1\rangle]$ (or simply the $[0,1]$ basis) to the basis $[(|0\rangle+|1\rangle)/\sqrt{2},(|0\rangle-|1\rangle)/\sqrt{2}]$ (which we will often call the $[|+\rangle,|-\rangle]$ or the $[+,-]$ basis). The unitary matrix that performs this operation is called the Hadamard transformation (or gate):

$$\mathsf{H} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{\mathsf{X}+\mathsf{Z}}{\sqrt{2}} = \frac{\boldsymbol{\sigma}_x+\boldsymbol{\sigma}_z}{\sqrt{2}}\,. \tag{60}$$

We see that H is self-adjoint, so that $\mathsf{H}^2 = \mathbf{I}$; a second application of the Hadamard transformation brings us back to the original basis.

Express the Hadamard transformation in the general forms (37) and (52).

It is sometimes useful to write the effect of the Hadamard transformation on a basis state $|j\rangle$, where $j = 0$ or 1, as,

$$\mathsf{H}|j\rangle = \frac{|0\rangle + (-1)^j|1\rangle}{\sqrt{2}} = \frac{|0\rangle + e^{i\pi j}|1\rangle}{\sqrt{2}}\,. \tag{61}$$

(h) Show that the Pauli spin matrices and the Hadamard transformation are related by identities such as,

$$\boldsymbol{\sigma}_x^\alpha = \mathsf{H}\boldsymbol{\sigma}_z^\alpha\mathsf{H}, \tag{62}$$
$$\boldsymbol{\sigma}_y^\alpha = \boldsymbol{\sigma}_z^{1/2}\boldsymbol{\sigma}_x^\alpha\boldsymbol{\sigma}_z^{-1/2}, \tag{63}$$
$$\mathsf{H}^\alpha = \boldsymbol{\sigma}_y^{1/4}\boldsymbol{\sigma}_z^\alpha\boldsymbol{\sigma}_y^{-1/4}, \tag{64}$$

which implies that all of $\boldsymbol{\sigma}_x^\alpha$, $\boldsymbol{\sigma}_y^\alpha$, $\boldsymbol{\sigma}_z^\alpha$ and H^α can be constructed from only H and $\boldsymbol{\sigma}_z^\alpha$.

(i) For use later in the course, it will be useful to know a relation between the rotation $\mathsf{R}_{\hat{\mathbf{u}}}(\theta) = e^{i\frac{\theta}{2}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}}$ by an arbitrary angle θ about an axis $\hat{\mathbf{u}}$ and a rotation $\mathsf{R}_{\hat{\mathbf{v}}}(\theta) = e^{i\frac{\theta}{2}\hat{\mathbf{v}}\cdot\boldsymbol{\sigma}}$ by the same angle about an axis $\hat{\mathbf{v}}$ that is perpendicular to $\hat{\mathbf{u}}$. Show that,

$$\mathsf{H}^{-1/2}\; e^{i\frac{\theta}{2}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}}\; \mathsf{H}^{1/2} \tag{65}$$

corresponds to a rotation by angle θ about an axis $\hat{\mathbf{v}}$ (whose components you are to find in terms of those of $\hat{\mathbf{u}}$), and deduce a condition on $\hat{\mathbf{u}}$ such that $\hat{\mathbf{u}}\cdot\hat{\mathbf{v}} = 0$.

(j) **Basis States for the Hadamard Transformation**

From the definition (60) of the Hadamard transformation H with respect to the $[|0\rangle, |1\rangle]$ basis, we have that,

$$\mathsf{H}|0\rangle = |+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}\,, \qquad \text{and} \qquad \mathsf{H}|1\rangle = |-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}\,. \tag{66}$$

From the fact that $\mathsf{H}^2 = \mathbf{I}$ we then find,

$$\mathsf{H}|+\rangle = |0\rangle = \frac{|+\rangle + |-\rangle}{\sqrt{2}}\,, \qquad \text{and} \qquad \mathsf{H}|-\rangle = |1\rangle = \frac{|+\rangle - |-\rangle}{\sqrt{2}}\,. \tag{67}$$

Show that the set of all orthonormal bases $[|\psi\rangle, |\phi\rangle]$ for which,

$$\mathsf{H}|\psi\rangle = \frac{|\psi\rangle + |\phi\rangle}{\sqrt{2}}\,, \qquad \text{and} \qquad \mathsf{H}|\phi\rangle = \frac{|\psi\rangle - |\phi\rangle}{\sqrt{2}}\,. \tag{68}$$

describe a cone on the unit Bloch sphere whose axis is at 45° to the x and z axes in the x-z plane.

If we define the basis state $|\psi\rangle$ according to eq. (42),

$$|\psi\rangle = e^{i\gamma}\left[\cos\frac{\alpha}{2}|0\rangle + e^{i\beta}\sin\frac{\alpha}{2}|1\rangle\right], \tag{42}$$

where (α, β) are the polar and azimuthal angles of the vector $|\psi\rangle$ on the Bloch sphere, then the orthogonal vector $|\phi\rangle$ has polar angle $\pi - \alpha$ and azimuthal angle $\beta + \pi$. Thus, we can write,

$$|\phi\rangle = e^{i\delta}\left[\sin\frac{\alpha}{2}|0\rangle - e^{i\beta}\cos\frac{\alpha}{2}|1\rangle\right], \tag{69}$$

where only the difference between phases γ and δ has physical significance. *Hence, you can, for example, set phase γ to zero in your solution without loss of generality, but you cannot necessarily set both γ and δ to zero.*

5. **Measurements**

In a computational context, the most common kind of measurement we will make on a Qbit $|\psi\rangle = \psi_0|0\rangle + \psi_1|1\rangle$ is a determination of whether that Qbit is a $|0\rangle$ or a $|1\rangle$.

For example, if we wish to find out whether $|\psi\rangle$ is $|0\rangle$ we could apply the operator,

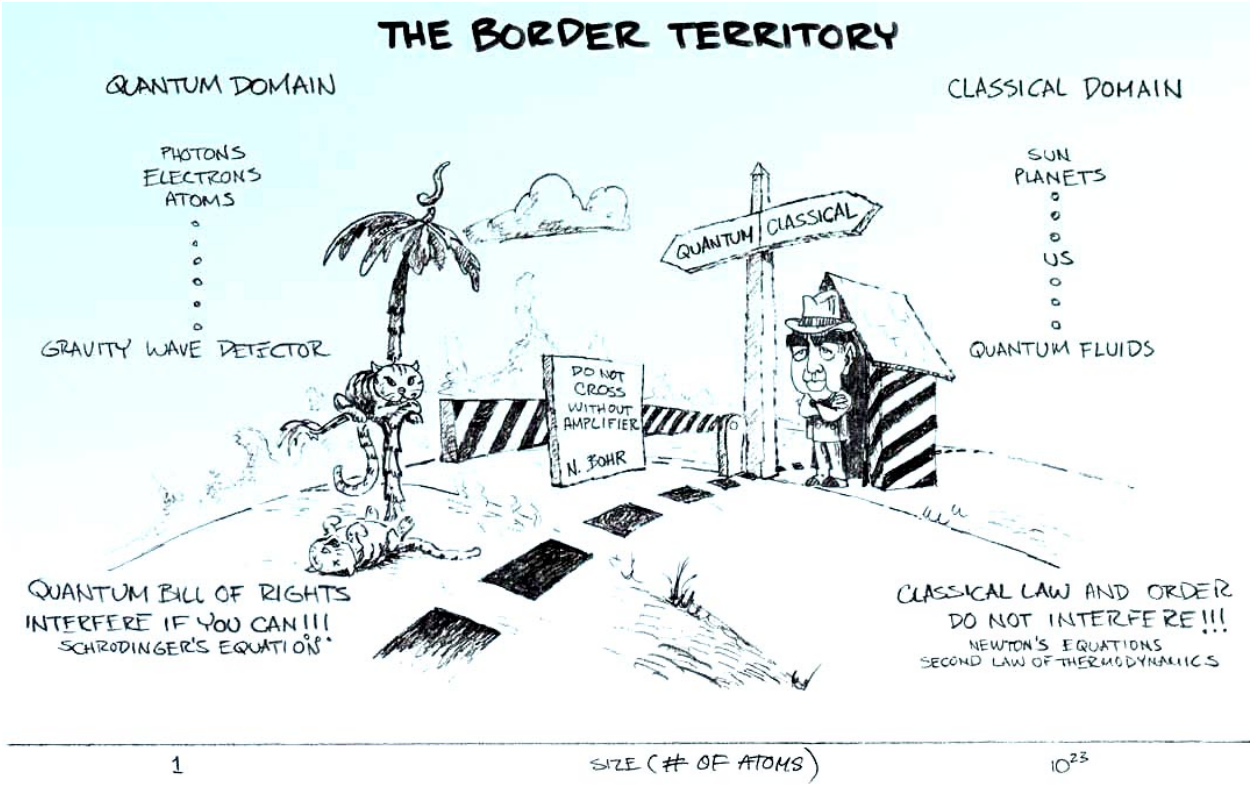
$$\mathsf{P}_0 = |0\rangle\langle 0| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \tag{70}$$

to it, with the result,

$$\mathsf{P}_0|\psi\rangle = \psi_0|0\rangle. \tag{71}$$

What can this mean in practice?

That the quantum state $|\psi\rangle$ can be subject to an operator P_0 implies that there is more to the universe than state $|\psi\rangle$ itself. The universe must contain additional entities that provide the physical implementation of the operator P_0, as well as some mechanism for recording the result of that operation. We do not wish to embark now on the lengthy detour to explore the full meaning of the previous sentence. We note that Bohr and Heisenberg considered that the system which makes the measurement must have a "classical" component, but they remained somewhat vague as to the nature of the divide between the classical and quantum parts of the system. I too will be somewhat vague now, but we will return to this topic in prob. 20.



We presume that the operator P_0 can be implemented in such a way that its effect on the entire system is either
1. The state $|\psi\rangle$ is projected into state $|0\rangle$ and the larger system is left with a record that $|0\rangle$ was found to be (or projected into, if you prefer) the state $|0\rangle$,
or,
2. The larger system is left with no record that state $|0\rangle$ was found to be $|0\rangle$, and the state $|\psi\rangle$ is left as $|\psi'\rangle = (|\psi\rangle - \psi_0|0\rangle)/\langle\psi'|\psi'\rangle$, which is simply $|\psi'\rangle = |1\rangle$ in the present example.
If we had many copies of $|\psi\rangle$ on which to make measurements, the probability that state $|\psi\rangle$ was found to be $|0\rangle$ would be $|\psi_0|^2 = \langle\psi|\mathsf{P}_0^\dagger\mathsf{P}_0|\psi\rangle$.

Similarly, we associate operator,

$$\mathsf{P}_1 = |1\rangle\langle 1| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \tag{72}$$

with a measurement of to what extent state $|\psi\rangle$ is state $|1\rangle$.

Neither the projection operator P_0 nor P_1 represents a measurement by itself. Rather, a measurement should provide us with information as to what extent state $|\psi\rangle$ is in either of its basis states $|0\rangle$ and $|1\rangle$. We express this more formally by defining a measurement M to determine the "value" of Qbit $|\psi\rangle$ by associating the outcome with the ("classical") quantity 0 if $|\psi\rangle$ is found to be $|0\rangle$ and with 1 if $|\psi\rangle$ is found to be $|1\rangle$. We will write the measurement M as,

$$\mathsf{M} = 0 \cdot \mathsf{P}_0 + 1 \cdot \mathsf{P}_1, \tag{73}$$

where we place the symbol $\cdot$ between the ("classical") value v_j and its corresponding projection operator $|j\rangle\langle j|$ to remind us that the effect of the operation $v_j \cdot |j\rangle\langle j|$ is to project state $|\psi\rangle$ into state $|j\rangle$ while returning value v_j for the measurement.

Both operators P_0 and P_1 are hermitian, since $\mathsf{P}_0^\dagger = \mathsf{P}_0$ and $\mathsf{P}_0^\dagger = \mathsf{P}_1$. This illustrates a general quantum rule: measurable quantities (observables) are associated with hermitian operators.

We now generalize to states other than a single Qbit, but where a quantum state $|\psi\rangle$ can be represented via a set of orthonormal basis states $\{|\phi_j\rangle\}$,

$$|\psi\rangle = \sum_j \psi_j |j\rangle, \qquad \text{where} \qquad \langle j|k\rangle = \delta_{jk}. \tag{74}$$

We will only consider so-called projective measurements for which each basis state $|j\rangle$ is an eigenvector of a hermitian operator M_j with eigenvalue m_j:

$$\mathsf{M}_j|j\rangle = m_j \cdot |j\rangle. \tag{75}$$

The hermitian operator,

$$\mathsf{M} = \sum_j \mathsf{M}_j, \tag{76}$$

is associated with a measurement to determine the value of the variable m whose possible values are the set $\{m_j\}$.[25] Comparing with the discussion for the projection operators (70) and (72), we see that the jth measurement operator, M_j, is simply related to the projection operator P_j, defined by,

$$\mathsf{P}_j = |j\rangle\langle j|, \qquad \text{for which} \qquad \mathsf{P}_j^\dagger = \mathsf{P}_j^2 = \mathsf{P}_j. \tag{77}$$

That is,

$$\mathsf{M}_j = m_j \cdot \mathsf{P}_j = m_j \cdot |j\rangle\langle j|. \tag{78}$$

Hence, in the $[|j\rangle]$ basis, the operator,

$$\mathsf{M} = \sum_j \mathsf{M}_j = \sum_j m_j \cdot \mathsf{P}_j = \sum_j m_j \cdot |j\rangle\langle j|, \tag{79}$$

is diagonal, when expressed as a matrix.

The hermitian measurement operator (79) projects the state $|\psi\rangle$ onto one of the basis states $|j\rangle$, and returns the value m_j in that case. The probability that the result of the projection is state $|j\rangle$ is,

$$P_j = \langle\psi|\mathsf{P}_j^\dagger\mathsf{P}_j|\psi\rangle = \langle\psi|\mathsf{P}_j|\psi\rangle = \langle\psi|j\rangle\langle j|\psi\rangle = |\langle j|\psi\rangle|^2 . \tag{80}$$

The total probability of finding some result of the measurement is, of course, unity:

$$1 = \sum_j P_j = \sum_j \langle\psi|\mathsf{P}_j^\dagger\mathsf{P}_j|\psi\rangle = \sum_j |\langle j|\psi\rangle|^2 . \tag{81}$$

A measurement $\mathsf{M}|\psi\rangle$ can only return a single value, say m_j, in which case the final state of $|\psi\rangle$ (the value m_j is recorded elsewhere in the measuring apparatus) is,

$$\frac{\mathsf{P}_j|\psi\rangle}{\sqrt{\langle\psi|\mathsf{P}_j^\dagger\mathsf{P}_j|\psi\rangle}} = \frac{\mathsf{P}_j|\psi\rangle}{|\langle j|\psi\rangle|}, \tag{82}$$

rather than $\mathsf{P}_j|\psi\rangle$. A measurement is a physical process, and so must also be associated with a transformation that preserves the normalization of a state, as does the form (82).[26]

The probable value (or expectation value) of variable m for state $|\psi\rangle$ is thus,

$$\langle m\rangle = \sum_j m_j P_j = \sum_j m_j\langle\psi|\mathsf{P}_j|\psi\rangle = \langle\psi|\sum_j m_j \cdot \mathsf{P}_j|\psi\rangle = \langle\psi|\mathsf{M}|\psi\rangle. \tag{83}$$

This important relation is still true when the state $|\psi\rangle$ is expressed in some other basis than the $[|j\rangle]$ basis, in which the hermitian matrix M is no longer diagonal.

[25] If the index j in eq. (76) does not run over a complete set of basis states the operator M represents only a partial measurement. A common example of this is a filter, such as a polarizing filter than can absorb (measure) photons of one polarization while transmitting photons of the orthogonal polarization.

[26] An aspect of the "measurement problem" of quantum mechanics is that the transformation of state $|\psi\rangle$ to state (82) is not reversible, and so is not expressible as a unitary transformation (even though total probability is preserved).

Nonorthogonal Quantum States Cannot Be Reliably Distinguished.

Another important fact is that two quantum states that are nonorthogonal (but not identical) cannot be reliably distinguished. That is, if states $|\psi_1\rangle$ and $|\psi_2\rangle$ are such that $\langle\psi_1|\psi_2\rangle \neq 0$ or 1, then measurements with the projection operator $\mathsf{P}_1 = |\psi_1\rangle\langle\psi_1|$ return the value 1 when applied to state $|\psi_1\rangle$, but when applied to state $|\psi_2\rangle$ the value 1 is returned with probability $|\langle\psi_1|\psi_2\rangle|^2$ and the value 0 with probability $1-|\langle\psi_1|\psi_2\rangle|^2$. Likewise, ambiguous results would be obtained by use of the projection operator $\mathsf{P}_2 = |\psi_2\rangle\langle\psi_2|$ on the state $|\psi_1\rangle$.

Measurement Requires Entanglement; Measurement Takes Time.

We follow an argument due to von Neumann[27] to get a sense of how a particular quantum system, say, one or more Qbits, might interact with a larger system to implement a measurement described by hermitian operator M that acts on the particular system.

We suppose that there exists an intermediary object, which we will call the pointer that can interact with the particular quantum system, and which is also very heavy so that the position of the pointer is "well defined." By the latter, we mean that the position of the pointer can be determined to sufficient accuracy, as defined below, by apparatus whose behavior is "classical" enough that we can leave the apparatus out of the quantum part of the analysis.

The goal is to establish a quantum correlation between the measurable property of the particular quantum state and the position of the pointer, and then to use a "classical" measurement of the position of the pointer to infer the result of the quantum correlation/measurement. Thus, the argument of von Neumann straddles the "quantum border" shown in the cartoon on p. 17.

To describe von Neumann's argument we need to know something about the time evolution of a quantum system. Since the total probability of the quantum system to be in some state remains constant over time, the time evolution of a quantum state $|\Psi(t)\rangle$ is described by a unitary operator,

$$|\Psi(t')\rangle = \mathsf{U}(t,t')|\Psi(t)\rangle. \tag{84}$$

Over a short time interval, $t' - t = \delta t$, the unitary operator U cannot differ much from the identity operator,

$$\mathsf{U}(t, t+\delta t) \approx \mathbf{I} + \mathsf{u}(t)\delta t. \tag{85}$$

That is,

$$|\Psi(t+\delta t)\rangle = \mathsf{U}(t,t+\delta t)|\Psi(t)\rangle \approx |\Psi(t)\rangle + \mathsf{u}(t)|\Psi(t)\rangle\delta t, \tag{86}$$

which implies that,

$$\frac{\partial|\Psi\rangle}{\partial t} = \mathsf{u}|\Psi\rangle. \tag{87}$$

[27] The argument we give is based on the last few pages of *Mathematical Foundations of Quantum Mechanics*, J. von Neumann, (Princeton U. Press, 1955); the German original was written in 1932, three years before Schrödinger coined the term "entanglement." See also sec. 3.1.1 of Preskill's Lectures.
http://kirkmcd.princeton.edu/examples/QM/vonneumann_55.pdf

The famous insight of Schrödinger is that if we write,

$$\mathsf{u} = -\frac{i}{\hbar}\mathcal{H} = -i\mathsf{h}, \tag{88}$$

then the operator $\mathcal{H} = \hbar\mathsf{h}$ is not the Hadamard transformation but is related to the Hamiltonian of the system in a well-defined manner. Thus, eq. (87) becomes Schrödinger's equation,

$$i\frac{d|\Psi\rangle}{dt} = \mathsf{h}|\Psi\rangle. \tag{89}$$

Since the operator $\mathsf{U} \approx \mathbf{I} - i\mathsf{h}\,\delta t$ *is unitary,* $\mathsf{U}^{-1} = \mathsf{U}^\dagger \approx \mathbf{I} + i\mathsf{h}^\dagger\delta t$*. Then,*

$$1 = \mathsf{U}^{-1}\mathsf{U} \approx (\mathbf{I} + i\mathsf{h}^\dagger\delta t)(\mathbf{I} - i\mathsf{h}\delta t) \approx \mathbf{I} + i(\mathsf{h}^\dagger - \mathsf{h})\delta t, \tag{90}$$

so that we must have $\mathsf{h}^\dagger = \mathsf{h}$*, i.e., the Hamiltonian operator is hermitian.*

Returning to the case of a particular quantum system plus the pointer, we take the Hamiltonian of the combined system to be of the form,

$$\mathsf{h} = \mathsf{h}_0 + \frac{\mathsf{p}^2}{2m} + \lambda\mathsf{M}\mathsf{p} \approx \lambda\mathsf{M}\mathsf{p}, \tag{91}$$

where h_0 is the Hamiltonian of the particular system when in isolation, $\mathsf{p} = -i\partial/\partial x$ is the momentum operator of the pointer (which can move only in the x direction), m is the (large) mass of the pointer, λ is a coupling constant, and M is the measurement operator that applies to the particular quantum system. The approximate form of the Hamiltonian follows on noting that mass m is large, and that during the measurement the effect of the interaction term $\lambda\mathsf{M}\mathsf{p}$ is much larger than that of isolated Hamiltonian h_0 (otherwise the measurement could not produce a crisp result[28]).

The state of the particular system to be measured is,

$$|\psi\rangle = \sum_j \psi_j|j\rangle, \tag{92}$$

and the initial state of the pointer is $|\phi(x)\rangle$, which is a Gaussian wave packet centered on, say, $x = 0$, normalized such that $\int |\phi(x)|^2\,dx = 1$. Since the pointer particle is heavy, its wave packet $|\phi(x)\rangle$ is narrow (but not so narrow that the wave packet spreads significantly during the measurement). The initial state of the combined system is the direct product,

$$|\Psi(0)\rangle = |\psi\rangle \otimes |\phi(x)\rangle = \left(\sum_j \psi_j|j\rangle\right) \otimes |\phi(x)\rangle. \tag{93}$$

The basis $[|j\rangle]$ for the particular system has been chosen so that the each basis state $|j\rangle$ has a well-defined value m_j of the measurement. That is, the measurement operator has the projective form,

$$\mathsf{M} = \sum_j m_j \cdot |j\rangle\langle j|. \tag{94}$$

[28] See, for example, http://kirkmcd.princeton.edu/examples/QM/peres_prd_32_1968_85.pdf

The Hamiltonian $\mathsf{h} \approx \lambda \mathsf{Mp}$ is time independent, so Schrödinger's equation (89) has the formal solution,

$$|\Psi(t)\rangle = e^{-i\mathsf{h}t}|\Psi(0)\rangle. \tag{95}$$

Now,

$$\begin{aligned} e^{-i\mathsf{h}t} = \sum_{n=0}^{\infty} \frac{(-i\mathsf{h}t)^n}{n!} &= \sum_{n=0}^{\infty} \frac{1}{n!} \left[-i\lambda \sum_j m_j \cdot |j\rangle\langle j| \left(-i\frac{\partial}{\partial x} \right) t \right]^n \\ &= \sum_j |j\rangle\langle j| \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\lambda m_j t \frac{\partial}{\partial x} \right)^n, \end{aligned} \tag{96}$$

recalling that $\langle j|k\rangle = \delta_{jk}$, so that the lengthy products of bras and kets all collapse back down to the projections $|j\rangle\langle j|$. Inserting eqs. (93) and (96) into (95), we obtain,

$$\begin{aligned} |\Psi(t)\rangle &= \sum_j |j\rangle\langle j| \sum_k \psi_k |k\rangle \otimes \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\lambda m_j t \frac{\partial}{\partial x} \right)^n |\phi(x)\rangle \\ &= \sum_j \psi_j |j\rangle \otimes |\phi(x - \lambda m_j t)\rangle, \end{aligned} \tag{97}$$

noting that the Taylor expansion of $\phi(x - x_0)$ is,

$$\phi(x - x_0) = \sum_{n=0}^{\infty} \frac{1}{n!} \left(-x_0 \frac{\partial}{\partial x} \right)^n \phi(x). \tag{98}$$

The initial direct product state (93) has evolved into the entangled state (97) during the course of the measurement.

The supposition is that the position of the pointer at time t can be determined well enough to distinguish the j locations $\lambda m_j t$ from one another. This is more plausible for larger t: (accurate) measurements take time! If the pointer is found at (or near) position $\lambda m_j t$, the particular system $|\psi\rangle$ must be in state $|j\rangle$ and the value of the measurement is m_j. The probability that this is the outcome of the measurement is, of course, $\left|\psi_j\right|^2$ since $\int |\phi(x - \lambda m_j t)|\, dx = 1$.

This argument does a nice job of explaining how to entangle the state $|\psi\rangle = \sum \psi_j |j\rangle$ with a pointer such that different positions of the pointer are correlated with different basis states $|j\rangle$. However, it does not explain how the observation of the position of the pointer to be, say, $\lambda m_j t$ "collapses the wave function" of $|\psi\rangle$ to the basis state $|j\rangle$.[29]

Von Neumann's argument indicates that underlying every measurement process is the entanglement that bothered Einstein, Podolsky and Rosen (and Schrödinger, *etc.*) so much. This deserves further discussion, some of which will be given in prob. 20.

[29] The transformation from $|\Psi(0)\rangle$ to $|\Psi(t)\rangle$ is unitary/reversible as eq. (97) is valid for both increasing or decreasing t. The irreversible step in the measurement process is the "classical" reading of the position of the pointer at time $t_{\rm meas}$, which selects a value of $x \approx \lambda m_j t_{\rm meas}$ and leaves the system in the state,

$$|\Psi(t > t_{\rm meas})\rangle = \frac{|j\rangle \otimes |\phi(x - \lambda m_j t_{\rm meas})\rangle}{\sqrt{N}}, \tag{99}$$

where N is the number of possible positions of the pointer.

Problems.

(a) Consider the hermitian operator (that acts on a single Qbit),

$$\hat{\mathbf{v}} \cdot \boldsymbol{\sigma} = v_x \mathsf{X} + v_y \mathsf{Y} + v_z \mathsf{Z}, \tag{100}$$

where $\hat{\mathbf{v}}$ is a real unit vector, *i.e.*, $v_x^2 + v_y^2 + v_z^2 = 1$. What are the eigenvalues and eigenvectors of the operator $\hat{\mathbf{v}} \cdot \boldsymbol{\sigma}$? *Hint: recall eq. (42).*

What are the projection operators onto those eigenvectors?

What is the probability of obtaining the result $+1$ for a measurement of $\hat{\mathbf{v}} \cdot \boldsymbol{\sigma}$ on the state $|0\rangle$? What is the state after the measurement if the result $+1$ is obtained?

(b) Suppose the state $|\psi\rangle$ consists of two Qbits in the entangled form,

$$|\psi\rangle = \psi_{00}|0\rangle|0\rangle + \psi_{01}|0\rangle|1\rangle + \psi_{10}|1\rangle|0\rangle + \psi_{11}|1\rangle|1\rangle. \tag{101}$$

What is the state $|\psi'\rangle$ after a measurement to determine the value of the second bit? What is the probability that the second bit is found to have value 1 in this measurement?

Suppose we wish to measure the values of both the first and second bits. What is the appropriate measurement operator? How does this operator determine the probabilities that the two bits are $|0\rangle|0\rangle$, $|0\rangle|1\rangle$, $|1\rangle|0\rangle$ or $|1\rangle|1\rangle$?

(c) **Stern-Gerlach.** The argument of eqs. (91)-(98) can be generalized to the case of a pointer whose value is described by a coordinate q by use of a Hamiltonian that couples the system to be measured to the canonical momentum p that is conjugate to q. That is, consider $\mathsf{h} \approx \lambda \mathsf{M}\mathsf{p}$ where $\mathsf{p} = -i\partial/\partial q$.

Use this fact to describe how a Stern-Gerlach apparatus "measures" the z-component of the spin of a neutral spin-1/2 particle that is moving in the x direction through magnetic field $\mathbf{B} \approx B(z)\,\hat{\mathbf{z}}$. Recall that such an apparatus gives a "kick" to the particle in the z-direction whose sign depends on whether the spin is "up" or "down".

It suffices to display an appropriate Hamiltonian for the system.

(d) **Quantum Nondemolition Measurement.**[30] The prescription (97) for the entanglement of a state $|\psi\rangle$ with the pointer state $|\phi\rangle$ permits, in principle, the state $|\psi\rangle$ to be measured without being destroyed. Ideally, quantum measurement is a nondemolition process. However, in examples such as photons, the state to be measured is typically destroyed in the process. In such cases a nondemolition measurement could be made if the state $|\psi\rangle = \sum_j a_j|j\rangle$ were first entangled with another state $|\phi\rangle$ such that the result is $|\psi\rangle|\phi\rangle = \sum_j a_j|j\rangle|j\rangle$. Then, a destructive measurement of state $|\phi\rangle$ leaves state $|\psi\rangle$ in a known, and still existing, basis state.

Deduce the form of a symmetric, unitary 4×4 matrix U that operates on Qbits $|\psi\rangle = a_0|0\rangle + a_1|1\rangle$ and $|\phi\rangle = |0\rangle$ according to $\mathsf{U}|\psi\rangle|\phi\rangle = a_0|0\rangle|0\rangle + a_1|1\rangle|1\rangle$.

That we could not simply make a copy of the (unknown) quantum state $|\psi\rangle$ before measuring it is the topic of Prob. 6.

[30] We use the term *quantum nondemolition measurement* in a slightly different way than in which it was introduced historically. See, http://kirkmcd.princeton.edu/examples/QM/caves_rmp_52_341_80.pdf

6. Quantum Cloning, Quantum Teleportation

We have previously remarked that the amplitudes of the various eigenstates of a general quantum state $|\psi\rangle = \sum \psi_j |j\rangle$ cannot be determined by a single measurement, or even a finite set of measurements should many copies of the state be available. It might therefore be considered "obvious" that an exact copy of this quantum state cannot be made, unless the amplitudes ψ_j are known *a priori*. However, it appears that this fact was never explicitly noted prior to 1982.[31]

We demonstrate the no-cloning theorem by contradiction. We first suppose that a unitary cloning operator C exists that acts on the zero state $|0\rangle$ and an arbitrary Qbit $|\psi\rangle$ such that state $|\psi\rangle$ is unchanged while $|0\rangle$ turns into $|\psi\rangle$. Thus, C is defined by

$$\mathsf{C}|\psi\rangle|0\rangle = |\psi\rangle|\psi\rangle. \tag{102}$$

This should work for another state $|\phi\rangle$ as well:

$$\mathsf{C}|\phi\rangle|0\rangle = |\phi\rangle|\phi\rangle. \tag{103}$$

Since C is unitary, it preserves the inner product,

$$\langle\phi|\psi\rangle = \langle 0|\langle\phi|\psi\rangle|0\rangle = \langle 0|\langle\phi|\mathsf{C}^\dagger\mathsf{C}|\psi\rangle|0\rangle = \langle\phi|\langle\phi|\psi\rangle|\psi\rangle = (\langle\phi|\psi\rangle)^2. \tag{104}$$

However, eq. (104) can be true only if $\langle\phi|\psi\rangle = 0$ or 1. If $|\psi\rangle$ and $|\phi\rangle$ are not orthogonal the operator C cannot have successfully copied them both. Hence, the general cloning operator C does not exist.

We see that there is no problem copying an unknown single-bit state if it can only be $|0\rangle$ or $|1\rangle$. This is what classical copying does, which operation could, therefore, be implemented by a quantum device.

However, we see that the result of a quantum computation cannot be in the form of a general quantum state, if we are to copy it exactly for further use. The result must be expressed as one of a set of orthogonal states, in which case we could in principle measure/copy the result without altering it.[32]

Hence, some of the richness of information content of a general quantum state is inevitably lost at the end of a practical quantum computation. Nonetheless, quantum computation can still have many advantages over classical computation.

Quantum Money. S. Wiesner anticipated the no-cloning theorem to some extent in 1970,[33] when he proposed that "quantum money" could not be counterfeited if its serial number consisted of a string of bits each of whose base is randomly chosen at the "mint" to be either [0,1] or $[+,-]$. A counterfeiter who measured the bits before duplicating the "money" would, on average, be able to duplicate correctly only 50% of the bits. A "bank" must know the "mint's" choice of bases to detect the counterfeit.

[31] http://kirkmcd.princeton.edu/examples/QM/wootters_nature_299_802_82.pdf
http://kirkmcd.princeton.edu/examples/QM/dieks_pl_a92_271_82.pdf
http://kirkmcd.princeton.edu/examples/QM/milonni_pl_a92_321_82.pdf

[32] Compare with the lesson of Maxwell's Demon (prob. 2): to "know" a state means that we can copy it.

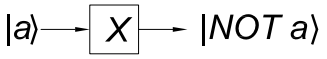
[33] http://kirkmcd.princeton.edu/examples/QM/wiesner_70.pdf

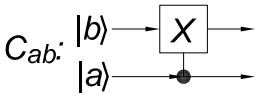
(a) Copying of a classical bit $|x\rangle$ onto a second Cbit $|y\rangle$ whose initial state is $|0\rangle$ can be accomplished by a unitary transformation C_{xy} that leaves bit $|y\rangle$ alone if $|x\rangle = 0$ but flips bit $|y\rangle$ if $|x\rangle = |1\rangle$. Express the two-bit operator C_{xy} as a 4×4 matrix that acts on the basis

$$|x\rangle|y\rangle = \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \otimes \begin{pmatrix} y_0 \\ y_1 \end{pmatrix} = \begin{pmatrix} x_0 y_0 \\ x_0 y_1 \\ x_1 y_0 \\ x_1 y_1 \end{pmatrix}, \quad \text{such that} \quad |0\rangle|1\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \textit{etc.} \tag{105}$$

The operator C_{xy} is called the Controlled-NOT because it flips the second bit, $|y\rangle$, only on the condition that the first bit, $|x\rangle$, is 1.[34]

We take this occasion to introduce a graphical representation of bit operations that will be used extensively throughout the course. The history of a bit is shown on a horizontal line, and bit operations are indicated in boxes. Thus, the figure on the left below indicates that the NOT operation, X, is applied to a Qbit $|a\rangle$. The Controlled-NOT operation C_{ab}, shown schematically in the righthand figure, involves 2 Qbits, $|a\rangle$ and $|b\rangle$. The first bit is called the control bit, whose effect is indicated by a solid circle with a vertical line connected to the box representing the NOT operation on the second bit (the target bit), meaning that target bit is subject to the NOT operation only if the control bit is in the $|1\rangle$ state.





I will use the convention that the flow of logic is from left to right in the bit-operation diagrams. Note, however, that some people use a right-to-left flow (without arrows to guide the eye), perhaps because in Dirac's bra-ket notation the input ket is on the right. Of course, since quantum bit operations are reversible, the flow of logic can be in either direction.

What is the state $|y\rangle$ obtained by applying the Controlled-NOT operation to $|x\rangle|0\rangle$ when $|x\rangle$ is the general Qbit $a|0\rangle + b|1\rangle$? For what states $|x\rangle$ is $|y\rangle$ an exact copy?

Despite the failure of the Controlled-NOT operator to copy successfully a general Qbit, it will become our favorite 2-bit operator.

Show that if we have two copies, $|x\rangle$ and $|y\rangle$, of a Cbit (perhaps obtained by use of the Controlled-NOT operation), then applying C_{xy} to the pair will delete the second bit, *i.e.*, transform it to the bit $|0\rangle$. Show, however, that the Controlled-NOT operation cannot be used to delete the second of two copies (obtaining by preparation rather than by copying!) of a general Qbit. This illustrates the no-deleting theorem.[35]

(b) **Successful Cloning Would Imply Faster-Than-Light Communication**

Show that exact cloning of an arbitrary quantum state could lead to a scheme for faster-than-light communication. In particular, consider an entangled state of

[34] In the language of classical computation, the Controlled-NOT operation is the XOR = exclusive OR operation (logic gate).

[35] http://kirkmcd.princeton.edu/examples/QM/pati_nature_404_164_00.pdf

two Qbits:

$$|\psi\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}} = \frac{|0\rangle_{\rm A}|0\rangle_{\rm B} + |1\rangle_{\rm A}|1\rangle_{\rm B}}{\sqrt{2}}, \tag{106}$$

such that after creation of this state the physical realizations of the first and second bits become separated in space. A famous example of this is the S-wave decay of an excited atom via two back-to-back photons.

Observer A (Alice) can chose to observe bit A in the basis $[|0\rangle, |1\rangle]$, or in the basis $[(|0\rangle + |1\rangle)/\sqrt{2}, (|0\rangle - |1\rangle)/\sqrt{2}] \equiv [|+\rangle, |-\rangle]$ (among the infinite set of bases for a single Qbit), and the choice of which basis she uses can be delayed; *i.e.*, Alice can wait to choose her observational basis until after the bits A and B have separated in space.

The "message" that Alice wishes to send to observer B (Bob) is her choice of basis for observation of bit A, and the result of her observation.

If the state (106) cannot be cloned, Bob can make only a single observation, and must choose a single basis (either $[|0\rangle, |1\rangle]$ or $[|+\rangle, |-\rangle]$) for this.

Describe the possible correlations between the results of measurements by Alice on bit A with the results of measurements by Bob on bit B, given that both Alice and Bob can choose to use either the $[|0\rangle, |1\rangle]$ or the $[|+\rangle, |-\rangle]$ bases. Can the measurements made by Bob, in the absence of classical communication from Alice as to the nature of her measurements, be interpreted by Bob as certain knowledge by him as to which choice of basis was made by Alice? To answer this, it is helpful to re-express state (106) in the $[|+\rangle, |-\rangle]$ basis. You may or may not find it helpful to construct and apply measurement operators to the appropriate representations of the entangled state of bits A and B.

Now suppose that Bob could clone his bit B in such a manner that the clones retained the entangled structure of state (106). Describe a set of measurements on these clones that would permit Bob to know with certainty (*i.e.*, with very high probability) what choice of basis had been made by Alice. Since Alice and Bob could, in principle, be separated by arbitrarily large distances at the times that they make their measurements, successful deduction by Bob as to Alice's choice of basis would imply an "instantaneous" communication from Alice

The above scenario was presented in 1982 by Herbert,[36] which appears to have been a strong motivation for the formulation of the no-cloning theorem.[37]

Note: You may not be able to reproduce Herbert's logic that led to his claim that faster-than-light communication is possible, as he never wrote down the form of

[36] http://kirkmcd.princeton.edu/examples/QM/herbert_fp_12_1171_82.pdf

[37] To me, the statement of the no-cloning theorem in 1982 marks the end of a somewhat nonproductive era in which numerous people used variants on the Einstein-Podolsky-Rosen argument (footnote 8, p. 12) to look for defects in quantum mechanics. A popular hope was that there might exist "hidden variables" that more completely characterized a quantum state than a description such as eq. (11). But if a more complete characterization existed, we might expect that cloning of an arbitrary quantum state would be possible. Hence, to me, the no-cloning theorem is a simple yet strong indication (much simpler than the convoluted arguments related to Bell's inequalities) that the search for hidden variables is misguided. After 1982 there has been a much healthier emphasis on uses of entanglement to explore the greater richness of quantum compared to classical phenomena.

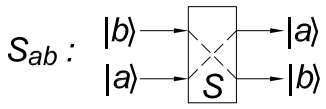
his quantum state including clones of bit B. He desired that all the clones of bit B be entangled with a single bit A.[38]

Consider the option that Bob makes a "copy" of bit B via a Controlled-NOT gate whose second input line, bit C, is initially $|0\rangle$. The resulting entangled state of bits A, B and C is as good a copy of state (106) as possible. Show, however, that measurements by Bob of bit B in the [0,1] basis and of bit C in the $[+,-]$ basis, as proposed by Herbert, do not add Bob's knowledge of bit A.[39]

(c) **Bit Swapping**

We cannot make an exact copy of an arbitrary quantum state, but we can swap two arbitrary states (that consist of the same number of Qbits).

Construct an operation S_{ab} (a 4×4 unitary matrix) that swaps two Cbits a and b. Show how S_{ab} can be implemented as a sequence of Controlled-NOT operations, noting that C_{ab} and C_{ba} are distinct.[40] Draw a bit-flow diagram for this sequence.[41]

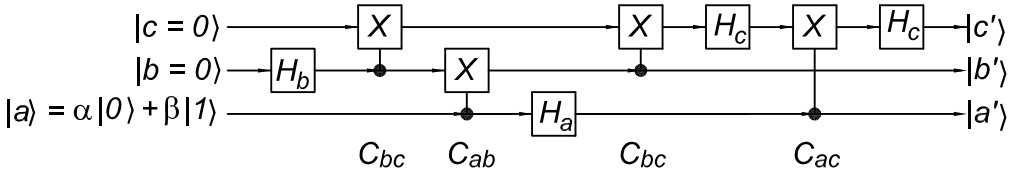


Verify that S_{ab} swaps two arbitrary Qbits $|a\rangle = \alpha|0\rangle + \beta|1\rangle$ and $|b\rangle = \gamma|0\rangle + \delta|1\rangle$.

(d) **Quantum Teleportation**

In 1993, Bennett *et al.*[42] developed a scheme for transforming an arbitrary Qbit in a very intriguing fashion that they described as quantum teleportation (although this provocative name perhaps overdramatizes the significance of this transformation). We illustrate this with a variant due to Brassard *et al.*[43]

Consider the sequence of operations on 3 Qbits sketched in the figure below. The initial Qbit $|a\rangle$ is arbitrary, while the initial Qbits $|b\rangle$ and $|c\rangle$ are both $|0\rangle$.



Show that despite the entangling effects of the Hadamard transformations, the final state is a direct product, $|a'b'c'\rangle = |a'\rangle|b'\rangle|c'\rangle$, where $|a'\rangle$ and $|b'\rangle$ are independent of $|a\rangle$, and $|c'\rangle = |a\rangle$. Thus, this process transfers the (unknown) character

[38] It is, of course, possible to prepare numerous copies of the entire state (106), each with its own set of bits A and B. There would be no correlations between these various copies, and measurements of the various copies of bits A and B would simply build up the probability distributions underlying the measurement of just one of these copies.

[39] For a historical survey of debates about faster-than-light effects in quantum theory, and a somewhat simpler paradox than Herbert's, see http://kirkmcd.princeton.edu/examples/epr/epr_colloq_81.pdf

[40] This decomposition is meant to illustrate how the Controlled-NOT operation is a logical building block for quantum computation. However, the hardware realization of a SWAP gate may be simpler than that of a single Controlled-NOT gate. Try (not for credit) making a Controlled-NOT gate out of SWAP gates.

[41] In the literature, the SWAP gate, here called S, is often symbolized as  The gate called S by Nielsen and Chuang will be called $\mathsf{Z}^{1/2} = \boldsymbol{\sigma}_z^{1/2}$ here.

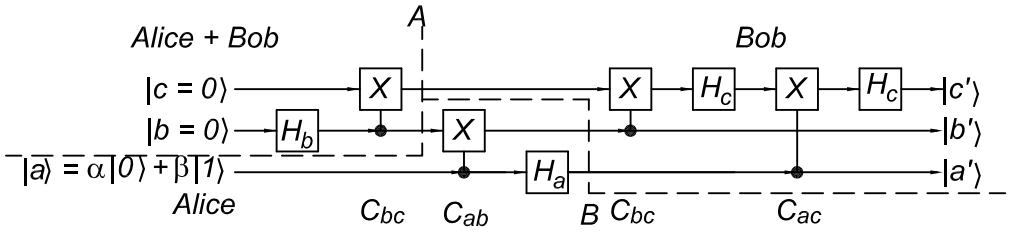
[42] http://kirkmcd.princeton.edu/examples/QM/bennett_prl_70_1895_93.pdf

[43] http://kirkmcd.princeton.edu/examples/QM/brassard_physica_d120_43_98.pdf

of state $|a\rangle$ to $|c'\rangle$. *Hint: Work out the successive effects of the 8 operations on the 3-Qbit state* $|abc\rangle$.

So far, the process shown above is just a more cumbersome way of swapping bits than you found in part (c). The subtlety is that the process may be split into 3 steps, that can in principle be carried out at different times and in widely separated places.

In the first step (corresponding to the region labeled "Alice + Bob" in the figure below) Qbits b and c are transformed from their $|0\rangle$ states into an entangled combination. At the end of this step (shown as the vertical line A), one observer (Alice) takes Qbit b with her, and the other observer (Bob) takes Qbit c with him. These two Qbits retain their quantum correlation over arbitrarily large spatial separation, so long as they are not "disturbed" or "measured".



In the second step (corresponding to the region labeled "Alice" in the figure above) Alice creates (or receives) the arbitrary Qbit a, which she further entangles with her already entangled Qbit b via the operation C-NOT$_{ab}$.

She then does something that might at first seem to destroy the quantum correlations: she measures Qbits a and b (at the position of the dashed line B above).

Alice then sends the classical results of her measurements to Bob (which transmission might be fast or slow, but surely not faster than the speed of light).

In the third step (corresponding to the region labeled "Bob" in the figure above) Bob uses the Cbits a and b that he has received from Alice, together with his previously entangled Qbit c to perform the remaining transformations.

Show algebraically that at the end, Bob's Qbit c (now in the state $|c'\rangle$) is an exact copy of the aribtrary initial Qbit a of Alice.[44] Since the original Qbit a was rendered classical by Alice's measurement, this operation is not cloning.

However, it is impressive that the full quantum state of the initial Qbit $|a\rangle$ can be reconstructed at $|c'\rangle$ via the transmission of two classical bits derived from $|a\rangle$.

This process is convoluted enough to deserve some special description, and so the term quantum teleportation has come into common use.

A generalization of the above scheme is the teleportation of a quantum gate[45] or an entire quantum computer.[46]

If Alice did not measure bits a and b at time B, then it might seem that the only

[44] For a graphical proof, see Fig. 2 of Chap. 6 of Mermin's lectures on quantum computation.

[45] http://kirkmcd.princeton.edu/examples/QM/gottesman_nature_402_390_99.pdf

[46] http://kirkmcd.princeton.edu/examples/QM/raussendorf_prl_86_5188_01.pdf
http://kirkmcd.princeton.edu/examples/QM/raussendof_pra_68_022312_03.pdf
http://kirkmcd.princeton.edu/examples/QM/nielsen_prl_93_040503_04.pdf
http://kirkmcd.princeton.edu/examples/QM/childs_quant-ph-0404132.pdf

action on bit a would be the Hadamard transformation H_a. Since $\mathsf{H}^2 = \mathbf{I}$, it might then appear that if Bob made a Hadamard transformation $\mathsf{H}_a|a'\rangle$ at the end of the process, bit a would be restored to its initial value $\alpha|0_a\rangle + \beta|1_a\rangle$? If so, we would have a scheme for exact copying of a Qbit. Could this be so?

(top, left) Richard Jozsa, William K. Wootters, Charles H. Bennett. (bottom, left) Gilles Brassard, Claude Crépeau, Asher Peres. Photo: André Berthiaume.

7. Quantum Optics

A properly prepared quantum system can exhibit interference, which adds to the richness of quantum computation. We illustrate this possibility with one or more photons in devices that contain various combinations of beam splitters, mirrors and phase-shifting plates.

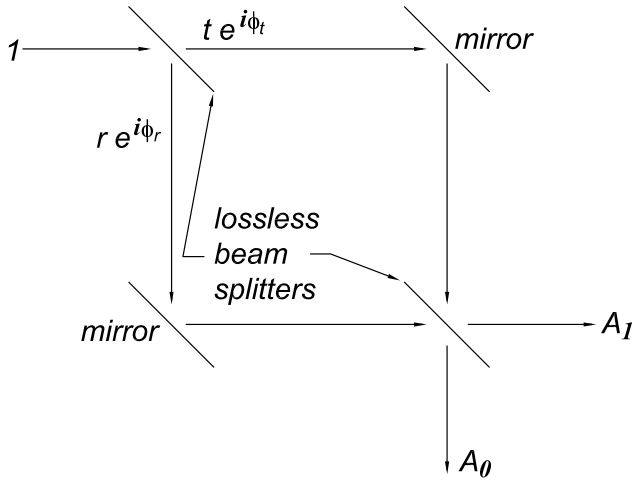
Dirac has written[47] "Each photon then interferes only with itself. Interference between two different photons never occurs." Indeed, a practical definition is that "classical" optics consists of phenomena due to the interference of photons only with themselves.[48] However, photons obey Bose statistics (which can be interpreted as a subtle kind of interference) which implies a "nonclassical" tendency for them to "bunch".

(a) Phase Shift in a Lossless Beam Splitter

Give a classical argument based on a version of a Mach-Zehnder interferometer, as shown in the figure below, that there is a 90° phase shift between the reflected and transmitted beams in a lossless, symmetric beam splitter. Then, following Dirac's dictum, your result will apply to a single photon.

A beam of light of unit amplitude is incident on the interferometer from the upper left. The reflected and transmitted amplitudes are $re^{i\phi_r}$ and $te^{i\phi_t}$, where the magnitudes r and t are real numbers. The condition of a lossless beam splitter is that,

$$r^2 + t^2 = 1. \tag{107}$$



The reflected and transmitted beams are reflected off mirrors and recombined in a second lossless beam splitter, identical to the first. The mirrors introduce an identical phase changes ϕ_m into both beams, which can be ignored in the analysis of interference between the two beams. Deduce a relation between the phase shifts ϕ_r and ϕ_t of the transmitted and reflected beams from the first splitter by noting that the amplitudes A_0 and A_1 of the beams out of the second splitter also obey $|A_0|^2 + |A_1|^2 = 1$.

[47] P.A.M. Dirac, *The Principles of Quantum Mechanics*, 4th ed. (Clarendon Press, London, 1958), p. 9, http://kirkmcd.princeton.edu/examples/QM/dirac_qm_58.pdf

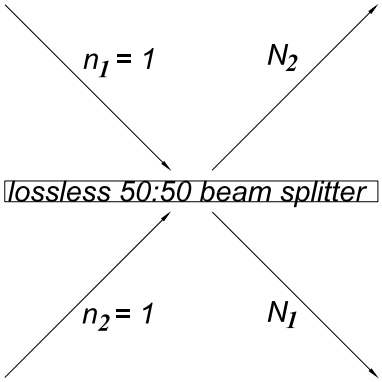
[48] For early experimental evidence of single-photon interference, see http://kirkmcd.princeton.edu/examples/QM/taylor_pcps_15_114_09.pdf

This relation will have two-fold ambiguity. A more detailed analysis[49] shows that $\phi_r = \phi_t + \pi/2$. In our applications we can ignore the overall phase of the output state, and we will use this freedom to define $\phi_t = 0$. Hence, we use $\phi_r = \pi/2$ in the following.

For the special case of 50:50 beam splitters in the interferometer, what are the relative intensities of output beams 1 and 2?

(b) **Bunching of Photons in a Beam Splitter**

As a simple example of nonclassical optical behavior, consider two photons of a single frequency that are simultaneously incident on two sides of a lossless, 50:50 beam splitter, as shown in the figure. Deduce the probability that $N_1 = 0$, 1 or 2 output photons are observed in the direction of beam 1. The key insight of a quantum analysis compared to a classical one is that photons obey Bose statistics, which means that the quantum probability that n indistinguishable photons are in a given state is $n!$ times the probability that n distinguishable particles would be in that state.[50]



Use the result of part (a) in your analysis, assuming that the two incident photons are in phase (at the midplane of the beam splitter). Compare your calculation for photons to a classical calculation for the output intensity on the two sides of the beam splitter in the case of two input light beams of intensities i_1 and i_2 which are not necessarily equal but which are in phase.

For use in part (e), deduce the probability, and probability amplitudes, for the fate of two in-phase photons that arrive simultaneously on opposite sides of a lossless beam splitter for which the probability of reflection of a single photon is $R = r^2$. The transmission probability is, of course, $T = 1 - R$.

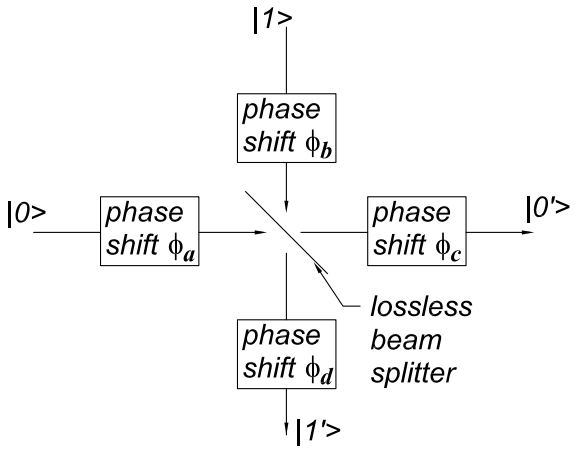
(c) **A Beam Splitter as a Quantum Processor**

A beam splitter can be regarded as a processor of Qbits in which the $|0\rangle$ state corresponds to the amplitude that a photon enters the splitter from one side, and the $|1\rangle$ state corresponds to the amplitude that the same photon enters from the other side. [This is called spatial encoding, which can be arranged by sending a single photon through a beam splitter prior to its arrival via two different paths at the splitter of interest.]

[49] See, for example, prob. 4(b) of http://kirkmcd.princeton.edu/examples/ph501set6.pdf

[50] See, for example, chap. 4, Vol. III of *The Feynman Lectures on Physics* (Addison-Wesley, 1965), http://feynmanlectures.caltech.edu/III_04.html.

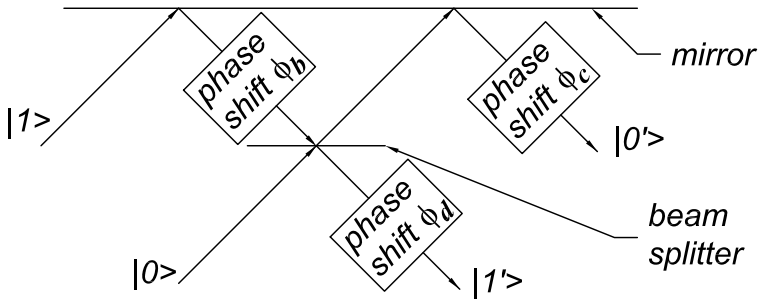
As shown in the figure below, label a photon that enters the splitter from the left as $|0\rangle$ and one that enters from above as $|1\rangle$. The reflection coefficient of the splitter is $R = \sin^2 \frac{\beta}{2}$.



Include up to four wave shifting plates at the entrance and exit ports of the splitter, also shown in the figure above. If these plates have index of refraction n, how thick does one have to be to introduce a phase shift δ (in radians) for a photon of wavelength λ?

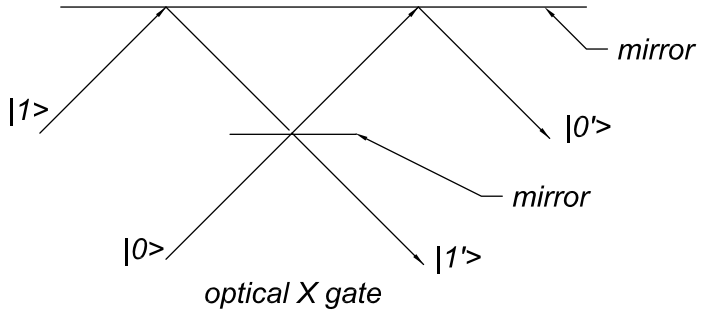
What is the 2×2 unitary matrix U that describes the effect of the splitter (+ wave shifting plates) on the general input state $\psi_0|0\rangle + \psi_1|1\rangle$?

Show that an arbitrary 2×2 unitary matrix, given by eqs. (44) and (56), can be represented by a beam splitter and associated wave-shifting plates for appropriate choices of the reflection coefficient and the phases ϕ_a, ϕ_b, ϕ_c and ϕ_d. Since a general 2×2 unitary matrix is described by 4 parameters (that we call α, β, γ and δ), you should find that one of the 4 phases ϕ_a, ϕ_b, ϕ_c and ϕ_d can be set to zero. With the convention that $\phi_a = 0$, our optical realization of an arbitrary 2×2 unitary matrix can also be drawn as



Thus, we have an example of how an arbitrary quantum transformation on a single Qbit can be accomplished with a passive system consisting of 4 (or 5, counting the large mirror at the top of the second figure) pieces of glass.

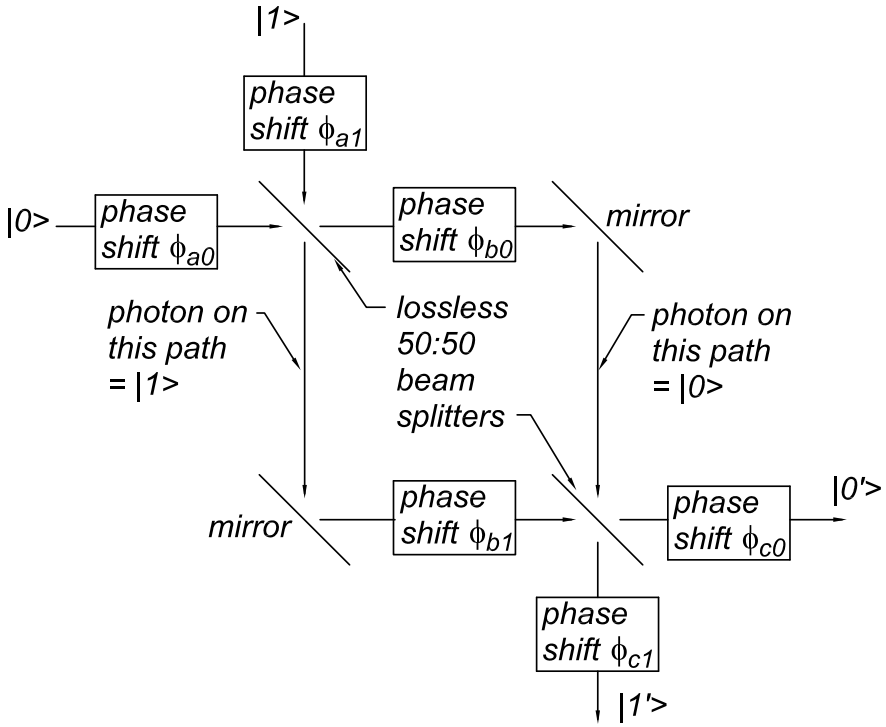
Note that a beam splitter with reflection coefficient $R = 1$ is essentially a two-sided mirror. Such a mirror, without any phase-shifting plates, implements the NOT ($= \mathsf{X}$) transformation by swapping the paths of input states $|0\rangle$ and $|1\rangle$, as shown on the next page.



Deduce the reflection coefficient R and the phase shifts ϕ_b, ϕ_c and ϕ_d needed to implement the $\mathsf{Z} = \boldsymbol{\sigma}_z$ gate and the Hadamard gate H. Present your results in figures such as that above.

(d) **A Mach-Zehnder Interferometer as a Quantum Processor**

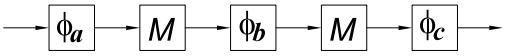
Show that an arbitrary 2×2 unitary matrix can also be realized by a Mach-Zehnder interferometer that uses lossless 50:50 beam splitters as well as 6 phase-shifting plates, as shown in the figure below.



A photon that enters the interferometer from the left can be defined to be the $|0\rangle$ input state, while a photon that enters from the top can be called the $|1\rangle$ input state. Similarly, we define a photon the emerges from the right of the interferometer to be in the $|0'\rangle$ output state and one that emerges from the bottom to be in the $|1'\rangle$ output state. Inside the interferometer, a photon that moves to the right and then down is said to be in the $|0\rangle$ state, while one that moves down and then to the right is said to be in the $|1\rangle$ state.

Deduce the form of a 2×2 unitary matrix M that describes the effect of a lossless 50:50 beam splitter on the input photon state. Likewise, define unitary matrices ϕ_a, ϕ_b and ϕ_c that describe the effect of the 3 pairs of phase-shifting plates. Then the interferometer, when used with a single input photon, can be considered to

be a single-Qbit processor of the form,



(e) **An Optical Controlled-NOT Gate**

Can beam splitters be used to implement 2-Qbit operations?

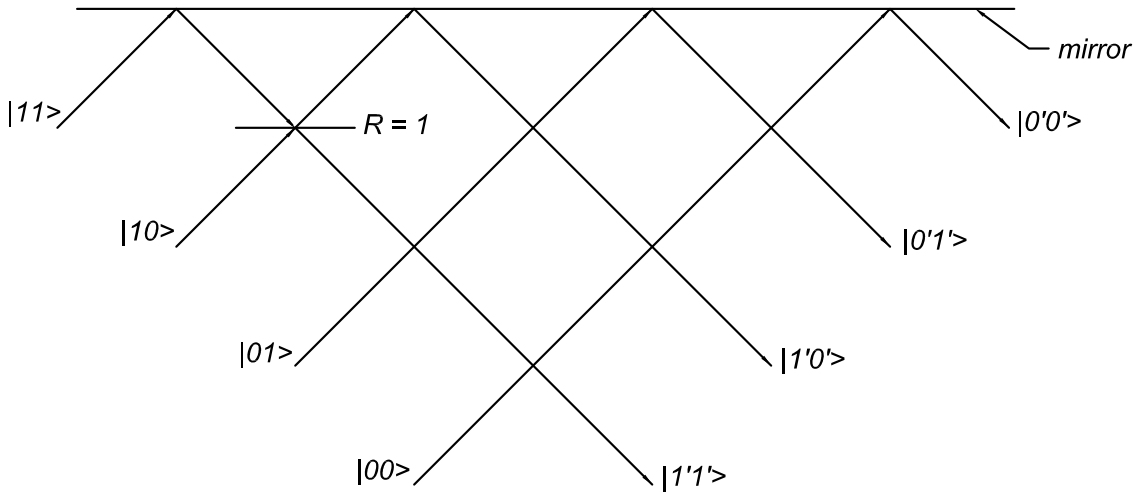
We explore this question by seeking an optical version of the Controlled-NOT gate, C_{xy}, that was introduced in problem 6(a). Recall that the effect of C_{xy} is to flip bit y if bit x is $|1\rangle$. The truth table for this operation is,

C_{xy} :

x	y	x'	y'
0	0	0	0
0	1	0	1
1	0	1	1
1	1	1	0

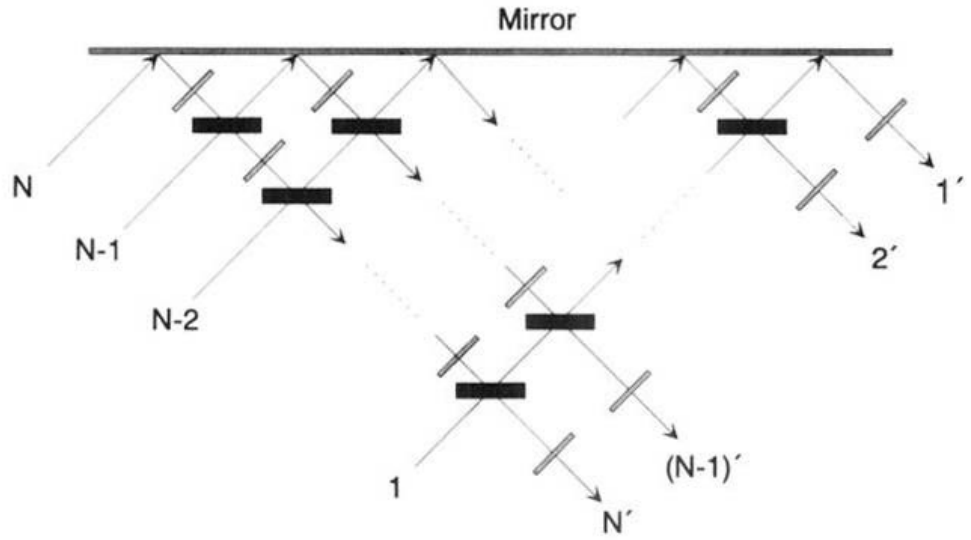
(108)

A simple implementation of the C_{xy} operation is shown in the figure below, which extrapolates from the construction of the NOT operation as discussed at the end of part (c). Again, only a single photon is used, but now that photon is "split" into 4 parts that are directed onto the incident paths labeled $|00\rangle$, $|10\rangle$, $|01\rangle$ and $|11\rangle$. A single beam splitter with reflection coefficient $R = 1$ swaps the paths of input states $|01\rangle$ and $|11\rangle$.



This is an example of a more general result that an arbitrary $N \times N$ unitary matrix can be realized by a set of at most $N(N-1)/2$ beam splitters and phase-shifting plates, together with a single photon that is directed onto N input paths,[51] as sketched in the figure on the next page.

[51] http://kirkmcd.princeton.edu/examples/QM/reck_prl_73_58_94.pdf
See also, http://kirkmcd.princeton.edu/examples/QM/cerf_pra_57_R1477_98.pdf



However, this method of realization of a large unitary matrix is not very practical. For example, if we wish to deal with a number that has n binary digits, its representation as a quantum vector involves 2^n states. Hence, if a quantum computation involving this number is reduced to a single unitary matrix, that matrix will have dimensions at least $2^n \times 2^n$. The number of beam splitters required to realize the needed unitary matrix will be of order 2^{2n}. If a beam splitter costs, say, \$1, then the cost of our optical quantum computer will be $\$2^{2n}$.

For example, to deal with a number of 100 decimal digits ($n = 100/\log 2$) would require a computer whose cost is $\$2^{200/\log 2} = \$10^{200}$.

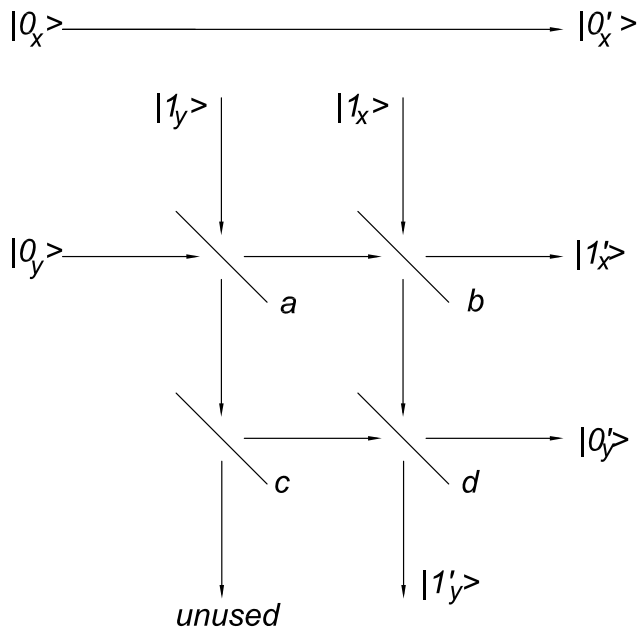
To avoid such high costs, we should not directly use large-dimension unitary matrices in our optical quantum computer. We defer until problems 11 and 12 a discussion of how large unitary matrices can be built up out of 2×2 matrices plus the 4×4 Controlled-NOT matrix.

A way to reduce the size of our optical quantum computer would be to represent different bits by different photons. It may be more practical to encode only one Qbit per photon (which still requires directing each photon along 2 paths), so that only $2n$ (input) optical paths are required to represent n Qbits.

These considerations are examples of more general concerns in computation, whereby algorithms/processes involving n bits that require an exponential (*i.e.*, 2^n) amount of resources are impractical, while those that require only a polynomial (*i.e.*, n^m) amount may be feasible.

So we now take up the challenge of realizing the 2-Qbit Controlled-NOT operation in a device that utilizes two photons, rather than one (conjecturing that an n-bit quantum computer utilizing such gates would require only n^m components).

The suggestion is to consider a Mach-Zehnder type of interferometer in which the mirrors are replaced by beam splitters, so that a second photon (x) can be made to interact with the first (y). In the operation C_{xy} of eq. (108) only the state $|1_x\rangle$ is to interact with photon y, so we consider an arrangement with 4 beam splitters, a, b, c and d, as shown on the next page.



We immediately see that this scheme cannot represent the Controlled-NOT operation in all cases, because it can happen that photon y emerges at the unused output, or photons x and y emerge in the states labeled $|0'_x\rangle$ and $|1'_x\rangle$, or in the states $|0'_y\rangle$ and $|1'_y\rangle$, or both photons emerge in the state $|1'_x\rangle$, *etc.*

However, we seek to show that when exactly one photon emerges in state $|0'_x\rangle$ or $|1'_x\rangle$ and the other in $|0'_y\rangle$ or $|1'_y\rangle$, then the operation is the Controlled-NOT. This success occurs only part of the time, but the success can be identified by the presence of exactly one final-state photon in the x' states and exactly one in the y' states.

First, deduce the values of the reflection coefficients R_a, R_b, R_c and R_d of the 4 beam splitters such that the two-photon initial state $|0_x\rangle|0_y\rangle$ cannot reach final state $|0'_x\rangle|1'_y\rangle$, and initial state $|0_x\rangle|1_y\rangle$ cannot reach final state $|0'_x\rangle|0'_y\rangle$. There is more that one solution for this; it suffices to consider only the simplest. Verify that the sum of the probabilities for all final states of the initial state $|0_x\rangle|0_y\rangle$ is unity.

Then, deduce additional constraints on the reflection coefficients such that the two-photon initial state $|1_x\rangle|0_y\rangle$ cannot reach final state $|1'_x\rangle|0'_y\rangle$ (and initial state $|1_x\rangle|1_y\rangle$ cannot reach final state $|1'_x\rangle|1'_y\rangle$).

What fraction of the time does this interferometer function successfully as a Controlled-NOT operation? *Answer:* $P_{\text{success}} = 1/9$.

8. **A Programmable Quantum Computer?**

In this problem you will show that there cannot be a completely general, programmable quantum computer, in contrast to the case for classical computation.

A classical n-to-n-bit function maps each of the 2^n initial n-bit words, into one of 2^n possible final words. Such a function is completely specified by $n2^n$ bits. A general n-bit classical computer (or programmable gate array) can be built that uses a list of $n2^n$ bits (the program) to implement any desired function.

In quantum computation, the general n-bit function is described by an $2^n \times 2^n$ unitary matrix U, which in turn can be described by 2^{2n} complex numbers subject to the 2^{2n} constraints that $\mathsf{U}^\dagger\mathsf{U} = \mathbf{I}$; *i.e.*, a total of 2^{2n} independent real numbers define U. Since a Qbit $|\psi\rangle = \psi_0|0\rangle + \psi_1|1\rangle$ is described by $4-1=3$ real numbers, it might seem that a list of 2^{2n-1} Qbits could be sufficient to serve as the program of a general programmable quantum gate array to implement an arbitrary $2^n \times 2^n$ unitary matrix U.

To this end, consider an m-Qbit program register $|p\rangle$ and an n-bit data register $|d\rangle$ in a direct-product state,

$$|p\rangle \otimes |d\rangle. \tag{109}$$

Quantum operations on this state are described by $(m+n)\times(m+n)$ unitary matrices V. A programmable quantum gate array requires the operator V to obey

$$\mathsf{V}[|p_U\rangle \otimes |d\rangle] = |p_U\rangle \otimes \mathsf{U}|d\rangle, \tag{110}$$

where U is an arbitrary $n \times n$ unitary matrix, and $|p_U\rangle$ is the state of the program register such that eq. (110) holds for each of the 2^n data words $|d\rangle$.

Consider two $n \times n$ unitary matrices U_p and U_q and their associated program register states $|p\rangle$ and $|q\rangle$,

$$\mathsf{V}[|p\rangle \otimes |d\rangle] = |p\rangle \otimes \mathsf{U}_p|d\rangle, \tag{111}$$
$$\mathsf{V}[|q\rangle \otimes |d\rangle] = |q\rangle \otimes \mathsf{U}_q|d\rangle, \tag{112}$$

to show that $|p\rangle$ and $|q\rangle$ are orthogonal whenever operators U_p and U_q are distinct.

Hint: Take the scalar product of eqs. (111) and (112).

There are only 2^m different orthogonal states in an m-Qbit register, but there are an infinite number of different unitary matrices U. Hence, a finite-sized quantum gate array can only be programmed for a small subset of possible quantum computations.

A consequence of this result is that the discussion of quantum computation in the remainder of this course will emphasize special-purpose quantum gate arrays, rather than programmable ones.

9. Designer Hamiltonians

This problem is something of a historical digression on early visions of quantum computation.

One of the first motivations for consideration of quantum computers may have been the desire for ever more compact processors, which leads us to contemplate processors on the atomic scale. The realization by Bennett and Landauer that computation can be intrinsically reversible, except for erasure, encouraged Benioff[52] to consider a spin-1/2 lattice as the basis of a computer. Benioff only considered algorithms based on Cbits, but with a quantum implementation. A possible worry was that classical systems are some kind of large-n limit of quantum systems, such that a quantum system might not successfully perform a "classical" algorithm. Benioff gave a formal argument that Hamiltonians for spin systems can be constructed such that the system behaves like a classical processor.

A largely independent approach was taken by Feynman[53] who appears to have been motivated by the challenge of computer simulation of quantum systems. For example, a system of n quantum subsystems which each have m states has a total of m^n states. A single "cycle" of such a system has m^{2n} possible transitions. Hence, a simulation of even a single cycle of this quantum system on a classical computer requires processing power that is exponential in the size n of the system. Clearly a classical computer has limited capability to simulate a quantum system. Perhaps a quantum computer would be more appropriate for this task.

However, Feynman's examples, like those of Benioff, were restricted to quantum implementation of classical algorithms. So, these early efforts served mainly to bring attention to the challenge of quantum computation.[54]

Both Benioff and Feynman emphasized the Hamiltonian of a quantum computer. This is, of course, inspired by Schrödinger's equation for the time development of a quantum state $|\psi\rangle$,

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \mathcal{H}(t)|\psi\rangle = \hbar\mathsf{h}(t)|\psi\rangle, \tag{113}$$

where $\mathcal{H}(t)$ is not the Hadamard transformation but is the Hamiltonian operator. A formal solution for the time dependence of state $|\psi\rangle$ is then,

$$|\psi(t)\rangle = e^{-i\int \mathsf{h}(t)dt}|\psi(0)\rangle = \mathsf{U}(t)|\psi(0)\rangle. \tag{114}$$

In case the Hamiltonian is time independent, as will be considered below, the time

[52] http://kirkmcd.princeton.edu/examples/QM/benioff_prl_48_1581_82.pdf

[53] Chap. 6 of *Feynman Lectures on Computation* (Addison-Wesley, 1996), which was written in 1985.
http://kirkmcd.princeton.edu/examples/QM/feynman_computation.pdf
See also http://kirkmcd.princeton.edu/examples/QM/feynman_fp_16_507_86.pdf
The challenge of simulation of quantum systems by other quantum systems was identified in the '70's by Poplavskii and by Manin, http://kirkmcd.princeton.edu/examples/QM/manin_quant-ph-9903008.pdf

[54] Somewhat surprisingly, Feynman wrote (p. 185 of the book cited above) that as to the question of the limitations to computing due to quantum mechanics and the uncertainty principle, "I have found that, aside from the obvious limitation to size if the working parts are to be made of atoms, there is no fundamental limits from these sources."

evolution becomes more simply,

$$|\psi(t)\rangle = e^{-i\mathsf{h}t}|\psi(0)\rangle. \tag{115}$$

The meaning of the formal expression (115) is clarified by making the Taylor expansion,

$$e^{-i\mathsf{h}t} = \sum_{n=0}^{\infty} \frac{(-i\mathsf{h}t)^n}{n!} = \mathbf{I} - i\mathsf{h}t - \frac{\mathsf{h}^2t^2}{2} + \frac{i\mathsf{h}^3t^3}{6} + \cdots \tag{116}$$

A way of thinking about this expansion is that during the time evolution of the state $|\psi\rangle$ the Hamiltonian operator is applied over and over in groups of various numbers n of repetitions (products of h), with the set of n repetitions being weighted by $(-it)^n/n!$.

Feynman addressed the task of finding a Hamiltonian h that implements a (reversible) quantum computation that is described by a unitary matrix M. That is, the desired computation is $|\psi'\rangle = \mathsf{M}|\psi\rangle$. In general, the computation is built up out of a sequence of m steps, each of which is described by a unitary matrix M_j. Then, the overall computation is represented by the product of the M_j,

$$\mathsf{M} = \prod_{j=1}^{m} \mathsf{M}_j = \mathsf{M}_m\mathsf{M}_{m-1}\cdots\mathsf{M}_2\mathsf{M}_1. \tag{117}$$

The computation involves, say, n Qbits, so that the matrices M_j are of size $2^n \times 2^n$.

If a Hamiltonian can be constructed so that its jth application executes operator M_j, then a group of m repetitions of this Hamiltonian would correspond to the desired computation. The time evolution operator (116) would include this action, among others. The remaining task would be to make some kind of measurement on the state $|\psi(t)\rangle$ that reveals the result of the computation.

Feynman increased the working size of the system to $m+n+1$ Qbits. He described the n Qbits on which the computation M is made as the data register, and the additional $m+1$ Qbits as the program counter.[55] The goal is to have the program counter in a readily identifiable state when the quantum computation has been completed.

Annihilation and Creation Operators.

To play with the Qbits in the program counter it is useful to introduce matrices that serve as annihilation and creation operators. The annihilation operator a for a single Qbit is defined as,

$$\mathsf{a} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad \text{(annihilation)}. \tag{118}$$

Its action on $|1\rangle$ is to turn it into $|0\rangle$, while it takes $|0\rangle$ to the "|vacuum⟩" state,

$$\mathsf{a}|1\rangle = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = |0\rangle, \quad \mathsf{a}|0\rangle = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = |\text{vac}\rangle, \tag{119}$$

[55] Feynman's quantum computer illustrates the result of prob. 8 that a program register of (at least) m bits is needed to program a computation that involves m different unitary operations.

It is conventional to say that the effect of the annihilation operator on the $|0\rangle$ state is to return the number 0:

$$\mathsf{a}|0\rangle = 0|0\rangle. \tag{120}$$

The hermitian conjugate (adjoint) of the annihilation operator a is the creation operator $\mathsf{a}^\dagger$,

$$\mathsf{a}^\dagger = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \qquad \text{(creation)}, \tag{121}$$

$$\mathsf{a}^\dagger|0\rangle = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = |1\rangle, \qquad \mathsf{a}^\dagger|1\rangle = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0. \tag{122}$$

The product $\mathsf{a}^\dagger\mathsf{a}$ is,

$$\mathsf{a}^\dagger\mathsf{a} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \equiv \mathsf{n}, \tag{123}$$

where n is the number operator,

$$\mathsf{n}|1\rangle = 1|1\rangle, \qquad \mathsf{n}|0\rangle = 0|0\rangle, \tag{124}$$

according to the convention (120). The product $\mathsf{a}\mathsf{a}^\dagger$ is,

$$\mathsf{a}\mathsf{a}^\dagger = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \mathbf{I} - \mathsf{n} \equiv \bar{\mathsf{n}}, \tag{125}$$

and clearly,

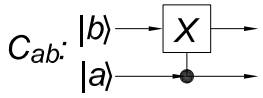
$$\mathsf{a}\mathsf{a}^\dagger + \mathsf{a}^\dagger\mathsf{a} = \mathbf{I}. \tag{126}$$

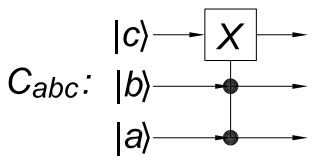
Any 2×2 matrix can be built up out of linear combinations of a, $\mathsf{a}^\dagger$, $\mathsf{a}\mathsf{a}^\dagger$ and $\mathsf{a}^\dagger\mathsf{a}$, *i.e.*, out of sums and products of a and $\mathsf{a}^\dagger$. For example, the other classical unitary 2×2 matrix, X (NOT), can be written as,

$$\mathsf{X} = \mathsf{a} + \mathsf{a}^\dagger. \tag{127}$$

(a) Express the Controlled-NOT operator C_{ab} (that flips bit b if bit $a = |1\rangle$) in terms of the annihilation and creation operators a and $\mathsf{a}^\dagger$ for bit a and b and $\mathsf{b}^\dagger$ for bit b. Use a tensor-product notation.

(b) Express the Controlled-Controlled-NOT operator (also called the Toffoli gate) C_{abc} (that flips bit c only if both bits a and b are $|1\rangle$) in terms of the annihilation and creation operators a, $\mathsf{a}^\dagger$, b, $\mathsf{b}^\dagger$, c and $\mathsf{c}^\dagger$. Your result should show the symmetry of C_{abc} with respect to bits a and b. For compactness, you need not expand unit matrices in terms of annihilation and creation operators. What is the 8×8 unitary matrix representation of the Toffoli gate?





Feynman suggested that we construct a time-independent Hamiltonian h based on the m unitary operations M_j with the aid of the annihilation and creation operators p_k and $\mathsf{p}_k^\dagger$ for the $m+1$ program counter bits according to,

$$\mathsf{h} = \sum_{j=0}^{m-1} \mathsf{p}_{j+1}^\dagger \otimes \mathsf{p}_j \otimes \mathsf{M}_{j+1} + \mathsf{p}_{j+1} \otimes \mathsf{p}_j^\dagger \otimes \mathsf{M}_{j+1}^\dagger. \tag{128}$$

The presence of the hermitian conjugate of each term insures that the Hamiltonian is hermitian even though each term of the sum is not. The tensor-product notation reminds us that the three factors of each term operate on different bits in our working space of $m+n+1$ bits.

This Hamiltonian is designed such that only one of the program counter bits is nonzero at any time.

For example, if all the program counter bits are $|0\rangle$, then the h acting on this system produces 0, since every term of eq. (128) contains an annihilation operator.

And, if exactly one program counter bit is equal to $|1\rangle$, say bit k (and the others are all $|0\rangle$), then only the term $\mathsf{p}_{k+1}^\dagger \otimes \mathsf{p}_k \otimes \mathsf{M}_{k+1}$ among the terms containing the M_j produces a nonzero result which is to flip bit k to $|0\rangle$, flip bit $k+1$ to $|1\rangle$ and apply operation M_{k+1} to the n data-register bits. The final result still has exactly one of the program counter bits equal to $|1\rangle$.

We see that for the program counter bit k to have been $|1\rangle$ the last previous operation on the register bits must have been M_k, which occured when program counter bit $k-1$ was $|1\rangle$. By induction, after the Hamiltonian has operated when bit k is $|1\rangle$ the register has been subject to the sequence of operations $\mathsf{M}_{k+1}\mathsf{M}_k \cdots \mathsf{M}_1$. Thus, repeated applications of the Hamiltonian increment the position of the nonzero program counter bit, and build up the computation on the register.

Once program-counter bit $m+1$ has been set to $|1\rangle$ (and all the other program-counter bits are $|0\rangle$), further applications of the terms of the Hamiltonian that contain the M_j make no changes to the system.

However, the Hamiltonian h also includes the term $\mathsf{p}_k \otimes \mathsf{p}_{k-1}^\dagger \otimes \mathsf{M}_k^\dagger$, which is needed for h to be hermitian. So if program counter bit k is $|1\rangle$, this term sets bit k back to $|0\rangle$, sets bit $k-1$ to $|1\rangle$ and applies operator $\mathsf{M}_k^\dagger$ to the register. Now, for bit k to have been $|1\rangle$ the last previous operation on the register must have been M_k. So, the combined action of the last two operations on the register is $\mathsf{M}_k^\dagger \mathsf{M}_k = \mathbf{I}$. This brings the register back to the state after operator M_{k-1} was applied to it. In brief, the effect of the term $\mathsf{p}_k \otimes \mathsf{p}_{k-1}^\dagger \otimes \mathsf{M}_k^\dagger$ is simply to set the computation back one step.

Since the Hamiltonian contains the sum of operators $\mathsf{p}_{k+1}^\dagger \otimes \mathsf{p}_k \otimes \mathsf{M}_{k+1}$ and $\mathsf{p}_k \otimes \mathsf{p}_{k-1}^\dagger \otimes \mathsf{M}_k^\dagger$, the register ends up in a quantum superposition of the computation up through the $k+1$st step and the $k-1$st step.

Indeed, after the Hamiltonian has been applied to the system a large number of times, the system is in a quantum superposition of all intermediate states of the system, including the final state that contains the desired answer to the computation.

To extract a result, we observe the state of program-counter bit $m+1$. If it is found to be $|0\rangle$, we were unlucky in projecting out one of the components of the quantum state

that did not correspond to a completed computation. But if we find program-counter bit $m+1$ to be $|1\rangle$, then we have observed the system in a state of completion, and we can look at the register bit to obtain the result of the computation.

Feynman's quantum computer therefore proceeds as follows:

- Construct the Hamiltonian h for the computation M according to eq. (128), and construct the time-evolution operator (116) from this.
- Initialize the program-counter bits to $|0\rangle$, except for the first counter bit, which is set to $|1\rangle$. Initialize the data-register bits as appropriate. This defines the quantum state $|\psi(0)\rangle$.
- Apply the time evolution operator (116) to the initial state $|\psi(0)\rangle$.
- Observe program-counter bit $m+1$ at suitable time intervals until it is found to have a value of $|1\rangle$. Then observe the data-register bits to read off the result of the computation.

A classical computer simulation of the quantum computation of a classical one-bit computation of the NOT operator is available as a Mathematica .nb file, http://kirkmcd.princeton.edu/examples/QM/Williams/WINDOWS/NBOOKS3/FEYNMAN.NB for Windows PC's. Versions for MAC and Linux are available in related directories.[56]

This notebook is a tutorial about constructing and running a simulation of a Feynman quantum computer for the computation $\mathsf{NOT} = \sqrt{\mathsf{NOT}} \cdot \sqrt{\mathsf{NOT}}$. The register contains only a single bit, and the program counter contains $2+1$ bits, since the computation is actually made in 2 steps. The computer requires a total of 4 Qbits, so the Hamiltonian and the time-evolution operator are 16×16 matrices.

Work through the tutorial at your leisure. Then complete exercises (c)-(h), using the `.nb` file as you find convenient.

Do these exercises in the spirit of Benioff and Feynman, *i.e.*, use only classical unitary operators to perform the computations themselves. In particular, use only the NOT, the Controlled-NOT, and the Controlled-Controlled-NOT operators, which are implemented in the `.nb` file as `NOTGate[i,m]`, `CNGate[i,j,m]`, and `CCNGate[i,j,k,m]`, respectively. Index `m` describes the total number of bits in the register, while bits `i`, `j` and `k` indicate which bits are acted upon by a particular gate. The bits `i`, `j` and `k` need not be in sequential order.

Although the exercises involve the construction of classical gates, this is to be done with quantum gates that can function on arbitrary Qbits. Remember that an arbitrary Qbit cannot be cloned, but a Controlled-NOT gate can be used to make a copy of a Cbit state.

For each exercise, provide a diagram of your circuit, and use the `.nb` file to present the `TruthTable` for the circuit. The circuits may involve some bits that are initialized to zero. Include only those lines of output from the `TruthTable` in which those bits

[56] To initialize the notebook, place the cursor to the right of the first executable command, `Off[General::spell1]`, and press Shift-Enter rather than merely Enter. Answer YES to the initialization popup window. To execute any of the other commands in the notebook, also press Shift-Enter after placing to cursor to the right of that command.

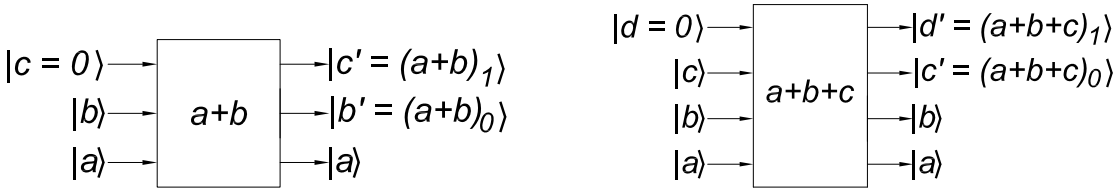
are initially zero. If your circuit involves a sequence of gates, say, $\mathsf{G}_1\mathsf{G}_2\mathsf{G}_3$, note that `TruthTable[G3 . G2 . G1]` is not the same as `TruthTable[G1 . G2 . G3]`.

We defer until problem 11 the question of whether the following operations could be implemented using only 1-bit and 2-bit gates.

(c) Construct the AND gate for 2 Cbits, which is equivalent to constructing a gate that multiplies two Cbits. Devise a measurement operator M to observe the output bit. What is the effect of the multiplier circuit on general input Qbits $|a\rangle = a_0|0\rangle + a_1|1\rangle$ and $|b\rangle = b_0|0\rangle + b_1|1\rangle$ as inferred from the measurement? For example, what are the probabilities that the output bit is 0 or 1?

(d) Construct the OR gate for 2 Cbits.

(e) Construct a circuit that adds two Cbits. The simplest circuit to do this will involve one ancillary bit. Since the sum of two bits is a two-bit number, one of the initial bits will be overwritten by the simplest circuit. Also construct a circuit with two ancillary bits that adds two Cbits, leaving these input bits unaltered.



Explore the effect of your circuits on general initial Qbits. Devise a measurement operator A to observe the two output bits. What are the probabilities that these bits corresponds to the sum being 0, 1 or 10? Note that the effect of your adder circuits on Qbits is not the same as the addition of Qbits as vectors. Note also that the result of a measurement of the output sum of your two circuits is in effect the same, even though it is not possible to clone an arbitrary Qbit in the sense discussed in problem 6. Hence, the no-cloning theorem is not quite as restrictive for quantum computation as it may have first appeared.

(f) Construct a circuit that adds three Cbits. This circuit will be needed for "carrying" during the addition of two multiple-bit numbers. In a particularly simple circuit, one of the two bits of the sum overwrites, say, the third input bit.

Hint: Since bit c is to be overwritten, start with a circuit that adds bits b and c, and append one that adds bits a and c.

(g) Construct a circuit that adds two 2-Cbit numbers a and b, overwriting b in the process.

It would be good to have a circuit that could be generalized to add two n-Cbit numbers. The circuit from part (f) can be an ingredient if we consider bit a to be the "carry" bit from the previous addition(s) of the lower-order bits of $a = \sum_{j=0} a_j 2^j$ *and* $b = \sum_{j=0} b_j 2^j$.[57] *The high-order output bit d then becomes the "carry" bit for the subsequent addition(s) of higher-order bits of a and b.*

[57] In decomposing an n-bit binary number as $a = \sum_{j=0}^{n-1} a_j 2^j$, we indicate the lowest-order bit as a_0 and the highest-order bit at a_{n-1}. Unfortunately, many authors, including Nielsen and Chuang, use a reverse convention in which a_1 is the highest-order bit and a_n is the lowest-order bit: $a = \sum_{j=1}^{n} a_j 2^{n-j}$.

The "obvious" solution has the defect that at the end of the addition, the "carry" lines may not be in their initial $|0\rangle$ states, which would prevent the circuit from being reused to perform other additions. Therefore, it is more elegant to augment the "obvious" solution with a group of gates that undoes the calculation of the lower-order "carry" bit(s), without changing the state of the b bits (which contain the sum $a + b$). Show that a minor variation of the circuit of part (f) does this job.

The hints are encouraging you to build a kind of "ripple adder", which uses roughly $3n$ bits to add two n-bit numbers. The tacit assumption is that in a quantum computer bits are expensive, but speed is not an issue. The converse is true in classical computation today, and addition is typically done by "carry-save" circuits that require of order n^2 bits to add two n-bit numbers.

(h) Construct a circuit that multiplies two 2-Cbit numbers.

Again, minimize the number of bits required, and restore any ancillary bits to their initial state.

Multiplication requires lots of bits. If bits are expensive in a quantum computer, it will be preferable to perform algorithms that multiply n-bit numbers modulo N, where $N < n$, since in this case the product is at most n bits long. As we will see in prob. 17, one of the most important algorithms in quantum computation is of this type.

Besides the artificial restriction by Feynman that his computer use only classical gates for the computation itself, it is also overly constrictive to embed a unitary computation M inside the Hamiltonian (128) that runs the computation forwards and backwards in a probabilistic manner. If we can actually implement the computation M in suitable hardware, it will be more efficient simply to apply M once, directly to the data-register state $|\psi_d\rangle$ to obtain the result $|\psi\text{‘}_d\rangle = \mathsf{M}|\psi_d\rangle$.

As Feynman stated, his computer was designed more to draw attention to the option of performing computations with quantum-scale devices that to indicate how this might best be done.[58]

[58] An early review of the quantum computers of Benioff and Feynman is given in http://kirkmcd.princeton.edu/examples/QM/peres_pra_32_3266_85.pdf

10. Deutsch's Algorithm

The first example of a quantum computation that could have a speed advantage over the corresponding classical version was given by Deutsch in 1985.[59]

A single classical processor can only evaluate a function of a single variable for one value of that variable at a time. In contrast a quantum processor that can evaluate a function for a general quantum state, *i.e.*, a superposition of orthogonal states, obtains information about that function for all orthogonal states during a single evaluation. In an important sense, a quantum processor is intrinsically a parallel processor.

To maximize the information that can be obtained from a single evaluation of $f(x)$ for a Qbit $|x\rangle$, we choose that Qbit to be a mixture of bits $|0\rangle$ and $|1\rangle$, namely

$$|x\rangle = |+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} = \mathsf{H}|0\rangle, \tag{129}$$

where H is the Hadamard transformation introduced in eq. (60).

This trick can be generalized to the case of functions of n Qbits by applying the n-fold tensor product of Hadamard transformations to the 0 state of n Qbits, $|0\rangle_n = |000\ldots0\rangle$,

$$\begin{aligned}
\mathsf{H}^{\otimes n}|0\rangle_n &= \mathsf{H}\otimes\mathsf{H}\otimes\cdots\otimes\mathsf{H}|000\ldots0\rangle = \mathsf{H}|0\rangle\mathsf{H}|0\rangle\cdots\mathsf{H}|0\rangle \\
&= \frac{|0\rangle+|1\rangle}{\sqrt{2}}\frac{|0\rangle+|1\rangle}{\sqrt{2}}\cdots\frac{|0\rangle+|1\rangle}{\sqrt{2}} \\
&= \frac{1}{2^{n/2}}\left(|000...00\rangle + |000...01\rangle + |000...10\rangle + \ldots + |111...11\rangle\right) \\
&= \frac{1}{2^{n/2}}\left(|0\rangle_n + |1\rangle_n + |2\rangle_n + \ldots + |2^n-1\rangle_n\right) \\
&= \frac{1}{2^{n/2}}\sum_{j=0}^{2^n-1}|j\rangle_n.
\end{aligned} \tag{130}$$

The state (130) is still a direct product ($\prod_{l=0}^{n-1}|+_l\rangle$), but any subsequent operation on this state that does not affect all Qbits identically will result in an entangled state.

A single evaluation of a function f for the quantum state (130) will contain information as to all of the values $f(|j\rangle_n)$, which is touted by Deutsch as "quantum parallelism".

Of course, to learn about the result we must make a measurement, which will project the state $f(\mathsf{H}^{\otimes n}|0\rangle_n)$ onto a single real number whose relation to the $f(|j\rangle_n)$ will not be known. Only with additional cleverness will it be advantageous to use a quantum processor.

Deutsch illustrated this for a function f of a single bit. Classically, there are only four possibilities for the function f:

---

[59] http://kirkmcd.princeton.edu/examples/QM/deutsch_prsl_a400_97_85.pdf
http://kirkmcd.princeton.edu/examples/QM/deutsch_prsl_a439_553_92.pdf
http://kirkmcd.princeton.edu/examples/QM/cleve_ptrsl_454_339_98.pdf
Online lectures by Deutsch: http://cam.qubit.org/video_lectures/index.php

$$f_0: \quad f(0)=0, \quad f(1)=0, \tag{131}$$
$$f_1: \quad f(0)=0, \quad f(1)=1, \tag{132}$$
$$f_2: \quad f(0)=1, \quad f(1)=0, \tag{133}$$
$$f_3: \quad f(0)=1, \quad f(1)=1. \tag{134}$$

To increase the amount of information that can be obtained from a single evaluation of function f, we suppose that this function can be implemented as a quantum processor of a single Qbit $|x\rangle$ such that the time for a single evaluation of the function is that same whether the argument is a Cbit or a Qbit.[60]

Since the output of a quantum processor is a quantum state, we can only obtain information as to the nature of this state by a measurement, which projects the quantum state onto a classical state. This would seem to imply that we can obtain only one piece of information per evaluation of a quantum processor.

However, by a sufficiently clever use of the quantum processor, we can obtain more information from a single evaluation than might have been expected.[61] To demonstrate this, we use a second Qbit $|y\rangle$ and construct the unitary transformation

$$\mathsf{U}_f(|x\rangle|y\rangle) = |x\rangle|y \oplus f(x)\rangle, \tag{135}$$

where the binary operation $\oplus$ is addition modulo 2 (for which $1 \oplus 1 = 0$). Note that $y \oplus f(x)$ leaves bit y alone if $f(x) = 0$ and flips bit y if $f(x) = 1$.

(a) For each of the four possible versions (131)-(134) of function f, express the unitary transformation U_{f_j} as a 4×4 matrix that acts on the basis

$$|x\rangle|y\rangle = \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \otimes \begin{pmatrix} y_0 \\ y_1 \end{pmatrix} = \begin{pmatrix} x_0y_0 \\ x_0y_1 \\ x_1y_0 \\ x_1y_1 \end{pmatrix}, \quad \text{such that} \quad |0\rangle|1\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \textit{etc.} \tag{136}$$

You will find that all four unitary matrices U_{f_j} are real and symmetric, so that $\mathsf{U}_{f_j}^{-1} = \mathsf{U}_{f_j}^{\dagger} = \mathsf{U}_{f_j}$.

(b) Show further that all four two-bit functions (131)-(134) can be expressed in terms the unit matrix, or only one nontrivial two-bit matrix called the Controlled-NOT,

$$\mathsf{C}_{xy} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \left(\begin{array}{c|c} \mathbf{I} & 0 \\ \hline 0 & \mathsf{X} \end{array}\right), \tag{137}$$

[60] Since a general Qbit has the form $a|0\rangle + b|1\rangle$, a *classical* simulation of the quantum processor would be $af(0) + bf(1)$, which involves two evaluations of function f. We suppose, along with Deutsch, that a true quantum processor could calculate the Qbit $|f(a|0\rangle + b|1\rangle)\rangle$ in a single evaluation of f.

[61] Indeed, so much cleverness is required that the improved version of Deutsch's algorithm presented here was developed only 14 years after his original paper.

or products of the Controlled NOT with operators that act only on bit $|x\rangle$ or only on bit $|y\rangle$.

Hint: Note that the two-bit operator X_x *that simply flips bit* $|x\rangle$ *obeys*

$$\begin{aligned} \mathsf{X}_x|0\rangle|0\rangle &= |1\rangle|0\rangle, \\ \mathsf{X}_x|0\rangle|1\rangle &= |1\rangle|1\rangle, \\ \mathsf{X}_x|1\rangle|0\rangle &= |0\rangle|0\rangle, \\ \mathsf{X}_x|1\rangle|1\rangle &= |0\rangle|1\rangle, \end{aligned} \tag{138}$$

and hence,

$$\mathsf{X}_x = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} = \left(\begin{array}{c|c} 0 & \mathsf{X} \\ \hline \mathsf{X} & 0 \end{array}\right); \tag{139}$$

and that the two-bit operator X_y *that simply flips bit* $|y\rangle$ *obeys*

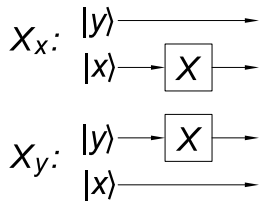
$$\begin{aligned} \mathsf{X}_y|0\rangle|0\rangle &= |0\rangle|1\rangle, \\ \mathsf{X}_y|0\rangle|1\rangle &= |0\rangle|0\rangle, \\ \mathsf{X}_y|1\rangle|0\rangle &= |1\rangle|1\rangle, \\ \mathsf{X}_y|1\rangle|1\rangle &= |1\rangle|0\rangle, \end{aligned} \tag{140}$$

and hence,

$$\mathsf{X}_y = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \left(\begin{array}{c|c} \mathsf{X} & 0 \\ \hline 0 & \mathsf{X} \end{array}\right). \tag{141}$$

Part (b) illustrates a general result that the 2-bit Controlled-NOT operator is universal in the sense that all multiple-bit operations can be built up out of it, the unit matrix, and single-bit operations. This theme will be pursued further in problems 11-12.

(c) The two-bit operations X_x and X_y that flip only one bit of the pair $|x\rangle|y\rangle$ can be represented as



Draw bit-operation diagrams for the four unitary transformations U_{f_j} used in Deutsch's algorithm.

Suppose we implement Deutsch's algorithm with $|y\rangle = (|0\rangle - |1\rangle)/\sqrt{2} = \mathsf{H}|1\rangle = \mathsf{HX}|0\rangle$. Then we have

$$|y \oplus f(x)\rangle = \frac{|0 \oplus f(x)\rangle - |1 \oplus f(x)\rangle}{\sqrt{2}} = (-1)^{f(x)}\frac{|0\rangle - |1\rangle}{\sqrt{2}} = (-1)^{f(x)}|y\rangle. \qquad (142)$$

This is a remarkable trick, whereby the state $|y\rangle$ is unchanged except for a phase factor that depends on the value of $f(x)$.

The trick can be exploited by choosing state $|x\rangle$ to be $\mathsf{H}|0\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. Then,

$$|x\rangle|y \oplus f(x)\rangle = \frac{(-1)^{f(0)}|0\rangle + (-1)^{f(1)}|1\rangle}{\sqrt{2}}\frac{|0\rangle - |1\rangle}{\sqrt{2}}. \qquad (143)$$

Another remarkable thing has happened here. The phase factor that appeared in eq. (142) is now a shared property of the direct product state $|x\rangle|y \oplus f(x)\rangle = \mathsf{U}_f|x\rangle|y\rangle$. And since the state of $|y\rangle$ did not change for the particular case that $|y\rangle = \mathsf{HX}|0\rangle$, this phase factor is now more a property of the state $|x\rangle$ than of $|y\rangle$. Yet, the operator U_f ostensibly did not change $|x\rangle$. This is one of the features that will permit (some) quantum computations to be more efficient than their classical versions.

To learn about the state (143) we don't need to measure the second bit since it is unchanged from its initial value of $(|0\rangle - |1\rangle)/\sqrt{2}$. So, we only measure the first bit. It is advantageous to measure this bit in the basis $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$, which leads to the result

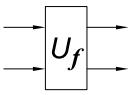
$$|x\rangle = \begin{cases} |+\rangle, & \text{if } f(0) = f(1), \\ |-\rangle, & \text{if } f(0) \neq f(1). \end{cases} \qquad (144)$$

While we have not learned both the values $f(0)$ and $f(1)$ from a single measurement, we have learned whether or not $f(0)$ equals $f(1)$. To obtain the latter knowledge by a classical computation would require two evaluations of function f, whereas we have learned this from a single quantum evaluation of f.

(d) If the transformation (135) is used with classical bits, the first bit would be unchanged and in general the second bit would be changed. Note that when we applied this transformation to a particular set of quantum bits, the first bit was changed in general, while the second bit was not. Clearly the behavior of a quantum processor is more subtle than that of a classical processor.

According to Bennett, it should be possible to reverse the quantum computation and restore the bits to their initial states. For the particular set of initial bits, $|x\rangle = |+\rangle$ and $|y\rangle = |-\rangle$, what is the procedure to restore these bits following application of transformation (135) and the measurement of the first bit?

(e) Suppose we prefer to measure in the $[|0\rangle, |1\rangle]$ basis rather than in the $[|+\rangle, |-\rangle]$ basis as was done for the state (143). Devise a (simple) unitary transformation of state (143) such that measurement in the $[|0\rangle, |1\rangle]$ tells us whether $f(0) = f(1)$ or not. Draw a diagram of the revised process, with input states $|x\rangle = |y\rangle = 0$. You may indicate the transformation U_f by the symbol



(f) **The Deutsch-Jozsa Algorithm.**

Design a circuit that generalizes Deutsch's algorithm to address the somewhat academic problem of determining whether a 1-bit function $f(x)$, where x consists of n bits, is balanced or constant when all we know about f is that it is one of these two types. The classical version of function f returns only 0 or 1. The function is balanced if it is 0 for exactly half of all values of x and 1 for the other half. And of course, the function is constant if it is either 0 for all values of x or 1 for all values.

Hint 1: The spirit of this part is to make the "obvious" generalization of Deutsch's algorithm for an n-to-1 bit function, and then to figure out what problem this algorithm solves.

Hint 2: Demonstrate that eq. (130) may be generalized as

$$\mathsf{H}^{\otimes n}|j\rangle_n = \frac{1}{2^{n/2}} \sum_{k=0}^{2^n-1} (-1)^{j\odot k}|k\rangle_n, \tag{145}$$

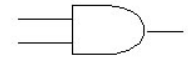
where $|j\rangle_n$ and $|k\rangle_n$ are n-bit, direct-product basis states as in eq. (16), and $j \odot k = \sum_l j_l k_l$ (mod 2) is the scalar product of j and k, modulo 2, when they are considered to be n-dimensional vectors whose elements are only 0's and 1's.

11. **Universal Gates for Classical Computation**

Classical computations are built up from computations on pairs of Cbits. A basic classical gate has two Cbits as input and one Cbit as output. The truth table for such a gate has 4 entries, corresponding to the input bits having values (0,0), (0,1), (1,0) and (1,1). Since the output bit can take on two values, 0 and 1, there are $2^4 = 16$ different classical 2-bit gates. The historical name for 2-bit classical gates (logic gates) is Boolean functions.[62]

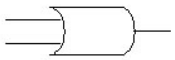
Symbols and truth tables for several 2-bit classical gates are shown below.

AND



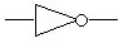
A	B	A • B
0	0	0
0	1	0
1	0	0
1	1	1

OR



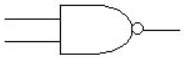
A	B	A + B
0	0	0
0	1	1
1	0	1
1	1	1

NOT



A	$\overline{A}$
0	1
1	0

NAND



A	B	$\overline{(A \bullet B)}$
0	0	1
0	1	1
1	0	1
1	1	0

NOR

A	B	$\overline{(A + B)}$
0	0	1
0	1	0
1	0	0
1	1	0

XOR

A	B	A ⊕ B
0	0	0
0	1	1
1	0	1
1	1	0

XNOR

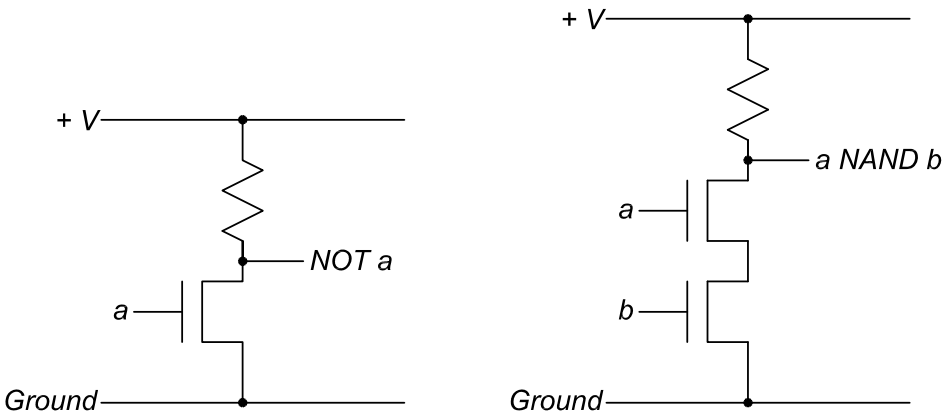
A	B	$\overline{(A \oplus B)}$
0	0	1
0	1	0
1	0	0
1	1	1

The question arises as to whether some subset of the 16 2-bit classical gates are universal in the sense that the remaining gates (and hence all classical computations) can be constructed from products of the universal gates.

It turns out that there are 19 different minimal sets of universal 2-bit classical gates. Two of these sets contain only a single gate, 7 contains two gates, and 12 contain three gates. The NAND and NOR gates are each universal for classical computation.

[62] An entertaining survey of classical computation is *The New Turing Omnibus* by A.K. Dewdney (Computer Science Press, New York, 1993).

Technical aside: The 1-bit NOT *gate and the 2-bit* NAND *gate are particularly simple to implement with transistors, and so have a favored status. The figure below sketches implementations of these gates with one and two transistors, respectively.*



Bits a and b are implemented by voltages with bit value 0 given by 0 volts, and bit value 1 given by a positive voltage large enough to make the transistor conduct when applied to the gate (= base). Thus in the NOT *circuit, the output voltage is high when the input voltage is low, and vice versa. Similarly, in the* NAND *circuit the output voltage is high unless both input voltages are high.*

(a) **The 2-Bit NAND Gate is Universal for Classical Computation.**

Illustrate the classical universality of the 2-bit NAND gate by showing how to use it to implement the AND, OR and NOT gates. If you will accept that the latter three gates form a universal set, you have then shown that the NAND gate alone is universal. (Otherwise, show that the remaining 11 2-bit classical gates can also be implemented via NAND gates.)

(b) **The Controlled-Controlled-NOT Gate is Universal for Classical Computation.**

Turning now to implementations of computation via quantum gates, we recall that these gates must be unitary (and so have the same number of outputs as inputs). We have seen in Prob. 9 that the Controlled-Controlled-NOT is useful in implementing various computational functions. Show that this 3-bit gate is universal for classical computations by implementing the 2-bit NAND gate with it.

If we add the requirement for quantum circuits that any ancillary bits must have their initial state as $|0\rangle$, then we must relax our definition of classical universality of a quantum gate to mean that it, together with any needed 1-bit quantum gates, suffices to implement any classical computation.

*See sec. 6.1.3 of Preskill's lecture notes for a discussion of how an arbitrary n-Cbit transformation can be built up out of Controlled-Controlled-*NOT *gates.*

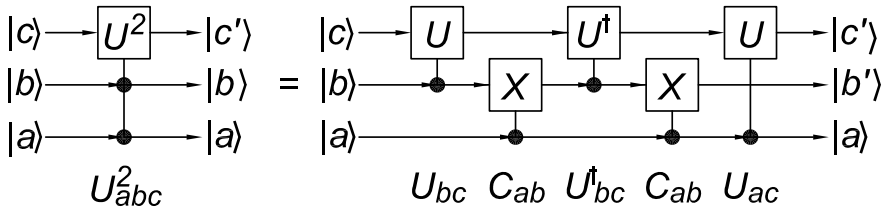
(c) **The Controlled-NOT Gate is Universal for Classical Computation.**

A 3-bit quantum gate requires the coupling of three input Qbits, which is much more difficult than coupling only two Qbits. Hence, it is advantageous if there is a set of 2-bit quantum gates that is universal for classical computation.

We have seen in Prob. 3(d) that all 24 2-bit classical unitary gates implement linear transformations of the input bits. However, the Controlled-Controlled-NOT gate C_{abc} is nonlinear in that it produces the product of bits a and b if bit c is initially $|0\rangle$. Hence, gate C_{abc} cannot be implemented by any combination of the 24 2-bit classical unitary gates (try it!). So, no subset of these gates is universal for classical computation (if they can be combined only with the 1-bit NOT gate, which is the only nontrivial 1-bit classical unitary gate).

Show that the 2-bit Controlled-NOT gate is universal for classical computation when combined with appropriate 1-bit quantum unitary gates.

i. First show that given a 2-bit Controlled-U gate (U_{ab}) where U is any 1-bit unitary operator, you can construct a 3-bit Controlled-Controlled-U^2 gate (U^2_{abc}) with the following circuit,

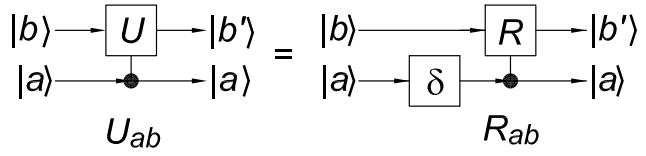


So, in particular, if $\mathsf{U} = \sqrt{\mathsf{NOT}} = \sqrt{\mathsf{X}}$ then the Controlled-Controlled-NOT gate can be constructed from 2-bit unitary gates, and therefore some set of 2-bit unitary gates is universal for classical computation.

ii. Next, show that the 2-bit Controlled-U operator, where U is a general 1-bit unitary operator of the form $\mathsf{U} = e^{i\delta}\mathsf{R}$ and R is a rotation operator, can be constructed from a 1-bit unitary gate with matrix form

$$\boldsymbol{\delta} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\delta} \end{pmatrix}, \tag{146}$$

and a 2-bit Controlled-R gate as in the figure below.

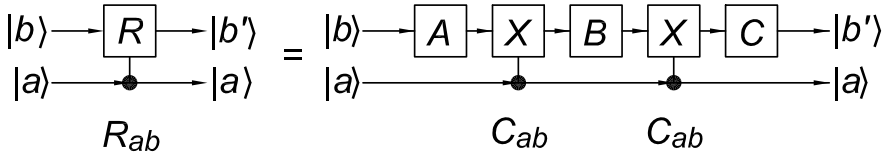


Show also that

$$\boldsymbol{\delta} = e^{i\delta\mathsf{n}}, \tag{147}$$

where n is the number operator introduced in eq. (123).

iii. Finally, show that the 2-bit Controlled-R gate (R_{ab}, where $\mathsf{R}(\alpha,\beta,\gamma)$ is a rotation operator) can be constructed from the 2-bit Controlled-NOT gate and three 1-bit unitary gates A, B and C with the following circuit,



What conditions must the 1-bit gates A, B and C obey for the circuit to function as a Controlled-R gate?
Deduce forms for operators A, B and C such that the above circuit is valid. *Hint: Recall eqs. (52) and (53), and express operators* A, B *and* C *as products of appropriate rotations about the y and z axes.*

iv. What 1-bit gates A, B and C and $\boldsymbol{\delta}$ are needed to construct the Controlled-U operator as in sec. ii-iii when $\mathsf{U} = \sqrt{\mathsf{NOT}} = \sqrt{\mathsf{X}}$?

Thus, the only 2-bit unitary gate needed to implement the circuit for the Controlled-Controlled-NOT gate is the Controlled-NOT gate. But this implementation requires the use of nonclassical 1-bit unitary gates.

While we can now, in principle, perform all classical computations using only one type of 2-bit quantum gate, it appears that we need a large variety of 1-bit quantum gates as well. The question arises as to what is the minimal set of 1-bit quantum gates required for our universal scheme for classical computation.

A further question is whether an arbitrary n-bit quantum gate can be built up from some minimal set of low-order-bit quantum gates.

12. **Universal Gates for Quantum Computation**

This problem concerns the impressive formal result that all unitary transformations ($2^n \times 2^n$ matrices) on a set of n Qbits can be built up from products of the 2-Qbit Controlled-NOT gate and powers (*i.e.*, sums of products) of two 1-Qbit gates. This result has been anticipated by Prob. 7(e) and Prob. 11. We consider it here for its intellectual interest, rather than any practical relevance to laboratory realization of quantum computation.

We proceed in four steps. Part (a) shows that any 1-Qbit unitary gate can be well approximated by (irrational) powers of the 1-Qbit gates H and $\boldsymbol{\sigma}_z^{1/4}$. Part (b) introduces some useful generalizations of the Controlled-NOT operator. Part (c) shows that any "two-level" $2^n \times 2^n$ unitary matrix can be represented by a set of 2-bit Controlled-NOT gates plus 1-Qbit gates. Part (d) shows that any $2^n \times 2^n$ unitary matrix can be decomposed into a product of "two-level" $2^n \times 2^n$ unitary matrices.

(a) **All 1-Bit Quantum Gates Can Be Built from the H and $\boldsymbol{\sigma}_z^{1/4}$ Gates.**
We have seen in prob. 4(d) than an arbitrary 1-bit quantum gate with unit determinant is equal to the product of 3 gates that correspond to rotations about a pair of orthogonal axes, there called y and z,

$$\mathsf{U} = \mathsf{R}_z(\gamma)\mathsf{R}_y(\beta)\mathsf{R}_z(\alpha). \tag{56}$$

Here, we wish to find a minimal set of 1-bit gates to implement, at least approximately, the decomposition (56). For this, we need a way to represent arbitrary rotations about 2 orthogonal axes using a minimal set of gates.
We have already seen in Prob. 4(i) that given a rotation represented by $e^{i\frac{\theta}{2}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}}$ where $u_z = -u_x$, then the operation $\mathsf{H}^{-1/2}\ e^{i\frac{\theta}{2}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}}\ \mathsf{H}^{1/2}$ represents a rotation by the same angle θ about an axis orthogonal to $\hat{\mathbf{u}}$.
We have also seen in prob. 4(h) that the operation $\mathsf{H}^{1/2}$ can be composed from the two gates H and $\boldsymbol{\sigma}_z^{1/4}$ via the relations

$$\mathsf{H}^{1/2} = \boldsymbol{\sigma}_y^{1/4}(\boldsymbol{\sigma}_z^{1/4})^2\boldsymbol{\sigma}_y^{-1/4}, \tag{64}$$

$$\boldsymbol{\sigma}_y^{1/4} = (\boldsymbol{\sigma}_z^{1/4})^2\boldsymbol{\sigma}_x^{1/4}(\boldsymbol{\sigma}_z^{-1/4})^2, \tag{63}$$

$$\boldsymbol{\sigma}_x^{1/4} = \mathsf{H}\boldsymbol{\sigma}_z^{1/4}\mathsf{H}. \tag{62}$$

This suggests that we look for a gate that can be built from H and $\boldsymbol{\sigma}_z^{1/4}$ whose representation in the form $e^{i\frac{\theta}{2}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}}$ corresponds to an axis of rotation with $u_z = -u_x$.
Verify that the combination

$$\boldsymbol{\sigma}_z^{-1/4}\boldsymbol{\sigma}_x^{1/4} = \mathsf{R}_u(\theta) \tag{148}$$

is of the required form. Show that the corresponding angle θ obeys a transcendental equation, which implies that its value is irrational (and not an integer multiple of π, although you need not show this). Deduce the orthogonal vector $\hat{\mathbf{v}}$ that corresponds to the operation

$$\mathsf{H}^{-1/2}\ \boldsymbol{\sigma}_z^{-1/4}\boldsymbol{\sigma}_x^{1/4}\ \mathsf{H}^{1/2} = e^{i\frac{\theta}{2}\hat{\mathbf{v}}\cdot\boldsymbol{\sigma}} = \mathsf{R}_v(\theta). \tag{149}$$

The prescription to construct an arbitrary 1-bit operator U (with unit determinant) from the gates H and $\boldsymbol{\sigma}_z^{1/4}$ is now reasonably straightforward. Since axes $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$ are orthogonal, there exists a decomposition

$$\mathsf{U} = \mathsf{R}_v(\gamma')\mathsf{R}_u(\beta')\mathsf{R}_v(\alpha') = e^{i\frac{\gamma'}{2}\hat{\mathbf{v}}\cdot\boldsymbol{\sigma}}\; e^{i\frac{\beta'}{2}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}}\; e^{i\frac{\alpha'}{2}\hat{\mathbf{v}}\cdot\boldsymbol{\sigma}}. \tag{150}$$

As the angle θ is irrational and not an integer multiple of π, the set of values $\{(n\theta)_{\mathrm{mod}(2\pi)},\ n = 1, 2, 3, ...\}$ includes a member that is arbitrarily close to any number on the interval $[0, 2\pi]$. Hence, we can find integers l, m and n such that $(l\theta)_{\mathrm{mod}(2\pi)}$, $(m\theta)_{\mathrm{mod}(2\pi)}$ and $(n\theta)_{\mathrm{mod}(2\pi)}$ approximate angles α', β' and γ' to any desired accuracy. Then,

$$\mathsf{U} \approx e^{i\frac{n\theta}{2}\hat{\mathbf{v}}\cdot\boldsymbol{\sigma}}\; e^{i\frac{m\theta}{2}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}}\; e^{i\frac{l\theta}{2}\hat{\mathbf{v}}\cdot\boldsymbol{\sigma}} = \mathsf{R}_v^n(\theta)\; \mathsf{R}_u^m(\theta)\; \mathsf{R}_v^l(\theta), \tag{151}$$

which involves only powers of the two gates H and $\boldsymbol{\sigma}_z^{1/4}$,

$$\mathsf{H} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \qquad \boldsymbol{\sigma}_z^{1/4} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix}. \tag{152}$$

Strictly speaking, we have only shown that 2×2 unitary gates with unit determinants (members of SU(2)) can be constructed from powers of H and $\boldsymbol{\sigma}_z^{1/4}$. To construct any 2×2 unitary gate we need to be able to multiple an arbitrary rotation gate R by a phase gate of the form

$$\begin{pmatrix} e^{i\delta} & 0 \\ 0 & e^{i\delta} \end{pmatrix}. \tag{153}$$

This matrix is not (I believe) constructible from powers of H and $\boldsymbol{\sigma}_z^{1/4}$, but note that

$$(\boldsymbol{\sigma}_z^{1/4})^{4\delta/\pi} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\delta} \end{pmatrix}. \tag{154}$$

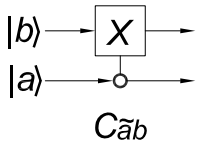
Then, recalling Prob. 11(c)ii-iii, we see that we can implement the gate $\mathsf{U} = e^{i\delta}\mathsf{R}$ in the form of a Controlled-U operation where the control bit is always set to $|1\rangle$, using 1-bit gates that are powers of H and $\boldsymbol{\sigma}_z^{1/4}$ together with the 2-bit Controlled-NOT gate.

Since we now want to show that this combination of gates can be used to implement an arbitrary $2^n \times 2^n$ unitary matrix, it is no loss of generality to have first implemented an arbitrary 2×2 unitary matrix with these gates.

(b) **Generalized Controlled-NOT Operations.**

To aid in our construction of arbitrary unitary operators from Controlled-NOT operators, its is useful to have available several generalization of the latter.

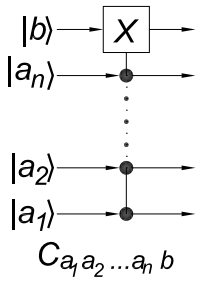
First, we desire an operator $\mathsf{C}_{\tilde{a}b}$ which flips bit b only when bit a is $|0\rangle$. We will symbolize this operator in bit-flow diagrams as shown on the next page, where the open circle on the control bit a indicates that the target operation is applied only when the control bit is $|0\rangle$.



Show how the 2-bit operator $\mathsf{C}_{\tilde{a}b}$ can be implemented using the 2-bit Controlled-NOT operator and suitable 1-bit operators. Your solution should be readily generalizable to produce, for example, operators $\mathsf{C}_{\tilde{a}bc}$, $\mathsf{C}_{a\tilde{b}c}$ and $\mathsf{C}_{\tilde{a}\tilde{b}c}$ from the Controlled-Controlled-NOT operator C_{abc}.

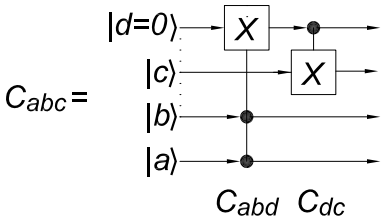
For completeness, give 4×4 matrix representations for the four variants of 2-bit Controlled-NOT operators, C_{ab}, C_{ba}, $\mathsf{C}_{\tilde{a}b}$ and $\mathsf{C}_{\tilde{b}a}$.

We will also wish to use Controlled-NOT operators with an arbitrary number of control bits a_1, a_2, ... a_n, all of which must be $|1\rangle$ for the target bit $|b\rangle$ to be flipped.



The operator $\mathsf{C}_{a_1 a_2 \ldots a_n b}$ can in principle be constructed by a generalization of the procedure of Prob. 11(c)i. However, it suffices for you to demonstrate that n-control-bit Controlled-NOT operator can be built out of a suitable set of Controlled-Controlled-NOT operators $\mathsf{C}_{a_j a_k b}$.

*Hint: The 3-bit Controlled-Controlled-*NOT *operator* C_{abc} *can be implemented, as shown below, in a fashion that seems wasteful as it utilizes an auxiliary bit. However, this scheme is readily generalizable to a construction for n control bits together with* $n - 1$ *auxiliary bits that uses an appropriate sequence of 3-bit Controlled-Controlled-*NOT *gates.*



Note that if we wished to restore the auxiliary bit d to its initial state, we should symmetrize this circuit about the operation C_{dc}.

Of course, we could use the first exercise of this part to implement a Controlled-NOT gate in which the target bit flips only when some or all of the control bits are $|0\rangle$ rather than $|1\rangle$.

(c) **Two-Level Unitary Matrices.**

An $m \times m$ unitary matrix U is called a "two-level" matrix if all diagonal elements are 1's and all off-diagonal elements are 0's except for the 4 elements

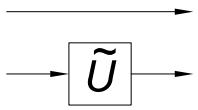
$$\mathsf{U}_{jj} = a,\ \mathsf{U}_{jk} = b,\ \mathsf{U}_{kj} = c \text{ and } \mathsf{U}_{kk} = d,$$

which together form a unitary 2×2 matrix $\tilde{\mathsf{U}}$ with unit determinant,

$$\mathsf{U}_{\rm two-level} = \begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & a & \dots & 0 & \dots & b & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & c & \dots & 0 & \dots & d & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix}, \qquad \tilde{\mathsf{U}} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \quad (155)$$

We would now like to construct the n-Qbit operation U by applying the 1-Qbit operation $\tilde{\mathsf{U}}$ along with any appropriate generalized Controlled-NOT operators (which can be constructed from 2-bit Controlled-NOT operators).

It is instructive to begin with the case that $m = 4$, *i.e.*, when U is a 2-bit operator. Note that the matrix corresponding to the 2-bit process



is not a two-level operator, since its matrix representation is

$$\mathbf{I} \otimes \tilde{\mathsf{U}} = \begin{pmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & a & b \\ 0 & 0 & c & d \end{pmatrix}. \quad (156)$$

Rather, two-level matrices result from various forms of Controlled-$\tilde{\mathsf{U}}$ operations,

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & a & b \\ 0 & 0 & c & d \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & a & 0 & b \\ 0 & 0 & 1 & 0 \\ 0 & c & 0 & d \end{pmatrix}, \quad \begin{pmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} a & 0 & b & 0 \\ 0 & 1 & 0 & 0 \\ c & 0 & d & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (157)$$

However, there are two more 2-bit, two-level operators based on $\tilde{\mathsf{U}}$,

$$\begin{pmatrix} a & 0 & 0 & b \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ c & 0 & 0 & d \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & a & b & 0 \\ 0 & c & d & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (158)$$

Deduce transformations of one (or more) of the four basic two-level matrices (157) into the forms (158). Recall that our goal is to perform such transformations using only Controlled-NOT operations.

A generalization of this type of procedure to construct higher-dimensional, two-level operators such as that of eq. (155) is described in sec. 4.5.2 of Nielsen and Chuang.

(d) **Any Unitary Operator is a Product of Two-Level Matrices.**

We can show by "brute force" than any $n \times n$ unitary matrix $\mathsf{U} = U_{jk}$ can be multiplied on the left by a sequence of $n(n-1)/2$ two-level unitary matrices U_l such that

$$\mathsf{U}_{n(n-1)/2} \cdots \mathsf{U}_l \cdots \mathsf{U}_2 \mathsf{U}_1 \mathsf{U} = \mathbf{I}, \quad \text{and so,} \quad \mathsf{U} = \mathsf{U}_1^\dagger \mathsf{U}_2^\dagger \cdots \mathsf{U}_l^\dagger \cdots \mathsf{U}_{n(n-1)/2}^\dagger. \tag{159}$$

The procedure has $n-1$ major steps such that after m steps the first m rows and columns of the product $\cdots \mathsf{U}_2 \mathsf{U}_1 \mathsf{U}$ have 1's on the diagonal and 0's off the diagonal.

If suffices to show how the first row and column of an $m \times m$ unitary matrix can be transformed via multiplication by $m-1$ two-level matrices into a matrix L with $L_{1j} = L_{j1} = 0$) except for the first diagonal element which is unity ($L_{11} = 1$).

This task is performed using $m-1$ two-level matrices of the form

$$\mathsf{U}_l = \begin{pmatrix} \frac{L_{11}^\star}{\sqrt{|L_{11}|^2 + |L_{l+1,1}|^2}} & 0 & \ldots & 0 & \frac{L_{l+1,1}^\star}{\sqrt{|L_{11}|^2 + |L_{l+1,1}|^2}} & 0 & \ldots & 0 \\ 0 & 1 & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \ldots & 1 & 0 & \ldots & \ldots & \ldots \\ \frac{L_{l+1,1}}{\sqrt{|L_{11}|^2 + |L_{l+1,1}|^2}} & 0 & \ldots & \ldots & -\frac{L_{11}}{\sqrt{|L_{11}|^2 + |L_{l+1,1}|^2}} & 0 & \ldots & \ldots \\ 0 & \ldots & \ldots & \ldots & 0 & 1 & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & 1 \end{pmatrix}, \tag{160}$$

where

$$\mathsf{L} = \mathsf{U}_l \cdots \mathsf{U}_1 \mathsf{U} = \begin{pmatrix} L_{11} & L_{12} & \ldots \\ 0 & L_{22} & \ldots \\ \ldots & \ldots & \ldots \\ 0 & L_{l+1,2} & \ldots \\ L_{l+2,1} & L_{l+2,2} & \ldots \\ \ldots & \ldots & \ldots \end{pmatrix}. \tag{161}$$

That is, each additional multiplication by an appropriate two-level matrix zeroes out one more element in the first column of the product matrix L.

A technicality is that if $U_{1j} = 0$ for $j < l$ then element L_{11} is actually zero until step l. Since a unitary matrix cannot have all elements of a column equal to zero, sooner or later we obtain $L_{11} \neq 0$.

After $m-1$ such steps, the product matrix L has all zeroes in its first column except for element L_{11}. Recalling that all rows and columns of a unitary matrix must be unit vectors, we see that $L_{11}=1$ now. This also implies that all elements of the first row are zero (except for L_{11} of course). Thus, after $m-1$ steps we have

$$\mathsf{L}=\mathsf{U}_{m-1}\cdots\mathsf{U}_1\mathsf{U}=\begin{pmatrix}1 & 0 & ... & 0\\ 0 & L_{22} & ... & L_{2m}\\ ... & ... & ... & ...\\ 0 & L_{m2} & ... & L_{mm}\end{pmatrix}. \tag{162}$$

Beginning with the $n\times n$ unitary matrix U, we follow the above procedure through $n-2$ major steps (involving a total of $(n-1)+(n-2)+\cdots+2=n(n-1)/2-1$ steps), at which point the product matrix L is the two-level unitary matrix

$$\mathsf{L}=\mathsf{U}_{n(n-1)/2-1}\cdots\mathsf{U}_1\mathsf{U}=\begin{pmatrix}1 & 0 & ... & ... & ... & 0\\ 0 & 1 & ... & ... & ... & 0\\ ... & ... & ... & ... & ... & ...\\ 0 & ... & ... & 1 & 0 & 0\\ 0 & ... & ... & 0 & L_{n-1,n-1} & L_{n-1,n}\\ 0 & ... & ... & 0 & L_{n,n-1} & L_{nn}\end{pmatrix}. \tag{163}$$

Choosing the last two-level matrix to be $\mathsf{U}_{n(n-1)/2}=L^\dagger$, we arrive at eq. (159). This completes our sketch of the universality for quantum computation of the 2-bit Controlled-NOT gate together with the 1-bit gates H and $\boldsymbol{\sigma}_z^{1/4}$.

There is no assigned problem in part (d).

Entanglement is Needed for Universality.

It turns out that "almost any" 2-Qbit gate is universal for quantum computation.[63]

However, the very useful SWAP gate S that was introduced in prob. 6(c) is an example of a 2-Qbit gate that is not universal.[64] Dodd *et al.*[65] have shown that any 2-Qbit gate that causes entanglement is universal for quantum computation. Then, Bremner *et al.*[66] have shown that the 2-Qbit gates which are not universal for quantum computation are of only two types, neither of which creates entanglement:

(i) Gates that are direct products of 1-Qbit gates.

(ii) Gates U that are equivalent to the SWAP gate S, meaning that

$$\mathsf{U}=(\mathsf{A}\otimes\mathsf{B})\mathsf{S}(\mathsf{C}\otimes\mathsf{D}), \tag{164}$$

where A, B, C and D are 1-Qbit gates.

[63] http://kirkmcd.princeton.edu/examples/QM/lloyd_prl_75_346_95.pdf
http://kirkmcd.princeton.edu/examples/QM/deutsch_prsl_a449_669_95.pdf

[64] I conjecture that practical quantum computers will involve more SWAP gates than any other kind. But SWAP gates alone are not sufficient for nontrivial quantum computation.

[65] http://kirkmcd.princeton.edu/examples/QM/dodd_pra_65_040301_02.pdf

[66] http://kirkmcd.princeton.edu/examples/QM/bremner_prl_89_247902_02.pdf

Entanglement is not necessary for all types of quantum computation (since by definition any quantum transformation is a computation), but entanglement is required for any approach that has the ambition to encompass all possible types of quantum computation (*i.e.*, all possible types of quantum phenomena).[67]

[67] Thus, "classic" courses in quantum mechanics that barely mention entanglement exclude the conceptual bulk of quantum phenomena, typically illustrating only one-particle systems and the small subclass of multiparticle interactions involving direct products.

13. The Bernstein-Vazirani Problem

We now have the vision of a quantum computation as any transformation on a set of Qbits as performed by a unitary operator. In this sense, quantum computation is a solution looking for a problem.

We again take up the quest, begun in prob. 10, for problems that can be solved by a quantum computation in a manner that affords potential advantages compared to classical computation.

A second example of a quantum computation that is faster than a corresponding classical computation has been given by Bernstein and Vazirani.[68] This problem is somewhat artificial, but instructive.

The task is to determine the value of an n-bit integer a by appropriate use of a 1-bit function

$$f_a(x) = a \cdot x \pmod 2 \equiv a \odot x. \tag{165}$$

That is, $f_a(x)$ computes the scalar product of a and x when they are regarded as vectors with binary coefficients in an n-dimensional space, and performs the summation modulo 2.

To determine a via classical computations of f_a we would use n applications of it, the mth of which sets $x = 2^m$ and thereby determines the mth bit of a.

The claim is that a quantum computation can be constructed that determines a in a single application of f_a.

(a) Give a bit-flow diagram for $f_a(x)$ as a quantum computation U_{f_a}. The circuit will have n lines for the n-Qbit input state $|x\rangle_n$ and another line, say $|y\rangle$, for the 1-Qbit output of $f(x)$. Since the function f depends on the value of the n-bit number a, you may find it useful to represent this dependence with an additional n lines representing the state $|a\rangle_n$.

Hint: Recall prob. 9(c), and note that addition modulo 2 of a set of bits can be implemented by a succession of bit flips, supposing that the output bit starts as $|y\rangle = |0\rangle$.

The output bit could, of course, begin as a $|1\rangle$. Convince yourself that the effect of your circuit can be summarized as

$$\mathsf{U}_{f_a}|a\rangle|x\rangle|y\rangle = |a\rangle|x\rangle|y \oplus f_a(x)\rangle, \tag{166}$$

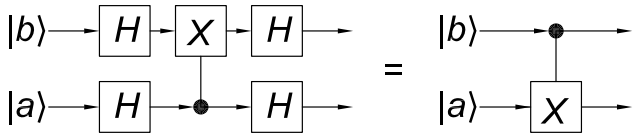
which reminds us of the Deutsch-Jozsa algorithm (prob. 10(f)).

(b) As in the Deutsch-Jozsa algorithm, it is advantageous to perform a quantum computation of $f_a(x)$ for the n-Qbit state

$$|x\rangle_n = \mathsf{H}^{\otimes n}|0\rangle_n = \frac{1}{2^{n/2}} \sum_{j=0}^{2^n-1} |j\rangle_n. \tag{167}$$

[68] http://kirkmcd.princeton.edu/examples/QM/bernstein_siamjc_26_1411_97.pdf

To appreciate the merits of eq. (167) for the present application, first demonstrate that the effect of surrounding a Controlled-NOT operation by Hadamard transformations, as in the figure below, is to swap the control and target bits.



Extend this trick to give a bit-flow diagram for a circuit that solves the Bernstein-Vazirani problem in a single quantum computation of the function $f_a(x)$.

(c) As we address more difficult problems in the future, we will rely more and more on algebraic understanding of the quantum computation. Use eq. (167) and various results from prob. 10(f) to give an algebraic analysis of your solution to the Bernstein-Vazirani problem.

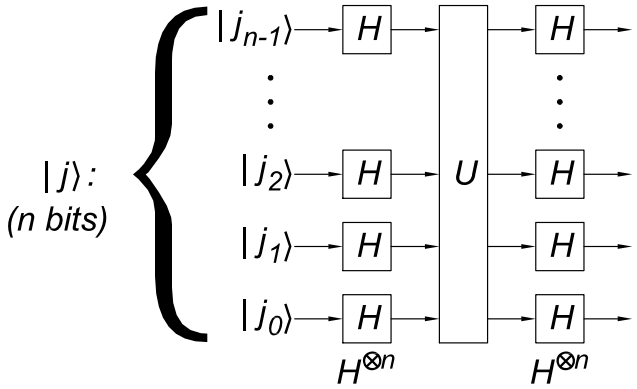
You may or may not find it useful to use/demonstrate the identity

$$\frac{1}{2^n}\sum_{j=0}^{2^n-1}(-1)^{j\odot k}(-1)^{a\odot j}=\delta_{ak}, \tag{168}$$

for which eqs. (16)-(17) might be helpful.

14. **Simon's Problem**

In both problems 10(f) and 13 we have found it advantageous to apply Hadamard gates on all lines of n-Qbit state before and after some operation U is performed.



An additional insight as to why this procedure might be useful is obtained from eq. (145). Since $-1 = e^{i\pi}$, we can write

$$\mathsf{H}^{\otimes n}|j\rangle_n = \frac{1}{2^{n/2}}\sum_{k=0}^{2^n-1}(-1)^{j\odot k}|k\rangle_n = \frac{1}{2^{n/2}}\sum_{k=0}^{2^n-1}e^{i\pi(j\odot k)}|k\rangle_n. \tag{169}$$

The second form of eq (169) is reminiscent of a Fourier transform.[69] Based on our experience with classical Fourier analysis, we anticipate that a quantum version of a Fourier transform will be a very powerful computational tool.

A further step towards quantum Fourier analysis was taken by Simon[70] who showed that a quantum algorithm is much faster than a classical one for the determination of a kind of periodicity of a function.

Suppose we have an n-to-n-bit function $f_a(x)$ that is not a one-to-one map, but rather a two-to-one map according to

$$f_a(x) = f_a(y) \qquad \text{iff} \qquad y = x \oplus a, \tag{170}$$

where the binary operation $\oplus$ as applied to multiple-bit states means bitwise addition modulo 2. We can say that the function f_a is periodic with period a in the restricted sense of eq. (170).

Simon's problem is to deduce the value of the "period" a in the minimum number of applications of the function f (assuming that the function f is a "black box" with the value of a hidden inside).

You may wish to convince yourself (without it being part of the written assignment) that $y = x \oplus a$ implies that $x \oplus y = a$, and that $x \oplus a \cdots \oplus a$ is either x or y depending on whether there are an even or odd number of additions.

[69] Indeed, this operation was called a discrete quantum Fourier transform in http://kirkmcd.princeton.edu/examples/QM/bernstein_siamjc_26_1411_97.pdf but we now reserve that name for the operation studied in prob. 17.

[70] http://kirkmcd.princeton.edu/examples/QM/simon_siamjc_26_1474_97.pdf

(a) Approximately how many evaluations of the classical n-bit function f_a would be required to deduce the value of the n-bit number a?

(b) Turning to a quantum computation involving Simon's n-bit function $f_a(x)$, we note that this function cannot be represented by a unitary transformation, since it is a two-to-one function. However, we can extend the trick of Deutsch's algorithm to embed f_a inside a unitary transformation U_{f_a} on $2n$ Qbits according to

$$\mathsf{U}_{f_a}|x\rangle_n|y\rangle_n = |x\rangle_n|y \oplus f_a(x)\rangle_n, \tag{171}$$

where $|x\rangle_n$ and $|y\rangle_n$ are each n-Qbit states.

You can readily convince yourself that U_{f_a} of eq. (171) is a reversible operation, and hence unitary.

Extrapolating from our experience with Deutsch's algorithm and the Bernstein-Vazirani problem, we anticipate that it will be useful to begin with states $|x\rangle_n$ and $|y\rangle_n$ as $|0\rangle_n$, and apply Hadamard transformations before and after operation U_{f_a} to all of the x and/or the y lines.

Explore the choice of applying the Hadamard transformations only to the x lines. Note that

$$\mathsf{U}_{f_a} \sum_{j=0}^{2^n-1} |j\rangle_n|0\rangle_n = \sum_{j=0}^{2^n-1} |j\rangle_n|f(j)\rangle_n, \tag{172}$$

and since f_a is a two-to-one function such that $f_a(x) = f_a(x \oplus a)$, we can also write eq. (172) as

$$\mathsf{U}_{f_a} \sum_{j=0}^{2^n-1} |j\rangle_n|0\rangle_n = \frac{1}{\sqrt{2}} \sum_{j=0}^{2^n-1} (|j\rangle_n + |j \oplus a\rangle_n)|f_a(j)\rangle_n. \tag{173}$$

Analyze the transformation

$$\mathsf{H}_x^{\otimes n}\mathsf{U}_{f_a}\mathsf{H}_x^{\otimes n}|0\rangle_x|0\rangle_y \tag{174}$$

to deduce a procedure to determine the value of a, based on repetitions of this operation, each ending with a measurement of the output state. That is, Simon's problem is solved by a probabilistic quantum computation, in contrast to the Deutsch-Jozsa problem and the Bernstein-Vazirana problem which can be solved by deterministic computation.

You should find that it would take only of order n repetitions, and hence the quantum solution to Simon's problem is exponentially faster than the classical solution. This result suggests that there can be spectacular advantages of quantum computation over classical computation for some types of problems.

The premise of this problem is that the function f_a has a period in the sense of eq. (170). A related problem would be to show whether or not a given n-to-n-bit function has such a period.

15. Grover's Search Algorithm

While Simon's problem shows that a quantum computation can be exponentially faster than its classical counterpart, one would probably not be willing to pay money to solve that particular problem. An example of a computational problem that is significant enough to attract funding, and for which a quantum solution is faster than a classical one, is the task of searching a list (a database) for a particular entry whose location is not known.[71]

We will cast this problem into the form of a "randomized" list of n-bit integers (a phone book?) in which we need to locate the integer a. We suppose that we have a function f_a that can recognize the desired integer according to

$$f_a(x) = \begin{cases} 0, & x \neq a, \\ 1, & x = a. \end{cases} \tag{175}$$

To locate the number a in our randomized list with certainty via classical computation using function f_a, we need to examine every entry in the list, *i.e.*, we must make 2^n evaluations of f_a.[72] If we make $m < n$ classical evaluations, the probability of success is, of course, only $P = m/n$.

(a) It should be no surprise by now that a quantum search for the number a will be assisted by the unitary function U_{f_a} defined by

$$\mathsf{U}_{f_a}|x\rangle_n|y\rangle = |x\rangle_n|y \oplus f_a(x)\rangle, \tag{176}$$

where $|x\rangle_n$ is an n-Qbit state and $|y\rangle$ is a single Qbit.

Give a bit-flow diagram of a possible implementation of the operation U_{f_a}. The circuit should include n lines for $|a\rangle$, so that it could be "programmed" to search for different values of a. That is, expand eq. (176) to

$$\mathsf{U}_{f_a}|a\rangle_n|x\rangle_n|y\rangle = |a\rangle_n|x\rangle_n|y \oplus f_a(x)\rangle. \tag{177}$$

A circuit for this exists that uses no ancillary lines.

(b) To search through all values of x as quickly as possible, we will certainly want to use the trick of eq. (130),

$$|\phi\rangle_n = \mathsf{H}^{\otimes n}|0\rangle_n = \frac{1}{2^{n/2}} \sum_{j=0}^{2^n-1} |j\rangle_n, \tag{130}$$

by initializing $|x\rangle$ to $|0\rangle_n$ and applying Hadamard transformations to all of the input lines.

---

[71] http://kirkmcd.princeton.edu/examples/QM/grover_prl_79_325_97.pdf
http://kirkmcd.princeton.edu/examples/QM/grover_prl_79_4709_97.pdf
http://kirkmcd.princeton.edu/examples/QM/grover_prl_80_4329_98.pdf

[72] More precisely, we need $2^n - 1$ classical evaluations of f to locate a with certainty, since if we have examined all but one of the entries in the list and not found a, we are certain that the remaining entry is a, provided we have the additional knowledge that the list is complete and without duplications.

Recall Deutsch's algorithm (Prob. 10) to show that an appropriate initialization of the auxiliary bit $|y\rangle$ leads to no change in $|y\rangle$ but to a "marking" of the desired state $|j = a\rangle_n$ after U_{f_a} is applied to the direct product state $|\phi\rangle_n|y\rangle$.

This kind of marking is a nonclassical effect that arises because the amplitude of a direct-product quantum state is a shared property of all of the components of that state. Of course, we cannot determine the details of an unknown quantum state in a small number of operations, so we must still find a way to exploit the "marking" in an efficient manner.

You should find that the "marking" consists of changing the sign of the amplitude of the state $|a\rangle_n$ in eq. (130), leaving all other $|j\rangle_n$ unchanged. A useful way to think about the "marking" process is that when $|y\rangle$ is properly prepared, the effect of U_{f_a} on $|x\rangle_n$ is a reflection in a 2^n-dimensional space about the (hyper)plane perpendicular to the desired state $|a\rangle_n$. That is, if we write

$$|\phi\rangle_n = \frac{1}{2^{n/2}}|a\rangle_n + \beta|b\rangle_n, \tag{178}$$

where $|b\rangle_n$ is orthogonal to $|a\rangle_n$, then

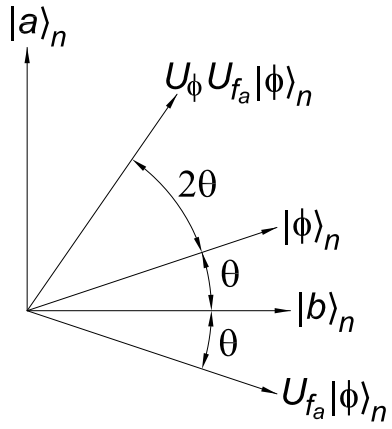
$$\mathsf{U}_{f_a}|\phi\rangle_n|y\rangle = \left[-\frac{1}{2^{n/2}}|a\rangle_n + \beta|b\rangle_n\right]|y\rangle. \tag{179}$$

If we observe either of the states $|\phi\rangle_n$ or $\mathsf{U}_{f_a}|\phi\rangle_n$,[73] the probability that we find it to be the desired state $|a\rangle_n$ is only $1/2^n$, so we have not immediately found a good search algorithm.

A geometric view of what we have accomplished is also very helpful.[74] The desired state $|a\rangle_n$ is nearly orthogonal to the state $|\phi\rangle_n$, since

$$\langle a|\phi\rangle_n = \frac{1}{2^{n/2}} = \cos(\pi/2 - \theta) = \sin\theta, \tag{180}$$

where the angle θ is a measure of the separation between the state $|\phi\rangle_n$ and the state $|b\rangle_n$ that is orthogonal to $|a\rangle_n$, as shown in the figure. The state $\mathsf{U}_{f_a}|\phi\rangle_n$ is the reflection of state $|\phi\rangle_n$ about the $|b\rangle_n$ axis.



[73] We suppress mention of the auxiliary bit $|y\rangle$ in much of the remainder of this problem.

[74] The geometric view given here of Grover's algorithm first appeared in the useful review
http://kirkmcd.princeton.edu/examples/QM/aharonov_quant-ph-9812037.pdf
based on an algebraic argument given in
http://kirkmcd.princeton.edu/examples/QM/boyer_fphys_46_493_98.pdf

(c) Grover realized that further reflections in the state-space of $|x\rangle_n$ provide a means of strengthening the "marking" of the desired state $|a\rangle_n$, such that after $2^{n/2}$ reflections a measurement of the evolved state of $|x\rangle_n$ will result in state $|a\rangle_n$ with high probability.

Indeed, we see from the figure that if we reflect the state $\mathsf{U}_{f_a}|\phi\rangle_n$ about the $|\phi\rangle_n$ axis, using the operation $\mathsf{U}_{|\phi\rangle_n}$, the final state has been rotated by angle 2θ towards the $|a\rangle_n$ axis.

If we iterate this pair of reflections about the $|b\rangle_n$ and $|\phi\rangle_n$ axes m times, then the state $|\phi\rangle_n$ will have been rotated to angle $(2m+1)\theta$ with respect to the $|b\rangle_n$ axis. If the final angle is close to $\pi/2$, then a measurement of the state will yield $|a\rangle_n$ with high probability, as desired. Thus, we need

$$(2m+1)\theta \approx \frac{2m+1}{2^{n/2}} \approx \frac{\pi}{2} \approx 2, \tag{181}$$

That is, with $m = 2^{n/2} = \sqrt{N}$ iterations of the transformation $\mathsf{U}_{|\phi\rangle_n}\mathsf{U}_{f_a}$, where $N = 2^n$ is the size of the list, a measurement of the state $(\mathsf{U}_{|\phi\rangle_n}\mathsf{U}_{f_a})^m|\phi\rangle_n$ is almost certain to be $|a\rangle_n$. The probability of failure is approximately $1/2^n$, so if we fail to find the state $|a\rangle_n$ at the end of the first set of iterations, we repeat the entire procedure. The probability of failing twice is $1/2^{2n}$, *etc.* Thus, Grover's search algorithm reduces the task of finding an item in a randomized list of length N from $\approx N$ classical samplings of the list to only $\approx \sqrt{N}$ quantum samplings.

Grover's procedure does not converge for the case that $n = 1$, i.e., when the list contains only 2 items. But then, we can find an item classically with only a single look into the list. However, for $n = 2$, i.e., a list of 4 items, $\sin\theta = 1/\sqrt{2^2} = 1/2$ so $\theta = 30°$. We see from the figure on the previous page that a single iteration brings the state $|\phi\rangle_2$ exactly onto $|a\rangle_2$. Grover's quantum search finds 1 item in a list of 4 in a single "enquiry".

It remains to construct the reflection $\mathsf{U}_{|\phi\rangle_n}$, based on the state $|\phi\rangle_n$, which has the effect on a general state

$$|\psi\rangle_n = \sum_{j=0}^{N-1} \psi_j |j\rangle_n, \tag{182}$$

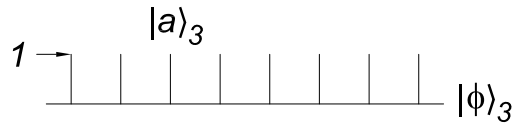
where $N = 2^n$, of reflecting it about the $|\phi\rangle_n$ "axis". Decompose the state (182) into its component "along" state $|\phi\rangle_n$ and the remainder (which is thereby orthogonal to $|\phi\rangle_n$), introducing the notation

$$\left\langle \psi_j \right\rangle = \frac{1}{N}\sum_{j=0}^{N-1} \psi_j = \frac{1}{\sqrt{N}}\langle\phi|\psi\rangle_n. \tag{183}$$

Then, consider the form of $\mathsf{U}_\phi|\psi\rangle_n$ to express $\mathsf{U}_{|\phi\rangle_n}$ as a kind of projection operator involving the state $|\phi\rangle_n$. Show that the effect of reflecting $|\psi\rangle_n$ about $|\phi\rangle_n$ is to change the sign of each amplitude $\psi_j - \left\langle \psi_j \right\rangle$, which was called a "reflection about the mean" by Grover.

Sketch the steps of the appropriate set of iterations for the case that $n = 3$, starting with the state $|\phi\rangle_3$ represented as a "comb" of 8 basis states of equal

amplitude, as shown below. What is the probability of success of Grover's search algorithm in this case?



Show also that the transformation U_{f_a} can be expressed via a projection operator involving state $|a\rangle_n$.

(d) While we have expressed the transformation $\mathsf{U}_{|\phi\rangle_n}$ as a projection operator, this does not immediately tell us how we could construct this operator out of elementary quantum gates.

Since operator $\mathsf{U}_{|\phi\rangle_n}$ is a reflection about $|\phi\rangle_n$ and

$$|\phi\rangle_n = \mathsf{H}^{\otimes n}|0\rangle_n, \tag{130}$$

we can relate the reflection about $|\phi\rangle_n$ to the reflection $\mathsf{U}_{|0\rangle_n}$ about $|0\rangle_n$ according to

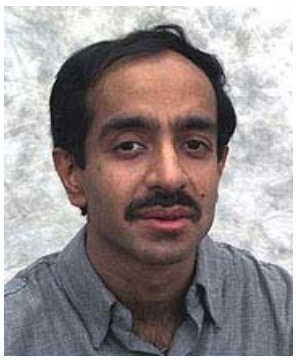
$$\mathsf{U}_{|\phi\rangle_n} = \mathsf{H}^{\otimes n}\mathsf{U}_{|0\rangle_n}\mathsf{H}^{\otimes n}. \tag{184}$$

Equation (184) may be more self evident using the representation of $\mathsf{U}_{|\phi\rangle_n}$ *in terms of the projection operator that you have deduced in part (c).*

Further, Grover's search algorithm is unaffected if the amplitudes of all states are multiplied by -1. That is, the operator $-\mathsf{U}_{|\phi\rangle_n}$ could be used instead of $\mathsf{U}_{|\phi\rangle_n}$. Since the reflection $\mathsf{U}_{|\phi\rangle_n}$ changes the sign of the amplitude of any state that is orthogonal to $|\phi\rangle_n$ while leaving the amplitude of the state $|\phi\rangle_n$ unchanged, operator $-\mathsf{U}_{|\phi\rangle_n}$ changes the sign of the amplitude of the state $|\phi\rangle_n$ but does not change the amplitude of any state that is orthogonal to $|\phi\rangle_n$. Thus, $-\mathsf{U}_{|\phi\rangle_n}$ is a reflection about the plane perpendicular to state $|\phi\rangle_n$.

Combining this insight with eq. (184), we see that we need the operator that describes a reflection about the plane perpendicular to the state $|0\rangle_n$. Recalling that the operator U_{f_a} of part (a) is the reflection about the plane perpendicular to the state $|a\rangle_n$ ($\mathsf{U}_{f_a} = -\mathsf{U}_{|a\rangle_n}$), we see that using operator U_{f_0} in eq. (184) provides us with the desired construction of $\mathsf{U}_{|\phi\rangle_n}$.

Simplify your construction of U_{f_a} for the particular case that $|a\rangle_n = |0\rangle_n$. Summarize briefly how the m iterations of Grover's search algorithm can be built up from an appropriate initial state, followed by a minimal set of products of operators.



16. **Parity of a Function**

At this point we might be optimistic that all classical computations could be performed significantly faster by an appropriate quantum computation. However, this appears not to be so,[75] as we will illustrate in this problem.

Consider an n-bit to 1-bit function $f(x)$, *i.e.*, $f(x) = 0$ or 1. The parity Π_f of function f is defined as[76]

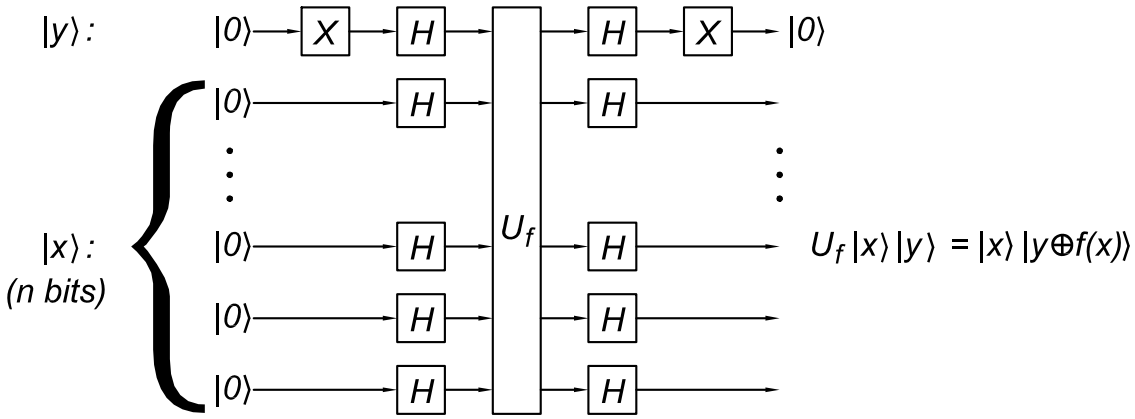
$$\Pi_f = \prod_{j=0}^{2^n-1} (-1)^{f(j)} = (-1)^{\sum_{j=0}^{2^n-1} f(j)}. \tag{185}$$

A classical computation of Π_f requires 2^n evaluations of the n-bit function f.

For a quantum computation of Π_f it seems appropriate to introduce the unitary function U_f defined by

$$\mathsf{U}_f|x\rangle_n|y\rangle = |x\rangle_n|y \oplus f(x)\rangle,) \tag{186}$$

that we first encountered in the Deutsch-Jozsa algorithm (prob. 10(f)).



Show that use of the above circuit permits a determination of the parity of f in a single use of U_f for the case that $n = 1$. This is a good start, since 2 evaluations of f are needed for a classical computation of Π_f when $n = 1$.

Hint: Note that $(-1)^f = (-1)^{-f}$ *when* $f = 0$ *or 1 only. Thus,*

$$(-1)^{f(1)} = (-1)^{f(0)}(-1)^{f(1)-f(0)} = (-1)^{f(0)}(-1)^{f(1)+f(0)} = (-1)^{f(0)}\Pi_f. \tag{187}$$

The example of $n = 1$ affords a glimpse that an application of U_f can be used to group the phase factors $(-1)^{f(x)}$ into pairs. To build up Π_f in this manner, the phase $(-1)^{f(0)}$ can be grouped together with the phases $(-1)^{f(1)}.....(-1)^{f(2^{n-1}-1)}$ during 2^{n-1} applications of U_f (together with a shift operator S_n) into the amplitude of the first

[75] http://kirkmcd.princeton.edu/examples/QM/farhi_prl_81_5442_98.pdf
http://kirkmcd.princeton.edu/examples/QM/beals_quant-ph-9802049.pdf

[76] Note that there are two technical meanings to the term parity. The computational definition, used in this problem, relates to whether a bit string contains an even or odd numbers of 1's. The quantum mechanical definition (due to Wigner) relates to the sign of the eigenvalue of the transformation of space inversion ($\mathbf{r} \to -\mathbf{r}$) that was mentioned in prob. 4(e).

2^{n-1} basis states, while at the same time the product of the second 2^{n-1} phase factors becomes the amplitude of the second 2^{n-1} basis states:

$$|0\rangle_n \to |\psi\rangle_n = \frac{1}{2^{n/2}} \prod_{k=0}^{2^{n-1}-1} (-1)^{f(k)} \sum_{j=0}^{2^{n-1}-1} |j\rangle_n + \frac{1}{2^{n/2}} \prod_{k=2^{n-1}}^{2^n-1} (-1)^{f(k)} \sum_{j=2^{n-1}}^{2^n-1} |j\rangle_n. \quad (188)$$

We see that if $\Pi_f = 1$, then

$$|\psi\rangle_n = \frac{1}{2^{n/2}} \sum_{j=0}^{2^n-1} |j\rangle_n = \mathsf{H}^{\otimes n}|0\rangle_n = |\phi\rangle_n, \qquad (\Pi_f = 1), \quad (189)$$

recalling eq. (130), so that a final Hadamard transformation will return this state to $|0\rangle_n$,

$$\mathsf{H}^{\otimes n}|\psi\rangle_n = |0\rangle_n, \qquad (\Pi_f = 1). \quad (190)$$

On the other hand, if $\Pi_f = -1$, then $|\psi_n\rangle$ is orthogonal to $|\phi\rangle_n$, and the final Hadamard transformation will bring it to some state orthogonal to $|0\rangle_n$. Thus, a measurement of the final state reveals the parity of function f with certainty ($\Pi_f = 1$ if the final state is observed to be $|0\rangle_n$, and $\Pi_f = 1$ otherwise), but only after 2^{n-1} evaluations of the function f.

For details, and a demonstration that Π_f cannot be determined in less than 2^{n-1} evaluations of f, see the references in the footnote on the previous page, sec. 6.7 of Preskill's lectures, or sec. 6.7 of Nielsen and Chuang.

For the record, the needed shift operator S_n (which is a generalization of the two-bit SWAP operator S_{ab} of prob. 6(c)) obeys

$$\mathsf{S}_n|j\rangle_n = \begin{cases} |j+1\rangle_n, & j = 0, ..., 2^{n-1}-2, \\ |0\rangle_n, & j = 2^{n-1}-1, \\ |j+1\rangle_n, & j = 2^{n-1}, ..., 2^n-2, \\ |2^{n-1}\rangle_n, & j = 2^n-1, \end{cases} \quad (191)$$

and the full algorithm can be summarized as

$$(\mathsf{H}_x^{\otimes n} \otimes \mathsf{X}_y\mathsf{H}_y)(\mathsf{U}_f\mathsf{S}_n)^{2^{n-1}-1}\mathsf{U}_f(\mathsf{H}_x^{\otimes n} \otimes \mathsf{H}_y\mathsf{X}_y)|0\rangle_x|0\rangle_y = \frac{1+\Pi_f}{2}|0\rangle_n|0\rangle_y + \frac{1-\Pi_f}{2}|a\rangle_n|0\rangle_y, \quad (192)$$

where $|a\rangle_n$ is a normalized state orthogonal to $|0\rangle_n$.

Bottom line: There are some computations for which quantum algorithms provide no advantage classical ones.[77]

[77] This negative conclusion makes the success of Grover's search algorithm (for a list of 2^n entries) all the more impressive. Part of that success lies in the fact that Grover's algorithm is only (highly) probably successful after its $m \approx 2^{n/2}$ iterations. A guarantee of success, however, cannot be given unless there are at least 2^{n-1} evaluations of the function $f(x)$, which is essentially the same guarantee as available in a classical computation.

17. Quantum Fourier Transform, Shor's Period-Finding Algorithm

We now elaborate on the theme of a quantum Fourier transform, which was mentioned in prob. 14.[78]

The Hadamard transformation

$$\mathsf{H}^{\otimes n}|j\rangle_n = \frac{1}{2^{n/2}}\sum_{k=0}^{2^n-1}(-1)^{j\odot k}|k\rangle_n = \frac{1}{2^{n/2}}\sum_{k=0}^{2^n-1}e^{i\pi(j\odot k)}|k\rangle_n. \tag{169}$$

acts something like a Fourier transform, taking a single state $|j\rangle_n$ into a weighted sum of all basis states.

We recall that the Fourier transform $\tilde{f}_k$ of a continuous (complex) function $f(x)$ on an interval [0,a] is given by

$$f(x) = \frac{1}{\sqrt{a}}\sum_k \tilde{f}_k e^{-2\pi ikx/a}, \qquad \tilde{f}_k = \frac{1}{\sqrt{a}}\int_0^a f(x)e^{2\pi ikx/a}dx. \tag{193}$$

If the function f is instead an 2^n-dimensional vector f_j, $j = 0, 1, ..., 2^n - 1$, then we consider the discrete Fourier transform, which can be extrapolated from eq. (193) to be

$$f_j = \frac{1}{2^{n/2}}\sum_{k=0}^{2^n-1}\tilde{f}_k e^{-2\pi ijk/2^n}, \qquad \tilde{f}_k = \frac{1}{2^{n/2}}\sum_{j=0}^{2^n-1}f_j e^{2\pi ijk/2^n}. \tag{194}$$

A quantum version of eq. (194) considers the function f to be an n-bit state $|f\rangle_n$,

$$|f\rangle_n = \sum_{j=0}^{2^n-1} f_j|j\rangle_n, \tag{195}$$

that is acted upon by the quantum Fourier transform $\mathbf{\Phi}$ (which surely must be a unitary operator) to produce

$$\begin{aligned} |\tilde{f}\rangle_n &= \sum_{k=0}^{2^n-1}\tilde{f}_k|k\rangle_n = \sum_{k=0}^{2^n-1}\frac{1}{2^{n/2}}\sum_{j=0}^{2^n-1}f_j e^{2\pi ijk/2^n}|k\rangle_n \\ &= \mathbf{\Phi}|f\rangle_n = \sum_{j=0}^{2^n-1} f_j\mathbf{\Phi}|j\rangle_n \end{aligned} \tag{196}$$

which implies that the effect of $\mathbf{\Phi}$ on a basis state $|j\rangle_n$ is

$$\mathbf{\Phi}|j\rangle_n = \frac{1}{2^{n/2}}\sum_{k=0}^{2^n-1}e^{2\pi ijk/2^n}|k\rangle_n. \tag{197}$$

(a) Re-express the transformation (197) in terms of basis states written as direct products, using eqs. (16)-(17) and the expansion of an n-bit binary number k and $\sum_0^{n-1} k_l 2^l$ where $k_l = 0$ or 1.

[78] A good self-contained discussion of Shor's algorithm is given in
http://kirkmcd.princeton.edu/examples/QM/gerjuoy_ajp_73_521_05.pdf

(b) Transcribe the result of part (a) into a bit-flow diagram. The input lines can be "overwritten" to become the output lines.

Hint: Your result from part (a) should show a simple relation between the highest-order bit of the output state and the lowest-order bit of the input state. It suffices to "read" the output lines in reverse order to the input lines, eliminating a tedious n-bit SWAP *operation.*[79] *Recall eq. (61), and use a small variation on the spirit of prob. (12) to express any needed 1-bit gates as powers of* H *and* $\mathsf{Z} = (\boldsymbol{\sigma}_z^{1/4})^4$.

(c) Give a bit-flow diagram for the operation $\boldsymbol{\Phi}^{-1}$ that performs the inverse Fourier transform,

$$\boldsymbol{\Phi}^{-1}|\tilde{f}\rangle_n = |f\rangle_n. \tag{198}$$

(d) **Shor's Period-Finding Algorithm.**[80]

While the quantum Fourier transform requires fewer gates than the classical discrete Fourier transform, and so is faster in a sense, its output is a quantum state whose value cannot be well determined in a small number of repetitions. The quantum Fourier transform will never be an all-purpose replacement for the classical discrete Fourier transform. We are left with the task of finding special problems of interest for which the quantum Fourier transform is a better solution than the classical Fourier transform.

As a first example, consider the problem of finding the period a of an n-to-n-bit function $f(x)$ that obeys

$$f(y) = f(x), \qquad \text{if} \qquad y = x + ma, \tag{199}$$

where a and m are integers ($< 2^{n-2}$). Unlike, say, a violin, the function f has only a single period. The simplicity of this function may permit a quantum analysis to be advantageous.

This problem is similar to Simon's problem (prob. 14), except that now the condition of periodicity involves ordinary addition, which is more complicated than addition modulo 2.

As in Simon's problem, consider the unitary operator U_f that links two n-bit states $|x\rangle_n$ and $|y\rangle_n$ together with the function $f(x)$ according to

$$\mathsf{U}_f|x\rangle_n|y\rangle_n = |x\rangle_n|y \oplus f(x)\rangle_n. \tag{171}$$

And also as in Simon's problem, we begin by applying Hadamard transformations to the x lines to produce the state

$$\mathsf{U}_f\mathsf{H}_x^{\otimes n}|0\rangle_x|0\rangle_y = \frac{1}{2^{n/2}}\sum_{j=0}^{2^n-1}|j\rangle_x|f(j)\rangle_y. \tag{200}$$

Because function f is periodic with period a, there are only a distinct states of $|f(j)\rangle_y$. Therefore, the sum in eq. (200) can be rearranged into a groups each

[79] In the following problem we consider an alternative algorithm that builds in the SWAP operation such that all two-Qbit gates involve neighboring lines.

[80] http://kirkmcd.princeton.edu/examples/QM/shor_ieeefcs_35_124_94.pdf
http://kirkmcd.princeton.edu/examples/QM/shor_siamjc_26_1484_97.pdf

being the sum of b states, where

$$b = \text{int}(2^n/a) + c, \qquad c = \begin{cases} 0, & \text{if } a = 2^d \text{ for integer } d < n, \\ 1, & \text{otherwise.} \end{cases} \tag{201}$$

Thus,

$$\mathsf{U}_f \mathsf{H}_x^{\otimes n} |0\rangle_x |0\rangle_y = \frac{1}{\sqrt{a}} \sum_{j=0}^{a-1} \frac{1}{\sqrt{b}} \sum_{m=0}^{b-1} |j + ma\rangle_x |f(j)\rangle_y. \tag{202}$$

Rather than applying a second Hadamard transformation to the state (202) (as in Simon's problem), we now apply the quantum Fourier transform $\mathbf{\Phi}_x$ to the x lines, and then measure the result.

Expand the operation

$$\mathbf{\Phi}_x \mathsf{U}_f \mathsf{H}_x^{\otimes n} |0\rangle_x |0\rangle_y \tag{203}$$

using eq. (197), rearranging the terms so that the sum over index m can be performed explicitly. Express the remaining coefficients as phase factors $e^{i\phi}$ times real numbers. Show that a measurement of the resulting state of $|x\rangle_n$ yields numerical values of x with relatively high probability only if x has a simple relation to the period a of the function f, such that the period can be deduced with only a few repetitions of the operation (203).

You need not worry about the special case that the period a is a power of 2 (unless, of course, you insist).

The RSA Public Encryption Protocol.[81]

The RSA (Rivest, Shamir, Adelman) encryption protocol is based on two (large) prime integers p and q whose product $N = pq$ is less than 2^n.

The codemaster Bob publicizes the integer N and another integer $c < N$ that is coprime with $(p-1)(q-1)$, meaning that c has no common factors with $p-1$ or $q-1$. Bob is then able to calculate, but he does not reveal, the number d that obeys

$$cd = 1 \pmod{(p-1)(q-1)} = 1 + e(p-1)(q-1), \tag{204}$$

for some integer e. *This follows from the fact that the integers that are coprime with another integer M form a group under multiplication modulo M, which implies in particular that every member of the group has a multiplicative inverse modulo M, as in eq. (204).*[82]

Anyone, say, Alice, who wishes to send a "secret" message to Bob, converts her message to a (series of) number(s) $j < N$ by some well-known scheme (such as ASCII encoding). *If the number j is not coprime with N, Alice adds 1 to it so that it is. The determination of whether two numbers, such as j and N have common factors is readily determined classically by a classic algorithm due to Euclid.* Alice further codes the message j according to

$$k = j^c \pmod{N}, \tag{205}$$

[81] Two popular books on codes are *The Codebreakers* by David Kahn, and *The Code Book* by Simon Singh.
[82] More details are given in Appendices 4 and 5 of Nielsen and Chuang.

and makes the number k public.

Alice and Bob, both being cryptographers, are aware of Fermat's Little Theorem which states that if j is not a multiple of p then

$$j^{p-1} = 1 \pmod{p} \qquad \text{(Fermat)}. \tag{206}$$

Since Alice chose j to be coprime with $N = pq$, neither p nor q are among its factors, and so we also have

$$j^{q-1} = 1 \pmod{q}. \tag{207}$$

Therefore

$$\begin{aligned} j^{(p-1)(q-1)} &= (j^{q-1})^{p-1} = 1 \pmod{p} = 1 + lp \\ &= (j^{p-1})^{q-1} = 1 \pmod{q} = 1 + mq \\ &= 1 + hpq = 1 \pmod{pq} \\ &= 1 \pmod{N}, \end{aligned} \tag{208}$$

since we have $lp = mq$ where p and q are prime, and so we must have that $l = hq$ and $m = hp$ for some integer h.

Bob takes the number k of eq. (205) and performs the calculation

$$\begin{aligned} k^d \pmod{N} &= (j^c)^d \pmod{N} = j^{cd} \pmod{N} \\ &= j^{1+e(p-1)(q-1)} \pmod{N} = j(j^{(p-1)(q-1)})^e \pmod{N} \\ &= j \pmod{N}, \end{aligned} \tag{209}$$

noting eq. (208). Thus, Bob has decoded Alice's message using his knowledge of the prime factors p and q of N.

This is the RSA public key encryption scheme, which is supposed to be practically secure for large values of p and q because of the difficulty of factoring large prime numbers. For example, it is claimed that factoring a 1000-digit number pq using present classical computers would take longer than the age of the Universe.

However, if we consider the function

$$f(x) = k^x \pmod{N}, \tag{210}$$

based on Alice's public message k and Bob's public number N, we see that it has period a, which is the smallest integer such that

$$k^a = 1 \pmod{N}. \tag{211}$$

That is, $f(x + a) = k^{x+a} \pmod{N} = k^x k^a \pmod{N} = k^x \pmod{N} = f(x)$.

A marvelous group-theoretic factoid is that the relation $k = j^c \pmod{N}$ implies that

$$j^a = 1 \pmod{N}. \tag{212}$$

So, a codebreaker (or eavesdropper, commonly designated as Eve) could use Shor's period-finding algorithm to deduce the number a from public information. Eve, who

also knows Bob's public number c, can now use a classical computer to calculate the number d' that obeys

$$cd' = 1 \text{ (mod } a) = 1 + ga, \tag{213}$$

for some integer g. She then calculates

$$\begin{aligned} k^{d'} \text{ (mod } N) &= (j^c)^{d'} \text{ (mod } N) = j^{cd'} \text{ (mod } N) \\ &= j^{1+ga} \text{ (mod } N) = j(j^a)^g \text{ (mod } N) \\ &= j \text{ (mod } N), \end{aligned} \tag{214}$$

noting eq. (212). Thus, Eve, with her quantum computer, has broken the supposedly unbreakable RSA encryption scheme.

Of course, Eve's quantum computer needs twice as many bits as there are in the number N. If N has 1000 decimal digits, it has 3322 binary digits, so the quantum computer need 6644 Qbits. The largest quantum computer to date has 7 Qbits.

Furthermore, the quantum Fourier transform requires two-bit gates that couple Qbits that are far removed from one another.

A "natural" homework assignment would be to construct the operation U_f now that we know what the relevant function f is. However, this exercise turns out to be considerably more involved that the construction of $\mathbf{\Phi} = \mathsf{U}_{\rm FT}$, and we defer it to the following problem.

Factoring

Shor's algorithm is often described as providing a fast way to factor large numbers. But so far we have avoided the explicit task of factoring Bob's number N, even though we have succeeded in breaking the RSA code based on it. Here we give a very brief sketch of how Alice (or Eve) might factor Bob's number $N = pq$.[83]

Alice again starts with a number j that is coprime with N, but now she simply chooses this at random. With her quantum computer she deduces the period a of the function $f(x) = j^x \text{ mod } N$. As in eq. (211), the period a is the smallest integer such that

$$j^a = 1 \text{ (mod } N). \tag{215}$$

If a turns out to be odd, she tries another j until the corresponding a is found to be even. She then checks the value of $j^{a/2}$ (mod N), and iterates until this value is not -1, *i.e.*, not $N-1$. The claim is that the probability of choosing a "good" value of j on the first try is greater than 50%.

The meaning of "good" is illustrated by the fact that the following identity is now nontrivial,

$$0 = j^a - 1 \text{ (mod } N) = (j^{a/2} - 1)(j^{a/2} + 1) \text{ (mod } N). \tag{216}$$

Since a is the smallest number such that eq. (215) holds, neither $j^{a/2} - 1$ nor $j^{a/2} + 1$ is a multiple of N, but their product is. As $N = pq$ where p and q are prime, it must be that p is a factor of $j^{a/2} - 1$ and q is a factor of $j^{a/2} + 1$, or vice versa. Then, p (or

[83] Greater detail is given in sec. 5.3 of Nielsen and Chuang, and in Appendix 3 to Mermin's Lecture 3.

q) is the greatest common divisor of N and $j^{a/2}-1$, as can be extracted efficiently by a classical algorithm due to Euclid.

Accuracy of the Quantum Fourier Transform

The quantum Fourier transform $\mathbf{\Phi}|f\rangle_n$ of an n-Qbit function f, where $|f\rangle_n = \sum_{j=0}^{2^n-1} f_j|j\rangle_n$, is

$$\mathbf{\Phi}|f\rangle_n = \sum_{k=0}^{2^n-1} \frac{1}{2^{n/2}} \sum_{j=0}^{2^n-1} f_j e^{2\pi i jk/2^n}|k\rangle_n = \sum_{k=0}^{2^n-1} \frac{1}{2^{n/2}} \sum_{j=0}^{2^n-1} f_j e^{i\phi_{jk}}|k\rangle_n, \tag{196}$$

which involves 2^{2n} phases $\phi_{jk} = 2\pi jk/2^n$ most of which are extremely tiny. In practice it will be impossible to construct accurately the quantum phase gates $\mathsf{Z}^{1/2^m}$ for m of order n as needed to implement these phase factors in the quantum Fourier transform (according to your solution to part (b).).

However, as discussed in part (d), Shor's use of the quantum Fourier transform is only to find the period of a function with high probability. In particular, if the function f has a period a, the probability of observing a state $|k\rangle_n$ is large (and of order $1/a$) only for a set of a of these states. So, the practical question is to what extent "errors" in the phase factors of eq. (197) alter the probability

$$P(k) = \frac{1}{2^n}\left|\sum_{j=0}^{2^n-1} f_j e^{i\phi_{jk}}\right|^2, \tag{217}$$

that a state $|k\rangle_n$ would be observed during the measurement of the Fourier transformed state (196) at the end of the period-finding algorithm.

Suppose that each of the phases ϕ_{jk} is altered by a small amount δ_{jk}, and δ is the maximum of the $|\delta_{jk}|$ Then, $e^{i\phi_{jk}} \to e^{i(\phi_{jk}+\delta_{jk})} \approx e^{i\phi_{jk}}(1+i\delta_{jk})$, and

$$\begin{aligned} P(k) \;&\to\; \frac{1}{2^n}\left|\sum_{j=0}^{2^n-1} f_j e^{i(\phi_{jk}+\delta_{jk})}\right|^2 \approx \frac{1}{2^n}\left|\sum_{j=0}^{2^n-1} f_j e^{i\phi_{jk}}(1+i\delta_{jk})\right|^2 \\ &\approx\; P(k) + \frac{2}{2^n} Im\left[\sum_{j=0}^{2^n-1} \delta_{jk} f_j^\star e^{-i\phi_{jk}} \sum_{j'=0}^{2^n-1} f_{j'} e^{i\phi_{j'k}}\right]. \end{aligned} \tag{218}$$

Now,

$$\begin{aligned} \frac{2}{2^n}\left|Im\left[\sum_{j=0}^{2^n-1} \delta_{jk} f_j^\star e^{-i\phi_{jk}} \sum_{j'=0}^{2^n-1} f_{j'} e^{i\phi_{j'k}}\right]\right| \;&\le\; \frac{2}{2^n}\left|\sum_{j=0}^{2^n-1} \delta_{jk} f_j^\star e^{-i\phi_{jk}} \sum_{j'=0}^{2^n-1} f_{j'} e^{i\phi_{j'k}}\right| \\ &\le\; \frac{2\delta}{2^n}\left|\sum_{j=0}^{2^n-1} f_j e^{i\phi_{jk}}\right|^2 = 2\delta P(k), \end{aligned} \tag{219}$$

so the effect of the "errors" δ_{jk} is to multiply the relevant probabilities by a factor bounded by $1 \pm 2\delta$.

A corollary is that it is not necessary to implement the (controlled) phase gates $\mathsf{Z}^{1/2^m}$ for $m > 10$ if we are content with the probabilities of the period-finding algorithm to be accurate to 1%.[84]

[84] This was first discussed by D. Coppersmith,
http://kirkmcd.princeton.edu/examples/QM/coppersmith_ibm_rc19642_94.pdf

18. Nearest-Neighbor Algorithms

The interest in codebreaking is so great that substantial effort has gone into exploration of variations on Shor's period-finding algorithm and into its potential implementation via quantum gates. In this course we do not have time to review such related themes as discrete logarithms, hidden subgroups, order-finding, and phase estimation, which are treated in Nielsen and Chuang.

Turning to the issue of implementation of Shor's algorithm, there are two broad choices:

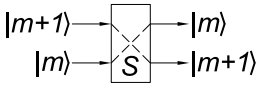
A. "Fast" algorithm's that involve a minimal number of steps in the computation, i.e., smaller depth of the quantum circuit, typically at the expense of requiring larger numbers of Qbits.[85]

B. Low-Qbit-count algorithms, which typically require greater circuit depth.[86]

An additional issue is that the "standard" algorithms that we have explored for n-bit problems all require two-bit gates that relate "distant" Qbits.

While the optical Controlled-NOT gates studied in prob. 7(e) might indeed work equally well for any pair of Qbits, this is not the case for a large class of proposed quantum computers in which Qbits are encoded on the spins of electrons, atoms or molecules. In the spin-based systems, interactions between Qbits are due to their magnetic dipole couplings, whose strength falls off with distance as $1/r^3$. If these systems are to serve as viable quantum computers, the computations they perform should involve only nearest-neighbor gates.

In general, a two-bit operation $\mathsf{U}_{j,j+k}$ involving bits j and $j+k$ that are members of a linearly ordered set could be performed by a sequence of $k-1$ nearest-neighbor SWAP operations, $\mathsf{S}_{m,m+1}$ (prob. 6(c)),



followed by the nearest-neighbor operation $\mathsf{U}_{j,j+1}$,

$$\mathsf{U}_{j,j+k} = \mathsf{S}_{j+k-1,j+k}\mathsf{S}_{j+k-2,j+k-1}\cdots\mathsf{S}_{j+1,j+2}\mathsf{U}_{j,j+1}. \tag{220}$$

The swapping of locations of the bits might increase the distance between pairs of bits used in later two-bit operators, requiring ever larger numbers of SWAP operations if further operations are to involve only neighboring bits. Fortunately, this is not the case for circuits that implement Shor's algorithm.

[85] http://kirkmcd.princeton.edu/examples/QM/zalka_quant-ph-9806084.pdf
http://kirkmcd.princeton.edu/examples/QM/gossett_quant-ph-9808061.pdf
http://kirkmcd.princeton.edu/examples/QM/cleve_quant-ph-0006004.pdf

[86] http://kirkmcd.princeton.edu/examples/QM/beauregard_quant-ph-0205095.pdf
http://kirkmcd.princeton.edu/examples/QM/fowler_qic_4_237_04.pdf

(a) **Fourier Transform.**

Give a bit-flow diagram of a circuit for the quantum Fourier transform $\mathbf{\Phi}$ (prob. 17(b)) that involves only nearest-neighbor gates. By associating one nearest-neighbor SWAP operation with each of the needed Controlled-Z^p gates, you should be able to implement the latter as nearest-neighbor gates also. And, you should be able to arrange the swaps so that the order of the output lines is reversed, bringing them into alignment with the input lines.

Begin by giving a diagram for the case of 3 Qbits, and verify that your circuit works well for 4 Qbits.

Your diagram should show groups of gates on adjacent pairs of lines. It will be possible in some hardware realizations of the circuit to coalesce each of these logical groups into a single physical two-Qbit gate, so the nearest-neighbor algorithm will not have excessive depth.

(b) **Fourier Addition.**

In prob. 9(g) we considered a quantum circuit that adds two numbers a and b. Some (slight) complexity of that circuit was needed to deal with "carrying" when adding bits that are both $|1\rangle$. Expand the quantum Fourier transform $\mathbf{\Phi}|a+b\rangle_n$ in a manner similar to that for $\mathbf{\Phi}|j\rangle_n$ in prob. 17(a) and draw a nearest-neighbor circuit that we will call $\mathbf{\Phi}_a^+(b)$, which uses $|a\rangle_n$ and $\mathbf{\Phi}|b\rangle_n$ as its inputs, overwriting the latter to become $\mathbf{\Phi}|a+b\rangle_n$ without any "carry" operations. Note that in this procedure, numbers a and b must have only $n-1$ bits, while their sum $a+b$ may have n bits.

Thus, at the expense of performing quantum Fourier transforms and their inverse, we obtain an adder circuit that uses $2n+2$ bits to add two n-bit numbers, in contrast to the use of $3n+1$ bits in the circuit of prob. 9(g).

It suffices to give a circuit for the case that a and b are 3-bit numbers. Of course, the sum then has 4 bits. Before deducing a nearest-neighbor version of the circuit, it may be useful to give a non-nearest-neighbor circuit modeled after your solution to prob. 17(a).

(c) **Fourier Subtraction.**

By $\mathbf{\Phi}_a^-(b)$ we mean the operation that uses $|a\rangle_n$ and $\mathbf{\Phi}|b\rangle_n$ as inputs, where a and b are $(n-1)$-bit numbers, and applies the inverse of the gates of operation $\mathbf{\Phi}_a^+(b)$ in reverse order. Using your expansion for the Fourier transform $\mathbf{\Phi}|a+b\rangle_n$, deduce in what sense operation $\mathbf{\Phi}_a^-(b)$ performs a Fourier subtraction.

Subtraction can result in negative numbers, which have been avoided in this course up until now. When dealing with binary subtraction, we recall the concepts of 1's and 2's complement arithmetic. *While the sum of two $(n-1)$-bit numbers can involve n bits, that sum can never reach the largest n-bit number, namely $\sum_{l=0}^{n-1} 2^l = 2^n - 1$. This permits the negative of an $(n-1)$-bit binary number $a = \sum_{l=0}^{n-2} a_l 2^l$, where $a_l = 0$ or 1, to be consistently defined as the n-bit number*

$$-a \equiv 2^{n-1} + \sum_{l=0}^{n-2} (1-a_l) 2^l \qquad \textit{(1's complement).} \tag{221}$$

Then $a + (-a) = \sum_{l=0}^{n-1} 2^l = 2^n - 1 =$ the largest n-bit number, which we can

define as equivalent to zero without contradiction so long as the inputs of the addition and subtraction operations are $(n-1)$-bit numbers.

However, having two forms of zero is not elegant, and is avoided in 2's complement arithmetic. Noting that $-a_{1\text{'s complement}} = 2^n - 1 - a$, we are led to define

$$-a \equiv 2^n - a \qquad \textit{(2's complement).} \tag{222}$$

Now, $a + (-a) = 2^n$, which is the $n+1$ bit number with a leading 1 followed by all zeroes. So, if we simply ignore the $n+1$th bit in the result of 2's complement subtraction of two positive $n-1$-bit numbers, i.e., $b - a \to b + (-a)_{2\text{'s complement}}$, we recover the desired n-bit sum in all cases.

Note that for $(n-1)$-bit numbers a and b, the nth bit of $(b-a)_{2\text{'s complement}}$ is 0 if $b \geq a$ and 1 if $b < a$.

(d) **Controlled Fourier Addition.**

It will shortly be desirable to have a circuit for Controlled Fourier addition (or subtraction), in which the addition $\mathbf{\Phi}_a^+(b)$ is performed only if the control bit is $|1\rangle$. For this, the Controlled-Z^p operators used in the Fourier-addition algorithm will become Controlled-Controlled-Z^p operators.

First, show how the circuit for a Controlled-Controlled-U^2 operation that was given in prob. 11(c) can be converted to a nearest-neighbor circuit. Then, give a circuit for Controlled-Fourier addition symbolizing the nearest-neighbor implementation of a Controlled-Controlled-Z^p operation as a single gate.

I found it convenient to group together all such operations on a given bit of $\mathbf{\Phi}(b)$. This may be simpler to implement if the control bit $|c\rangle$ is between $|a\rangle$ and $|\mathbf{\Phi}(b)\rangle$, and the order of the bits of $|a\rangle$ is reversed.

We now consider how to implement the operation

$$\mathsf{U}_f|j\rangle_x|0\rangle_y = |j\rangle_x|f(j)\rangle_y. \tag{171}$$

where

$$f(j) = k^j \pmod N, \tag{210}$$

as needed in Shor's period-finding algorithm.

Writing the binary number j as

$$j = \sum_{l=0}^{n-1} j_l 2^l, \tag{223}$$

where $j_l = 0$ or 1, eq. (210) becomes

$$f(j) = \prod_{l=0}^{n-1} k^{j_l 2^l} \pmod N = \prod_{l=0}^{n-1} (k^{2^l})^{j_l} \pmod N. \tag{224}$$

Since $j_l = 0$ or 1, the lth factor in the product (224) is nontrivial only if $j_l = 1$. That is, the function $f(k)$ involves multiplication by k^{2^l} conditional on bit l of number j.

The implementation of eq. (224) that we pursue here follows Fowler *et al.*, based on earlier work of Beauregard, of Draper, and of Vedral *et al.*[87]

If we have completed the first m multiplications in eq. (224), we have the quantity

$$\Pi_m = \prod_{l=0}^{m-1} (k^{2^l})^{j_l} \pmod N, \tag{225}$$

which we next wish to multiply by $k^{2^m} \pmod N$ only if $j_m = 1$.

A clever way to accomplish this involves use of two registers of n Qbits, where Π_m is in the first register, and the second register is $|0\rangle_n$ at the beginning of each cycle. That is, we start from the state $|\Pi_m, 0\rangle_{2n}$. If $j_m = 0$, then $\Pi_{m+1} = \Pi_m$, and we can proceed to the next cycle. If $j_m = 1$, we perform the operation

$$|\Pi_m, 0\rangle_{2n} \to |\Pi_m, 0 + k^{2^m} \Pi_m \pmod N\rangle_{2n} = |\Pi_m, \Pi_{m+1}\rangle_{2n}, \tag{226}$$

followed by the swap

$$|\Pi_m, \Pi_{m+1}\rangle_{2n} \to |\Pi_{m+1}, \Pi_m\rangle_{2n}, \tag{227}$$

followed by a subtraction to restore the second register to $|0\rangle_n$,

$$|\Pi_{m+1}, \Pi_m\rangle_{2n} \to |\Pi_{m+1}, \Pi_m - k^{-2^m} \Pi_{m+1} \pmod N\rangle_{2n} = |\Pi_{m+1}, 0\rangle_{2n}. \tag{228}$$

The arithmetic operations on the second register in steps (226) and (228) can be written out as

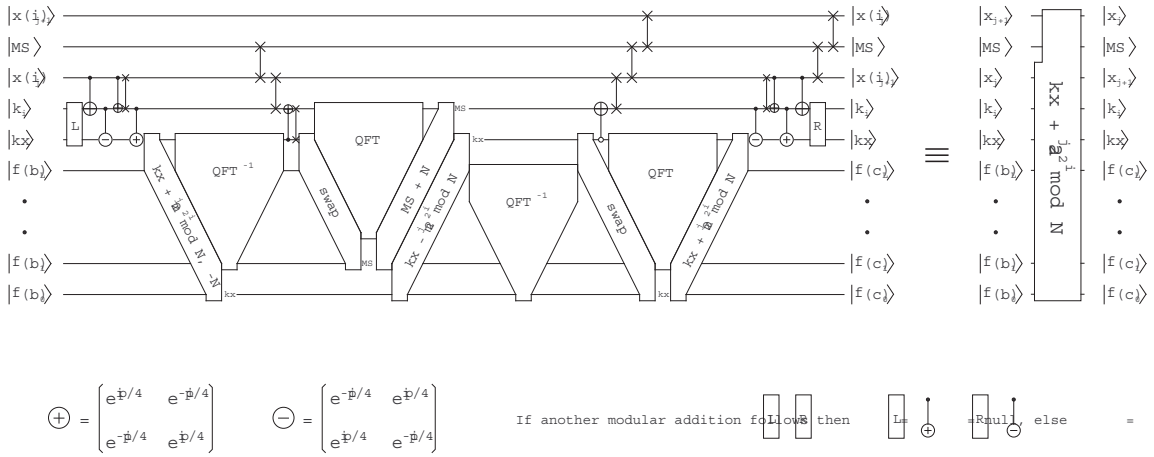
$$0 + k^{2^m} \Pi_m \pmod N = 0 + \sum_{l=0}^{m-1} k^{2^m} \Pi_{m,l} 2^l \pmod N, \tag{229}$$

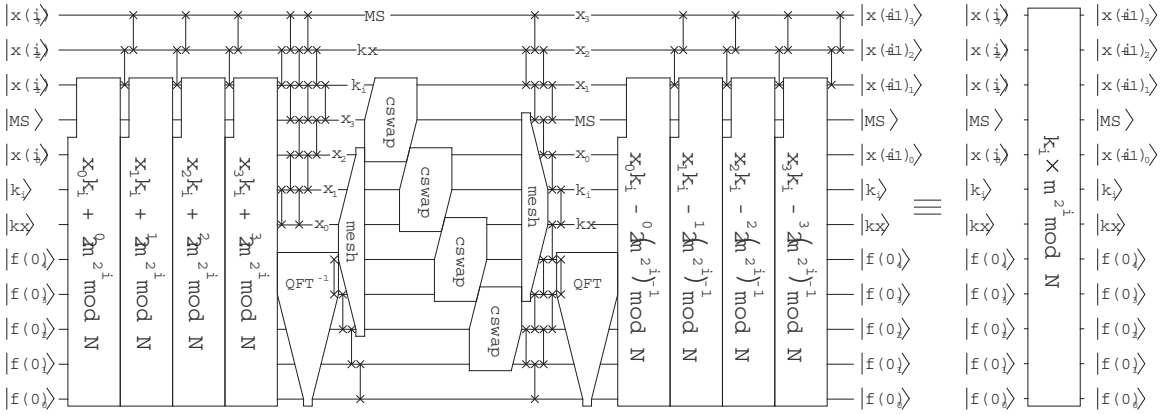
$$\Pi_m - k^{-2^m} \Pi_{m+1} \pmod N = \Pi_m - \sum_{l=0}^{m-1} k^{-2^m} \Pi_{m+1,l} 2^l \pmod N, \tag{230}$$

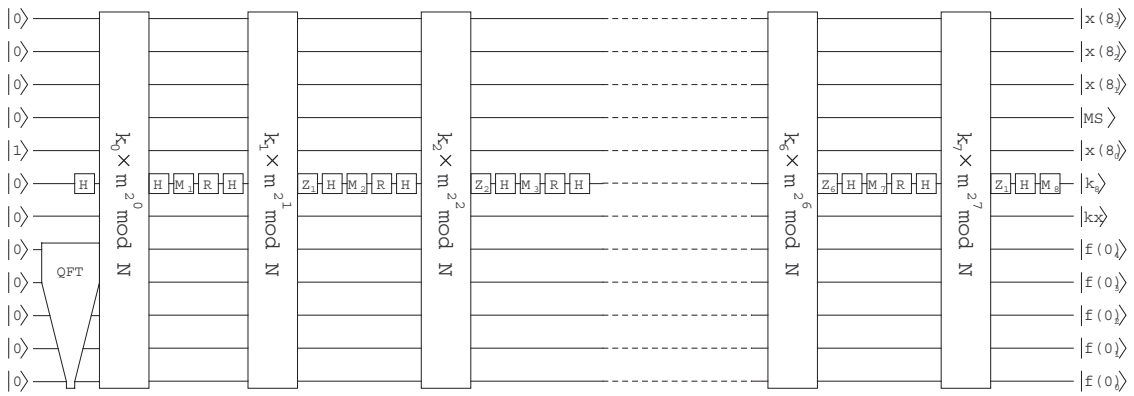
where $\Pi_{m,l} = 0$ or 1 is the lth bit of the partial product Π_m. Hence, the lth addition (or subtraction) needs to be performed only if $\Pi_{m,l} = 1$ (or $\Pi_{m+1,l} = 1$). And, the entire sequence (226)-(228) needs to be performed only if $j_m = 1$.

We have developed essentially all of the circuit ingredients needed for the modular exponentiation algorithm using only nearest-neighbor gates. However, the implementation is rather massive, and we content ourselves with a look at three levels of circuitry as proposed by Fowler *et al.*, shown on the next page.

[87] http://kirkmcd.princeton.edu/examples/QM/fowler_qic_4_237_04.pdf
http://kirkmcd.princeton.edu/examples/QM/beauregard_quant-ph-0205095.pdf
http://kirkmcd.princeton.edu/examples/QM/draper_quant-ph-0008033.pdf
http://physics.princeton.edu/~mcdonald/examples/QM/vedral_pra_54_147_96.pdf


QFT
QFT ⁻¹
swap
MS
kx
L
R
If another modular addition follows then
else
=


MS
kx
cswap
mesh
QFT
QFT ⁻¹


QFT
H
R
Reset: |0⟩ → |0⟩ |1⟩ → |0⟩
ith measurement
Phase: |0⟩ → |0⟩ |1⟩ → e^{iπM3M2M1/2}|1⟩

19. **Spin Control**[88]

We again take up the question, considered briefly in prob. 7, of possible hardware implementations of quantum computation.

David DiVincenzo has given a set of five criteria that an acceptable hardware implementation of quantum computation should satisfy:[89]

(a) A scalable physical system with well characterized Qbits.

(b) The ability to initialize the Qbits to $|0\rangle$.

(c) A mechanism to "read" (measure) the Qbits.

(d) A universal set of quantum gates (such as one-Qbit gates H, Z^p and the two-Qbit conditional gate C_{xy}).

(e) Gate operation times that are short compared to the "lifetime" of the Qbits.

The scheme mentioned at the beginning of prob. 7(e), in which Qbits are encoded onto the spatial behavior of a single photon, satisfies all of DiVincenzo's criteria except (a). The related scheme in which each Qbit is encoded on the spatial behavior of a separate photon satisfies all criteria except (d).[90]

The present state of affairs is that no scheme satisfies all five criteria, and therefore hardware realizations of large-scale quantum computation remain a distant goal.

Nonetheless, it may be useful to explore some of the basic features of quantum computation based on Qbits that are encoded onto the energy, rather than position, of simple quantum systems. This type of Qbit is typically based on the spin of an electron, nucleus, atom or ion, whose energy depends on the interaction of the spin magnetic moment with a magnetic field. Such Qbits are typically stationary in space, and require a mechanism for confining (or trapping) the host particles at well-defined sites. Besides trapping particles in "free" space with combinations of electric, magnetic and/or electromagnetic (laser) fields, trapping can occur at bonds of large molecules, inside solid-state quantum dots, or in the "macroscopic" quantum states (Cooper pairs) of superconductors.

There is no perfect scheme for (re)setting a spin-based Qbit to $|0\rangle$ = the low-energy state when the spin is aligned with respect to the reference magnetic field. The best that can be done is to cool the Qbit to a temperature T such that $kT \ll$ the energy $\hbar\gamma B_0$ (see part (a)) to flip the spin in the magnetic field. Even so, the probability that the bit is in a $|1\rangle$ state remains $e^{-\hbar\gamma B_0/kT}$, where k is Boltzmann's constant. And, the time required to (re)set a spin-based Qbit to $|0\rangle$ can be relatively long, as determined by the rate of residual interactions of the spin with its environment.[91]

[88] For a discussion that emphasizes geometric rather than algebraic arguments, see
http://kirkmcd.princeton.edu/examples/QM/laflamme_quant-ph-0207172.pdf
Another application of spin control besides quantum computation is to laser phenomena. Some notes on this from a classical perspective are at
http://kirkmcd.princeton.edu/examples/ph501lecture25/ph501lecture25.pdf

[89] http://kirkmcd.princeton.edu/examples/QM/divincenzo_quant-ph-0002077.pdf

[90] Some benefit of the doubt has been awarded to these schemes regarding criteria (c) and (e).

[91] An alternative is to measure all Qbits at the beginning of a calculation, and apply a NOT operation to those that are found to be in the $|1\rangle$ state.

Schemes based on nuclear spins have the additional limitation that the energy difference between the spin-up and spin-down states in practical magnetic fields is too small to permit the readout of individual bits. One must construct many copies of a nuclear-spin quantum computer, operate them simultaneously, and then perform a readout of the entire ensemble.

The spin-based Qbit schemes face a general class of conflicting tendencies. Qbits that are well isolated from interaction with their surrounding environment tend to participate slowly in the needed conditional gate operations, while Qbits that interact readily with one another tend also to interact detrimentally with the trapping system.[92]

Thus, while we have argued that quantum computation is a "natural" conceptual extension of quantum behavior, quantum systems do not appear to perform "interesting" computational algorithms readily.

In this problem, you will explore some of the basic manipulations of simple spin-based systems, as needed for possible implementation of quantum computation.

You may assume that the spin-1/2 particles that serve as Qbits are at rest, and that their only interaction is with an external, possibility time-dependent, magnetic field $\mathbf{B}$, and with one another.

(a) **Single-Qbit Gates via Pulsed Magnetic Fields**

First, consider a single spin-1/2 particle in a static magnetic field along the z-axis, $\mathbf{B} = B_0\hat{\mathbf{z}}$. Then, the magnetic energy is $U = -\boldsymbol{\mu}\cdot\mathbf{B}$, where $\boldsymbol{\mu} = \Gamma\mathbf{s} = \hbar\Gamma\boldsymbol{\sigma}/2$ is the particle's magnetic moment and $\boldsymbol{\sigma} = (\boldsymbol{\sigma}_x, \boldsymbol{\sigma}_y, \boldsymbol{\sigma}_z)$ is the Pauli-matrix vector. The (reduced) interaction Hamiltonian $\mathsf{h} = \mathcal{H}/\hbar$ of the particle with the magnetic field is

$$\mathsf{h} = -\frac{\Gamma}{2}\boldsymbol{\sigma}\cdot\mathbf{B} = -\frac{\Gamma B_0}{2}\boldsymbol{\sigma}_z = -\frac{\omega_0}{2}\boldsymbol{\sigma}_z \tag{231}$$

where $\hbar\omega_0 = \hbar\Gamma B_0$ is the energy required to "flip" the spin. The frequency ω_0 is often called the Larmor frequency.

Schrödinger's equation, $i\partial_t|\psi\rangle = \mathsf{h}|\psi\rangle$, for the time dependence of the Qbit whose initial state is

$$|\psi(t=0)\rangle = e^{i\gamma}\left(\cos\frac{\alpha}{2}|0\rangle + \sin\frac{\alpha}{2}e^{i\beta}|1\rangle\right) \tag{232}$$

has the immediate solution

$$\begin{aligned} |\psi(t)\rangle &= e^{i\gamma}\left(\cos\frac{\alpha}{2}e^{i\omega_0 t/2}|0\rangle + \sin\frac{\alpha}{2}e^{i\beta}e^{-i\omega_0 t/2}|1\rangle\right) \\ &= e^{i(\gamma+\omega_0 t/2)}\left(\cos\frac{\alpha}{2}|0\rangle + \sin\frac{\alpha}{2}e^{i(\beta-\omega_0 t)}|1\rangle\right). \end{aligned} \tag{233}$$

If we think of this Qbit as a unit vector in the Bloch sphere, with initial polar angle α and initial azimuthal angle β, then the constant magnetic field $B_0\hat{\mathbf{z}}$ causes this vector to precess about the z axis with angular velocity $-\omega_0$ (*i.e.*, counterclockwise as viewed from the $+z$-axis), while also changing the overall phase of the Qbit.

[92] Recent results which suggest that "dephasing" can be suppressed in certain quantum dot systems are reported in http://kirkmcd.princeton.edu/examples/QM/johnson_nature_435_925_05.pdf

The magnetic field has no effect on our interpretation of this Qbit in the [0,1] basis, so it is convenient to adopt a viewpoint that renders this Qbit as, in effect, constant. This can be done by viewing the Qbit from a frame that rotates about the z axis of Bloch space with angular velocity $-\omega_0$.

We now want to transform the Qbit in various ways, always interpreting the effect of the transformations in the rotating frame. For example, a general unitary 1-Qbit operation can be represented by a product of 3 rotations plus an overall phase change. To perform a rotation of the Qbit by angle λ about some axis $\hat{\mathbf{u}}$ (in the rotating frame), we can use the insights of eqs. (231)-(233) and turn on a uniform magnetic field $B_u\hat{\mathbf{u}}$ (in the rotating frame), which will cause the Qbit to precess about the $\hat{\mathbf{u}}$ with angular velocity $\omega_u = \Gamma B_u$. So, if we turn off the field B_u after time $t = \theta/\omega_u$, we will have accomplished the desired rotation (with respect to the rotating frame). The quantity ω_u is often called the Rabi frequency.

We can establish a formal relation between the lab- and rotating-frame descriptions by considering eq. (232) to be the rotating-frame Qbit $|\psi_{\rm rot}\rangle$. Using eq. (233) for the lab-frame version, $|\psi_{\rm lab}\rangle$ of this Qbit, we have

$$\begin{aligned} |\psi_{\rm lab}\rangle &= e^{i\omega_0 t/2}\begin{pmatrix} 1 & 0 \\ 0 & e^{-i\omega_0 t}\end{pmatrix}|\psi_{\rm rot}\rangle = e^{i\omega_0 t/2}\boldsymbol{\sigma}_z^{-\omega_0 t/\pi}|\psi_{\rm rot}\rangle \\ &= e^{i\omega_0 t/2}e^{-i\omega_0 t/2}e^{i\frac{\omega_0 t}{2}\boldsymbol{\sigma}_z}|\psi_{\rm rot}\rangle = e^{i\frac{\omega_0 t}{2}\boldsymbol{\sigma}_z}|\psi_{\rm rot}\rangle, \end{aligned} \tag{234}$$

recalling eqs. (54)-(55).

We can deduce a version of Schrödinger's equation for the rotating-frame Qbit $|\psi_{\rm rot}\rangle$ as follows,

$$\begin{aligned} i\partial_t|\psi_{\rm lab}\rangle &= i\partial_t e^{i\frac{\omega_0 t}{2}\boldsymbol{\sigma}_z}|\psi_{\rm rot}\rangle = -\frac{\omega_0}{2}\boldsymbol{\sigma}_z e^{i\frac{\omega_0 t}{2}\boldsymbol{\sigma}_z}|\psi_{\rm rot}\rangle + e^{i\frac{\omega_0 t}{2}\boldsymbol{\sigma}_z}(i\partial_t|\psi_{\rm rot}\rangle) \\ &= \mathsf{h}_{\rm lab}|\psi_{\rm lab}\rangle = \mathsf{h}_{\rm lab}e^{i\frac{\omega_0 t}{2}\boldsymbol{\sigma}_z}|\psi_{\rm rot}\rangle, \end{aligned} \tag{235}$$

which can be rearranged as[93]

$$i\partial_t|\psi_{\rm rot}\rangle = e^{-i\frac{\omega_0 t}{2}\boldsymbol{\sigma}_z}\left(\mathsf{h}_{\rm lab} + \frac{\omega_0}{2}\boldsymbol{\sigma}_z\right)e^{i\frac{\omega_0 t}{2}\boldsymbol{\sigma}_z}|\psi_{\rm rot}\rangle = \mathsf{h}_{\rm rot}|\psi_{\rm rot}\rangle. \tag{236}$$

Now consider the case that the lab-frame magnetic field includes a static field along the z-axis, $B_0\hat{\mathbf{z}}$, and a time-dependent field $B_u\hat{\mathbf{u}}$ that rotates (precesses) about the z-axis with angular velocity $-\omega_0$,

$$\mathbf{B} = B_0\hat{\mathbf{z}} + B_u[u_x(\cos\omega_0 t\ \hat{\mathbf{x}} - \sin\omega_0 t\ \hat{\mathbf{y}}) + u_y(\sin\omega_0 t\ \hat{\mathbf{x}} + \cos\omega_0 t\ \hat{\mathbf{y}}) + u_z\ \hat{\mathbf{z}}], \tag{237}$$

where the components (u_x, u_y, u_z) of unit vector $\hat{\mathbf{u}}$ are for time $t = 0$ (and hence are the components of $\hat{\mathbf{u}}$ in the rotating frame).

Show that the Hamiltonian in the rotating frame has the time-independent form

$$\mathsf{h}_{\rm rot} = -\frac{\omega_u}{2}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}, \qquad \text{where} \qquad \omega_u = \Gamma B_u. \tag{238}$$

[93] Experts will recognize eqs. (234) and (236) as an example of the effect of a basis transformation by a unitary operator U, namely $|\psi'\rangle = \mathsf{U}|\psi\rangle$ and $\mathsf{h}' = \mathsf{U}^{-1}\mathsf{h}\mathsf{U}$.

Then, from our analysis of eqs. (231)-(234) we see that the effect of a pulse of magnetic field $B_u\hat{\mathbf{u}}$ of duration t is to perform the transformation

$$|\psi'_{\rm rot}\rangle = e^{i\frac{\omega_u t}{2}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}_z}|\psi_{\rm rot}\rangle = e^{i\frac{\lambda}{2}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}_z}|\psi_{\rm rot}\rangle. \tag{239}$$

Recalling eq. (49), we see that this rotates the Qbit $|\psi_{\rm rot}\rangle$ by angle $-\lambda = -\omega_u t$ about the $\hat{\mathbf{u}}$ axis, as viewed in the rotating frame. Since any 2×2 unitary transformation can be represented, up to a phase, as the product of three rotations, we can implement any single-Qbit transformation on a spin-based Qbit by a sequence of pulses of magnetic fields. In these pulses, the magnetic field components in the x-y plane have a carrier frequency ω_0 that is equal to the Larmor frequency of the spin in the "background" field B_0, which condition is often called magnetic resonance.

To perform a sequence of single-Qbit transformations, an appropriately timed sequence of magnetic field pulses must be applied. In particular, to accomplish a rotation about the z-axis, which creates phase shift of the $|1\rangle$ state relative to that of the $|0\rangle$ state, we should turn on an additional magnetic field in the z direction from some time t. However, the background field $B_0\hat{\mathbf{z}}$ is already causing such phase shifts. So, a rotation about the z-axis can be accomplished simply by taking no action for an appropriate time interval, and then shifting back the origin of time in the rotating frame by this amount.

Verify the preceding analysis by a lab-frame calculation when the magnetic field is

$$\mathbf{B} = B_0\hat{\mathbf{z}} + B_x(\cos\omega t\ \hat{\mathbf{x}} - \sin\omega t\ \hat{\mathbf{y}}), \tag{240}$$

and the frequency ω of the magnetic field that rotates in the x-y plane is not necessarily equal to the Larmor frequency. That is, solve Schrödinger's equation in the lab frame, $i\partial_t|\psi\rangle = \mathsf{h}|\psi\rangle$, for the time dependence of the Qbit whose initial state is $|\psi(t=0)\rangle = a|0\rangle + b|1\rangle$.

Hints: If $B_x = 0$, the solution is $|\psi(t)\rangle = ae^{i\omega_0 t/2}|0\rangle + be^{-i\omega_0 t/2}|1\rangle$. This suggests that you seek a solution of the form

$$|\psi(t)\rangle = Ae^{i\alpha t}|0\rangle + Be^{-i\beta t}|1\rangle. \tag{241}$$

You may or may not find it helpful to re-express the time-dependent part of the Hamiltonian in terms of the annihilation and creation operators a *and* $\mathsf{a}^\dagger$ *of prob. 9 by using the exponential forms of the cosine and sine. In any case, you should find two solutions of the form (241), so that the general solution is a superposition of these two. You may also find it useful to define* $\Omega = \sqrt{(\omega_0-\omega)^2+\omega_x^2}$ *where* $\omega_x = \Gamma B_x$.

Deduce a condition on the frequency of the oscillatory field B_x such that the Qbit can be flipped with 100% probability if that field is applied for a characteristic time (which you should also deduce). This magnetic resonance phenomenon provides a realization of a NOT gate for spin-based Qbits.

Application of the field B_x for half the characteristic time realizes the $\sqrt{\mathsf{NOT}}$ operation, *etc.*

(b) **Two-Qbit Coupling**

The interaction of two magnetic dipoles $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ whose separation is $\mathbf{r}_{12}$ leads to a configuration energy,

$$U_{12} = \frac{\boldsymbol{\mu}_1 \cdot \boldsymbol{\mu}_2 - 3(\boldsymbol{\mu}_1 \cdot \hat{\mathbf{r}}_{12})(\boldsymbol{\mu}_2 \cdot \hat{\mathbf{r}}_{12})}{r_{12}^3}\,. \tag{242}$$

In many realizations of spin-based Qbits, the 2-Qbit interaction energy is well approximated by the simpler form

$$U_{12} = \frac{\boldsymbol{\mu}_1 \cdot \boldsymbol{\mu}_2}{r_{12}^3}\,. \tag{243}$$

The corresponding lab-frame interaction Hamiltonian can be written

$$\mathsf{h}_{12} = \frac{\omega_{12}}{2}\boldsymbol{\sigma}^{(1)} \cdot \boldsymbol{\sigma}^{(2)}, \tag{244}$$

where the coupling strength ω_{12} is (relatively) large only for nearest-neighbor Qbits. This Hamiltonian has the same functional form with respect to, say, the first Qbit as does the interaction (231) of that Qbit with an external magnetic field. Therefore, if we view the two Qbits in their respective rotating frames, the effect of the Hamiltonian (244) during a time t is to evolve the 2-Qbit state $|\psi_R\rangle_{12}$ according to[94]

$$\begin{aligned} |\psi'_{\rm rot}\rangle_{12} &= e^{-i\frac{\omega_{12}t}{2}\boldsymbol{\sigma}^{(1)}\cdot\boldsymbol{\sigma}^{(2)}}|\psi_{\rm rot}\rangle_{12} \\ &= e^{-i\frac{\omega_{12}t}{2}\boldsymbol{\sigma}_x^{(1)}\boldsymbol{\sigma}_x^{(2)}}e^{-i\frac{\omega_{12}t}{2}\boldsymbol{\sigma}_y^{(1)}\boldsymbol{\sigma}_y^{(2)}}e^{-i\frac{\omega_{12}t}{2}\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}}|\psi_{\rm rot}\rangle_{12}. \end{aligned} \tag{245}$$

Show that the interaction of two Qbits for a time $t = \pi/2\omega_{12}$ according to eq. (244) performs the SWAP operation S_{12} to within a phase.

This result gives an additional perspective as to why a collection of spins is not spontaneously a "useful" quantum computer. Namely, the dipole-dipole coupling between any pair of spins results in continual swapping of their states. The rate of swapping falls off with the distance between the pair, but at some rate each spin swaps its state with all other spins. The overall behavior is somewhat "chaotic".

A possible way to gain control over the swapping is to place each spin in a "trap" which is fairly well isolated from all other spins. Then, when an interaction is desired between a particular pair of nearest-neighbor spins, the "trapping" forces are adjusted to bring those two spins closer together. After a time sufficient for the desired interaction to occur, the spins are returned to their (relatively) isolated initial positions.

As we discussed at the end of prob. 12, a SWAP operation is not a sufficient building block for all 2-Qbit interactions. However, if we allow a pair of spins to interact for only 1/2 the time needed for a SWAP operation, we will have achieved

[94] The second form of eq. (245) clarifies the meaning of the exponential of the product of two Pauli matrices that act on different Qbits. That is, $e^{i\alpha\boldsymbol{\sigma}_j^{(1)}\boldsymbol{\sigma}_j^{(2)}} = \cos\alpha\,\mathbf{I} + i\sin\alpha\,\boldsymbol{\sigma}_j^{(1)}\boldsymbol{\sigma}_j^{(2)}$ for $j = x, y, z$, but $e^{i\alpha\boldsymbol{\sigma}^{(1)}\cdot\boldsymbol{\sigma}^{(2)}}$ does not equal $\cos\alpha\,\mathbf{I} + i\sin\alpha\,\boldsymbol{\sigma}^{(1)}\cdot\boldsymbol{\sigma}^{(2)}$.

the $\sqrt{\mathsf{S}_{12}}$ operation, which is a universal 2-Qbit gate. Thus, we have one scheme whereby, in principle, we can construct a spin-based quantum computer capable of executing any quantum algorithm.

It turns out that in some spin systems the 2-Qbit coupling is well approximated by a further simplification of eq. (244), namely

$$\mathsf{h}_{12} = \frac{\omega_{12}}{2}\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}, \tag{246}$$

when the system is immersed in a "background" magnetic field along the z-axis. Show that a Controlled-NOT operation C_{12} can be represented by the following sequence of magnetic-field pulses (to within an overall phase),

$$\mathsf{C}_{12} \propto e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}} e^{i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(2)}} e^{i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(2)}} e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}} e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(2)}}. \tag{247}$$

Given eq. (247), we see at once that

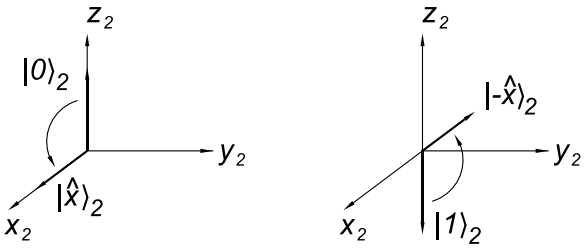
$$\mathsf{C}_{21} \propto e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(2)}} e^{i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(1)}} e^{i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}} e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}} e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(1)}}. \tag{248}$$

The sequence of operations (247) can be given a geometrical interpretation. Recall from eq. (49) that an operator $e^{i\frac{\phi}{2}\boldsymbol{\sigma}_j}$ corresponds to a rotation of a Qbit in Bloch space by angle $-\phi$ about axis j.

i. The first operation, $e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(2)}}$, rotates the target Qbit $|\psi\rangle_2$ by $+90°$ about the y_2 axis. This means that $|0\rangle_2$ rotates from $\hat{\mathbf{z}}_2$ to $\hat{\mathbf{x}}_2$, while $|1\rangle_2$ rotates from $-\hat{\mathbf{z}}_2$ to $-\hat{\mathbf{x}}_2$.

$\|\psi\rangle_2$	Initial direction	Direction after step i
$\|0\rangle$	$\hat{\mathbf{z}}_2$	$\hat{\mathbf{x}}_2$
$\|1\rangle$	$-\hat{\mathbf{z}}_2$	$-\hat{\mathbf{x}}_2$

(249)



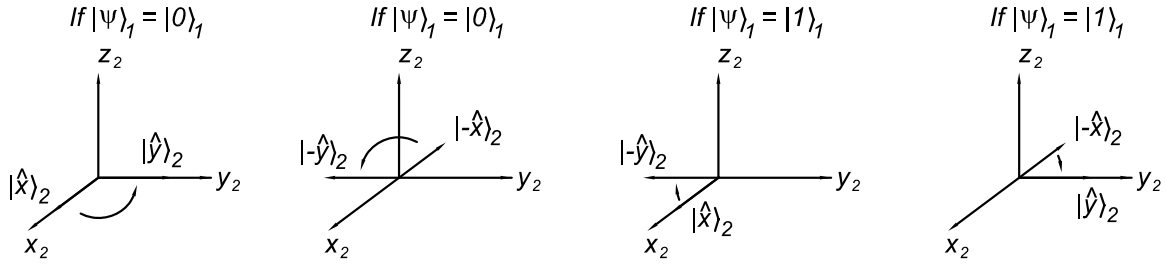
ii. The second operation, $e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}}$, provides the key conditional action. It rotates Qbit 2 by $+90°$ about the z_2 axis if Qbit 1 is $|0\rangle$, but it rotates Qbit 2 by $-90°$ about the z_2 axis if Qbit 1 is $|1\rangle$. Algebraically,

$$e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}} = \frac{\mathbf{I} - i\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}}{\sqrt{2}} = \begin{cases} \frac{\mathbf{I}-i\boldsymbol{\sigma}_z^{(2)}}{\sqrt{2}} = e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(2)}} & \text{if } |\psi\rangle_1 = |0\rangle_1, \\ \frac{\mathbf{I}+i\boldsymbol{\sigma}_z^{(2)}}{\sqrt{2}} = e^{i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(2)}} & \text{if } |\psi\rangle_1 = |1\rangle_1. \end{cases} \tag{250}$$

Alternatively, the operation $e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}}$ *rotates Qbit 1 by* $+90°$ *about the* z_1 *axis if Qbit 2 is* $|0\rangle$*, but it rotates Qbit 1 by* $-90°$ *about the* z_1 *axis if Qbit 2 is* $|1\rangle$*. Since we seek to understand the fate of the target bit 2, it is more relevant to give the previous interpretation to the operator* $e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}}$*. Furthermore, we only list the final states of Qbit 2 in the following discussion.*

Initial $\lvert\psi\rangle_1$	Initial $\lvert\psi\rangle_2$	$\lvert\psi\rangle_2$ after step ii
$\lvert 0\rangle$	$\lvert 0\rangle$	$\hat{\mathbf{y}}_2$
$\lvert 0\rangle$	$\lvert 1\rangle$	$-\hat{\mathbf{y}}_2$
$\lvert 1\rangle$	$\lvert 0\rangle$	$-\hat{\mathbf{y}}_2$
$\lvert 1\rangle$	$\lvert 1\rangle$	$\hat{\mathbf{y}}_2$

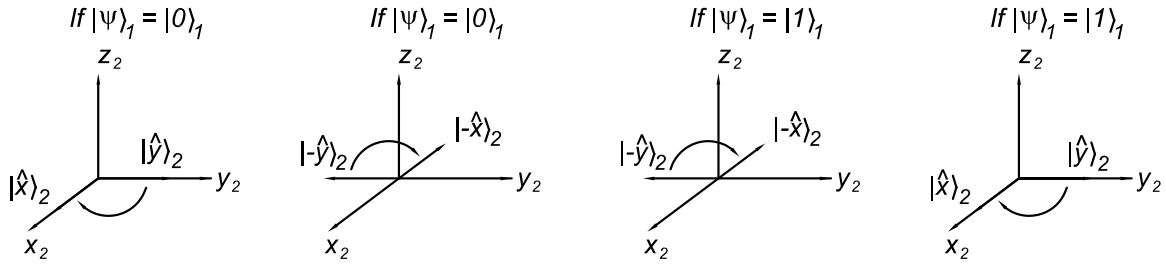
(251)



iii. The third operation, $e^{i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(2)}}$, rotates Qbit 2 by $-90°$ about the z_2 axis.

Initial $\lvert\psi\rangle_1$	Initial $\lvert\psi\rangle_2$	$\lvert\psi\rangle_2$ after step iii
$\lvert 0\rangle$	$\lvert 0\rangle$	$\hat{\mathbf{x}}_2$
$\lvert 0\rangle$	$\lvert 1\rangle$	$-\hat{\mathbf{x}}_2$
$\lvert 1\rangle$	$\lvert 0\rangle$	$-\hat{\mathbf{x}}_2$
$\lvert 1\rangle$	$\lvert 1\rangle$	$\hat{\mathbf{x}}_2$

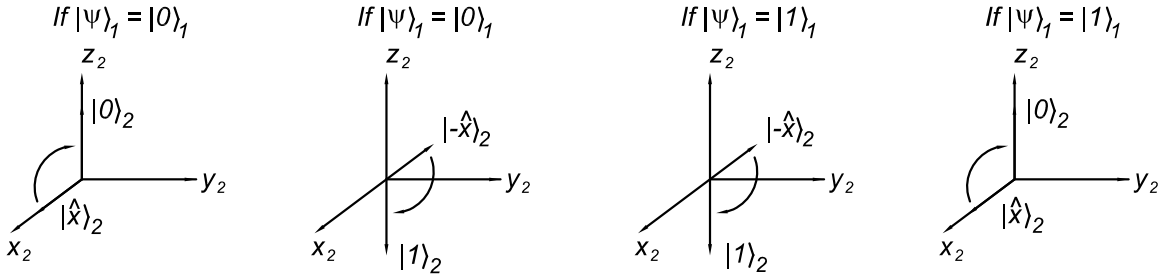
(252)



iv. The fourth operation, $e^{i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(2)}}$, rotates Qbit 2 by $-90°$ about the y_2 axis. This restores Qbit 2 to its initial direction if Qbit 1 is $|0\rangle$, but inverts Qbit 2 if Qbit 1 is $|1\rangle$.

Initial $\|\psi\rangle_1$	Initial $\|\psi\rangle_2$	$\|\psi\rangle_2$ after step iv
$\|0\rangle$	$\|0\rangle$	$\hat{\mathbf{z}}_2$
$\|0\rangle$	$\|1\rangle$	$-\hat{\mathbf{z}}_2$
$\|1\rangle$	$\|0\rangle$	$-\hat{\mathbf{z}}_2$
$\|1\rangle$	$\|1\rangle$	$\hat{\mathbf{z}}_2$

(253)



It looks like we are done, but if Qbit 1 was not in a basis state, things are slightly confused. This is fixed by the final step.

v. The fifth operation, $e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}}$, rotates Qbit 1 by $+90°$ about the z_1 axis. This tidies up the phases of the Qbits, as you will demonstrate in the algebraic exercise assigned above, so the complete set of steps performs the desired Controlled-NOT to within an overall phase.

Initial $\|\psi\rangle_1$	Initial $\|\psi\rangle_2$	Final $\|\psi\rangle_1$	Final $\|\psi\rangle_2$
$\|0\rangle$	$\|0\rangle$	$\|0\rangle$	$\|0\rangle$
$\|0\rangle$	$\|1\rangle$	$\|0\rangle$	$\|1\rangle$
$\|1\rangle$	$\|0\rangle$	$\|1\rangle$	$\|1\rangle$
$\|1\rangle$	$\|1\rangle$	$\|1\rangle$	$\|0\rangle$

(254)

This analysis could have been carried out in more detail, keeping track of the phase of Qbit 1 at each step. But then there would be little difference between the "geometric" approach and an algebraic approach.

Give a similar "geometric" argument to show how the operation

$$e^{-i\frac{\pi}{4}\boldsymbol{\sigma}^{(1)}\cdot\boldsymbol{\sigma}^{(2)}} = e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_x^{(1)}\cdot\boldsymbol{\sigma}_x^{(2)}} e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(1)}\cdot\boldsymbol{\sigma}_y^{(2)}} e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}\cdot\boldsymbol{\sigma}_z^{(2)}} \tag{255}$$

swaps Qbits 1 and 2.

It suffices to show that this operation performs a SWAP *on each of the 2-Qbit basis states* $|0\rangle_1|0\rangle_2$, $|0\rangle_1|1\rangle_2$, $|1\rangle_1|0\rangle_2$ *and* $|1\rangle_1|1\rangle_2$ *up to a phase which varies from*

state to state. Your earlier algebraic argument shows that the phase change is actually independent of the 2-Qbit state.

The first step of operation (255) is $e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}}$ *which performs conditional rotations by* $\pm 90°$ *about the z-axes. Since all of our initial states are aligned along the z axes, the first step merely changes the phases of the initial states, but not their directions. In the geometric view, this step has no effect.*

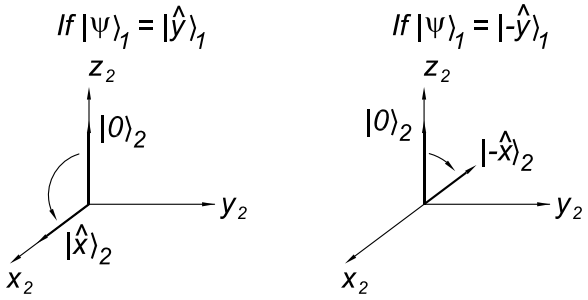
The second step of operation (255) is $e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(1)}\boldsymbol{\sigma}_y^{(2)}}$ *which performs a conditional rotation of bit 2 by* $\pm 90°$ *about the* y_2*-axis depending on the state of bit 1 as projected onto the* y_1*-axis (or equivalently, a conditional rotation of bit 1 by* $\pm 90°$ *about the* y_1*-axis depending on the state of bit 2 as projected onto the* y_2*-axis). This will result in the initial states, which were aligned along the z axes, being transformed into various combinations of states readily expressed with one bit along its x-axis and the other bit along its y-axis.*

The third step of operation (255) is $e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_x^{(1)}\boldsymbol{\sigma}_x^{(2)}}$ *which performs conditional rotations* $\pm 90°$ *about the x-axes. Since the input states to this operation are simply expressed with one of the two bits aligned along its x-axis, it is relatively straightforward to keep track of this step.*

Here we illustrate the case when the initial state is $|0\rangle_1|0\rangle_2$*, which can be expressed in various equivalent ways,*

$$|\psi_0\rangle = |0\rangle_1|0\rangle_2 = |\hat{\mathbf{z}}\rangle_1|\hat{\mathbf{z}}\rangle_2 = \frac{|\hat{\mathbf{y}}\rangle_1 + |-\hat{\mathbf{y}}\rangle_1}{\sqrt{2}}|\hat{\mathbf{z}}\rangle_2 = |\hat{\mathbf{z}}\rangle_1\frac{|\hat{\mathbf{y}}\rangle_2 + |-\hat{\mathbf{y}}\rangle_2}{\sqrt{2}}\,. \qquad (256)$$

To determine the rotation of the second Qbit by the operation $e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(1)}\boldsymbol{\sigma}_y^{(2)}}$*, which is conditional on the state of the first Qbit, we need the first Qbit to be described relative to the* y_1 *axis. Recalling from eq. (49) that the operator* $e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(2)}}$ *is a rotation of bit 2 by* $+90°$ *about the* y_2*-axis, we learn that when bit 1 is* $|\hat{\mathbf{y}}\rangle_1$ *bit 2 is rotated from* $|\hat{\mathbf{z}}\rangle_2$ *to* $|\hat{\mathbf{x}}\rangle_2$*, and that when bit 1 is* $|-\hat{\mathbf{y}}\rangle_1$ *bit 2 is rotated from* $|\hat{\mathbf{z}}\rangle_2$ *to* $|-\hat{\mathbf{x}}\rangle_2$*.*



That is,

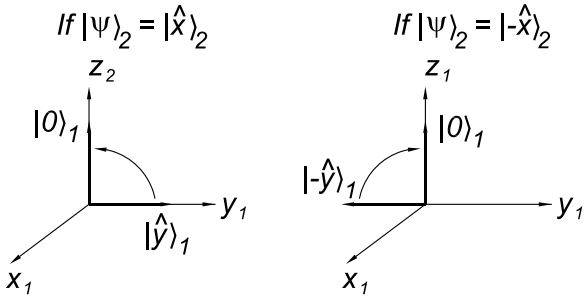
$$e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(1)}\boldsymbol{\sigma}_y^{(2)}}|0\rangle_1|0\rangle_2 = e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(1)}\boldsymbol{\sigma}_y^{(2)}}\frac{|\hat{\mathbf{y}}\rangle_1 + |-\hat{\mathbf{y}}\rangle_1}{\sqrt{2}}|\hat{\mathbf{z}}\rangle_2 = \frac{|\hat{\mathbf{y}}\rangle_1|\hat{\mathbf{x}}\rangle_2 + |-\hat{\mathbf{y}}\rangle_1|-\hat{\mathbf{x}}\rangle_2}{\sqrt{2}}\,. \qquad (257)$$

Alternatively, we could consider the 2nd bit to be the control bit, in which case the transformation can be written

$$|0\rangle_1|0\rangle_2 = |\hat{\mathbf{z}}\rangle_1\frac{|\hat{\mathbf{y}}\rangle_2 + |-\hat{\mathbf{y}}\rangle_2}{\sqrt{2}} \to \frac{|\hat{\mathbf{x}}\rangle_1|\hat{\mathbf{y}}\rangle_2 + |-\hat{\mathbf{x}}\rangle_1|-\hat{\mathbf{y}}\rangle_2}{\sqrt{2}}\,, \qquad (258)$$

which you may wish to verify is actually the same as eq. (257).

To understand the effect of the conditional operation $e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_x^{(1)}\boldsymbol{\sigma}_x^{(2)}}$*, one of the two bits should be expressed in terms of its projections onto the* x*-axis. Both eqs. (257) and (258) are already of the desired form, so we can use either. Specifically, the transformation of eq. (257) is that when bit 2 is* $|\hat{\mathbf{x}}\rangle_2$ *bit 1 is rotated from* $|\hat{\mathbf{y}}\rangle_1$ *to* $|\hat{\mathbf{z}}\rangle_1$*, and that when bit 2 is* $|-\hat{\mathbf{x}}\rangle_2$ *bit 1 is rotated from* $|-\hat{\mathbf{y}}\rangle_1$ *to* $|\hat{\mathbf{z}}\rangle_1$*.*



Then,

$$e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_x^{(1)}\boldsymbol{\sigma}_x^{(2)}}\frac{|\hat{\mathbf{y}}\rangle_1|\hat{\mathbf{x}}\rangle_2+|-\hat{\mathbf{y}}\rangle_1|-\hat{\mathbf{x}}\rangle_2}{\sqrt{2}}=\frac{|\hat{\mathbf{z}}\rangle_1|\hat{\mathbf{x}}\rangle_2+|\hat{\mathbf{z}}\rangle_1|-\hat{\mathbf{x}}\rangle_2}{\sqrt{2}}=|0\rangle_1|0\rangle_2, \quad (259)$$

as expected for $\mathsf{SWAP}|0\rangle_1|0\rangle_2$*.*

In systems where the spin-spin coupling is well described by the Hamiltonian (246), so that the corresponding time evolution of a pair of spins is

$$|\psi'_{\rm rot}\rangle_{12}=e^{-i\frac{\omega_{12}t}{2}\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}}|\psi_{\rm rot}\rangle_{12}, \quad (260)$$

there is another way to arrange that this interaction is effective only during specified time intervals. Namely, any interval in which it is desired that there be no net effect of the spin-spin coupling, a pair of magnetic pulses can be applied at the beginning and middle of the interval such that the evolution of the spin during the first half interval is reversed during the second half interval. This is a quantum computer version of the spin echo technique.[95]

The formal basis of this technique are the identities,

$$\boldsymbol{\sigma}_x\boldsymbol{\sigma}_z\boldsymbol{\sigma}_x=\boldsymbol{\sigma}_y\boldsymbol{\sigma}_z\boldsymbol{\sigma}_y=-\boldsymbol{\sigma}_z, \quad \text{and} \quad \boldsymbol{\sigma}_x=-ie^{i\frac{\pi}{2}\boldsymbol{\sigma}_x}, \quad \boldsymbol{\sigma}_y=-ie^{i\frac{\pi}{2}\boldsymbol{\sigma}_y}, \quad (261)$$

which imply that

$$\begin{aligned}
e^{i\frac{\pi}{2}\boldsymbol{\sigma}_x^{(1)}}e^{-i\frac{\omega_{12}t}{2}\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}}e^{i\frac{\pi}{2}\boldsymbol{\sigma}_x^{(1)}} &= -\boldsymbol{\sigma}_x^{(1)}e^{-i\frac{\omega_{12}t}{2}\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}}\boldsymbol{\sigma}_x^{(1)} \\
&= -\boldsymbol{\sigma}_x^{(1)}\left(\cos\frac{\omega_{12}t}{2}-i\sin\frac{\omega_{12}t}{2}\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}\right)\boldsymbol{\sigma}_x^{(1)} \\
&= -\left(\cos\frac{\omega_{12}t}{2}+i\sin\frac{\omega_{12}t}{2}\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}\right) \\
&= -e^{i\frac{\omega_{12}t}{2}\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}}, \quad (262)
\end{aligned}$$

[95] http://kirkmcd.princeton.edu/examples/QM/hahn_pr_80_580_50.pdf

and therefore,

$$e^{-i\frac{\omega_{12}t}{2}\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}}e^{i\frac{\pi}{2}\boldsymbol{\sigma}_x^{(1)}}e^{-i\frac{\omega_{12}t}{2}\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}}e^{i\frac{\pi}{2}\boldsymbol{\sigma}_x^{(1)}} = -\mathbf{I}. \tag{263}$$

We could also have used the operation $e^{i\frac{\pi}{2}\boldsymbol{\sigma}_x}$ at the middle and end of the interval to cancel the effect of the spin-spin coupling during that interval. Or, we could have used the operation $e^{i\frac{\pi}{2}\boldsymbol{\sigma}_y}$ instead of $e^{i\frac{\pi}{2}\boldsymbol{\sigma}_x}$. Or, we could have applied the magnetic pulses to spin 2 rather than to spin 1, *etc.*

20. **Dephasing**[96]

In prob. 19 we alluded to the difficulty that Qbits interact with their environment in ways that are detrimental to quantum computation. This problem explores methods of formalizing our understanding of these difficulties.

The situation is a version of Schrödinger's cat: we desire that our Qbits remain in a pure quantum state inside our quantum computer, at least for the duration of the calculation. But the worry is that these ideal quantum states do not survive that long, and are converted into Cbits even before we "look" at the results of our computation.

The interactions of the Qbits with their environment may perform some kind of "measurement" on them, the results of which we do not know in detail because we haven't yet observed the Qbits ourselves. So, we desire a description of quantum states that includes the possibility that they are in one of several possible basis states as a result of interactions beyond our control, as well as the possibility that the Qbits are still in a coherent superposition of basis states as we have assumed in this course until now.

Such a description was provided independently(?) in 1927 by Landau and by von Neumann, and is based on the concept of the density operator, also called the density matrix.

Wave Function of a Pure State

If a quantum system is in an idealized pure state, we have characterized this by a wave function

$$|\psi\rangle = \sum_j \psi_j |j\rangle, \tag{2}$$

that is a weighted sum of basis states $|j\rangle$. The time evolution of state $|\psi\rangle$ has been described by a unitary transformation $\mathsf{U}(t,t')$ such that

$$|\psi(t')\rangle = \mathsf{U}(t,t')|\psi(t)\rangle. \tag{84}$$

If the state $|\psi\rangle$ is observed via a (hermitian) measurement operator M whose eigenvectors are the basis states $|j\rangle$ with corresponding eigenvalues m_j, then we can write

$$\mathsf{M} = \sum_j \mathsf{M}_j = \sum_j m_j \cdot \mathsf{P}_j = \sum_j m_j \cdot |j\rangle\langle j|, \tag{81}$$

and the probability that the result of the measurement is that state $|\psi\rangle$ is found in basis state $|j\rangle$ is

$$P_j = \langle\psi|\mathsf{P}_j^\dagger \mathsf{P}_j|\psi\rangle = \langle\psi|\mathsf{P}_j|\psi\rangle = \langle\psi|j\rangle\langle j|\psi\rangle = |\langle j|\psi\rangle|^2 . \tag{78}$$

The probable value (or expectation value) of variable m for state $|\psi\rangle$ is thus

$$\langle m\rangle = \sum_j m_j P_j = \sum_j m_j \langle\psi|\mathsf{P}_j|\psi\rangle = \langle\psi|\sum_j m_j \cdot \mathsf{P}_j|\psi\rangle = \langle\psi|\mathsf{M}|\psi\rangle. \tag{82}$$

[96] Problem 20 is covered in sec. 2.4 and chap. 8 of Nielsen and Chuang.

Density Matrix of a Pure State

The density operator $\boldsymbol{\rho}$ of a pure state (2) is simply its corresponding projection operator,

$$\boldsymbol{\rho} = |\psi\rangle\langle\psi| = \sum_{j,k} \psi_j \psi_k^\star |j\rangle\langle k| \qquad \text{(pure state)}. \tag{264}$$

Clearly, the operator $\boldsymbol{\rho}$ can be represented by the hermitian matrix whose elements are

$$\boldsymbol{\rho}_{jk} = \psi_j \psi_k^\star. \tag{265}$$

Examples:

$$\boldsymbol{\rho}(|0\rangle) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \boldsymbol{\rho}(|1\rangle) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{\rho}(|\pm\rangle) = \boldsymbol{\rho}\left(\frac{|0\rangle \pm |1\rangle}{\sqrt{2}}\right) = \frac{1}{2}\begin{pmatrix} 1 & \pm 1 \\ \pm 1 & 1 \end{pmatrix}, \tag{266}$$

$$\boldsymbol{\rho}(|00\rangle) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \boldsymbol{\rho}\left(\frac{|00\rangle \pm |11\rangle}{\sqrt{2}}\right) = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & \pm 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \pm 1 & 0 & 0 & 1 \end{pmatrix}, \tag{267}$$

$$\boldsymbol{\rho}\left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) = \boldsymbol{\rho}\left(\frac{|00\rangle - |01\rangle + |10\rangle - |11\rangle}{2}\right) = \frac{1}{4}\begin{pmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \end{pmatrix}. \tag{268}$$

Basis states have only a single nonzero (diagonal) element to their density matrices. A pure state that is a superposition of basis states has nonzero off-diagonal elements to its density matrix.

Other properties of density matrices follow immediately: The square of the density operator of a pure state is itself,

$$\boldsymbol{\rho}^2 = \boldsymbol{\rho} \qquad \text{(pure state)}. \tag{269}$$

The trace of the density matrix is 1,

$$\mathrm{tr}(\boldsymbol{\rho}) = \sum_j \boldsymbol{\rho}_{jj} = \sum_j \psi_j \psi_j^\star = \sum_j \left|\psi_j\right|^2 = 1, \tag{270}$$

since quantum states are normalized to unit probability. An alternative derivation of this is

$$\mathrm{tr}(\boldsymbol{\rho}) = \mathrm{tr}(|\psi\rangle\langle\psi|) = \sum_j \langle j|\psi\rangle\langle\psi|j\rangle = \sum_j \langle\psi|j\rangle\langle j|\psi\rangle = \langle\psi|\psi\rangle = 1. \tag{271}$$

The time evolution of the density operator follows from (84) as

$$\boldsymbol{\rho}(t') = |\psi(t')\rangle\langle\psi(t')| = \mathsf{U}(t,t')|\psi(t)\rangle\langle\psi(t)|\mathsf{U}^\dagger(t,t') = \mathsf{U}(t,t')\boldsymbol{\rho}(t)\mathsf{U}^\dagger(t,t'). \tag{272}$$

The probability that state $|\psi\rangle$ is found in basis state $|j\rangle$ as the result of a measurement follows from (78) as

$$\begin{aligned} P_j &= |\langle j|\psi\rangle|^2 = \langle\psi|j\rangle\langle j|\psi\rangle = \langle j|\psi\rangle\langle\psi|j\rangle = \langle j|j\rangle\langle j|\psi\rangle\langle\psi|j\rangle = \sum_k \langle k|j\rangle\langle j|\psi\rangle\langle\psi|k\rangle \\ &= \mathrm{tr}(|j\rangle\langle j|\psi\rangle\langle\psi|) = \mathrm{tr}(\mathsf{P}_j\boldsymbol{\rho}), \end{aligned} \tag{273}$$

which is a special case of the general result for an operator O that

$$\langle\psi|\mathsf{O}|\psi\rangle = \mathrm{tr}(\mathsf{O}\boldsymbol{\rho}). \tag{274}$$

The expectation value $\langle m\rangle$ for a measurement of state $|\psi\rangle$ using operator $\mathsf{M} = \sum_j m_j{\cdot}\mathsf{P}_j$ of eq. (79) follows from eqs. (83) and (274) as

$$\langle m\rangle = \sum_j m_j\langle\psi|\mathsf{P}_j|\psi\rangle = \sum_j m_j\mathrm{tr}(\mathsf{P}_j\boldsymbol{\rho}) = \mathrm{tr}(\mathsf{M}\boldsymbol{\rho}). \tag{275}$$

Density Matrix of a Mixed State

These results show that the density-matrix description of a pure quantum state recovers all the features of the usual description. However, there does not yet appear to be any advantage to the use of density matrices. That advantage lies in the ease with which the density-matrix description can be extended to include so-called mixed states in which the quantum state is one of a set of pure states $|\psi_i\rangle$ with probability P_i, where the total probability is, of course, unity: $\sum_i P_i = 1$. In this case, we define

$$\boldsymbol{\rho} = \sum_i P_i|\psi_i\rangle\langle\psi_i| = \sum_i P_i\boldsymbol{\rho}_i \qquad \text{(mixed state)}. \tag{276}$$

We readily set that the mixed-state density matrix (276) obeys all of the properties (270)-(275). However,

$$\boldsymbol{\rho}^2 \neq \boldsymbol{\rho} \qquad \text{(mixed state)}, \tag{277}$$

which provides a means of determining whether a given density matrix describes a pure state or a mixed state.

Example: A 50:50 mixture of states $|00\rangle$ and $|11\rangle$ has density matrix

$$\boldsymbol{\rho} = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{278}$$

A mixture of basis states has no off-diagonal elements in its density matrix.

Example: A 50:50 mixture of states $|0\rangle$ and $|1\rangle$ has the same density matrix as a 50:50 mixture of states $|+\rangle$ and $|-\rangle$. From eq. (266) we have,

$$\boldsymbol{\rho} = \frac{\boldsymbol{\rho}(|0\rangle) + \boldsymbol{\rho}(|1\rangle)}{2} = \frac{\boldsymbol{\rho}(|+\rangle) + \boldsymbol{\rho}(|-\rangle)}{2} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{\mathbf{I}}{2}. \tag{279}$$

Indeed, the rotation (46) of basis states $|0\rangle$ and $|1\rangle$ by angle θ about the y-axis in Bloch space leads to the new basis states $\cos\frac{\theta}{2}|0\rangle - \sin\frac{\theta}{2}|1\rangle$ and $\sin\frac{\theta}{2}|0\rangle + \cos\frac{\theta}{2}|1\rangle$. Hence, a 50:50 mixture of these new basis states also has density matrix (279).

A historic debate about the meaning of the quantum wave functions concerns whether they reflect that Nature is intrinsically probabilistic or that the probabilities merely reflect our ignorance of some underlying well-defined "reality". We argue that a pure state is one for which probabilities are intrinsic.[97] In contrast, a mixed state (276) can be regarded as actually being in one of its component pure states $|\psi_i\rangle$, but we don't know which.[98] The probabilities P_i in eq. (276) summarize our ignorance/knowledge of which pure states are present, while the coefficients ψ_j in eq. (2) represent intrinsic probabilities (strictly, probability amplitudes) as to what can be observed of the pure state $|\psi\rangle$.

Mixed states and their density-matrix description are therefore useful in quantum statistical mechanics in which we are ignorant of details of the state of our system or ensemble of systems.

Density Matrix of a Composite System

The density-matrix description is also useful when dealing with a system for which we have different qualities of information about its component subsystems.

Consider a system with two subsystems A and B for which the density matrix of the whole system is $\boldsymbol{\rho}_{AB}$. If our knowledge of system B is limited, we may wish to consider what we can say about system A only. That is, we desire the density matrix $\boldsymbol{\rho}_A$.

The claim is that the appropriate procedure is to calculate

$$\boldsymbol{\rho}_A = \mathrm{tr}_B(\boldsymbol{\rho}_{AB}), \tag{280}$$

where the trace over subsystem B can be accomplished with the aid of the definition

$$\begin{aligned} \mathrm{tr}_B(|A_1B_1\rangle\langle A_2B_2|) &= \mathrm{tr}_B(|A_1\rangle\langle A_2| \otimes |B_1\rangle\langle B_2|) = |A_1\rangle\langle A_2|\ \mathrm{tr}_B(|B_1\rangle\langle B_2|) \\ &= |A_1\rangle\langle A_2|\ \langle B_1|B_2\rangle, \end{aligned} \tag{281}$$

recalling eq. (271).

If subsystems A and B have no nontrivial couplings, then $\boldsymbol{\rho}_{AB} = \boldsymbol{\rho}_A \otimes \boldsymbol{\rho}_B$, so that $\mathrm{tr}(\boldsymbol{\rho}_{AB}) = \mathrm{tr}(\boldsymbol{\rho}_A \otimes \boldsymbol{\rho}_B) = \boldsymbol{\rho}_A\ \mathrm{tr}(\boldsymbol{\rho}_B) = \boldsymbol{\rho}_A$, as expected.

Of greater interest is the case when subsystems A and B are entangled. For example, consider the entangled 2-Qbits states described by the righthand case in eq. (267). To apply eq. (281) it is easier to rewrite the density matrix (267) as a density operator,

$$\boldsymbol{\rho}_{AB}\left(\frac{|00\rangle \pm |11\rangle}{\sqrt{2}}\right) = \frac{|00\rangle \pm |11\rangle}{\sqrt{2}}\frac{\langle 00| \pm \langle 11|}{\sqrt{2}} = \frac{|00\rangle\langle 00| \pm |00\rangle\langle 11| \pm |11\rangle\langle 00| + |11\rangle\langle 11|}{2}. \tag{282}$$

[97] For a recent review of the Kochen-Specker theorem that a quantum-mechanical spin-1/2 state cannot "really" have simultaneous definite values of its spin vector along three orthogonal axes, see http://kirkmcd.princeton.edu/examples/QM/cassinello_ajp_73_272_05.pdf

[98] The example of eq. (279) reminds us that a given mixed-state density matrix corresponds to different mixtures in different bases, so considerable quantum subtlety remains even for mixed states.

Then,

$$\begin{aligned}\boldsymbol{\rho}_A &= \mathrm{tr}_B(\boldsymbol{\rho}_{AB}) = \frac{|0\rangle\langle 0|\ \langle 0|0\rangle \pm |0\rangle\langle 1|\ \langle 0|1\rangle \pm |1\rangle\langle 0|\ \langle 1|0\rangle + |1\rangle\langle 1|\ \langle 1|1\rangle}{2} \\ &= \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{\mathbf{I}}{2}. \end{aligned} \tag{283}$$

This noteworthy result helps us understand why Alice could not extract any information about Bob's observations of the second Qbit of an entangled, but spatially separated 2-Qbit system, as considered in prob. 6(b), by her observation of the first Qbit. Since the two Qbits are spatially separated, Alice is ignorant of the state of Bob's Qbit, and her density-matrix description of the system is obtained by taking the trace over the second Qbit. Although the original system was prepared as a pure state, Alice's knowledge of that system is as if her Qbit was prepared as a 50:50 mixed state of $|0\rangle_A$ and $|1\rangle_A$. Nothing Bob does changes her understanding of the first Qbit, and when she measures it, she finds it to be a $|0\rangle$ with 50% probability, or a $|1\rangle$ with 50% probability, independent of the history of second Qbit.

Skeptics, however, might infer from the result (283) that the claim (280) is incorrect. For further justification of its validity, see Box 2.6, p. 107 of Nielsen and Chuang.[99]

Dephasing

As a Qbit interacts with its environment it can be perturbed in various ways. The information content of the Qbit can become distributed over, or entangled with, that environment, such that the quality of information of the Qbit is effectively reduced. In principle, the interactions with the environment could be reversed, and the original information recovered (as in the spin-echo example given at the end of prob. 19).

But from a practical point of view, the quality of the Qbit has suffered. A quantitative measure of this is obtained by taking a partial trace over the environment of the density matrix of Qbit + environment. Analyses of this type are given the name decoherence.[100]

As an example, suppose the interaction of the Qbit $|\psi\rangle = |\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$ with its environment "merely" changes the relative phases of the amplitudes of $|0\rangle$ and $|1\rangle$, which causes no change in the energy of the Qbit.

Rather than construct the full density matrix of the system, we simply suppose that a phase-changing interaction with the environment can be approximated as a rotation (operator) $\mathsf{R}_z(\theta)$ by a small angle θ that is applied to state $|\psi\rangle$. A single such interaction transforms the density matrix of the Qbit according to eqs. (47) and (272) as,

$$\boldsymbol{\rho}' = \mathsf{R}_z(\theta)|\psi\rangle\langle\psi|\mathsf{R}_z^\dagger(\theta) = \frac{1}{2}\begin{pmatrix} e^{i\theta/2} & 0 \\ 0 & e^{-i\theta/2} \end{pmatrix}\begin{pmatrix} 1 & \pm 1 \\ \pm 1 & 1 \end{pmatrix}\begin{pmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{pmatrix}$$

[99] The density-matrix explanation of why we cannot expect controversial results from observation of one of two entangled subsystems could have been given by Bohr as an answer to the EPR "paradox" in 1935, but it was not. Some enthusiasts of EPR's argument obliquely acknowledge the impact of the density matrix by referring to it as the "destiny matrix".

[100] The spokesperson for decoherence is W. Zurek,
http://kirkmcd.princeton.edu/examples/QM/zurek_prd_24_1516_81.pdf
http://kirkmcd.princeton.edu/examples/QM/zurek_quant-ph-0306072.pdf
http://kirkmcd.princeton.edu/examples/QM/zurek_rmp_75_715_03.pdf

$$= \frac{1}{2}\begin{pmatrix} 1 & \pm e^{i\theta} \\ \pm e^{-i\theta} & 1 \end{pmatrix}. \tag{284}$$

Typically there will be a distribution of possible phase shifts, which we approximate with a Gaussian of variance 4λ,

$$P(\theta) = \frac{e^{-\theta^2/4\lambda}}{\sqrt{4\pi\lambda}}, \qquad \text{so that} \qquad \left\langle e^{\pm i\theta} \right\rangle = \frac{1}{\sqrt{4\pi\lambda}} \int e^{\pm i\theta} e^{-\theta^2/4\lambda} d\theta = e^{-\lambda}, \tag{285}$$

and hence the expectation of the density matrix (284) after one scatter is

$$\boldsymbol{\rho}' = \frac{1}{2}\begin{pmatrix} 1 & \pm e^{-\lambda} \\ \pm e^{-\lambda} & 1 \end{pmatrix}. \tag{286}$$

After a few times $1/\lambda$ such scatters, the phase information in the off-diagonal elements of the density matrix has been lost, and the originally pure state has become a mixed state,

$$\boldsymbol{\rho}' \to \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{\mathbf{I}}{2}. \tag{287}$$

If the state $|\pm\rangle$ were an input state to one of the quantum algorithms considered in previous problems, and the dephasing (287) occurred before the main operation U of the algorithm has been applied, then the effect of the algorithm would be to produce

$$\boldsymbol{\rho}'' = \mathsf{U}\boldsymbol{\rho}'\mathsf{U}^\dagger = \mathsf{U}\frac{\mathbf{I}}{2}\mathsf{U}^\dagger = \frac{\mathbf{I}}{2}, \tag{288}$$

just as if the algorithm had never been applied.

The major challenge in laboratory realization of quantum computation today is to increase the dephasing time to be longer than that needed for the computation U.

(a) Show that the density matrix for a Qbit can be written as

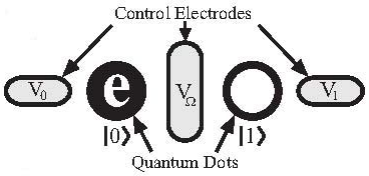
$$\boldsymbol{\rho} = \frac{\mathbf{I} + \mathbf{r}\cdot\boldsymbol{\sigma}}{2}, \tag{289}$$

where $\mathbf{r}$ is a real 3-vector with $|\mathbf{r}| \le 1$, and the maximum holds only if the Qbit is in a pure state. What is the unit vector $\hat{\mathbf{r}}$ that corresponds to the pure state

$$|\psi\rangle = e^{i\gamma}\left[\cos\frac{\alpha}{2}|0\rangle + e^{i\beta}\sin\frac{\alpha}{2}|1\rangle\right]? \tag{42}$$

(b) In prob. 7 we considered a spatially encoded Qbit for which one path of a photon was called state $|0\rangle$ and another path was called state $|1\rangle$. Another type of spatially encoded Qbit consists of a pair (or a quartet) of quantum dots [= regions in a thin silicon layer where electrodes define a potential minimum that can "trap" electrons (Earnshaw's theorem applies in three dimensions, but not in two)]; the

states $|0\rangle$ and $|1\rangle$ are defined by the presence of an electron on one or the other of the two quantum dots, as sketched in the figure below.[101]



We can also think of a spatially encoded Qbit as consisting of a pair of Qbits. Suppose that state $|0\rangle$ ($|1\rangle$) corresponds to the presence of a particle in region A (B). Then we can consider the first Qbit to have the two states $|0\rangle_A$ and $|\text{vac}\rangle_\text{A}$, where the "vacuum" state occurs when there is no particle in region A. Similarly, the second Qbit consists of the two states $|1\rangle_B$ and $|\text{vac}\rangle_\text{B}$. That is,

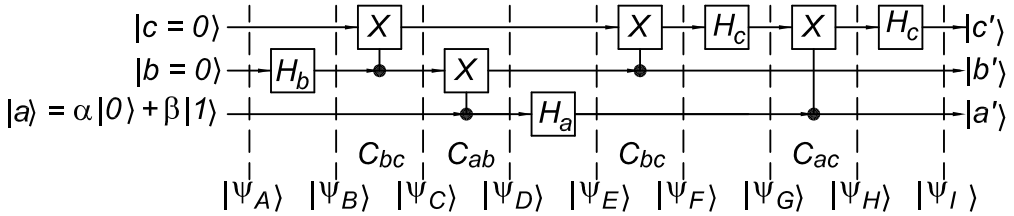
$$|0\rangle = |0\rangle_A|\text{vac}\rangle_\text{B}, \qquad \text{and} \qquad |1\rangle = |\text{vac}\rangle_\text{A}|1\rangle_\text{B}, \tag{290}$$

The system AB also supports the 2-bit states $|\text{vac}\rangle_\text{A}|\text{vac}\rangle_\text{B}$ and $|0\rangle_A|1\rangle_B$, but these are not to be used for our spatially encoded Qbit.

Write down the density operator ρ for the pure states $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$ in terms of the 2-bit states (290). What is the reduced density operator $\rho_A = \text{tr}_B(\rho)$, that summarizes the knowledge of an observer who is only aware of subsystem A? Does this reduced density operator correspond to a pure state or to a mixed state?

(c) Recall that in the circuit for quantum teleportation, discussed in prob. 6(d), Alice makes her measurements of bits $|a\rangle$ and $|b\rangle$ when the wave function of the system is

$$|\psi_E\rangle = \alpha\frac{|000\rangle + |100\rangle + |011\rangle + |111\rangle}{2} + \beta\frac{|010\rangle - |110\rangle + |001\rangle - |101\rangle}{2}. \tag{582}$$



What is the reduced density matrix of the system at this time (when $|\psi\rangle = |\psi_E\rangle$) from Bob's point of view? Recall that Bob has only bit $|c\rangle$ at this time, so his knowledge of the system is described by tracing over bits $|a\rangle$ and $|b\rangle$ in the full density matrix.

Your result should convince you that at this time Bob does not have knowledge of the initial state of bit $|a\rangle$, and must await receipt of Alice's results of her measurements of bits $|a\rangle$ and $|b\rangle$ of $|\psi_E\rangle$ before he can reconstruct the initial state of $|a\rangle$.

[101] See, for example,
http://kirkmcd.princeton.edu/examples/QM/oi_quant-ph-0412122.pdf

(d) **Other Types of Bit Errors**

A more general view of interactions of a Qbit $|\psi\rangle$ with its environment begins by supposing that initially both the Qbit and the environment are pure states, and that the entire system is in a direct product state. Then, we write the initial state of the environment as

$$|e_i\rangle_{\rm env}, \tag{291}$$

and the initial state of our system as

$$|\Psi\rangle = |e_i\rangle_{\rm env} \otimes |\psi\rangle. \tag{292}$$

The density operator $\boldsymbol{\rho}_\Psi$ of the whole system is then

$$\boldsymbol{\rho}_\Psi = |\Psi\rangle\langle\Psi| = |e_i\rangle_{\rm env} \otimes |\psi\rangle\langle\psi| \otimes \langle e_i|_{\rm env} = \boldsymbol{\rho}_{\rm env} \otimes \boldsymbol{\rho}_\psi, \tag{293}$$

where $\boldsymbol{\rho}_\psi = |\psi\rangle\langle\psi|$ and $\boldsymbol{\rho}_{\rm env} = |e_i\rangle_{\rm env}\langle e_i|_{\rm env}$.

The time evolution of the entire system is described by a unitary operator U according to eq. (272),

$$\boldsymbol{\rho}'_\Psi = \mathsf{U}\boldsymbol{\rho}_\Psi\mathsf{U}^\dagger = \mathsf{U}|e_i\rangle_{\rm env}\langle e_i|_{\rm env} \otimes \boldsymbol{\rho}_\psi\mathsf{U}^\dagger. \tag{294}$$

If, as will be typical, the state of the environment is not accessible to us, we should describe the state of the Qbit $|\psi'\rangle = |\psi(t)\rangle$ by the density operator $\boldsymbol{\rho}'_\psi$ that is the result of taking the partial trace over the environment of the density operator $\boldsymbol{\rho}_\Psi$ of the whole system. The initial density operator $\boldsymbol{\rho}_\psi$ is, of course, $|\psi\rangle\langle\psi|$, as can be obtained by taking the partial trace of eq. (293). The time evolution of $\boldsymbol{\rho}'_\psi$ follows from eq. (294),

$$\begin{aligned}\boldsymbol{\rho}'_\psi &= {\rm tr}_{\rm env}(\boldsymbol{\rho}'_\Psi) = \sum_k \langle e_k|_{\rm env}\boldsymbol{\rho}'_\Psi|e_k\rangle_{\rm env} = \sum_k \langle e_k|_{\rm env}\mathsf{U}|e_i\rangle_{\rm env}\langle e_i|_{\rm env} \otimes \boldsymbol{\rho}_\psi\mathsf{U}^\dagger|e_k\rangle_{\rm env} \\ &\equiv \sum_k \mathsf{E}_k\boldsymbol{\rho}_\psi\mathsf{E}_k^\dagger,\end{aligned} \tag{295}$$

where the $|e_k\rangle_{\rm env}$ are basis states of the environment, and the operators E_k are given by

$$\mathsf{E}_k = \langle e_k|_{\rm env}\mathsf{U}|e_i\rangle_{\rm env}, \tag{296}$$

which are neither unitary nor hermitian, in general. However, they do obey the condition

$$\sum_k \mathsf{E}_k^\dagger\mathsf{E}_k = \mathbf{I}, \tag{297}$$

which follows from taking the trace of eq. (295),

$$1 = {\rm tr}(\boldsymbol{\rho}'_\psi) = {\rm tr}\left(\sum_k \mathsf{E}_k\boldsymbol{\rho}_\psi\mathsf{E}_k^\dagger\right) = {\rm tr}\left(\sum_k \mathsf{E}_k^\dagger\mathsf{E}_k\boldsymbol{\rho}_\psi\right), \tag{298}$$

and noting that eq. (297) must be true for eq. (298) to hold for arbitrary $\boldsymbol{\rho}_\psi$.

The operators E_k can be represented by 2×2 matrices, so we can write

$$\mathsf{E}_k = a_0\mathbf{I} + \mathbf{a}\cdot\boldsymbol{\sigma} = a_0\mathbf{I} + a_x\boldsymbol{\sigma}_x + a_y\boldsymbol{\sigma}_y + a_z\boldsymbol{\sigma}_z, \tag{299}$$

recalling eq. (413). The complex coefficients a_j depend on the index k, and do not, in general, satisfy the conditions that permitted us to write a 2×2 unitary matrix in the form (37).

The form (299) leads to say that the Pauli operators $\boldsymbol{\sigma}_j$ describe three classes of "errors" which the environment can induce upon the Qbit $|\psi\rangle$:

i. The Pauli operator $\boldsymbol{\sigma}_z$ is associated with a phase-flip error, $\psi_0|0\rangle + \psi_1|1\rangle \to \psi_0|0\rangle - \psi_1|1\rangle$.
ii. The Pauli operator $\boldsymbol{\sigma}_x$ is associated with a bit-flip error, $\psi_0|0\rangle + \psi_1|1\rangle \to \psi_1|0\rangle + \psi_0|1\rangle$.
iii. The Pauli operator $\boldsymbol{\sigma}_y$ is associated with a bit-phase-flip error, $\psi_0|0\rangle + \psi_1|1\rangle \to -i\psi_1|0\rangle + i\psi_0|1\rangle$.

We will use this language in prob. 21 when we discuss methods of quantum error correction.

Now, the problem:

Suppose that only one of the three basic types of bit errors occurs, say that associated with Pauli operator $\boldsymbol{\sigma}_j$. Then it suffices to characterize the environment by only two states, its initial state $|e_i\rangle_{\rm env}$, and the state $|e_j\rangle_{\rm env}$ of the environment that results when our Qbit $|\psi\rangle$ suffers an error due to its interaction with the environment. The error transformation (295) now contains only two terms,

$$\boldsymbol{\rho}'_\psi(\boldsymbol{\sigma}_j) = \mathsf{E}_i \boldsymbol{\rho}_\psi \mathsf{E}_i^\dagger + \mathsf{E}_j \boldsymbol{\rho}_\psi \mathsf{E}_j^\dagger. \tag{300}$$

The first term corresponds to no error, so we can write $\mathsf{E}_i = a\mathbf{I}$ for some complex number a, while the second term corresponds to an error associated with the Pauli matrix $\boldsymbol{\sigma}_j$, so we write $\mathsf{E}_j = b\boldsymbol{\sigma}_j$ for some complex number b. The condition (297) requires that $|a|^2 + |b|^2 = 1$. There is no loss of generality to take a and b real, with $b = \sqrt{p}$ and $a = \sqrt{1-p}$, where p is the probability that the bit error occurs. The bit evolution (300) for a single type of error can now be written

$$\boldsymbol{\rho}'_\psi(p, \boldsymbol{\sigma}_j) = (1-p)\boldsymbol{\rho}_\psi + p\ \boldsymbol{\sigma}_j \boldsymbol{\rho}_\psi \boldsymbol{\sigma}_j. \tag{301}$$

The set of density operators $\boldsymbol{\rho}_\psi$ for all possible initial pure-state Qbits can be described as a sphere of unit radius in Bloch space, according to eq. (289). What is the corresponding surface in Bloch space for the density operator $\boldsymbol{\rho}'_\psi(p, \boldsymbol{\sigma}_j)$ for $j = x$, y and z?

Error transformations in which all three types of bit errors occur with equal strength have special interest. In particular, consider the case that

$$\boldsymbol{\rho}'_\psi = \left(1 - \frac{3p}{4}\right)\boldsymbol{\rho}_\psi + \frac{p}{4}(\boldsymbol{\sigma}_x \boldsymbol{\rho}_\psi \boldsymbol{\sigma}_x + \boldsymbol{\sigma}_y \boldsymbol{\rho}_\psi \boldsymbol{\sigma}_y + \boldsymbol{\sigma}_z \boldsymbol{\rho}_\psi \boldsymbol{\sigma}_z). \tag{302}$$

Show that this form can be also be written as

$$\boldsymbol{\rho}'_\psi = (1-p)\,\boldsymbol{\rho}_\psi + p\,\frac{\mathbf{I}}{2}, \tag{303}$$

which implies that the initial (pure) state is either left alone with probability $1-p$, or turned into the mixed state $\mathbf{I}/2$ with probability p. This is called a depolarizing error.

21. Quantum Error Correction[102]

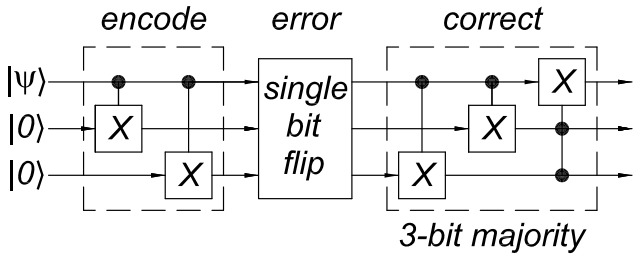
Quantum error correction is an extrapolation of classical procedures wherein the contents of a bit are redundantly coded to permit recovery from some types of errors. *The spirit of this type of error correction is* safety in numbers. *For a different approach to the correction of certain types of errors, see prob. 22(d).*

The simplest classical procedure[103] is to encode the Cbit states (written $|\bar{0}\rangle$ and $|\bar{1}\rangle$ where the bar indicates the presence of an error-correcting code) as triplets of ordinary Cbits,

$$|\bar{0}\rangle_{\text{bit flip}} = |0\rangle|0\rangle|0\rangle = |000\rangle, \qquad |\bar{1}\rangle_{\text{bit flip}} = |1\rangle|1\rangle|1\rangle = |111\rangle. \tag{304}$$

There is only one type of classical error for Cbits (assuming that the bit is not destroyed), namely a bit flip. The coding (304) is robust against the occurrence of exactly one (or zero) bit flips, by setting a damaged Cbit to the state consistent with the majority of its constituent bits. That is, if we find a coded Cbit to be $|100\rangle$ we set it to $|000\rangle$, but if we find $|110\rangle$ we set it to $|111\rangle$.

A quantum version of this 3-bit code is shown below.[104]



If p is the probability of an individual Cbit flip, then the probability that the error correction fails is $3p^2(1-p)+p^3 = 3p^2-2p^3 \approx 3p^2$. Hence, if we desired a failure rate of the coded bits of less than, say, 10^{-15}, the probability of an individual bit failure must be less than 1.7×10^{-7}, *etc.* If the Cbit coding were based on 5 bits, the probability of a coding failure would be $10p^3 + ...$, and a failure rate of 10^{-15} would be achieved if $p < 4.6 \times 10^{-3}$.

Shor's 9-Bit Quantum Error Correction Code[105]

As we saw in prob. 20(d), Qbits can suffer three different basics types of errors, called bit flip, bit-phase flip, and phase flip. Shor noted that coding of the type (304) offers no protection against phase flip (or against bit-phase flip), but that if the only possible error were a phase flip, then this could be corrected using a 3-Qbit code as shown on the right in the figure on the top of the next page.

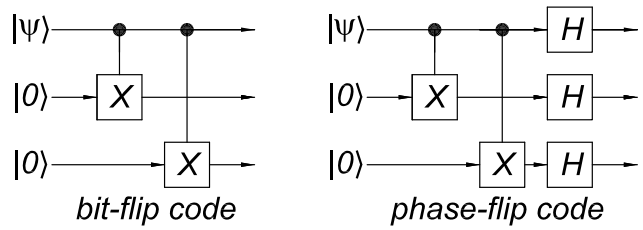
---

[102] Problem 21 is covered in chap. 10 of Nielsen and Chuang.

[103] von Neumann (1944),
http://kirkmcd.princeton.edu/examples/QM/von-neumann_logics.pdf

[104] The 3-bit majority circuit was introduced in prob. 9(g). Note that the ancillary bits are left in an unknown state, so if we wish to re-encode the state $|\psi\rangle$ we must either use new ancillary bits, or apply energy-consuming resets as discussed in prob. 2.

[105] http://kirkmcd.princeton.edu/examples/QM/shor_pra_52_R2493_95.pdf



The coded basis states are

$$|\bar{0}\rangle_{\text{phase flip}} = \frac{|0\rangle + |1\rangle}{\sqrt{2}} \frac{|0\rangle + |1\rangle}{\sqrt{2}} \frac{|0\rangle + |1\rangle}{\sqrt{2}} = |+\!+\!+\rangle, \tag{305}$$

$$|\bar{1}\rangle_{\text{phase flip}} = \frac{|0\rangle - |1\rangle}{\sqrt{2}} \frac{|0\rangle - |1\rangle}{\sqrt{2}} \frac{|0\rangle - |1\rangle}{\sqrt{2}} = |-\!-\!-\rangle. \tag{306}$$

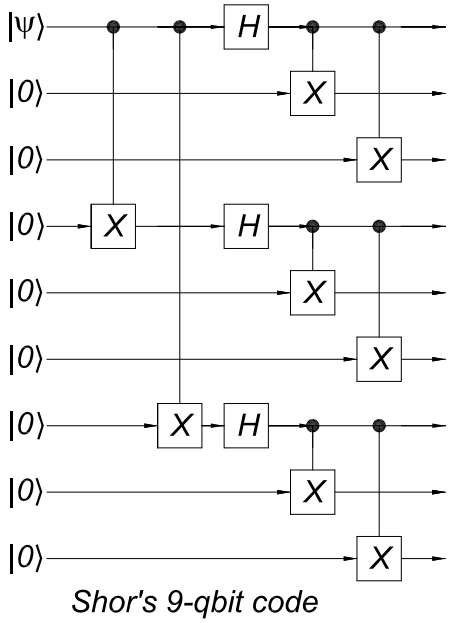
A phase-flip error turns the superposition $|0\rangle + |1\rangle$ into $|0\rangle - |1\rangle$, and *vice versa*. Shor noted that both of the logical Qbits (305)-(306) are eigenstates of the operators $\mathsf{X}_1\mathsf{X}_2$ and $\mathsf{X}_2\mathsf{X}_3$ with eigenvalue $+1$. If a phase-flip error had occurred on, say, the first bit of either $|\bar{0}\rangle_{\text{phase flip}}$ or $|\bar{1}\rangle_{\text{phase flip}}$, then the "damaged" state is still an eigenstate of the operators $\mathsf{X}_1\mathsf{X}_2$ and $\mathsf{X}_2\mathsf{X}_3$, but now the eigenvalues are -1 and $+1$, respectively. So if we apply these operators to the coded Qbits, those Qbits are not changed but we learn about the presence of a phase-flip error from the pattern of the eigenvalues.

Shor then proposed "tripling" the phase-flip coding to the 9-bit forms

$$|\bar{0}\rangle_{\text{Shor}} = \frac{1}{2\sqrt{2}}(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle), \tag{307}$$

$$|\bar{1}\rangle_{\text{Shor}} = \frac{1}{2\sqrt{2}}(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle), \tag{308}$$

which provide protection against a single bit-flip or phase-flip error. Shor's code could be realized by the circuit below.



We cannot, however, simply measure the state of Shor's encoded Qbit (as a first step in the error correction procedure) since this would destroy the superposition in eqs. (307)-(308), and in any case Shor's states $|\bar{0}\rangle$ and $|\bar{1}\rangle$ cannot be distinguished by measurement.[106]

Shor's suggestion is to measure the effect of various operators on the coded Qbit,

$$|\bar{\psi}\rangle = a|\bar{0}\rangle + b|\bar{1}\rangle, \tag{309}$$

where the operators are chosen such that states $|\bar{0}\rangle$ and $|\bar{1}\rangle$, as well as any state obtained from these via a single Qbit error, are eigenstates of the operators. (See also prob. 22(c).) Further, the eigenvalues should be the same for $|\bar{0}\rangle$ and $|\bar{1}\rangle$, and the same for any pair of states obtained from these by a single Qbit error. Then, the coded state $|\bar{\psi}\rangle$ is unchanged, except for a possible overall phase, by the measurement. If the results of the measurements permits us to deduce what kind of single-Qbit error, if any, has occurred, then that error can be corrected and the coded state restored to form (309) for any values of a and b.

As a first example, consider a single bit-flip error that occurs in the first triplet of the 9-bit states. Then, the desired triplet combinations will be altered according to

$$|000\rangle \pm |111\rangle \to |100\rangle \pm |011\rangle, \quad |010\rangle \pm |101\rangle, \quad \text{or} \quad |001\rangle \pm |110\rangle \qquad \text{(bit flip)}. \tag{310}$$

Suitable operators to diagnose these errors are $\mathsf{Z}_1\mathsf{Z}_2$, $\mathsf{Z}_2\mathsf{Z}_3$ and $\mathsf{Z}_1\mathsf{Z}_3$, and it suffices to apply any two of them. Thus,

$$\mathsf{Z}_1\mathsf{Z}_2(|000\rangle \pm |111\rangle) = +(|000\rangle \pm |111\rangle), \qquad \mathsf{Z}_2\mathsf{Z}_3(|000\rangle \pm |111\rangle) = +(|000\rangle \pm |111\rangle), \tag{311}$$
$$\mathsf{Z}_1\mathsf{Z}_2(|100\rangle \pm |011\rangle) = -(|100\rangle \pm |011\rangle), \qquad \mathsf{Z}_2\mathsf{Z}_3(|100\rangle \pm |011\rangle) = +(|100\rangle \pm |011\rangle), \tag{312}$$
$$\mathsf{Z}_1\mathsf{Z}_2(|010\rangle \pm |101\rangle) = -(|010\rangle \pm |101\rangle), \qquad \mathsf{Z}_2\mathsf{Z}_3(|010\rangle \pm |101\rangle) = -(|010\rangle \pm |101\rangle), \tag{313}$$
$$\mathsf{Z}_1\mathsf{Z}_2(|001\rangle \pm |110\rangle) = +(|001\rangle \pm |110\rangle), \qquad \mathsf{Z}_2\mathsf{Z}_3(|001\rangle \pm |110\rangle) = -(|001\rangle \pm |110\rangle). \tag{314}$$

The four possible sets of results of measurements of operators

$$\mathsf{Z}_1\mathsf{Z}_2 \qquad \text{and} \qquad \mathsf{Z}_2\mathsf{Z}_3 \tag{315}$$

corresponds to the four possible error conditions in the first bit triplet: $(+,+)$ implies no error, $(-,+)$ implies bit 1 was flipped, $(-,-)$ implies bit 2 was flipped, and $(+,-)$ implies bit 3 was flipped.

The error recovery procedure is to flip bits 1, 2 or 3 corresponding to the results $(-,+)$, $(-,-)$ or $(+,-)$, respectively. Of, we do nothing when we obtain the result $(+,+)$.

However, if two or three bits have been flipped in the first triplet of state (309), the pattern of eigenvalues of the operators (315) is:

$$\mathsf{Z}_1\mathsf{Z}_2(|110\rangle \pm |001\rangle) = +(|110\rangle \pm |001\rangle), \qquad \mathsf{Z}_2\mathsf{Z}_3(|110\rangle \pm |001\rangle) = -(|110\rangle \pm |001\rangle), \tag{316}$$
$$\mathsf{Z}_1\mathsf{Z}_2(|101\rangle \pm |010\rangle) = -(|101\rangle \pm |010\rangle), \qquad \mathsf{Z}_2\mathsf{Z}_3(|101\rangle \pm |010\rangle) = -(|101\rangle \pm |010\rangle), \tag{317}$$
$$\mathsf{Z}_1\mathsf{Z}_2(|011\rangle \pm |100\rangle) = -(|011\rangle \pm |100\rangle), \qquad \mathsf{Z}_2\mathsf{Z}_3(|011\rangle \pm |100\rangle) = +(|011\rangle \pm |100\rangle), \tag{318}$$
$$\mathsf{Z}_1\mathsf{Z}_2(|111\rangle \pm |000\rangle) = +(|111\rangle \pm |000\rangle), \qquad \mathsf{Z}_2\mathsf{Z}_3(|111\rangle \pm |000\rangle) = +(|111\rangle \pm |000\rangle). \tag{319}$$

[106] Measurement is a key step in most quantum error-correction procedures. This is a kind of variant on the quantum Zeno effect whereby an unstable state can be kept from decaying if it is observed often enough. http://kirkmcd.princeton.edu/examples/QM/misra_jmp_18_756_77 See also sec. 12.5 Of *Introduction to Quantum Mechanics*, 2nd ed., by D.J. Griffiths (Prentice Hall, 2005).

Indeed, we see that the result $(+,+)$ corresponds to either no error or to all 3 bits having flipped, while the result $(+,-)$ occurs when either bit 3 has flipped or when both bits 1 and 2 have flipped, *etc.* If we apply the error-recovery procedure described above, then the results are $|000\rangle \pm |111\rangle \rightarrow |111\rangle \pm |000\rangle = \pm(|000\rangle \pm |111\rangle)$, which could be described as a phase flip error in the coded bits $|\bar{0}\rangle$ and $|\bar{1}\rangle$.

If the only type of Qbit errors were bit flips, a 3-Qbit code would suffice. As we will see below, corrections of phase-flip errors requires use of more than 3 Qbits. In Shor's 9-bit code, (307)-(308), additional bit-flip errors can occur in the additional bits, but we can use the two operators

$$\mathsf{Z}_4\mathsf{Z}_5 \qquad \text{and} \qquad \mathsf{Z}_5\mathsf{Z}_6, \tag{320}$$

to detect a single error in the second bit triplet, and the two operators

$$\mathsf{Z}_7\mathsf{Z}_8 \qquad \text{and} \qquad \mathsf{Z}_8\mathsf{Z}_9, \tag{321}$$

to spot a single error in the third triplet.

Turning to the issue of phase-flips errors, we see that the effect of such an error on any bit of a triplet state is

$$|000\rangle \pm |111\rangle \rightarrow |000\rangle \mp |111\rangle \qquad \text{(phase flip)}. \tag{322}$$

To diagnose this error, we again seek operators for which both states $|\bar{0}\rangle$ and $|\bar{1}\rangle$ are eigenvectors with the same eigenvalue, and for which the damaged versions of these states are also eigenvectors with the same eigenvalue, but the eigenvalues are different for damaged and undamaged states.

Shor noted that the operator $\mathsf{X}_1\mathsf{X}_2\mathsf{X}_3$ plays the same role for his 9-Qbit code as the operator X_1 does for the 3-Qbit phase-flip code (305)-(306):

$$\mathsf{X}_1\mathsf{X}_2\mathsf{X}_3(|000\rangle \pm |111\rangle) = |111\rangle \pm |000\rangle = \pm(|000\rangle \pm |111\rangle). \tag{323}$$

Operators for which the logical Qbits $|\bar{0}\rangle$ and $|\bar{1}\rangle$ are eigenstates with eigenvalue $+1$, and which also diagnose the presence of a single phase-flip error among the 9 physical Qbits, can be constructed out of six X_j's:

$$\mathsf{X}_1\mathsf{X}_2\mathsf{X}_3\mathsf{X}_4\mathsf{X}_5\mathsf{X}_6, \qquad \mathsf{X}_4\mathsf{X}_5\mathsf{X}_6\mathsf{X}_7\mathsf{X}_8\mathsf{X}_9, \qquad \text{or} \qquad \mathsf{X}_1\mathsf{X}_2\mathsf{X}_3\mathsf{X}_7\mathsf{X}_8\mathsf{X}_9. \tag{324}$$

(a) Describe a procedure for diagnosis and correction of a single phase-flip error in Shor's coded Qbits (307)-(308).

(b) Verify that the procedures for diagnosis and correction of a single bit-flip error and a single phase-flip error in Shor's coded Qbits automatically correct a single bit-phase-flip error as well. *It suffices to consider a phase-flip error in the first of the nine Qbits of a coded Qbit.*

(c) If p_x, p_y and p_z are the probabilities of single-Qbit errors of types bit flip, bit-phase flip, and phase flip, respectively, what are the leading probabilities of failure of Shor's error-correcting scheme?

(d) We saw in prob. 20(d) that a general error on a single (physical) Qbit is described by a set of operators $\{\mathsf{E}_k\}$ each of the form

$$\mathsf{E}_k = a_0\mathbf{I} + \mathbf{a}\cdot\boldsymbol{\sigma} = a_0\mathbf{I} + a_x\mathsf{X} + a_y\mathsf{Y} + a_z\mathsf{Z}, \tag{299}$$

where index k refers to a possible final state of the environment after its error-inducing interaction with the Qbit. Verify that the procedures for diagnosis and correction of a single bit-flip error and a single phase-flip error in Shor's coded bits also correct for general single-bit errors. *Again, it suffices to consider only an error in the first of the nine Qbits of a coded Qbit.*

Error Correction without Measurement?[107]

The process of error diagnosis and correction of a coded state $|\bar{\psi}\rangle$ involves a set of measurements, described by (hermitian) measurement operators $\{\mathsf{M}_j\}$, and a corresponding set of unitary error-correction operators $\{\mathsf{U}_j\}$ that are applied conditionally on the results of the measurements.

We can avoid the need to perform the measurements, which may be physically awkward, at the expense of introducing an additional n ancillary Qbits that must be initially $|0\rangle_n$ at the time of each error-correction procedure.[108]

If there are m different measurement operators to be used, the number n of ancillary Qbits must be such that $m < 2^n$, so that we can associate a different, nonzero basis state of the ancillary Qbits to each measurement operator.

Then, we can create a new unitary operator U defined by

$$\mathsf{U}|\bar{\psi}\rangle|0\rangle_n \equiv \sum_{j\neq 0}(\mathsf{U}_j\mathsf{M}_j|\bar{\psi}\rangle)|j\rangle_n, \tag{325}$$

$$\mathsf{U}|\bar{\psi}\rangle|k\neq 0\rangle_n \equiv |\bar{\psi}\rangle|k\rangle_n. \tag{326}$$

(e) Verify that the operator U is indeed unitary, *i.e.*, it preserves inner products. *I could not actually verify this, and I worry that the claim is false.*

The operation on the state $|\bar{\psi}\rangle|0\rangle_n$ described by the first line of eq. (325) involves "virtual measurements" of all m types, followed by the corresponding error correction. Thus, the error is corrected without explicit measurements that force the state $|\bar{\psi}\rangle$ into one of its basis states.

However, at the end of the operation U, the ancillary Qbits are in an unknown, nonzero state. If the error correction procedure is to be applied another time (or many times as will be needed in general), we must either add new, initially zero, ancillary bits to the system, or reset the ancillary Qbits. To reset these Qbits, we must either measure them and apply the appropriate unitary reset operators, or we must invoke the energy- (and time-) consuming reset process of an unknown state that we considered in prob. 2.

[107] See Box 10.1, p. 439 of Nielsen and Chuang.

[108] The 3-Qbit circuit shown at the beginning of this problem corrects single bit-flip errors without measurement.

It appears that the use of quantum error correction will entail considerable costs: large numbers of Qbits are required, as well as pauses in the calculation to make either measurements or resets of ancillary Qbits. Or, if both measurements and resets are to be avoided, a new set of ancillary Qbits is required at every error-correction cycle.

Error Correction Codes with Fewer Than Nine Qbits

Following Shor's introduction of the 9-Qbit error correction code,[109] it was quickly realized that quantum error-correction codes could be implemented with as few as 5 physical Qbits per logical Qbit.[110] Subsequent studies suggest that a 7-bit code may be the best compromise between size and realizability.[111] These codes are more complex than Shor's 9-Qbit code. For example, the logical Qbits of Steane's 7-Qbit code are

$$\begin{aligned}
|\bar{0}\rangle_{\rm Steane} &= \frac{1}{2\sqrt{2}}(\mathbf{I}+\mathsf{X}_4\mathsf{X}_5\mathsf{X}_6\mathsf{X}_7)(\mathbf{I}+\mathsf{X}_2\mathsf{X}_3\mathsf{X}_6\mathsf{X}_7)(\mathbf{I}+\mathsf{X}_1\mathsf{X}_3\mathsf{X}_5\mathsf{X}_7)|0000000\rangle \\
&= \frac{1}{2\sqrt{2}}(\mathbf{I}+\mathsf{X}_1\mathsf{X}_3\mathsf{X}_5\mathsf{X}_7+\mathsf{X}_2\mathsf{X}_3\mathsf{X}_6\mathsf{X}_7+\mathsf{X}_4\mathsf{X}_5\mathsf{X}_6\mathsf{X}_7 \\
&\qquad +\mathsf{X}_1\mathsf{X}_2\mathsf{X}_5\mathsf{X}_6+\mathsf{X}_1\mathsf{X}_3\mathsf{X}_4\mathsf{X}_6+\mathsf{X}_2\mathsf{X}_3\mathsf{X}_4\mathsf{X}_5+\mathsf{X}_1\mathsf{X}_2\mathsf{X}_4\mathsf{X}_7) \\
&\equiv \frac{1}{2\sqrt{2}}(\mathbf{I}+\mathsf{M}_1+\mathsf{M}_2+\mathsf{M}_3+\mathsf{M}_4+\mathsf{M}_5+\mathsf{M}_6+\mathsf{M}_7)|0000000\rangle \equiv \bar{\mathbf{0}}_{\rm Steane}|0000000\rangle \\
&= \frac{1}{2\sqrt{2}}(|0000000\rangle+|1010101\rangle+|0110011\rangle+|0001111\rangle \\
&\qquad +|1100110\rangle+|1011010\rangle+|0111100\rangle+|1101001\rangle), \qquad (327) \\
|\bar{1}\rangle_{\rm Steane} &= \frac{1}{2\sqrt{2}}(\mathbf{I}+\mathsf{X}_4\mathsf{X}_5\mathsf{X}_6\mathsf{X}_7)(\mathbf{I}+\mathsf{X}_2\mathsf{X}_3\mathsf{X}_6\mathsf{X}_7)(\mathbf{I}+\mathsf{X}_1\mathsf{X}_3\mathsf{X}_5\mathsf{X}_7)|1111111\rangle \\
&= \frac{1}{2\sqrt{2}}(|1111111\rangle+|0101010\rangle+|1001100\rangle+|1110000\rangle \\
&\qquad +|0011001\rangle+|0100101\rangle+|1000011\rangle+|0010110\rangle) \qquad (328) \\
&= \mathsf{X}_1\mathsf{X}_2\mathsf{X}_3\mathsf{X}_4\mathsf{X}_5\mathsf{X}_6\mathsf{X}_7|\bar{0}\rangle_{\rm Steane} \equiv \bar{\mathsf{X}}_{\rm Steane}|\bar{0}\rangle_{\rm Steane} = \bar{\mathsf{X}}_{\rm Steane}\bar{\mathbf{0}}_{\rm Steane}|0000000\rangle \\
&= \bar{\mathbf{0}}_{\rm Steane}|1111111\rangle = \bar{\mathbf{0}}_{\rm Steane}\bar{\mathsf{X}}_{\rm Steane}|0000000\rangle \equiv \bar{\mathbf{1}}_{\rm Steane}|0000000\rangle.
\end{aligned}$$

For the record, Steane's error-correction procedure involves measurement of the six operators

$$\mathsf{X}_4\mathsf{X}_5\mathsf{X}_6\mathsf{X}_7, \quad \mathsf{X}_2\mathsf{X}_3\mathsf{X}_6\mathsf{X}_7, \quad \mathsf{X}_1\mathsf{X}_3\mathsf{X}_5\mathsf{X}_7, \quad \mathsf{Z}_4\mathsf{Z}_5\mathsf{Z}_6\mathsf{Z}_7, \quad \mathsf{Z}_2\mathsf{Z}_3\mathsf{Z}_6\mathsf{Z}_7, \quad \mathsf{Z}_1\mathsf{Z}_3\mathsf{Z}_5\mathsf{Z}_7, \qquad (329)$$

for which both $|\bar{0}\rangle_{\rm Steane}$ and $|\bar{1}\rangle_{\rm Steane}$ are eigenvectors with eigenvalue $+1$ (while the states resulting from a single Qbit error in $|\bar{0}\rangle_{\rm Steane}$ and $|\bar{1}\rangle_{\rm Steane}$ are eigenstates of these operators with eigenvalues ± 1).[112] If the measured eigenvalues of the first three operators are all $+1$, then the pattern eigenvalues of the second three operators identifies

[109] Shor's code was independently developed by Steane,
http://kirkmcd.princeton.edu/examples/QM/steane_prl_77_793_96.pdf

[110] http://kirkmcd.princeton.edu/examples/QM/bennett_pra_54_3824_96.pdf
http://kirkmcd.princeton.edu/examples/QM/laflamme_prl_77_198_96.pdf

[111] http://kirkmcd.princeton.edu/examples/QM/calderbank_pra_54_1098_96.pdf
http://kirkmcd.princeton.edu/examples/QM/steane_prsl_a452_2551_96.pdf

[112] Indeed, $|\bar{0}\rangle_{\rm Steane}$ and $|\bar{1}\rangle_{\rm Steane}$ are eigenvectors with eigenvalue $+1$ of all seven operators M_j, as well as operator $\bar{\mathbf{0}}_{\rm Steane}$, introduced in eq. (327). Note also that all seven operators M_j commute with each other.

the presence of a bit-flip error. Similarly, if the measured eigenvalues of the second three operators are all $+1$, then the pattern eigenvalues of the first three operators identifies the presence of a phase-flip error. And, if there are eigenvalues -1 among the measurements of both the first three and the second three operators, the pattern of eigenvalues identifies the presence of a bit-phase-flip error.

We leave it to your discretion to verify these claims.

Error Codes That Can Recover from More Than One Error

In a recent paper, Crépeau *et al.* argue that approximate error correction codes can be constructed in which a logical Qbit is encoded on n physical Qbits in such a manner that excellent recovery is possible from $(n-1)/2$ errors.[113] That is, an appropriate 5-Qbit code can approximately correct for two errors, *etc.*

[113] http://kirkmcd.princeton.edu/examples/QM/crepeau_quant-ph-0503139.pdf

22. Fault-Tolerant Quantum Computation[114]

The possibility of errors in quantum computation is by no means restricted to errors in the Qbits themselves. Indeed, a more general type of error is the failure of a quantum operation (gate) to perform as desired. Failures of gates can be more troublesome than individual bit errors, since the action of a gate (such as the C-NOT) may be to distribute the information content of one bit among many others.

Fault-tolerant quantum computation follows the spirit of Qbit error correction introduced in prob. 21, which is to make multiple calculations of each quantum operation, and to take the majority result as the correct one. Typically, if the Qbit error-correction code involves n physical Qbits per logical Qbit, then n calculations of each quantum operation will be made.

Again, we seek safety in numbers. For each logical gate in our basic circuit, we utilize ≈ 20 physical gates in a fault-tolerant version of the circuit. Rules of thumb, optimistically called threshold theorems, then suggest that there are about 10^4 places where an error might occur in the fault-tolerant implementation of a quantum gate. So, if the probability of a single error is less that 10^{-4}, there will typically be 0 or 1 errors in the operation of the gate, and the case of exactly 1 error will be properly corrected.

In prob. 12 we saw that any quantum computation can be built up from only 3 types of gates, the Hadamard gate H, the $\mathsf{Z}^{1/4}$ gate, and the C-NOT gate. So, a sense of the methods of fault-tolerant quantum computation can be gained by demonstration of fault-tolerant versions of each of these 3 gates, following the initial work by Shor.[115] We will close with discussion of a recent suggestion by Grover for a different approach to fault-tolerant computation.

When we wish to implement a quantum operation U with logical Qbits $|\bar{0}\rangle$ and $|\bar{1}\rangle$, we need to construct an operator $\bar{\mathsf{U}}$ whose effect on the logical Qbits is the same as that of operator U on physical Qbits. The details of the construction of $\bar{\mathsf{U}}$ depend on the choice of the error-correction code. The examples given here will be based on Steane's code, eqs. (327)-(328), which has the merit that the constructions of many important gates $\bar{\mathsf{U}}_{\rm Steane}$ are straightforward.

(a) Fault-Tolerant Hadamard Gate

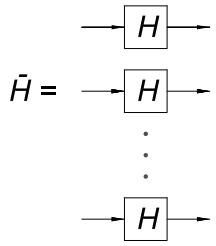
If the logical version $\bar{\mathsf{U}}$ of a single-Qbit operator U is obtained simply by applying the operator U to all n physical Qbits, we say that the logical gate $\bar{\mathsf{U}}$ is transverse. Not all single-Qbit operators $\bar{\mathsf{U}}_{\rm Steane}$ are transverse with respect to the Steane coding of logical Qbits. It turns out that the Pauli gates $\bar{\mathsf{X}}_{\rm Steame}$ (introduced in eq. (327) and $\bar{\mathsf{Z}}_{\rm Steane} = \mathsf{Z}_1\mathsf{Z}_2\mathsf{Z}_3\mathsf{Z}_4\mathsf{Z}_4\mathsf{Z}_5\mathsf{Z}_6\mathsf{Z}_7$, as well as the Hadamard gate $\bar{\mathsf{H}}_{\rm Steane} = (\bar{\mathsf{X}}_{\rm Steane} + \bar{\mathsf{Z}}_{\rm Steane})/\sqrt{2}$, are transverse for Steane's code (327)-(328) (in which the logical Qbit $|\bar{0}\rangle$ contains an even number of $|1\rangle$'s while the logical Qbit $|\bar{1}\rangle$ contains an odd number). However, the gates $\bar{\mathsf{Y}}_{\rm Steane}$ and $\bar{\mathsf{Z}}^{1/4}_{\rm Steane}$ are not transverse.

Here, we consider a transverse construction for the logical gate $\bar{\mathsf{H}}_{\rm Steane}$,

$$\bar{\mathsf{H}}_{\rm Steane} = \mathsf{H}_1\mathsf{H}_2\mathsf{H}_3\mathsf{H}_4\mathsf{H}_5\mathsf{H}_6\mathsf{H}_7. \tag{330}$$

[114] Problem 22 is covered sec. 10.6 of Nielsen and Chuang.

[115] http://kirkmcd.princeton.edu/examples/QM/shor_quant-ph-9605011.pdf



The 7-fold direct-product operator $\bar{\mathsf{H}}_{\rm Steane}$ *takes each of the eight 7-Qbit states of* $|\bar{0}\rangle$ *or* $|\bar{1}\rangle$ *into a superposition of* $2^7 = 128$ *states, for an overall superposition of 1024 states. It is straightforward but tedious to show explicitly that this superposition reduces to the sixteen 7-Qbit states contained in* $(|\bar{0}\rangle \pm |\bar{1}\rangle)/\sqrt{2}$.

This result is, of course, expected, since the relations

$$\bar{\mathsf{X}}_{\rm Steane}|\bar{0}\rangle_{\rm Steane} = |\bar{1}\rangle_{\rm Steane}, \qquad \bar{\mathsf{X}}_{\rm Steane}|\bar{1}\rangle_{\rm Steane} = |\bar{0}\rangle_{\rm Steane}, \tag{331}$$

$$\bar{\mathsf{Z}}_{\rm Steane}|\bar{0}\rangle_{\rm Steane} = |\bar{0}\rangle_{\rm Steane}, \qquad \bar{\mathsf{Z}}_{\rm Steane}|\bar{1}\rangle_{\rm Steane} = -|\bar{1}\rangle_{\rm Steane}, \tag{332}$$

imply that

$$\bar{\mathsf{H}}_{\rm Steane}|\bar{0}\rangle_{\rm Steane} = \frac{|\bar{0}\rangle_{\rm Steane} + |\bar{1}\rangle_{\rm Steane}}{\sqrt{2}}, \qquad \bar{\mathsf{H}}_{\rm Steane}|\bar{1}\rangle_{\rm Steane} = \frac{|\bar{0}\rangle_{\rm Steane} - |\bar{1}\rangle_{\rm Steane}}{\sqrt{2}}. \tag{333}$$

If a single error occurs on one of the physical Qbits during or after the use of the logical gate $\bar{\mathsf{H}}_{\rm Steane}$, the effect is to apply one of the operators X, Y or Z to that Qbit. If the logical gate $\bar{\mathsf{H}}_{\rm Steane}$ is immediately followed by (Steane's version of) Qbit error correction on all of the physical Qbits, the error will be corrected.

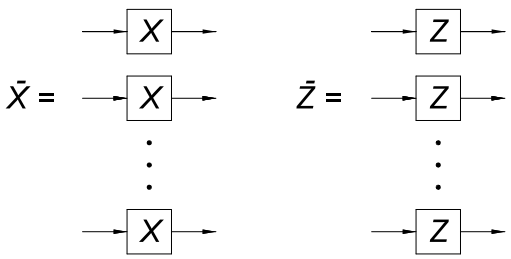
It could also happen that one of the physical Qbits suffers an error just before the gate $\bar{\mathsf{H}}_{\rm Steane}$ is applied. If so, the effective operation on that Qbit would be HX, HY or HZ. Recalling eq. (62), we have that

$$\mathsf{HX} = \mathsf{ZH}, \qquad \mathsf{HY} = \mathsf{YH}, \qquad \text{and} \qquad \mathsf{HZ} = \mathsf{XH}. \tag{334}$$

Thus, a single Qbit error before use of $\bar{\mathsf{H}}_{\rm Steane}$ is equivalent to a single Qbit error (of a different type in general) that occurs after $\bar{\mathsf{H}}_{\rm Steane}$, and hence this error would also be corrected during error correction of all Qbits following operation $\bar{\mathsf{H}}_{\rm Steane}$.

We therefore say that the logical gate $\bar{\mathsf{H}}_{\rm Steane}$ is fault tolerant, in that a single Qbit error associated with use of that gate will be corrected during a subsequent application of quantum error correction to all of the physical Qbits.

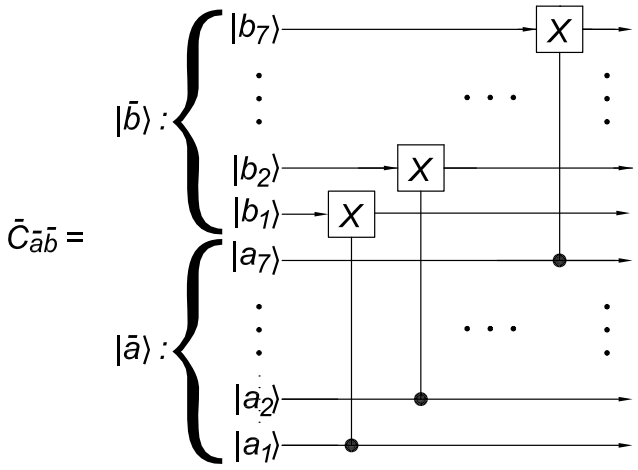
Show (briefly) that the transverse logical gates $\bar{\mathsf{X}}_{\rm Steane}$ and $\bar{\mathsf{Z}}_{\rm Steane}$ are also fault tolerant in this sense.



(b) **Fault-Tolerant C-NOT Gate**

It is very agreeable that the logical operator $\mathsf{C\text{-}NOT}_{\rm Steane} = \bar{\mathsf{C}}_{\bar{a}\bar{b}}$, for use with a pair of Steane's logical Qbits, $|\bar{a}\rangle$ and $|\bar{b}\rangle$, can be constructed transversely, as shown in the figure below.

Again, it is not self-evident that this construction works, but it is straightforward to verify explicitly that the superpositions of 64 states in the output states $|\bar{b}\rangle$ from $\bar{\mathsf{C}}_{\bar{a}\bar{b}}$ reduce to the desired superpositions of only 8 states according to the truth table for the C-NOT *operator.*



Give a (brief) analytic argument as to why the transverse construction of $\mathsf{C\text{-}NOT}_{\rm Steane}$ works, based on the operators $\bar{\mathbf{0}}_{\rm Steane}$ and $\bar{\mathbf{1}}_{\rm Steane}$ that were introduced in eqs. (327)-(328).

With the arguments of part (a) in mind, we readily see that the gate $\bar{\mathsf{C}}_{\bar{a}\bar{b}}$ is fault tolerant against a single Qbit error. The only difficult case is an error that occurs on one of the physical Qbits of $|\bar{a}\rangle$ before or during the C-NOT operation, because such an error propagates to an error in the final state of $|\bar{b}\rangle$ as well.

However, a single error in $|\bar{a}\rangle$ leads to, at most, one error in each of $|\bar{a}\rangle$ and $|\bar{b}\rangle$. When we apply separate error-correction procedures to both the logical Qbits $|\bar{a}\rangle$ and $|\bar{b}\rangle$ after the C-NOT operation, both of these errors are corrected.

Thus, the gate $\bar{\mathsf{C}}_{\bar{a}\bar{b}}$ is fault tolerant.

(c) **Fault-Tolerant $\mathsf{Z}^{1/4}$ Gate**

We first digress slightly to deduce the fractional powers p such that the gate $\bar{\mathsf{Z}}^p$ could be implemented transversely with Steane's 7-Qbit code. Recall that

$$\mathsf{Z}^p = \begin{pmatrix} 1 & 0 \\ 0 & (-1)^p \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & e^{ip\pi} \end{pmatrix}. \tag{335}$$

Hence, we must have

$$\bar{\mathsf{Z}}^p|\bar{0}\rangle = |\bar{0}\rangle, \qquad \text{and} \qquad \bar{\mathsf{Z}}^p|\bar{1}\rangle = e^{ip\pi}|\bar{1}\rangle. \tag{336}$$

It seems natural to seek a transverse construction of the form

$$\bar{\mathsf{Z}}^p_{\rm Steane} = \mathsf{Z}_1^q\mathsf{Z}_2^q\mathsf{Z}_3^q\mathsf{Z}_4^q\mathsf{Z}_5^q\mathsf{Z}_6^q\mathsf{Z}_7^q, \tag{337}$$

for some power q. Since $|\bar{0}\rangle_{\rm Steane}$ is a superposition of states with either 0 or 4 physical $|1\rangle$'s, while $|\bar{1}\rangle_{\rm Steane}$ is a superposition of states with either 3 or 7 physical $|1\rangle$'s, we must have

$$4q = 2m, \qquad \text{so that} \qquad \bar{\mathsf{Z}}^p|\bar{0}\rangle = |\bar{0}\rangle, \tag{338}$$

$$3q = p + 2n, \qquad \text{so that} \qquad \bar{\mathsf{Z}}^p|\bar{1}\rangle = e^{ip\pi}|\bar{1}\rangle, \tag{339}$$

for some integers m and n. Subtracting these, we find that $q = 2(m-n) - p$, and inserting this in either of eqs. (338) or (339) to eliminate q, we obtain

$$p = \frac{3m - 4n}{2}\,. \tag{340}$$

Then, for $(m, n) = (3, 2)$ we have $p = 1/2$, which is the only possible positive value of p less than one. In this case, $q = 3/2$, so we can make a transverse construction of the gate

$$\bar{\mathsf{Z}}^{1/2}_{\rm Steane} = \mathsf{Z}_1^{3/2}\mathsf{Z}_2^{3/2}\mathsf{Z}_3^{3/2}\mathsf{Z}_4^{3/2}\mathsf{Z}_5^{3/2}\mathsf{Z}_6^{3/2}\mathsf{Z}_7^{3/2}. \tag{341}$$

However, we cannot construct the desired gate $\bar{\mathsf{Z}}^{1/4}_{\rm Steane}$ in a like manner.[116] Instead, we invoke a somewhat indirect approach, due to Boykin *et al.*[117] They noted that IF we can construct the state

$$|\phi_0\rangle = \mathsf{Z}^{1/4}\mathsf{H}|0\rangle = \frac{|0\rangle + e^{i\pi/4}|1\rangle}{\sqrt{2}}\,, \tag{342}$$

that involves the operator $\mathsf{Z}^{1/4}$, and we combine it with a general Qbit $|\psi\rangle = a|0\rangle + b|1\rangle$ via a Controlled-NOT operation with $|\psi\rangle$ as the control Qbit and $|\phi_0\rangle$ as the target Qbit,[118] then we have

$$\begin{aligned}
\mathsf{C}_{\psi\phi_0}|\psi\rangle|\phi_0\rangle &= \mathsf{C}_{\psi\phi_0}(a|0\rangle + b|1\rangle)\frac{|0\rangle + e^{i\pi/4}|1\rangle}{\sqrt{2}} \\
&= \frac{a|0\rangle + e^{i\pi/4}b|1\rangle}{\sqrt{2}}|0\rangle + \frac{e^{i\pi/4}a|0\rangle + b|1\rangle}{\sqrt{2}}|1\rangle \\
&= (\mathsf{Z}^{1/4}|\psi\rangle)\frac{|0\rangle}{\sqrt{2}} + (\mathsf{Z}^{-1/4}|\psi\rangle)\frac{e^{i\pi/4}|1\rangle}{\sqrt{2}}\,.
\end{aligned} \tag{343}$$

If we now measure the second Qbit (which originally was $|\phi_0\rangle$), we will either find it to be $|0\rangle$ (which forces the first Qbit into the state $\mathsf{Z}^{1/4}|\psi\rangle$), or $|1\rangle$ (which forces the first Qbit into the state $\mathsf{Z}^{-1/4}|\psi\rangle$ to within a phase).

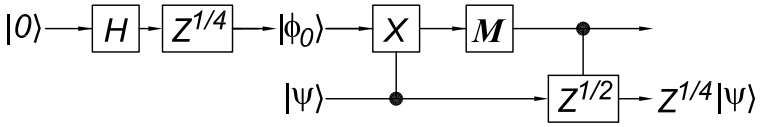
[116] In principle, whenever we need to execute the logical operation $\bar{\mathsf{Z}}^{1/4}$ on a logical Qbit $|\bar{\psi}\rangle$ we could convert the coding from Steane's version to another in which the $\bar{\mathsf{Z}}^{1/4}$ can be constructed transversely, and then convert back to Steane's coding afterwards (A. McDonald, 4/28/05). However, I am not aware of any error coding in which a transverse construction of $\bar{\mathsf{Z}}^{1/4}$ is possible. Since $(e^{i\pi/4})^8 = 1$, the coded state $|\bar{0}\rangle$ would have to involve 0 or 8 $|1\rangle$'s and $|\bar{1}\rangle$ would have to involve 1 or 9 $|1\rangle$'s. So it seems that the code would need to use at least 9 physical Qbits per logical Qbit.

[117] http://kirkmcd.princeton.edu/examples/QM/boykin_quant-ph-9906054.pdf

[118] Another version of this remarkable result is obtained if we use $|\phi_0\rangle$ as the control Qbit and $|\psi\rangle$ as the target Qbit.

We can bring the first Qbit to the state $\mathsf{Z}^{1/4}|\psi\rangle$ in both cases if we apply the operator $\mathsf{Z}^{1/2}$ to the first Qbit conditional on the result of the measurement of the second Qbit being $|1\rangle$.

This is a kind of quantum teleportation of the gate $\mathsf{Z}^{1/4}$ from the second Qbit to the first Qbit!



In the above figure, the gate M represents a measurement of the second Qbit.

To successfully use this trick, we must augment it in two ways: we must create a logical state $|\bar{\phi}_0\rangle_{\text{Steane}}$ that implements eq. (342) in the Steane code, and we must find a fault-tolerant version of the algorithm.

As the gate $\bar{\mathsf{Z}}^{1/4}_{\text{Steane}}$ cannot be constructed transversely, we cannot create the logical state $|\bar{\phi}_0\rangle_{\text{Steane}}$ in an obvious manner. Boykin *et al.* noted that the desired state $|\phi_0\rangle$ is an eigenstate of the operator U_ϕ that can be represented two ways,

$$\mathsf{U}_\phi = \mathsf{Z}^{1/4}\mathsf{X}\mathsf{Z}^{-1/4} = e^{-i\pi/4}\mathsf{Z}^{1/2}\mathsf{X} = \begin{pmatrix} 0 & e^{-i\pi/4} \\ e^{i\pi/4} & 0 \end{pmatrix}, \tag{344}$$

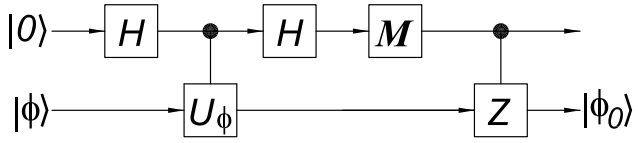
the second form of which can be implemented in the Steane code with transverse operators. The orthogonal eigenstate of U_ϕ is

$$|\phi_1\rangle = \mathsf{Z}^{1/4}\mathsf{H}|1\rangle = \frac{|0\rangle - e^{i\pi/4}|1\rangle}{\sqrt{2}} = \mathsf{Z}|\phi_0\rangle, \tag{345}$$

and we have $\mathsf{U}_\phi|\phi_j\rangle = (-1)^j|\phi_j\rangle$ for $j = 0, 1$.

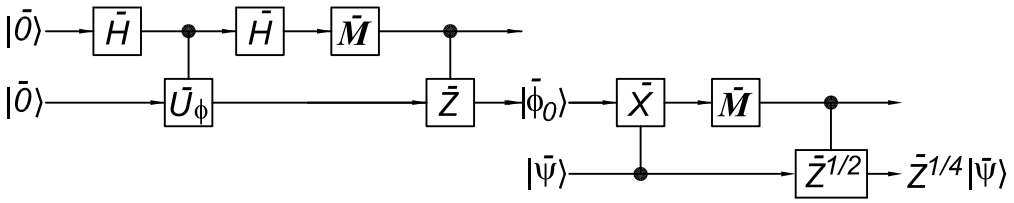
Now that we have an operator U_ϕ whose eigenstates include the desired state $|\phi_0\rangle$, we can make a measurement of that operator, following a procedure due to Shor,[119] which will force the result into one of the eigenstates.

Show that the circuit show below (where $|0\rangle$ and $|\phi\rangle$ are single physical Qbits) will take a general Qbit $|\phi\rangle$ into the eigenstate $|\phi_0\rangle$ of operator U_ϕ, provided its two eigenvalues are ± 1.



If, as holds in present case, the operator U_ϕ can be implemented transversely in Steane's code, the state $|\phi\rangle$ can be replaced by the 7-Qbit logical state $|\bar{\phi}\rangle$, and the output of the measurement circuit will be the logical state $|\bar{\phi}_0\rangle$. We can then use this state as the input to the Steane-code version of the previous circuit to obtain the gate $\bar{\mathsf{Z}}^{1/4}_{\text{Steane}}$:

[119] http://kirkmcd.princeton.edu/examples/QM/shor_quant-ph-9605011.pdf
Shor's measurement circuit was the prototype of the circuit of Boykin *et al.* shown at the top of the page.



Finally, to make this procedure fault-tolerant, we apply Steane's error correction procedure to the logical Qbits between every pair of logical gates.

Actually, it is more complicated than this to render the circuit fault tolerant. Extra effort must be made to protect the 7-fold measurement operator $\bar{M}$ from errors. In addition, we must also prepare the logical $|\bar{0}\rangle_{\text{Steane}}$ states in a fault-tolerant manner. However, we leave to you pursue these details at a later time.

(d) **Grover's Fault-Tolerant Procedure**[120]

Grover has recently proposed a procedure that addresses the possibility of a systematic error in an operator U. Here, the purpose of operator U is to transform a source state $|s\rangle$ into a target state $|t\rangle$. That is, the desired operation is

$$\mathsf{U}|s\rangle = |t\rangle. \tag{346}$$

However, due to defects in the construction of U, what actually happens is

$$\mathsf{U}|s\rangle = |t'\rangle \approx |t\rangle. \tag{347}$$

If the operator U is reliable in the sense that it always produces the same output for the same input, and we can also reliably construct the inverse operator $\mathsf{U}^{-1} = U^\dagger$, then we can make an iterative use of these operators such that we come ever closer to the desired final state $|t\rangle$.

To achieve this goal, Grover uses two additional operators that he calls R_s and R_t whose effect on a general state

$$|\psi\rangle = \alpha|s\rangle + \beta|u\rangle = \gamma|t\rangle + \delta|v\rangle, \tag{348}$$

where $|u\rangle$ is orthogonal to $|s\rangle$ and $|v\rangle$ is orthogonal to $|t\rangle$, is given by

$$\mathsf{R}_s|\psi\rangle = e^{i\pi/3}\alpha|s\rangle + \beta|u\rangle, \qquad \text{and} \qquad \mathsf{R}_t|\psi\rangle = e^{i\pi/3}\gamma|t\rangle + \delta|v\rangle. \tag{349}$$

That is, operators R_s and R_t make selective phase shifts on the components $|s\rangle$ and $|t\rangle$, respectively, of a general state.

Grover's iteration operator is

$$\mathsf{V} = \mathsf{U}\mathsf{R}_s\mathsf{U}^\dagger\mathsf{R}_t. \tag{350}$$

Then, we claim, $\mathsf{V}\mathsf{U}|s\rangle$ is closer to the desired target state $|t\rangle$ than is $\mathsf{U}|s\rangle$, and $\mathsf{V}^2\mathsf{U}|s\rangle$ is closer still. In particular, if $|\langle t|\mathsf{U}|s\rangle|^2 = 1-\epsilon$, then

$$\left|\langle t|\mathsf{U}\mathsf{R}_s\mathsf{U}^\dagger\mathsf{R}_t\mathsf{U}|s\rangle\right|^2 = 1-\epsilon^3, \tag{351}$$

[120] http://kirkmcd.princeton.edu/examples/QM/grover_quant-ph-0503205.pdf

and even one iteration greatly reduces the error in the calculation of the target state.

Show that

$$\mathsf{U}\mathsf{R}_s\mathsf{U}^\dagger\mathsf{R}_t\mathsf{U}|s\rangle = [e^{i\pi/3} + |\langle t|\mathsf{U}|\mathsf{s}\rangle|^2 (e^{i\pi/3}-1)^2]\mathsf{U}|s\rangle + \langle t|\mathsf{U}|\mathsf{s}\rangle(e^{i\pi/3}-1)|t\rangle, \quad (352)$$

As a measure of the deviation of this state from the target state $|t\rangle$, calculate the square of the amplitude of the part of eq. (352) that is orthogonal to $|t\rangle$, *i.e.*, calculate that probability that the result of the first iteration is not state $|t\rangle$. Show that this probability is

$$(1-|\langle t|\mathsf{U}|\mathsf{s}\rangle|^2)\left|e^{i\pi/3} + |\langle t|\mathsf{U}|\mathsf{s}\rangle|^2 (e^{i\pi/3}-1)^2\right|^2 = \epsilon^3, \quad (353)$$

if $|\langle t|\mathsf{U}|\mathsf{s}\rangle|^2 = 1-\epsilon$.

Grover notes that this procedure is a kind of quantum search algorithm during which one monotonically converges on the goal, in contrast to the search algorithm of prob. 15, in which one is close to the goal only at various stages of a quasi-cyclic process. This may help dispel the perception that quantum computation is more of an art than a science, as has been quipped: "*The quantum search algorithm is like baking a souffle....you have to stop at just the right time or it gets burnt.*"[121]

[121] http://kirkmcd.princeton.edu/examples/QM/brassard_science_275_627_97.pdf

23. **Quantum Cryptography**[122]

We have already discussed one aspect of quantum cryptography in problem 17, namely, how quantum computation might be used to decode so-called public keys based on large prime numbers. Here, we explore how quantum techniques might improve the security of private keys based on one-time pads.

The goal is the transmission of "secret" messages, which for computational purposes are taken to be strings of bits. A message is encoded (and subsequently decoded) by adding (modulo 2) each bit of the message to a corresponding bit from a private key. For maximal security, the key should consist of a random bit string of same length as the message, and the key should be used for one time only (Vernam, 1917). The difficulty with this scheme is that the key must be shared between the sender and receiver via a communication channel that is subject to "eavesdropping", *i.e.*, a public channel.

The challenge is to devise a procedure to generate a (random) private key that can be sent over a public channel.

As classical bit strings can be copied exactly (without altering the originals), the transmission of classical private keys over public channels is unsatisfactory. However, the no-cloning theorem (prob. 6) suggests that quantum cryptography may offer advantages, since an unknown quantum state cannot in general be copied exactly (and measurements of a large number of copies of an unknown quantum state are required to determine its character to good accuracy).

(a) Demonstrate the following variant on the no-cloning theorem. Suppose that Eve intercepts a quantum state $|\psi\rangle$ from, say, a public communication channel, and she wishes to determine if this state is distinct from another state $|\phi\rangle$. Show that this cannot be done, in general, without modification to state $|\psi\rangle$.

Hint: You may suppose that Eve tries to distinguish $|\psi\rangle$ from $|\phi\rangle$ with the aid of an ancillary state $|a\rangle$ and a unitary transformation U *such that*

$$\mathsf{U}|\psi\rangle|a\rangle = |\psi\rangle|b\rangle, \qquad \textit{and} \qquad \mathsf{U}|\phi\rangle|a\rangle = |\phi\rangle|c\rangle, \tag{354}$$

which leaves states $|\psi\rangle$ and $|\phi\rangle$ unaltered. Show, however, that if $\langle\psi|\phi\rangle \neq 0$, then $\langle b|c\rangle = 1$ and thus $|\psi\rangle$ and $|\phi\rangle$ cannot be distinguished by this procedure (unless $|\psi\rangle$ and $|\phi\rangle$ are orthogonal).

Problem (a) alerts us to another difficulty in cryptography that uses a public channel, namely that messages and keys may be altered by an eavesdropper, or by Nature via various forms of bit errors. To guard against the latter issue, we anticipate that practical quantum cryptography would be performed with logical Qbits based on error-correction codes (prob. 21). If the tampering of an n-bit key by an eavesdropper is known to be limited to at most $m < n$ bits, then a largely classical procedure called privacy amplification can be used to build a relatively secure key of length $< n - m$ bits.[123] Here, we turn directly to schemes for secure transmission of a private key over a public channel.

[122] See sec. 12.6 of Nielsen and Chuang.

[123] http://kirkmcd.princeton.edu/examples/QM/bennett_siamjc_17_210_88.pdf

The BB84 Scheme for Generation of a Private Key Using a Public Channel.[124]

i. Alice encodes an n-bit string in n Qbits where the basis of each Qbit is chosen at random among the [0,1] and $[+,-]$ bases. *In practice, these Qbits could be photons in an optical fiber, where the [0,1] basis corresponds to horizontal and vertical linear polarization, and the* $[+,-]$ *basis corresponds to* 45° *and* 135° *linear polarization; alternatively, the second basis could be left- and right-handed circular polarization related by* $|L,R\rangle = (|x\rangle \pm i|y\rangle)/\sqrt{2}$.

ii. Alice sends the n Qbits to Bob over a public communication channel. *For now, we ignore the possibility of bit errors and/or tampering by an eavesdropper.*

iii. Bob measures each of the Qbits in either the [0,1] or the $[+,-]$ basis, choosing the basis at random for each bit.

iv. Alice announces over the public channel in which basis she created each of the bits (but she does not announce the values of the bits).

v. Likewise, Bob announces over the public channel which basis he used to measure each of the bits (but he does not announce the results of the measurements)

vi. Alice and Bob now share a private key, of length roughly $n/2$ bits, consisting of those bits for which they used the same basis for creation/measurement.

(b) An eavesdropper Eve wishes to gain as much information as possible about the private key of Alice and Bob, so she performs a "nondemolition" measurement on each of the transmitted Qbits, as discussed in prob. 5(d). How much of the key can Eve learn, and what effect do her actions have on the key? *For your own satisfaction, you may want to verify that the results would be the same if Eve destructively measures each Qbit, and then sends on to Bob a new Qbit of whatever value she measured.*

The no-cloning theorem advises us that Eve cannot learn all details as to the private key of Alice and Bob. From prob. (a) we anticipate that Eve's gain in knowledge comes at the expense of alterations to the bits received by Bob. Thus, while Alice and Bob can be confident that their private key cannot be completely broken by Eve, they cannot be sure, without further checking, that their own versions of the private key are actually the same. However, as you have deduced in prob. (b), there is a bound on how much damage Eve does to the private key if she only tries to learn its details. Alice and Bob could verify the extent of the damage by publicly comparing/sacrificing some fraction of the bits of their not-quite-private key, and then use the technique of privacy amplification to generate a shorter, but more truly private key.

The E91 Variant[125]

Among the many variants proposed for quantum key distribution, one of the more interesting concepts is that due to Ekert, in which the key is derived from pairs of entangled bits generated by Alice (or by Bob or even by a third party) such that

[124] http://kirkmcd.princeton.edu/examples/QM/bennett_ieeecssp_175_84.pdf
[125] http://kirkmcd.princeton.edu/examples/QM/ekert_prl_67_661_91.pdf

Alice receives one bit of the pair and Bob the other. The entanglement could be of the type $(|0\rangle|0\rangle + |1\rangle|1\rangle)/\sqrt{2} = (|+\rangle|+\rangle + |-\rangle|-\rangle)/\sqrt{2}$ in which case Alice and Bob always find the same bit value provided they measure in the same basis, or of the type $(|0\rangle|1\rangle - |1\rangle|0\rangle)/\sqrt{2} = (|-\rangle|+\rangle - |+\rangle|-\rangle)/\sqrt{2}$ in which case Alice and Bob always find the opposite bit value provided they measure in the same basis.

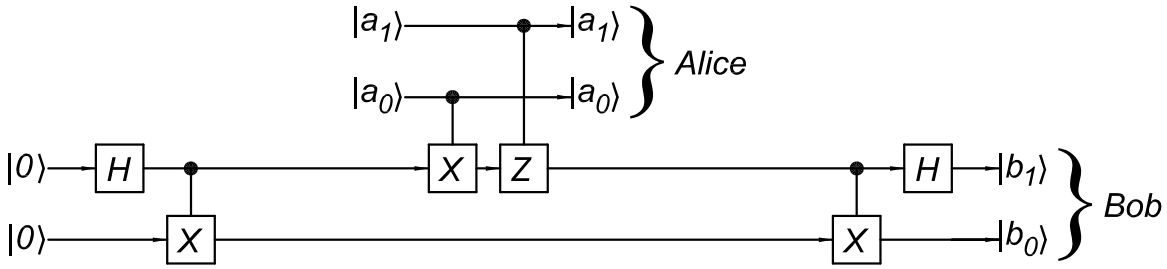
As in the BB84 scheme, Alice and Bob announce publicly the bases in which they have measured their various bits, and they use only those bits that were measured in the same basis to form their private key.

If Eve measures Alice's and/or Bob's bit before the latter do so, but in a different basis than that eventually chosen by Alice and Bob, there is a 50% probability that Alice and Bob's bits do not exhibit the expected correlation. So, if Alice and Bob compare/sacrifice a subset of their measured bits, they can detect Eve's tampering.

The particular interest in the E91 scheme is the possibility (not yet technically feasible, however) that Alice and Bob store their entangled bits without measuring them until just before they need to apply a private key.[126] In this way, the key does not come into existence until the last possible moment, which protects it against classical copying/theft before its use.

Quantum Dense Coding[127]

Bennett and Wiesner have proposed a variant of Ekert's variant, not for cryptography, but simply for information transfer. Their scheme, shown below, is called quantum dense coding in that Alice can transmit two bits of information to Bob while ostensibly interacting with only one of the two bits that Bob receives. The trick is that the bit of Bob with which Alice interacts is entangled with Bob's other bit, so that in effect Alice interacts with both of Bob's bits. Note that the initial preparation of Bob's entangled bits could be done by a third party at any desired distance from Alice and Bob.

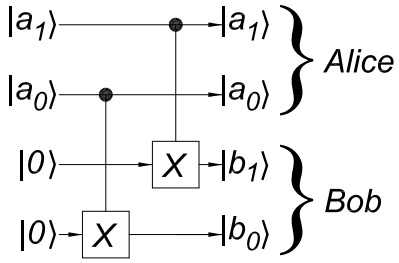


(c) Verify algebraicly that the effect of the above circuit permits Bob to receive exact copies of all four of Alice's two-bit basis states $|0\rangle|0\rangle$, $|0\rangle|1\rangle$, $|1\rangle|0\rangle$ and $|1\rangle|1\rangle$. However, the no-cloning theorem tells us that Bob cannot receive an exact copy of a general two-bit state of Alice. For example, what is the final state of the system in case Alice's initial state is $(|0\rangle|0\rangle + |1\rangle|1\rangle)/\sqrt{2}$?

[126] Some of the most dramatic examples of "quantum weirdness" cited in the popular literature are based on long-term storage of entangled bits. While this is possible in principle, present quantum reality is that entangled bits tend to decohere in tiny fractions of a second.

[127] http://kirkmcd.princeton.edu/examples/QM/bennett_prl_69_2881_92.pdf

The dense coding circuit is equivalent to the simpler circuit shown below,[128] and could be regarded as a "teleported" version of it.[129]



Impossibility of Secure Quantum Bit Commitment

This final topic is about the interplay between a negative result for quantum cryptography and the basic character of the quantum world.

The issue is whether is it possible to realize a secure quantum bit commitment scheme.

In a such a scheme Alice supplies an encoded bit to Bob, but in a manner that Bob cannot determine its value without additional information from Alice. Later, Alice announces the value of the bit, and gives Bob the information such that he can determine its value to be in agreement with Alice's claim.

The bit commitment scheme is secure if Bob cannot determine the value of the bit prior to receiving the additional information from Alice, and if the value that Bob finds after Alice's announcements is always in agreement with Alice's claim.

A secure bit commitment scheme is desirable in that it would permit reliable communication between two parties that do not entirely trust one another. An early discussion of quantum bit commitment is given in the BB84 paper cited above. A review has been given by Bub.[130]

Secure bit commitment is not possible classically. Once Alice commits to a classical value of her bit, she cannot completely isolate/hide it from Bob, who can always determine its value with enough effort.

It was conjectured that secure quantum bit commitment might be possible if Alice's bit is in the form of some kind of quantum superposition, the details of which are "protected" by the no-cloning theorem. However, if the bit is to have a well-defined value upon measurement according to Alice's revealed procedure, although not before, it must be entangled with some other bit. But, in this case Alice could always take a last-minute action to "force" the value of the bit to be the complement of her original commitment.

Thus, neither classical nor quantum secure bit commitment is possible.

This result closes out an interesting class of quantum cryptographic procedures, but also opens up a general vista on the quantum world. It has recently been argued by Clifton, Bub and Halvorson (the latter of Princeton) that the entire structure of

[128] For a graphical proof of this claim, see Fig. 1 of Chap. 6 of Mermin's course on quantum computation.
[129] The version of "teleportation" discussed in prob. 6(d) is somewhat more sophisticated than dense coding, and was invented by Bennett *et al.* a few months after they invented dense coding.
[130] http://kirkmcd.princeton.edu/examples/QM/bub_fp_31_735_01.pdf

quantum theory can be deduced from three negative-sounding information-theoretic constraints:[131]

- No superluminal information transfer.
- No broadcasting/cloning of an unknown (quantum) state.
- No secure bit commitment.

Of these three constraints, the first and the third hold in the classical as well as the quantum world. Hence, it appears that the no-cloning theorem represents the essential distinction between the classical and the quantum worlds (beyond such details as classical bits having values of only 0 or 1, while quantum bits can be a superposition of $|0\rangle$ and $|1\rangle$ with imaginary coefficients).

[131] http://kirkmcd.princeton.edu/examples/QM/clifton_quant-ph-0211089.pdf

24. **The End of Quantum Information?**

In prob. 1 we referred briefly to the possibility of very compact quantum computers – of the size of a 1-kg black hole. Here we return to the theme of black holes and quantum information, and comment on unresolved issues of potentially deep significance.

In 1974 Stephen Hawking had the important insight that the interaction of (classical) black holes with quantum fluctuations of the "vacuum" leads to "evaporation" of the black holes.[132] Hawking argued that the radiation from the evaporating black hole has a purely thermal spectrum, with temperature T related to the acceleration g of gravity at the Schwarzschild radius of the black hole by $kT = \hbar g/2\pi$, where k is Boltzmann's constant.

The existence of Hawking radiation makes a dramatic change in the character of an older paradox concerning black holes. If a black hole "swallows" a collection of particles with entropy S, that entropy is no longer accessible to the Universe exterior to the black hole.[133] Hence, it might seem that black holes violate the second law of thermodynamics. However, if the black hole is a stable object, one can say that it stores the entropy that it "swallowed", so that the total entropy of the Universe has not really decreased (although some of the entropy is not very accessible), and there is not actually a paradox.

But if/when the black hole evaporates, what has become of the entropy that was stored inside it? If the radiation that leads to the disappearance of the black hole is indeed thermal as Hawking claimed, this radiation cannot carry the information corresponding to the stored entropy. When the black hole has completely evaporated, the entropy of the Universe has decreased, and the paradox is severe.

The resolution of this paradox has been a major issue for theoretical physics over the past 30 years. Various solutions have been suggested, including[134]

- The second law of thermodynamics is indeed violated by black holes, so that new laws of physics are required to resolve the paradox.
- Black holes don't evaporate completely, but "remnants" continue to store the "swallowed" entropy (perhaps in a "baby" universe).
- Hawking radiation is not thermal, such that entropy "leaks" back out of the black hole all the time.

[132] http://kirkmcd.princeton.edu/examples/QED/hawking_nature_248_30_74.pdf
http://kirkmcd.princeton.edu/examples/QED/hawking_cmp_43_199_75.pdf
A rough model is that the gravitational potential energy of a black hole is converts electron-positron pairs in the QED vacuum (just outside the Schwarzschild radius of the black hole) from "virtual" to real particles, which then annihilate in to pair of real photons one of which is captured by the black hole and the other of which is radiated to "infinity".

[133] The entropy kA/L_P^2 (mentioned in prob. 1) that is associated with the surface area A of a black hole is much less, in general, than the amount of entropy that the black hole can have "swallowed".

[134] For a review of the situation in 1992 from an information theoretic perspective, see
http://kirkmcd.princeton.edu/examples/QM/preskill_hep-th-9209058.pdf
Hawking's recent views appear in
http://kirkmcd.princeton.edu/examples/QM/hawking_prd_72_084013_05.pdf

- Black holes do evaporate, but somehow the entropy comes out in a burst at the end.[135]
- Black holes do evaporate, and the entropy "leaks" out continuously into a quantum state surrounding the black hole, such that this entropy/information is inaccessible until a quantum key is released by the black hole at the very end of its life.[136]

The last two conjectures are in the spirit of John Wheeler's notion of "It from Bit." Your assignment (with no time limit) is to clarify and extend these ideas to help bring about a more unified understanding of our quantum Universe.[137]

[135] http://kirkmcd.princeton.edu/examples/QM/carlitz_prd_36_2336_87.pdf
http://kirkmcd.princeton.edu/examples/QM/horowitz_jhep_022004008.pdf
http://kirkmcd.princeton.edu/examples/QM/gottesman_hep-ph-0311269.pdf

[136] http://kirkmcd.princeton.edu/examples/QM/smolin_prl_96_081302_06.pdf

[137] A sketch of a partial answer by Preskill is given at
http://www.theory.caltech.edu/people/preskill/talks/GR100-Caltech-preskill.pdf.

1. **The Ultimate Laptop**

Energy Limits Speed

The time interval Δt during which a quantum process takes place is bounded by the energy scale E associated with that process by the uncertainty principle,

$$\Delta t \geq \frac{\hbar}{E}\,. \tag{355}$$

So if the laptop has energy E available for computation, the number N of processing cycles per second is limited by

$$N \approx \frac{1}{\Delta t} \leq \frac{E}{\hbar}\,. \tag{356}$$

In our ultimate laptop, we suppose that all of the rest energy is available to drive the computation. Then, according to Einstein,

$$E = m_{\rm laptop}c^2 = 1 \cdot (3 \times 10^8)^2 \approx 10^{17}\ \text{J}, \tag{357}$$

and the speed of the laptop is

$$N \approx \frac{E}{\hbar} \approx \frac{10^{17}\ \text{J}}{10^{-34}\ \text{J-s}} \approx 10^{51}\ \text{bit operations/s}. \tag{358}$$

However, for the full rest energy to be available, the laptop must be hot enough that the nucleons are unbound, and can be transformed into other forms of mass/energy. That is, the temperature must be related to the mass of a proton m_p by

$$kT \approx m_p c^2, \qquad \text{or} \qquad T \approx \frac{m_p c^2}{k} \approx \frac{10^{-27}(3\times 10^8)^2}{10^{-23}} \approx 10^{13}\text{K}. \tag{359}$$

For additional discussion, see, for example
http://kirkmcd.princeton.edu/examples/QM/margolis_physica_d120_188_98.pdf

Transit Time

The above argument ignores possible restrictions due to signal propagation from one part of the laptop to the other.

In 1 kg there are about 10^{27} nucleons. If there are 10^{51} operations per second in total, then each nucleon must perform about 10^{24} operations per second. Once the laptop has been converted to a plasma with temperature $T \approx 10^{13}$K, the average distance between nucleons is about 1 Angstrom $= 10^{-10}$ m. It takes light about 3×10^{-19} s to travel this distance, so each nucleon could communicate with its nearest neighbor only about 3×10^{18} times per second. Such communication is necessary for computation, so we should derate the ultimate laptop from 10^{51} operations per second to about 10^{45} operations per second.

Entropy Limits Memory

A memory bit is a two-state system. A set of M bits has a total of $W = 2^M$ possible states. Thus, memory size in bits is related to the number of states of the memory according to

$$M = \log_2 W. \tag{360}$$

But we know from the statistical interpretation of thermodynamic entropy, S, that

$$S = k \ln W, \qquad i.e., \qquad W = e^{S/k}, \tag{361}$$

where k is Boltzmann's constant. If we suppose that all the degrees of freedom of our laptop are used for its memory, the memory size is limited by

$$M = \frac{S}{k \ln 2}\,. \tag{362}$$

Now, the entropy of our laptop is a function of its energy E, and to a first approximation, this function is simply

$$S \approx \frac{E}{T}\,, \tag{363}$$

where T is the operating temperature of the laptop. In this approximation, the memory capacity of the laptop is given by

$$M \approx \frac{E}{kT \ln 2}\,. \tag{364}$$

Memory Size and Degrees of Freedom

The equipartition theorem of Maxwell tells us that the energy of the laptop at temperature T is related to the number m of its degrees of freedom,

$$E = \frac{mkT}{2}\,. \tag{365}$$

Comparing eqs. (364) and (365) we arrive at the reasonable conclusion that the ultimate memory capacity of the laptop is equal to its number of degrees of freedom,

$$M \approx m. \tag{366}$$

We also note that the relation between energy and temperature is often expressed in terms of the (temperature dependent) heat capacity $C(T)$,

$$E = CT. \tag{367}$$

In general, the heat capacity is a monotonic function of temperature, being small when $T \approx 0$, and approaching $3nk$ at high temperature when the laptop has become a gas of n fermions.

Thus, the memory size of our ultimate laptop is roughly equal to the number of nucleons in a kg,

$$M \approx m \approx 1000 N_A \approx 10^{27} \text{ bits}, \tag{368}$$

where N_A is Avagadro's number. This result could have been anticipated without the digression about entropy.

Ultimate Speed of a Room-Temperature Laptop

Backing away from a laptop whose state of matter is something like that in the early universe, we return to room temperature. Then, the ultimate memory size is still given by eq. (368). To estimate the speed according to eq. (356), we note that the available energy is roughly

$$E_{\text{room temp}} \approx mkT \approx 10^{27} \cdot (1/40) \text{ eV} \approx 10^{27} \cdot \frac{1}{40} \cdot 10^{-19} \text{ J} \approx 10^{6} \text{ J}. \tag{369}$$

The ultimate bit-operation speed of a room-temperature laptop is therefore

$$N_{\text{room temp}} \approx \frac{E_{\text{room temp}}}{\hbar} \approx \frac{10^6}{10^{-34}} \approx 10^{40}/\text{s}, \tag{370}$$

which evades our most optimistic transit-time limit.

The Black-Hole Laptop

Bekenstein's result that inspired this problem is that the entropy of a black hole is

$$S \approx \frac{kA}{L_P^2} \approx \frac{kR_S^2}{L_P^2}, \tag{371}$$

where the Schwarzschild radius of an object of mass m is

$$R_S = \frac{2Gm}{c^2}, \tag{372}$$

and the Planck length is

$$L_P = \sqrt{\frac{G\hbar}{c^3}}. \tag{373}$$

Combining these with eq. (362), the memory size M of a black hole is

$$M \approx \frac{R_S^2}{L_P^2} = \frac{Gm^2}{\hbar c}. \tag{374}$$

The memory size of a 1-kg black-hole laptop is

$$M_{\text{black-hole laptop}} \approx \frac{10^{-10} \cdot 1^2}{10^{-34} \cdot 10^8} \approx 10^{16} \text{ bits}. \tag{375}$$

While still quite large, this is a substantial reduction compared to the ultimate memory capacity (368) of a laptop of ordinary density.

See also,
http://kirkmcd.princeton.edu/examples/QM/gambini_quant-ph-0507262.pdf

2. **Maxwell's Demon**

(a) **Classical Erasure**

It is assumed that the partition can be removed and inserted without expenditure of energy (without any flow of heat).

The entropy change of a gas during the irreversible free expansion is that same as that during a slow, reversible isothermal expansion between the same initial and final states.

During an isothermal expansion of volume V to $2V$, the work done by the molecule is

$$W_{\text{by molecule}} = \int_V^{2V} P\, dV = kT \int_V^{2V} \frac{dV}{V} = kT \ln 2. \tag{376}$$

To keep the temperature, and hence the internal energy, of the molecule constant, heat must flow into the box from the heat bath, and in amount $Q = W = kT \ln 2$. Hence, the thermodynamic entropy change of the box during the isothermal expansion is $\Delta S_{\text{box, iso exp}} = \Delta Q_{\text{into box}}/T = k \ln 2$. The entropy change of the bath is equal and opposite, $\Delta S_{\text{bath, iso exp}} = -k \ln 2$.

During the free expansion to the same final state as of the isothermal expansion, the entropy change of the box is the same as during the isothermal expansion, but the thermal bath experiences no entropy change as it transfer no heat.

$$\Delta S_{\text{box, free exp}} = k \ln 2, \qquad \Delta S_{\text{bath, free exp}} = 0. \tag{377}$$

Then, during the isothermal compression, the entropy changes of the box and bath are just the opposite of those during the isothermal expansion,

$$\Delta S_{\text{box, iso comp}} = -k \ln 2, \qquad \Delta S_{\text{bath, iso comp}} = k \ln 2. \tag{378}$$

Thus, the total entropy changes during the erasure of the bit, consisting of a free expansion followed by an isothermal compression, are

$$\Delta S_{\text{box, erasure}} = 0, \qquad \Delta S_{\text{bath, erasure}} = k \ln 2, \tag{379}$$

and the total entropy change of the universe is

$$\Delta S_{\text{universe, erasure}} = k \ln 2. \tag{380}$$

Note that there is no net energy change in either the free expansion or the isothermal compression, so the erasure of a bit is accomplished at zero energy cost in this model; this is, of course, a consequence of conservation of total energy in the Universe. However, the agent that performs the isothermal compression does work $kT \ln 2$, which is the energy cost to that agent in performing the erasure of the bit. This is the sense of Landauer's claim (footnote 6, p. 2) that there is an energy cost of at least $kT \ln 2$ in erasing a bit.

(b) **Classical Copying of a Known bit**

The original bit box has its molecule in either the left half (0) or the right half (1), and we know which is the case. The copy box is initially in a particular state that we might as well take to be 0, *i.e.*, its molecule is in the left half.

The copying can be accomplished as follows:[138]

i. If the original bit is 0, do nothing to the copy bit, which already was 0.

ii. If the original bit is 1, rotate the copy box by 180° about an axis in its left-right midplane. After this, the molecule appears to be in the right half of the copy box, and is therefore in a 1 state as desired.

No energy is expended in any of these steps. No heat flows. Hence, there is no (thermodynamic) entropy change in either the computer or in the environment.

The rotation of the copy box by 180° is equivalent to the logical NOT operation. Thus, the copying procedure suggested above could be called a controlled-NOT operation, in which the NOT operation is performed only if a relevant control bit is in the 1 state. Since all computation involves changing 0's into 1's and *vice versa*, we get a preview of the important role of controlled-NOT operations in classical and quantum computation.

The trick of rotating a box by 180° *if it is in a 1-state to bring it to the 0-state is a possible process for classical erasure of a bit. However, this process requires knowledge of the initial state of the box, whereas the method or part (a) does not require such knowledge. So, rotation of the box is not a solution to part (a) as posed.*

Since the rotation of the box is a reversible process, it doesn't change entropy. Could it then be that the trick of rotating the box provides a means of erasure with no entropy cost?

The issue now whether the task of acquiring the knowledge as to the state of the both implies an increase of entropy, of at least $k \ln 2$.

This is a famous question, associated with the concept of negentropy *– that information is associated with a kind of negative entropy, and that the creation of information implies a corresponding increase of entropy somewhere in the larger system.*

A sense of this was noted already in 1868 by P.T. Tait, Sketch of Thermodynamics, *p. 100.*[139] *A longer discussion was given by Brilloun.*[140]

If the box has moment of inertia I *and we wish to accomplish the erasure in time* t*, we give the box constant angular acceleration* $\ddot{\theta} = 4\pi/t^2$ *for time* $t/2$ *and then the negative of this for an additional time* $t/2$ *to leave the box at rest after rotation by* π *radians. The torque required during this process is* $\tau = I\ddot{\theta} = 4\pi I/t^2$*, so the work done in rotating the box is* $W = \tau\Delta\theta = \pi\tau = 4\pi^2 I/t^2$.

For a box that consists of a pair of C_{60} *"buckyballs," of mass* $m = 60m_C \approx 720m_p$ *and radius* $r \approx 0.5$ *nm each, the moment of inertia is* $I \approx 2mr^2(1 + 2/3) = 10mr^2/3 \approx 6 \times 10^{-16} m_p \approx 10^{-42}$ *J. If the erasure to be accomplished in time* $t = 0.1$ *ns (10 GHz), the work done in rotating the box is* $W = 4\pi^2 I/t^2 \approx 4 \times 10^{-21}$ *J* $\approx 1/40$ *eV* $= kT$ *for room temperature.*

That is, even the rotating pair of buckyballs as a memory element obeys Landauer's claim that the energy cost of erasure is at least $kT \ln 2$ *for "practical" parameters.*

[138] A scheme involving two pistons and the (re)movable partition, but no rotation, also works.

[139] http://kirkmcd.princeton.edu/examples/statmech/tait_thermo_68_p100.txt

[140] L. Brillouin, *The Negentropy Principle of Information*, J. Appl. Phys. **24**, 1152 (1953), http://kirkmcd.princeton.edu/examples/statmech/brillouin_jap_24_1152_53

3. (a) 2×2 **Classical Unitary Matrices**

A classical 2×2 unitary matrix must transform the Cbits (3) into themselves. Hence, the matrix elements must be either 0 or 1. Each 2×2 matrix has 4 elements, so there are $2^4 = 16$ matrices whose elements are only 0's or 1's.

A further restriction is that each column and row must have exactly one 1. If a column or row were all 0's, then the determinant of the matrix would be zero, so the matrix would have no inverse and could not be unitary. If a column has more than one 1, then one of the Cbits (3) will be transformed into

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \tag{381}$$

which is not a valid Cbit. If one row has more than one 1, and the other has only one 1, then one of the columns will have more than one 1, and the preceding argument indicates that the matrix transformation will not produce valid Cbits.

If we consider a system of n Cbits, we could represent a state by a 2^n-dimensional vector, and operations by $2^n \times 2^n$ matrices. As noted just before eq. (14), the elements of a Cbit vector are all zero except of a single element with value one. This implies that a real matrix that transforms Cbits into Cbits can have only one 1 in any column or row, and all other matrix elements must be 0.

The lone 1 in the first column could be in any of 2^n positions, leaving $2^n - 1$ choices for placement of the lone 1 in the second column, *etc.* Thus, there are $(2^n)!$ possible matrices M for a system of n bits.

These matrices are unitary. To see this, note that if a matrix M has only one 1 in each row and column, then the columns form a set of orthogonal vectors, $\mathbf{c}_i$, $i = 1, ..., n$, each normalized to one. These vectors are the 2^n permutations of the $|0\rangle$ vector, whose lone 1 is the uppermost element. The transpose conjugate matrix $\mathsf{M}^\dagger$ then has rows equal to the columns of the original matrix, $\mathbf{r}_i^\dagger = \mathbf{c}_i$. Then the ij element of $\mathsf{M}\mathsf{M}^\dagger$ is

$$(\mathsf{M}\mathsf{M}^\dagger)_{ij} = \mathbf{c}_i \cdot \mathbf{r}_j^\dagger = \mathbf{c}_i \cdot \mathbf{c}_j = \delta_{ij}. \tag{382}$$

Thus, there are only $(2^1)! = 2$ unitary matrices that act on a single classical bit,

$$\mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \text{and} \qquad \mathsf{NOT} = \mathsf{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{383}$$

(b) **Square Root of** NOT

The $\sqrt{\mathsf{NOT}}$ operator is a 2×2 unitary matrix,

$$\sqrt{\mathsf{NOT}} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{384}$$

such that

$$\sqrt{\mathsf{NOT}}\sqrt{\mathsf{NOT}} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 + bc & ab + bd \\ ac + cd & bc + d^2 \end{pmatrix} = \mathsf{NOT} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{385}$$

Thus, we obtain

$$\begin{aligned} a^2 + bc &= 0, && (386)\\ ab + bd &= 1, && (387)\\ ac + cd &= 1, && (388)\\ bc + d^2 &= 0. && (389) \end{aligned}$$

Comparing eqs. (386) and (389) we see that $d = \pm a$. We cannot have $d = -a$, as then eq. (387) we read $1 = ab - ba = 0$. Hence,

$$d = a. \tag{390}$$

Equations (387) and (388) now read

$$2ab = 1 = 2ac, \tag{391}$$

and hence,

$$c = b. \tag{392}$$

Equation (386) then reads

$$a^2 + b^2 = 0. \tag{393}$$

The conditions that $\sqrt{\mathsf{NOT}}$ be unitary are now

$$\sqrt{\mathsf{NOT}}^\dagger \sqrt{\mathsf{NOT}} = \begin{pmatrix} a^\star & b^\star \\ b^\star & a^\star \end{pmatrix} \begin{pmatrix} a & b \\ b & a \end{pmatrix} = \begin{pmatrix} |a|^2 + |b|^2 & a^\star b + a b^\star \\ a b^\star + a^\star b & |a|^2 + |b|^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{394}$$

$$\begin{aligned} |a|^2 + |b|^2 &= 1, && (395)\\ a b^\star + a^\star b = 2Re(a b^\star) &= 0. && (396) \end{aligned}$$

We write the complex numbers a and b as

$$a = a_x + i a_y, \qquad \text{and} \qquad b = b_x + i b_y, \tag{397}$$

where a_x, a_y, b_x and b_y are real. Then, taking the real and imaginary parts of eqs. (391), (393), (395) and (396), we obtain

$$\begin{aligned} a_x b_x - a_y b_y &= 1/2, && (398)\\ a_x b_y &= -a_y b_x, && (399)\\ a_x^2 - a_y^2 + b_x^2 - b_y^2 &= 0, && (400)\\ a_x a_y &= -b_x b_y, && (401)\\ a_x^2 + a_y^2 + b_x^2 + b_y^2 &= 1, && (402)\\ a_x b_x &= -a_y b_y. && (403) \end{aligned}$$

Equations (399) and (403) imply that

$$b_y = -\frac{a_y}{a_x} b_x = -\frac{a_x}{a_y} b_x, \tag{404}$$

whence,

$$a_y = a_x, \quad b_y = -b_x, \qquad \text{or} \qquad a_y = -a_x, \quad b_y = b_x. \tag{405}$$

Equation (401) now tells us that $b_x = \pm a_x$. However, eqs. (398) and (405) permit only

$$a_x = b_x \tag{406}$$

The two cases are now

$$a_x = a_y \;=\; b_x = -b_y, \tag{407}$$
$$a_x = -a_y \;=\; b_x = b_y, \tag{408}$$

In both cases we find

$$|a_x| = |a_y| = |b_x| = |b_y| = \frac{1}{2}\,, \tag{409}$$

using eq. (402). The remaining equation (400) is now also satisfied. Thus, we find four representations of $\sqrt{\mathsf{NOT}}$,

$$\sqrt{\mathsf{NOT}} = \pm\frac{1}{2}\begin{pmatrix} 1+i & 1-i \\ 1-i & 1+i \end{pmatrix}, \qquad \pm\frac{1}{2}\begin{pmatrix} 1-i & 1+i \\ 1+i & 1-i \end{pmatrix}. \tag{410}$$

Are there a well-defined number of square roots of a 2×2 unitary matrix, such as 4? Apparently not. If we write a general 2×2 unitary matrix as

$$\mathsf{U} = a\mathbf{I} + b\mathsf{X} + c\mathsf{Y} + d\mathsf{Z}, \tag{411}$$

for X, Y, Z *as defined in eq. (415), then*

$$\begin{aligned} \mathsf{U}^2 &= (a^2+b^2+c^2+d^2)\mathbf{I} + 2a(b\mathsf{X}+c\mathsf{Y}+d\mathsf{Z}) + bc(\mathsf{XY}+\mathsf{YX}) + \cdots \\ &= (a^2+b^2+c^2+d^2)\mathbf{I} + 2a(b\mathsf{X}+c\mathsf{Y}+d\mathsf{Z}), \end{aligned} \tag{412}$$

using eq. (418). So, if $a = 0$ and $b^2+c^2+d^2 = 1$, then $\mathsf{U}^2 = \mathbf{I}$. The condition that U be unitary when $a = 0$ is simply that $|b|^2+|c|^2+|d|^2 = 1$, so any triplet of real numbers (b,c,d) such that $b^2+c^2+d^2 = 1$ leads to a square root of $\mathbf{I}$; i.e., there are an infinite number of square roots of $\mathbf{I}$. (C. Mugnolo, 2/8/05)

(c) **Arbitrary 2×2 Unitary Matrix**

A straightforward alternative to expansion (32) of a general 2×2 unitary matrix U that involves the unit matrix $\mathbf{I}$ and the NOT matrix X is

$$\begin{aligned} \mathsf{U} &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \\ &= \frac{a+d}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{a-d}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \frac{b+c}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \frac{-b+c}{2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ &= \frac{a+d}{2}\,\mathbf{I} + \frac{a-d}{2}\,\mathsf{Z} + \frac{b+c}{2}\,\mathsf{X} + \frac{-b+c}{2}\,\tilde{\mathsf{Y}}. \end{aligned} \tag{413}$$

The unitary matrices $\tilde{\mathsf{Y}}$ and Z have real matrix elements, which seems desirable at first glance. However, when multiplying the unitary matrices based on expansion (413), we find the products

$$\mathsf{X}\tilde{\mathsf{Y}} = \mathsf{Z}, \qquad \tilde{\mathsf{Y}}\mathsf{Z} = \mathsf{X}, \qquad \mathsf{Z}\mathsf{X} = -\tilde{\mathsf{Y}}. \tag{414}$$

A symmetric pattern of products is obtained, following Pauli, if we use the unitary matrix $\mathsf{Y} = i\tilde{\mathsf{Y}}$. Then,

$$\boldsymbol{\sigma}_x = \mathsf{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_y = \mathsf{Y} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_z = \mathsf{Z} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{415}$$

and

$$\mathsf{X}\mathsf{Y} = i\mathsf{Z}, \qquad \mathsf{Y}\mathsf{Z} = i\mathsf{X}, \qquad \mathsf{Z}\mathsf{X} = i\mathsf{Y}. \tag{416}$$

We can now write our expansion of a general 2×2 unitary matrix as

$$\mathsf{U} = a\ \mathbf{I} + \mathbf{b} \cdot \boldsymbol{\sigma}, \tag{417}$$

where a is a complex number (in general different from the a of eq. (413)), $\mathbf{b}$ is a triplet of complex numbers, and $\boldsymbol{\sigma}$ is the triplet $(\boldsymbol{\sigma}_x, \boldsymbol{\sigma}_y, \boldsymbol{\sigma}_z)$ of Pauli matrices. The Pauli matrices $\boldsymbol{\sigma}_j$ obey

$$\boldsymbol{\sigma}_j^2 = \mathbf{I}, \qquad \text{and} \qquad \boldsymbol{\sigma}_j \boldsymbol{\sigma}_k = i\epsilon_{jkl}\ \boldsymbol{\sigma}_l \quad \text{when } j \neq k, \tag{418}$$

where $\epsilon_{jkl} = 1$ for an even permutation of xyz, -1 for an odd permutation, and 0 otherwise. Thus,

$$\begin{aligned} (\mathbf{a} \cdot \boldsymbol{\sigma})(\mathbf{b} \cdot \boldsymbol{\sigma}) &= \sum_j a_j\ \boldsymbol{\sigma}_j \sum_k b_k\ \boldsymbol{\sigma}_k = \sum_{j=k} a_j b_k\ \boldsymbol{\sigma}_j \boldsymbol{\sigma}_k + \sum_{j \neq k} a_j b_k\ \boldsymbol{\sigma}_j \boldsymbol{\sigma}_k \\ &= (\mathbf{a} \cdot \mathbf{b})\ \mathbf{I} + i \sum_{j \neq k} a_j b_k \epsilon_{jkl}\ \boldsymbol{\sigma}_l \\ &= (\mathbf{a} \cdot \mathbf{b})\ \mathbf{I} + i\ \boldsymbol{\sigma} \cdot \mathbf{a} \times \mathbf{b}. \end{aligned} \tag{419}$$

The condition that matrix (417) be unitary can now be written

$$\begin{aligned} \mathbf{I} &= \mathsf{U}\mathsf{U}^\dagger = (a\ \mathbf{I} + \mathbf{b} \cdot \boldsymbol{\sigma})(a^\star\ \mathbf{I} + \mathbf{b}^\star \cdot \boldsymbol{\sigma}) \\ &= (|a|^2 + |\mathbf{b}|^2)\ \mathbf{I} + \boldsymbol{\sigma} \cdot [2Re(a\mathbf{b}^\star) + i\ \mathbf{b} \times \mathbf{b}^\star]. \end{aligned} \tag{420}$$

Hence, we need

$$(|a|^2 + |\mathbf{b}|^2) = 1, \tag{421}$$

$$0 = 2Re(a\mathbf{b}^\star) + i\ \mathbf{b} \times \mathbf{b}^\star = 2Re(a\mathbf{b}^\star) + 2Re(\mathbf{b}) \times Im(\mathbf{b}). \tag{422}$$

If $a \neq 0$, we write it as $a = a_0 e^{i\delta}$ where a_0 and δ are real. We also write $\mathbf{b} = e^{i\delta}(\mathbf{c} + i\mathbf{d})$ where $\mathbf{c}$ and $\mathbf{d}$ are real vectors. Then, we eq. (422) becomes

$$0 = Re(a\mathbf{b}^\star) + Re(\mathbf{b}) \times Im(\mathbf{b}) = a_0\ \mathbf{c} + \mathbf{c} \times \mathbf{d}, \tag{423}$$

which implies that $\mathbf{c} = 0$. Thus,

$$\mathbf{b} = ib_0 e^{i\delta}\ \hat{\mathbf{u}}, \tag{424}$$

where $b_0 = |d|$ and $\hat{\mathbf{u}} = \mathbf{d}/\,|d|$ is a real unit vector.
On the other hand, if $a = 0$ then eq. (422) requires that vector $Re(\mathbf{b})$ must be parallel to vector $Im(\mathbf{b})$, so the vector $\mathbf{b}$ can be written as

$$\mathbf{b} = Re(b)\ \hat{\mathbf{u}} + iIm(b)\ \hat{\mathbf{u}} = ib_0 e^{i\delta}\ \hat{\mathbf{u}}, \tag{425}$$

where b_0 and δ are real, and $\hat{\mathbf{u}}$ is a real unit vector.
Hence, in any case the general 2×2 unitary matrix (417) can be written

$$\mathsf{U} = e^{i\delta}(a_0\ \mathbf{I} + ib_0\ \hat{\mathbf{u}} \cdot \boldsymbol{\sigma}), \tag{426}$$

where the real numbers a_0 and b_0 obey

$$a_0^2 + b_0^2 = 1, \tag{427}$$

so that condition (421) is satisfied. We can formally express a_0 and b_0 in terms of an angle θ such that

$$a_0 = \cos\frac{\theta}{2}, \qquad b_0 = \sin\frac{\theta}{2}. \tag{428}$$

Then,

$$\mathsf{U} = e^{i\delta}\left(\cos\frac{\theta}{2}\ \mathbf{I} + i\sin\frac{\theta}{2}\ \hat{\mathbf{u}} \cdot \boldsymbol{\sigma}\right) = e^{i\delta}e^{i\frac{\theta}{2}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}}. \tag{429}$$

By the exponential e^{A} of an operator A we, of course, mean the Taylor series

$$e^{\mathsf{A}} = \sum_{n=0}^{\infty} \frac{\mathsf{A}^n}{n!}\,. \tag{430}$$

For two noncommuting operators A and B, in general $e^{\mathsf{A}+\mathsf{B}} = e^{\mathsf{B}+\mathsf{A}} \neq e^{\mathsf{A}}e^{\mathsf{B}} \neq e^{\mathsf{B}}e^{\mathsf{A}}$.
The validity of the exponential form in eq. (398) is confirmed by noting that

$$\begin{aligned} e^{i\frac{\theta}{2}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}} &= \sum_j \frac{(i\frac{\theta}{2}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma})^j}{j!} = \left[\mathbf{I} - \frac{(\frac{\theta}{2})^2(\hat{\mathbf{u}}\cdot\boldsymbol{\sigma})^2}{2} + \ldots\right] + i\left[\frac{\theta}{2}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma} - \frac{(\frac{\theta}{2})^3(\hat{\mathbf{u}}\cdot\boldsymbol{\sigma})^3}{6} + \ldots\right] \\ &= \left[1 - \frac{(\frac{\theta}{2})^2}{2} + \ldots\right]\mathbf{I} + i\left[\frac{\theta}{2} - \frac{(\frac{\theta}{2})^3}{6} + \ldots\right]\hat{\mathbf{u}}\cdot\boldsymbol{\sigma} = \cos\frac{\theta}{2}\ \mathbf{I} + i\sin\frac{\theta}{2}\ \hat{\mathbf{u}}\cdot\boldsymbol{\sigma}, \end{aligned} \tag{431}$$

via repeated uses of eq. (419) with $\mathbf{a} = \mathbf{b} = \hat{\mathbf{u}}$.
While the Pauli operators $\boldsymbol{\sigma}_j$ do not commute with one another, we see from eq. (431) that $e^{a\boldsymbol{\sigma}_j + b\boldsymbol{\sigma}_k} = e^{b\boldsymbol{\sigma}_k + a\boldsymbol{\sigma}_j}$. However, $e^{a\boldsymbol{\sigma}_j + b\boldsymbol{\sigma}_k} \neq e^{a\boldsymbol{\sigma}_j}e^{b\boldsymbol{\sigma}_k} \neq e^{b\boldsymbol{\sigma}_k}e^{a\boldsymbol{\sigma}_j}$ when $j \neq k$. In particular, $e^{i\frac{\theta}{2}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}} \neq e^{i\frac{\theta}{2}u_x\boldsymbol{\sigma}_x}e^{i\frac{\theta}{2}u_y\boldsymbol{\sigma}_y}e^{i\frac{\theta}{2}u_z\boldsymbol{\sigma}_z}$.
The matrix form of eq. (429) is

$$\mathsf{U} = e^{i\delta}\begin{pmatrix} \cos\frac{\theta}{2} + i\sin\frac{\theta}{2}\ u_x & i\sin\frac{\theta}{2}(u_x - iu_y) \\ i\sin\frac{\theta}{2}(u_x + iu_y) & \cos\frac{\theta}{2} - i\sin\frac{\theta}{2}\ u_x \end{pmatrix}, \tag{432}$$

so the determinant of U is

$$\Delta_{\mathsf{U}} = e^{2i\delta}\left[\cos^2\frac{\theta}{2} + \sin^2\frac{\theta}{2}\left(u_x^2+u_y^2+u_z^2\right)\right] = e^{2i\delta}. \tag{433}$$

Hence, the 2×2 special unitary operators (those for which $\Delta_{\mathsf{U}}=1$) are those with $\delta = 0$ or π,

$$\mathsf{U} = \pm\left(\cos\frac{\theta}{2}\,\mathbf{I} + i\sin\frac{\theta}{2}\,\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}\right) = \pm e^{i\frac{\theta}{2}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}}, \qquad \mathsf{U}\in \mathrm{SU}(2). \tag{434}$$

We note that the $\sqrt{\mathsf{NOT}}$ operator, eq. (410), has determinant $\pm i$, and so is not a special unitary operator.

As to the factoid related to eq. (38), whenever $\mathsf{A}^2=\mathbf{I}$ *we can make the Taylor expansion,*

$$\begin{aligned} e^{i\theta\mathsf{A}} &= \sum_{k\ \text{even}}^{\infty}\frac{(i\theta\mathsf{A})^k}{k!} + \sum_{k\ \text{odd}}^{\infty}\frac{(i\theta\mathsf{A})^k}{k!} = \sum_{k\ \text{even}}^{\infty}\frac{(-1)^{k/2}\theta^k}{k!}\mathbf{I} + i\sum_{k\ \text{odd}}^{\infty}\frac{(-1)^{(k-1)/2}\theta^k}{k!}\mathsf{A} \\ &= \cos\theta\,\mathbf{I} + i\sin\theta\,\mathsf{A}. \end{aligned} \tag{435}$$

(d) Applying the Z gate to classical bits $|0\rangle$ and $|1\rangle$ we have,

$$\mathsf{Z}|0\rangle = \begin{pmatrix}1&0\\0&-1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = |0\rangle, \qquad \mathsf{Z}|1\rangle = \begin{pmatrix}1&0\\0&-1\end{pmatrix}\begin{pmatrix}0\\-1\end{pmatrix} = -|1\rangle. \tag{436}$$

In a classical context, the state $-|1\rangle$ is either not defined, or could be considered as identical to $|1\rangle$. Either way, the Z operator does not produce a result on classical bits different from the two valid unitary operations $\mathbf{I}$ and X (N. Hu, 9/22/16).

In contrast, the effect of Z on a general Qbit $|\psi\rangle = a|0\rangle + b|1\rangle$ is nontrivial,

$$\mathsf{Z}|\psi\rangle = \begin{pmatrix}1&0\\0&-1\end{pmatrix}\begin{pmatrix}a\\b\end{pmatrix} = \begin{pmatrix}a\\-b\end{pmatrix}. \tag{437}$$

(e) **2-Bit Classical Unitary Operators**

In part (a) we deduced that there are $(2^n)!$ classical unitary operators that act on n bits. Hence, there are $(2^2)! = 4! = 24$ classical 2-bit unitary operators.

In the form

$$\begin{pmatrix}x\\y\end{pmatrix} \to \begin{pmatrix}x'\\y'\end{pmatrix} = \mathsf{M}\begin{pmatrix}x\\y\end{pmatrix} \oplus \begin{pmatrix}a\\b\end{pmatrix}, \tag{438}$$

M is a 2×2 invertible matrix, a and b are constant Cbits, and x, y, x' and y' are variable Cbits.

If the output quantities x' and y' of eq. (438) are to be Cbits, the elements of matrix M can only be 0's or 1's. With this constraint, the transformation

eq. (438) is automatically unitary, in that it takes a pair of Cbits into another pair of Cbits.[141]

Since the 2×2 matrix M has 4 elements, each 0 or 1, there are $2^4 = 16$ different M's. The additional requirement that M be invertible implies that the determinant of M is nonzero. This eliminates those M's in which the rows are identical, or in which the columns are identical, or in which a row or column contains only 0's. The excluded matrices are

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix},$$
$$\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \tag{439}$$

which leaves six viable versions of M, of which only the first two are unitary:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}. \tag{440}$$

There are, of course, four versions of the constant Cbit vector,

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{441}$$

Hence, there are $6 \times 4 = 24$ version of the linear, unitary transformation eq. (438), which exhausts the number of 2-bit classical unitary operators.

For the record, these 24 operators can be expressed as the 4×4 matrices

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix},$$
$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix},$$
$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix},$$
$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix},$$
$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix},$$

[141] The use of addition modulo 2 in eq. (438) implies that $1 + 1 = 0$, so that the sum of any two Cbits is still a Cbit.

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}. \tag{442}$$

However, the matrix forms (442) obscure the fact that the 2-bit classical unitary operators are linear functions of the bits.

4. **Rotation Matrices**

(a) **Pauli Spin Matrices and Rotations**

The NOT operation, $\mathsf{X} = \boldsymbol{\sigma}_x$, that "flips" a bit can be interpreted as a rotation by 180° of the Bloch-sphere state vector about the x-axis. Thus,

$$\boldsymbol{\sigma}_x \begin{pmatrix} \cos\frac{\alpha}{2} \\ e^{i\beta}\sin\frac{\alpha}{2} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \cos\frac{\alpha}{2} \\ e^{i\beta}\sin\frac{\alpha}{2} \end{pmatrix} = \begin{pmatrix} e^{i\beta}\sin\frac{\alpha}{2} \\ \cos\frac{\alpha}{2} \end{pmatrix}, \tag{443}$$

while a rotation $\mathbf{R}_x(180°)$ by 180° about the x-axis in our abstract spherical coordinate system takes α to $\pi - \alpha$ and β to $-\beta$,

$$\mathbf{R}_x(180°) \begin{pmatrix} \cos\frac{\alpha}{2} \\ e^{i\beta}\sin\frac{\alpha}{2} \end{pmatrix} = \begin{pmatrix} \cos\frac{\pi-\alpha}{2} \\ e^{-i\beta}\sin\frac{\pi-\alpha}{2} \end{pmatrix} = e^{-i\beta} \begin{pmatrix} e^{i\beta}\sin\frac{\alpha}{2} \\ \cos\frac{\alpha}{2} \end{pmatrix}. \tag{444}$$

Since the overall phase of a state does not affect its meaning, our prescription can be considered satisfactory thus far.

Can we interpret the operation $\boldsymbol{\sigma}_y$ as a rotation by 180° about the y-axis? On one hand,

$$\boldsymbol{\sigma}_y \begin{pmatrix} \cos\frac{\alpha}{2} \\ e^{i\beta}\sin\frac{\alpha}{2} \end{pmatrix} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} \cos\frac{\alpha}{2} \\ e^{i\beta}\sin\frac{\alpha}{2} \end{pmatrix} = \begin{pmatrix} -ie^{i\beta}\sin\frac{\alpha}{2} \\ i\cos\frac{\alpha}{2} \end{pmatrix}, \tag{445}$$

while a rotation $\mathbf{R}_y(180°)$ by 180° about the y-axis in our abstract spherical coordinate system takes α to $\pi - \alpha$ and β to $\pi - \beta$,

$$\mathbf{R}_y(180°) \begin{pmatrix} \cos\frac{\alpha}{2} \\ e^{i\beta}\sin\frac{\alpha}{2} \end{pmatrix} = \begin{pmatrix} \cos\frac{\pi-\alpha}{2} \\ e^{i(\pi-\beta)}\sin\frac{\pi-\alpha}{2} \end{pmatrix} = ie^{-i\beta} \begin{pmatrix} -ie^{i\beta}\sin\frac{\alpha}{2} \\ i\cos\frac{\alpha}{2} \end{pmatrix}. \tag{446}$$

Similarly, we interpret the operation $\boldsymbol{\sigma}_z$ as a rotation by 180° about the z-axis:

$$\boldsymbol{\sigma}_z \begin{pmatrix} \cos\frac{\alpha}{2} \\ e^{i\beta}\sin\frac{\alpha}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos\frac{\alpha}{2} \\ e^{i\beta}\sin\frac{\alpha}{2} \end{pmatrix} = \begin{pmatrix} \cos\frac{\alpha}{2} \\ -e^{i\beta}\sin\frac{\alpha}{2} \end{pmatrix}, \tag{447}$$

while a rotation $\mathbf{R}_z(180°)$ by 180° about the z-axis in our abstract spherical coordinate system takes α to α and β to $\pi + \beta$,

$$\mathbf{R}_z(180°) \begin{pmatrix} \cos\frac{\alpha}{2} \\ e^{i\beta}\sin\frac{\alpha}{2} \end{pmatrix} = \begin{pmatrix} \cos\frac{\alpha}{2} \\ e^{i(\pi+\beta)}\sin\frac{\alpha}{2} \end{pmatrix} = \begin{pmatrix} \cos\frac{\alpha}{2} \\ -e^{i\beta}\sin\frac{\alpha}{2} \end{pmatrix}. \tag{448}$$

(b) **Rotation and Pauli Spin Matrices**

The rotations (45)-(47) are readily seen to be exponentials of the Pauli matrices,

$$\mathsf{R}_x(\phi) = \begin{pmatrix} \cos\frac{\phi}{2} & i\sin\frac{\phi}{2} \\ i\sin\frac{\phi}{2} & \cos\frac{\phi}{2} \end{pmatrix} = \cos\frac{\phi}{2}\mathbf{I} + i\sin\frac{\phi}{2}\boldsymbol{\sigma}_x = e^{i\frac{\phi}{2}\boldsymbol{\sigma}_x}, \tag{449}$$

$$\mathsf{R}_y(\phi) = \begin{pmatrix} \cos\frac{\phi}{2} & \sin\frac{\phi}{2} \\ -\sin\frac{\phi}{2} & \cos\frac{\phi}{2} \end{pmatrix} = \cos\frac{\phi}{2}\mathbf{I} + i\sin\frac{\phi}{2}\boldsymbol{\sigma}_y = e^{i\frac{\phi}{2}\boldsymbol{\sigma}_y}, \tag{450}$$

$$\mathsf{R}_z(\phi) = \begin{pmatrix} e^{i\phi/2} & 0 \\ 0 & e^{-i\phi/2} \end{pmatrix} = \cos\frac{\phi}{2}\mathbf{I} + i\sin\frac{\phi}{2}\boldsymbol{\sigma}_z = e^{i\frac{\phi}{2}\boldsymbol{\sigma}_z}, \tag{451}$$

recalling eq. (37).

The αth power of the Pauli matrix $\boldsymbol{\sigma}_z$ follows from eq. (54) as

$$\begin{aligned} \boldsymbol{\sigma}_z^\alpha = \mathsf{Z}^\alpha &= (e^{i\frac{\pi}{2}}e^{-i\frac{\pi}{2}\boldsymbol{\sigma}_z})^\alpha = e^{i\frac{\pi\alpha}{2}}e^{-i\frac{\pi\alpha}{2}\boldsymbol{\sigma}_z} = e^{i\frac{\pi\alpha}{2}}\left[\mathbf{I}\cos\frac{-\pi\alpha}{2} + i\boldsymbol{\sigma}_z\sin\frac{-\pi\alpha}{2}\right] \\ &= e^{i\frac{\pi\alpha}{2}}\begin{pmatrix} \cos\frac{-\pi\alpha}{2} + i\sin\frac{-\pi\alpha}{2} & 0 \\ 0 & \cos\frac{\pi\alpha}{2} + i\sin\frac{\pi\alpha}{2} \end{pmatrix} \\ &= e^{i\frac{\pi\alpha}{2}}\begin{pmatrix} e^{-i\frac{\pi\alpha}{2}} & 0 \\ 0 & e^{i\frac{\pi\alpha}{2}} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi\alpha} \end{pmatrix}. \end{aligned} \tag{452}$$

(c) **More Square Roots of NOT**

If we take the NOT operation to mean flipping a bit with a possible phase change, we can accomplish this by a rotation by 180° about any axis in the x-y plane of the Bloch sphere. The corresponding $\sqrt{\mathsf{NOT}}$ operation would be a rotation by 90° about the same axis.

Consider an axis y_1 in the x-y plane that makes angle α to the y-axis. A prescription for rotation by angle ϕ about the y_1-axis, based on the general procedure (44), is to make a rotation by α about the z axis, followed by a rotation by ϕ about the y_1-axis, and finally a rotation by $-\alpha$ about the z' axis:

$$\mathsf{R}_{y_1}(\phi) = \mathsf{R}(\alpha, \phi, -\alpha) = \begin{pmatrix} \cos(\phi/2) & \sin(\phi/2)e^{-i\alpha} \\ -\sin(\phi/2)e^{i\alpha} & \cos(\phi/2) \end{pmatrix} \tag{453}$$

Note that using $\alpha = -\pi/2$ in eq. (453) gives the matrix (45) for rotations about the x-axis.

Putting $\phi = \pi$ and $\pi/2$ in eq. (453) we obtain new expressions for the NOT and $\sqrt{\mathsf{NOT}}$ operators:

$$\mathsf{NOT} = \mathsf{R}_{y_1}(\pi) = \begin{pmatrix} 0 & e^{-i\alpha} \\ -e^{i\alpha} & 0 \end{pmatrix}, \tag{454}$$

$$\sqrt{\mathsf{NOT}} = \mathsf{R}_{y_1}(\pi/2) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & e^{-i\alpha} \\ -e^{i\alpha} & 1 \end{pmatrix}. \tag{455}$$

For example, with $y_1 = x$, we have $\alpha = -\pi/2$ and

$$\mathsf{NOT} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \tag{456}$$

$$\sqrt{\mathsf{NOT}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}. \tag{457}$$

This version of the NOT operator is i times that of eq. (31), and this $\sqrt{\mathsf{NOT}}$ is $(1-i)/\sqrt{2}$ times the first form of eq. (410). Since the absolute magnitude of these factors is unity, we can say that the results the two sets of NOT's and $\sqrt{\mathsf{NOT}}$'s are equivalent.

Similarly, with $y_1 = y$, we use $\alpha = 0$ in eqs. (454) and (455) to obtain

$$\mathsf{NOT} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{458}$$

$$\sqrt{\mathsf{NOT}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}. \tag{459}$$

In this case, the new operators are not simply the originals, eqs. (31) and (410), multiplied by a phase factor. The NOT operator (456) takes state $|0\rangle$ to $-|1\rangle$ and state $|1\rangle$ to $|0\rangle$. These final states are, to within a phase factor, the results of "flipping" the initial states, so we can again say that the new representation is valid.

(d) **Rotation Matrices and the General Form (37)**

$$\mathsf{U} = e^{i\delta}\left(\cos\frac{\theta}{2}\,\mathbf{I} + i\sin\frac{\theta}{2}\,\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}\right) = e^{i\delta}e^{i\frac{\theta}{2}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}}. \tag{37}$$

The general rotation matrix (44) can be expanded as

$$\begin{aligned}
\mathsf{R}(\alpha,\beta,\gamma) &= \begin{pmatrix} \cos\frac{\beta}{2}e^{i(\alpha+\gamma)/2} & \sin\frac{\beta}{2}e^{i(-\alpha+\gamma)/2} \\ -\sin\frac{\beta}{2}e^{i(\alpha-\gamma)/2} & \cos\frac{\beta}{2}e^{-i(\alpha+\gamma)/2} \end{pmatrix} \\
&= \begin{pmatrix} \cos\frac{\beta}{2}(\cos\frac{\alpha+\gamma}{2} + i\sin\frac{\alpha+\gamma}{2}) & \sin\frac{\beta}{2}(\cos\frac{\alpha-\gamma}{2} - i\sin\frac{\alpha-\gamma}{2}) \\ -\sin\frac{\beta}{2}(\cos\frac{\alpha-\gamma}{2} + i\sin\frac{\alpha-\gamma}{2}) & \cos\frac{\beta}{2}(\cos\frac{\alpha+\gamma}{2} - i\sin\frac{\alpha+\gamma}{2}) \end{pmatrix} \\
&= \cos\frac{\beta}{2}\cos\frac{\alpha+\gamma}{2}\,\mathbf{I} - i\sin\frac{\beta}{2}\sin\frac{\alpha-\gamma}{2}\,\boldsymbol{\sigma}_x \\
&\quad + i\sin\frac{\beta}{2}\cos\frac{\alpha-\gamma}{2}\,\boldsymbol{\sigma}_y + i\cos\frac{\beta}{2}\sin\frac{\alpha+\gamma}{2}\,\boldsymbol{\sigma}_z.
\end{aligned} \tag{460}$$

To cast this in the general form (37) we set $\delta = 0$, and

$$\cos\frac{\theta}{2} = \cos\frac{\beta}{2}\cos\frac{\alpha+\gamma}{2}, \tag{461}$$

$$\sin\frac{\theta}{2} = \sqrt{1-\cos^2\frac{\beta}{2}\cos^2\frac{\alpha+\gamma}{2}} = \sqrt{\sin^2\frac{\beta}{2}+\cos^2\frac{\beta}{2}\sin^2\frac{\alpha+\gamma}{2}}, \quad (462)$$

$$\hat{\mathbf{u}} = \frac{\left(-\sin\frac{\beta}{2}\sin\frac{\alpha-\gamma}{2}, \sin\frac{\beta}{2}\cos\frac{\alpha-\gamma}{2}, \cos\frac{\beta}{2}\sin\frac{\alpha+\gamma}{2}\right)}{\sqrt{\sin^2\frac{\beta}{2}+\cos^2\frac{\beta}{2}\sin^2\frac{\alpha+\gamma}{2}}}, \quad (463)$$

The forms (461)-(463) are consistent for $0 < \theta < 2\pi$. However, for $-2\pi < \theta < 0$, the unit vector $\hat{\mathbf{u}}$ should be the negative of expression (463).

For example, $\alpha = -\gamma = -\pi/2$ leads to $\theta = \beta$, $\hat{\mathbf{u}} = (1,0,0) = \hat{\mathbf{x}}$, as expected from eq. (453). Likewise, $\alpha = \gamma = 0$ corresponds to $\theta = \beta$ and $\hat{\mathbf{u}} = (0,1,0) = \hat{\mathbf{y}}$, while $\beta = \gamma = 0$ corresponds to $\theta = \alpha$ and $\hat{\mathbf{u}} = (0,0,1) = \hat{\mathbf{z}}$.

For completeness, we record expressions for the rotation angles α, β and γ in terms of θ and $\hat{\mathbf{u}}$. From eq. (460) we also have

$$u_x \sin\frac{\theta}{2} = -\sin\frac{\beta}{2}\sin\frac{\alpha-\gamma}{2}, \quad (464)$$

$$u_y \sin\frac{\theta}{2} = \sin\frac{\beta}{2}\cos\frac{\alpha-\gamma}{2}, \quad (465)$$

$$u_z \sin\frac{\theta}{2} = \cos\frac{\beta}{2}\sin\frac{\alpha+\gamma}{2}. \quad (466)$$

Equations (461) and (465) give

$$\sin\frac{\beta}{2} = \sin\frac{\theta}{2}\sqrt{u_x^2+u_y^2}, \quad (467)$$

and so

$$\sin\frac{\alpha-\gamma}{2} = -\frac{u_x}{u_x^2+u_y^2}. \quad (468)$$

Equations (464) and (466) give

$$\tan\frac{\alpha+\gamma}{2} = u_z\tan\frac{\theta}{2}. \quad (469)$$

Hence,

$$\alpha = \tan^{-1}\left(u_z\tan\frac{\theta}{2}\right) - \sin^{-1}\frac{u_x}{\sqrt{u_x^2+u_y^2}}, \quad (470)$$

$$\beta = 2\sin^{-1}\left(\sin\frac{\theta}{2}\sqrt{u_x^2+u_y^2}\right), \quad (471)$$

$$\gamma = \tan^{-1}\left(u_z\tan\frac{\theta}{2}\right) + \sin^{-1}\frac{u_x}{\sqrt{u_x^2+u_y^2}}. \quad (472)$$

(e) **Double NOT**

We first write the unitary operator U as

$$\mathsf{U} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \quad (473)$$

Then

$$\mathsf{XUX} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} b & a \\ d & c \end{pmatrix} = \begin{pmatrix} d & c \\ b & a \end{pmatrix}. \tag{474}$$

Writing

$$\mathsf{U} = e^{i\delta}\mathsf{R}(\alpha,\beta,\gamma) = e^{i\delta} \begin{pmatrix} \cos(\beta/2)e^{i(\alpha+\gamma)/2} & \sin(\beta/2)e^{i(-\alpha+\gamma)/2} \\ -\sin(\beta/2)e^{i(\alpha-\gamma)/2} & \cos(\beta/2)e^{-i(\alpha+\gamma)/2} \end{pmatrix}, \tag{475}$$

and recalling eq. (460), we see that

$$\mathsf{XUX} = e^{i\delta} \begin{pmatrix} \cos(\beta/2)e^{-i(\alpha+\gamma)/2} & -\sin(\beta/2)e^{i(\alpha-\gamma)/2} \\ \sin(\beta/2)e^{i(-\alpha+\gamma)/2} & \cos(\beta/2)e^{i(\alpha+\gamma)/2} \end{pmatrix} = e^{i\delta}\mathsf{R}(-\alpha,-\beta,-\gamma). \tag{476}$$

(f) **Basis Change**

Since U is unitary, we have

$$\mathsf{U}^\dagger\mathsf{U} = \begin{pmatrix} a^\star & c^\star \\ b^\star & d^\star \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} |a|^2+|c|^2 & a^\star b + c^\star d \\ ab^\star + cd^\star & |b|^2+|d|^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{477}$$

Thus the states

$$|\psi\rangle = \mathsf{U}|0\rangle = a|0\rangle + c|1\rangle, \qquad |\phi\rangle = \mathsf{U}|1\rangle = b|0\rangle + d|1\rangle. \tag{478}$$

obey

$$\begin{aligned} \langle\psi|\psi\rangle &= |a|^2+|c|^2 = 1, \quad \text{or} \quad \langle\psi|\psi\rangle = \langle 0|\mathsf{U}^\dagger\mathsf{U}|0\rangle = \langle 0|0\rangle = 1, && (479)\\ \langle\phi|\phi\rangle &= |b|^2+|d|^2 = 1, \quad \text{or} \quad \langle\phi|\phi\rangle = \langle 1|\mathsf{U}^\dagger\mathsf{U}|1\rangle = \langle 1|1\rangle = 1, && (480)\\ \langle\psi|\phi\rangle &= a^\star b + c^\star d = 0, \quad \text{or} \quad \langle\psi|\phi\rangle = \langle 0|\mathsf{U}^\dagger\mathsf{U}|1\rangle = \langle 0|1\rangle = 0, && (481) \end{aligned}$$

and so they are orthonormal as claimed.

(g) **Hadamard Transformation**

The Hadamard transformation,

$$\mathsf{H} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{\mathsf{X}+\mathsf{Z}}{\sqrt{2}} = \frac{\boldsymbol{\sigma}_x + \boldsymbol{\sigma}_z}{\sqrt{2}}, \tag{482}$$

has determinant $\Delta_{\mathsf{H}} = -1 = e^{2i\delta}$, where $e^{i\delta}$ is the phase factor in the general forms (37) and (56),

$$\mathsf{U} = e^{i\delta}\left(\cos\frac{\theta}{2}\,\mathbf{I} + i\sin\frac{\theta}{2}\,\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}\right) = e^{i\delta}e^{i\frac{\theta}{2}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}}, \tag{37}$$

$$\mathsf{U} = e^{i\delta}\mathsf{R}(\alpha,\beta,\gamma) = e^{i\delta}\mathsf{R}_z(\gamma)\mathsf{R}_y(\beta)\mathsf{R}_z(\alpha). \tag{56}$$

Thus, $\delta = \pm\pi/2$.

We first examine that case that $\delta = \pi/2$. To relate eq. (482) to the form (37), we write

$$\begin{aligned} \mathsf{H} &= \frac{\boldsymbol{\sigma}_x + \boldsymbol{\sigma}_z}{\sqrt{2}} = e^{i\delta}(-i)\frac{\boldsymbol{\sigma}_x + \boldsymbol{\sigma}_z}{\sqrt{2}} = e^{i\delta}(-i\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}) \\ &= e^{i\delta}[\cos(-\pi/2)\ \mathbf{I} + i\sin(-\pi/2)\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}] = e^{i\delta}e^{i\frac{\theta}{2}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}}, \end{aligned} \tag{483}$$

where

$$\delta = \frac{\pi}{2}, \qquad \theta = -\pi, \qquad \text{and} \qquad \hat{\mathbf{u}} = \frac{(1,0,1)}{\sqrt{2}}. \tag{484}$$

Because $\theta = -\pi$, we must be careful to use the negative of expression eq. (463) for the unit vector $\hat{\mathbf{u}}$. First, we can set

$$\frac{\alpha - \gamma}{2} = \frac{\pi}{2} \tag{485}$$

to have $\hat{u}_y = 0$. Then to obtain $\hat{u}_x = \hat{u}_z = 1/\sqrt{2}$ we need

$$\frac{\sin\frac{\beta}{2}\sin\frac{\alpha-\gamma}{2}}{\sqrt{\sin^2\frac{\beta}{2} + \cos^2\frac{\beta}{2}\sin^2\frac{\alpha+\gamma}{2}}} = \frac{1}{\sqrt{2}} = -\frac{\cos\frac{\beta}{2}\sin\frac{\alpha+\gamma}{2}}{\sqrt{\sin^2\frac{\beta}{2} + \cos^2\frac{\beta}{2}\sin^2\frac{\alpha+\gamma}{2}}}. \tag{486}$$

This is satisfied by taking

$$\beta = \frac{\pi}{2} \tag{487}$$

to have $\cos\frac{\beta}{2} = 1/\sqrt{2}$, and

$$\frac{\alpha + \gamma}{2} = -\frac{\pi}{2}. \tag{488}$$

Combining eqs. (485) and (488) we arrive at

$$\alpha = 0, \qquad \beta = \frac{\pi}{2}, \qquad \gamma = -\pi, \qquad \text{and} \qquad \delta = \frac{\pi}{2}, \tag{489}$$

so that the Hadamard transformation can also be written as

$$\begin{aligned} \mathsf{H} &= e^{i\delta}\mathsf{R}(\alpha,\beta,\gamma) = e^{i\pi/2}\mathsf{R}_z(-\pi)\mathsf{R}_y\left(-\frac{\pi}{2}\right) \\ &= i\begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \end{aligned} \tag{490}$$

By setting $\delta = \pm\pi/2$ and $\alpha - \gamma = \pm\pi$ we can obtain 3 more representations of the Hadamard transformation,

$$\mathsf{H} = e^{i\pi/2}\mathsf{R}_y\left(-\frac{\pi}{2}\right)\mathsf{R}_z(-\pi) = e^{-i\pi/2}\mathsf{R}_y\left(-\frac{\pi}{2}\right)\mathsf{R}_z(\pi) = e^{-i\pi/2}\mathsf{R}_z(\pi)\mathsf{R}_y\left(\frac{\pi}{2}\right). \tag{491}$$

(h) The αth power of the Pauli matrix $\boldsymbol{\sigma}_x$ follows from eq. (54) as

$$\begin{aligned}
\boldsymbol{\sigma}_x^\alpha &= e^{i\frac{\pi\alpha}{2}}e^{-i\frac{\pi\alpha}{2}\boldsymbol{\sigma}_x} = e^{i\frac{\pi\alpha}{2}}\left[\mathbf{I}\cos(\pi\alpha/2) - i\boldsymbol{\sigma}_x\sin(\pi\alpha/2)\right] \\
&= \frac{e^{-i\frac{\pi\alpha}{2}}}{2}\left[\mathbf{I}(e^{i\frac{\pi\alpha}{2}} + e^{-i\frac{\pi\alpha}{2}}) - \boldsymbol{\sigma}_x(e^{i\frac{\pi\alpha}{2}} - e^{-i\frac{\pi\alpha}{2}})\right] \\
&= \frac{1}{2}\left[\mathbf{I}(e^{i\pi\alpha}+1) - \boldsymbol{\sigma}_x(e^{i\pi\alpha}-1)\right] \\
&= \frac{1}{2}\begin{pmatrix} 1+e^{i\pi\alpha} & 1-e^{i\pi\alpha} \\ 1-e^{i\pi\alpha} & 1+e^{i\pi\alpha}\end{pmatrix}. \qquad (492)
\end{aligned}$$

Meanwhile, we have

$$\begin{aligned}
\mathsf{H}\boldsymbol{\sigma}_z^\alpha\mathsf{H} &= \frac{1}{2}\begin{pmatrix}1&1\\1&-1\end{pmatrix}\begin{pmatrix}1&0\\0&e^{i\pi\alpha}\end{pmatrix}\begin{pmatrix}1&1\\1&-1\end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} 1+e^{i\pi\alpha} & 1-e^{i\pi\alpha} \\ 1-e^{i\pi\alpha} & 1+e^{i\pi\alpha}\end{pmatrix} = \boldsymbol{\sigma}_x^\alpha. \qquad (493)
\end{aligned}$$

Similarly,

$$\begin{aligned}
\boldsymbol{\sigma}_y^\alpha &= e^{i\frac{\pi\alpha}{2}}e^{-i\frac{\pi\alpha}{2}\boldsymbol{\sigma}_y} = \frac{1}{2}\left[\mathbf{I}(e^{i\pi\alpha}+1) - \boldsymbol{\sigma}_y(e^{i\pi\alpha}-1)\right] \\
&= \frac{1}{2}\begin{pmatrix} 1+e^{i\pi\alpha} & -i(1-e^{i\pi\alpha}) \\ i(1-e^{i\pi\alpha}) & 1+e^{i\pi\alpha}\end{pmatrix}, \qquad (494)
\end{aligned}$$

while

$$\begin{aligned}
\boldsymbol{\sigma}_z^{1/2}\boldsymbol{\sigma}_x^\alpha\boldsymbol{\sigma}_z^{-1/2} &= \frac{1}{2}\begin{pmatrix}1&0\\0&i\end{pmatrix}\begin{pmatrix} 1+e^{i\pi\alpha} & 1-e^{i\pi\alpha} \\ 1-e^{i\pi\alpha} & 1+e^{i\pi\alpha}\end{pmatrix}\begin{pmatrix}1&0\\0&-i\end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} 1+e^{i\pi\alpha} & -i(1-e^{i\pi\alpha}) \\ i(1-e^{i\pi\alpha}) & 1+e^{i\pi\alpha}\end{pmatrix} = \boldsymbol{\sigma}_y^\alpha. \qquad (495)
\end{aligned}$$

Using eqs. (483)-(484) we can write the Hadamard transformation in exponential form as[142]

$$\mathsf{H} = e^{i\frac{\pi}{2}}e^{-i\frac{\pi}{2\sqrt{2}}(\boldsymbol{\sigma}_x+\boldsymbol{\sigma}_z)} = e^{i\frac{\pi}{2}}e^{-i\frac{\pi}{2}\mathsf{H}}, \qquad (496)$$

so the αth power of this is

$$\begin{aligned}
\mathsf{H}^\alpha &= e^{i\frac{\pi\alpha}{2}}e^{-i\frac{\pi\alpha}{2}\mathsf{H}} = e^{i\frac{\pi\alpha}{2}}\left[\mathbf{I}\cos(\pi\alpha/2) - i\mathsf{H}\sin(\pi\alpha/2)\right] \\
&= \frac{e^{i\frac{\pi\alpha}{2}}}{2}\left[\mathbf{I}(e^{i\frac{\pi\alpha}{2}} + e^{-i\frac{\pi\alpha}{2}}) - \mathsf{H}(e^{i\frac{\pi\alpha}{2}} - e^{-i\frac{\pi\alpha}{2}})\right] \\
&= \frac{1}{2}\left[\mathbf{I}(1+e^{i\pi\alpha}) + \mathsf{H}(1-e^{i\pi\alpha})\right] \\
&= \frac{1}{2\sqrt{2}}\begin{pmatrix} \sqrt{2}+1+(\sqrt{2}-1)e^{i\pi\alpha} & 1-e^{i\pi\alpha} \\ 1-e^{i\pi\alpha} & \sqrt{2}-1+(\sqrt{2}+1)e^{i\pi\alpha}\end{pmatrix}, \qquad (497)
\end{aligned}$$

[142] This also follows from eq. (38).

while

$$\begin{aligned}\boldsymbol{\sigma}_y^{1/4}\boldsymbol{\sigma}_z^{\alpha}\boldsymbol{\sigma}_y^{-1/4} &= \frac{1}{8}\begin{pmatrix} \sqrt{2}+1+i & -1-i(\sqrt{2}-1) \\ 1+i(\sqrt{2}-1) & \sqrt{2}+1+i \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi\alpha} \end{pmatrix} \\ &\quad \begin{pmatrix} \sqrt{2}+1-i & 1-i(\sqrt{2}-1) \\ -1+i(\sqrt{2}-1) & \sqrt{2}+1-i \end{pmatrix} \\ &= \frac{1}{2\sqrt{2}}\begin{pmatrix} \sqrt{2}+1+(\sqrt{2}-1)e^{i\pi\alpha} & 1-e^{i\pi\alpha} \\ 1-e^{i\pi\alpha} & \sqrt{2}-1+(\sqrt{2}+1)e^{i\pi\alpha} \end{pmatrix} \\ &= \mathsf{H}^{\alpha}. \end{aligned} \tag{498}$$

(i) To characterize the product

$$\mathsf{H}^{-1/2}\; e^{i\frac{\theta}{2}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}}\; \mathsf{H}^{1/2}, \tag{499}$$

we note from eq. (497) that

$$\mathsf{H}^{\pm 1/2} = e^{\pm i\frac{\pi}{4}}e^{\mp i\frac{\pi}{4}\mathsf{H}} = e^{\pm i\frac{\pi}{4}}[\cos(\pi/4)\mathbf{I} \mp i\sin(\pi/4)\mathsf{H}] = \frac{e^{\pm i\frac{\pi}{4}}}{\sqrt{2}}[\mathbf{I}\mp i\mathsf{H}]. \tag{500}$$

Thus,

$$\begin{aligned}\mathsf{H}^{-1/2}\; e^{i\frac{\theta}{2}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}}\; \mathsf{H}^{1/2} &= \frac{1}{2}[\mathbf{I}+i\mathsf{H}][\cos(\theta/2)\mathbf{I} + i\sin(\theta/2)\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}][\mathbf{I}-i\mathsf{H}] \\ &= \cos(\theta/2)\mathbf{I} + \frac{i}{2}\sin(\theta/2)[\hat{\mathbf{u}}\cdot\boldsymbol{\sigma} + \hat{\mathbf{u}}\cdot\mathsf{H}\boldsymbol{\sigma}\mathsf{H} - i\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}\mathsf{H} + i\hat{\mathbf{u}}\cdot\mathsf{H}\boldsymbol{\sigma}]. \end{aligned} \tag{501}$$

We would now like to show that the quantity in brackets in the second line of eq. (501) can be written as $2\hat{\mathbf{v}}\cdot\boldsymbol{\sigma}$. For this, we recall that $\mathsf{H} = (\boldsymbol{\sigma}_x+\boldsymbol{\sigma}_z)/\sqrt{2}$ and the facts that $\boldsymbol{\sigma}_j^2 = \mathbf{I}$ and $\boldsymbol{\sigma}_j\boldsymbol{\sigma}_k = i\epsilon_{jkl}\boldsymbol{\sigma}_l$ when $j\neq k$ to accumulate the following relations:

$$\boldsymbol{\sigma}_x\mathsf{H} = \boldsymbol{\sigma}_x\frac{\boldsymbol{\sigma}_x+\boldsymbol{\sigma}_z}{\sqrt{2}} = \frac{\mathbf{I}-i\boldsymbol{\sigma}_y}{\sqrt{2}}, \qquad \mathsf{H}\boldsymbol{\sigma}_x = \frac{\mathbf{I}+i\boldsymbol{\sigma}_y}{\sqrt{2}},$$

$$\mathsf{H}\boldsymbol{\sigma}_x\mathsf{H} = \frac{\boldsymbol{\sigma}_x+\boldsymbol{\sigma}_z-i(\boldsymbol{\sigma}_x+\boldsymbol{\sigma}_z)\boldsymbol{\sigma}_y}{2} = \boldsymbol{\sigma}_z, \qquad \boldsymbol{\sigma}_x\mathsf{H}-\mathsf{H}\boldsymbol{\sigma}_x = -\sqrt{2}i\boldsymbol{\sigma}_y, \tag{502}$$

$$\boldsymbol{\sigma}_y\mathsf{H} = \boldsymbol{\sigma}_y\frac{\boldsymbol{\sigma}_x+\boldsymbol{\sigma}_z}{\sqrt{2}} = i\frac{-\boldsymbol{\sigma}_z+\boldsymbol{\sigma}_x}{\sqrt{2}}, \qquad \mathsf{H}\boldsymbol{\sigma}_y = i\frac{\boldsymbol{\sigma}_z-\boldsymbol{\sigma}_x}{\sqrt{2}},$$

$$\mathsf{H}\boldsymbol{\sigma}_y\mathsf{H} = i\frac{(\boldsymbol{\sigma}_x+\boldsymbol{\sigma}_z)(\boldsymbol{\sigma}_x-\boldsymbol{\sigma}_z)}{2} = -\boldsymbol{\sigma}_y, \qquad \boldsymbol{\sigma}_y\mathsf{H}-\mathsf{H}\boldsymbol{\sigma}_y = \sqrt{2}i(\boldsymbol{\sigma}_x-\boldsymbol{\sigma}_z), \tag{503}$$

$$\boldsymbol{\sigma}_z\mathsf{H} = \boldsymbol{\sigma}_z\frac{\boldsymbol{\sigma}_x+\boldsymbol{\sigma}_z}{\sqrt{2}} = \frac{i\boldsymbol{\sigma}_y+\mathbf{I}}{\sqrt{2}}, \qquad \mathsf{H}\boldsymbol{\sigma}_z = \frac{-i\boldsymbol{\sigma}_y+\mathbf{I}}{\sqrt{2}},$$

$$\mathsf{H}\boldsymbol{\sigma}_z\mathsf{H} = \frac{i(\boldsymbol{\sigma}_x+\boldsymbol{\sigma}_z)\boldsymbol{\sigma}_y+\boldsymbol{\sigma}_x+\boldsymbol{\sigma}_z}{2} = \boldsymbol{\sigma}_x, \qquad \boldsymbol{\sigma}_z\mathsf{H}-\mathsf{H}\boldsymbol{\sigma}_z = \sqrt{2}i\boldsymbol{\sigma}_y, \tag{504}$$

Using eqs. (502)-(504) in eq. (501) we find that

$$\begin{aligned}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}+\hat{\mathbf{u}}\cdot\mathsf{H}\boldsymbol{\sigma}\mathsf{H}-i\hat{\mathbf{u}}\cdot(\boldsymbol{\sigma}\mathsf{H}-\mathsf{H}\boldsymbol{\sigma}) &= u_x\boldsymbol{\sigma}_x+u_y\boldsymbol{\sigma}_y+u_z\boldsymbol{\sigma}_z\\ &\quad u_x\boldsymbol{\sigma}_z-u_y\boldsymbol{\sigma}_y+u_z\boldsymbol{\sigma}_x\\ &\quad +\sqrt{2}[-u_x\boldsymbol{\sigma}_y+u_y(\boldsymbol{\sigma}_x-\boldsymbol{\sigma}_x)+u_z\boldsymbol{\sigma}_y]\\ &= \boldsymbol{\sigma}_x(u_x+\sqrt{2}u_y+u_z)-\sqrt{2}\boldsymbol{\sigma}_y(u_x-u_z)\\ &\quad +\boldsymbol{\sigma}_z(u_x-\sqrt{2}u_y+u_z)\\ &= 2\hat{\mathbf{v}}\cdot\boldsymbol{\sigma}, \end{aligned}\tag{505}$$

where

$$\hat{\mathbf{v}}=\frac{(u_x+\sqrt{2}u_y+u_z,-\sqrt{2}(u_x-\ u_z),u_x-\sqrt{2}u_y+u_z)}{2}\,.\tag{506}$$

The vector $\hat{\mathbf{v}}$ is readily verified to be a unit vector, while

$$\hat{\mathbf{u}}\cdot\hat{\mathbf{v}}=(u_x+u_z)^2.\tag{507}$$

Hence,

$$\mathsf{H}^{-1/2}\,e^{i\frac{\theta}{2}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}}\,\mathsf{H}^{1/2}=e^{i\frac{\theta}{2}\hat{\mathbf{v}}\cdot\boldsymbol{\sigma}},\tag{508}$$

where $\hat{\mathbf{v}}$ is orthogonal to $\hat{\mathbf{u}}$ provided $u_z=-u_x$.

(j) **Basis States for the Hadamard Transformation**[143]

We seek all orthonormal bases $[|\psi\rangle,|\phi\rangle]$ for which

$$\mathsf{H}|\psi\rangle=\frac{|\psi\rangle+|\phi\rangle}{\sqrt{2}}\,,\qquad\text{and}\qquad\mathsf{H}|\phi\rangle=\frac{|\psi\rangle-|\phi\rangle}{\sqrt{2}}\,.\tag{68}$$

where the basis states are described as unit vectors on the Bloch sphere,

$$|\psi\rangle=\cos\frac{\alpha}{2}|0\rangle+e^{i\beta}\sin\frac{\alpha}{2}|1\rangle,\tag{42}$$

$$|\phi\rangle=e^{i\delta}\left[\sin\frac{\alpha}{2}|0\rangle-e^{i\beta}\cos\frac{\alpha}{2}|1\rangle\right].\tag{69}$$

Inserting eqs. (42) and (69) into the first of eq. (68), we find

$$\begin{aligned}&\left[\cos\frac{\alpha}{2}+e^{i\beta}\sin\frac{\alpha}{2}\right]\frac{|0\rangle}{\sqrt{2}}+\left[\cos\frac{\alpha}{2}-e^{i\beta}\sin\frac{\alpha}{2}\right]\frac{|1\rangle}{\sqrt{2}}\\ &=\left[\cos\frac{\alpha}{2}+e^{i\delta}\sin\frac{\alpha}{2}\right]\frac{|0\rangle}{\sqrt{2}}+\left[-e^{i(\beta+\delta)}\cos\frac{\alpha}{2}+e^{i\beta}\sin\frac{\alpha}{2}\right]\frac{|1\rangle}{\sqrt{2}}\,.\end{aligned}\tag{509}$$

Equating the coefficients of $|0\rangle$ we see that $\beta=\delta$, while from the coefficients of $|1\rangle$ we find

$$\cos\frac{\alpha}{2}(1+e^{2i\beta})=2e^{i\beta}\sin\frac{\alpha}{2}\,,\tag{510}$$

and hence,

$$\cos\beta=\tan\frac{\alpha}{2}=\frac{1-\cos\alpha}{\sin\alpha}\,.\tag{511}$$

[143] http://kirkmcd.princeton.edu/examples/QM/maitra_quant-ph-0505068.pdf

Use of the second of eq. (68) leads to the same result.

The solution to eq. (511) includes $|\psi\rangle = |0\rangle$ with $\alpha = 0$, $\beta = \pm\pi/2$, and $|\psi\rangle = |+\rangle$ with $\alpha = \pi/2$, $\beta = 0$. To see that the general solution, eq. (511), sweeps out a cone whose axis lies in the x-z plane at 45° to the x and z axes, it is helpful to write down the rectangular coordinates of the vector $|\phi\rangle$ on the unit Bloch sphere,

$$x = \sin\alpha\cos\beta = 1 - \cos\alpha, \tag{512}$$

$$y = \sin\alpha\sin\beta = \sqrt{2(\cos\alpha - \cos^2\alpha)}, \tag{513}$$

$$z = \cos\alpha. \tag{514}$$

If we rotate the coordinates by 45° about the y axis into the (x', y', z') system, we have

$$x' = \frac{x - z}{\sqrt{2}} = \frac{1 - 2\cos\alpha}{\sqrt{2}}, \tag{515}$$

$$y' = y = \sqrt{2(\cos\alpha - \cos^2\alpha)}, \tag{516}$$

$$z' = \frac{x + z}{\sqrt{2}} = \frac{1}{\sqrt{2}}. \tag{517}$$

Thus, the solution to eq. (511) in the (x', y', z') coordinate system is the intersection of the unit sphere with the plane $z' = 1/\sqrt{2}$, which is a circle of radius $r = \sqrt{x'^2 + y'^2} = 1/\sqrt{2}$. That is, the solution-states $|\psi\rangle$ sweep out a cone of half angle 45° about the z' axis.

5. **Measurements**

(a) Expressing the operator $\hat{\mathbf{v}} \cdot \boldsymbol{\sigma}$ as a matrix, we have

$$\hat{\mathbf{v}} \cdot \boldsymbol{\sigma} = \begin{pmatrix} v_z & v_x - iv_y \\ v_x + iv_y & -v_z \end{pmatrix}. \tag{518}$$

The eigenvalues λ satisfy the equation

$$\hat{\mathbf{v}} \cdot \boldsymbol{\sigma}|\psi\rangle = \lambda|\psi\rangle, \tag{519}$$

which implies that

$$0 = \Delta(\hat{\mathbf{v}} \cdot \boldsymbol{\sigma} - \lambda \mathbf{I}) = \begin{vmatrix} v_z - \lambda & v_x - iv_y \\ v_x + iv_y & -v_z - \lambda \end{vmatrix} = \lambda^2 - v_z^2 - v_x^2 - v_y^2 = \lambda^2 - 1, \tag{520}$$

so the eigenvalues are

$$\lambda = \pm 1. \tag{521}$$

To find the eigenvectors, we recall the relations between the Pauli spin matrices and rotations. Thus, we expect that the eigenvectors of operator $\hat{\mathbf{v}} \cdot \boldsymbol{\sigma}$ are the $|0\rangle$ and $|1\rangle$ states that correspond to the direction of the unit vector $\hat{\mathbf{v}}$. Recalling eq. (42), we anticipate that one of the eigenvectors is the state

$$|+\rangle = \cos(\alpha/2)|0\rangle + e^{i\beta}\sin(\alpha/2)|1\rangle, \tag{522}$$

where the angles α and β are given by

$$v_x = \sin\alpha\cos\beta, \qquad v_y = \sin\alpha\sin\beta, \qquad v_z = \cos\alpha. \tag{523}$$

We verify this, expressing the matrix (518) in terms of α and β,

$$\begin{aligned} \hat{\mathbf{v}} \cdot \boldsymbol{\sigma}|+\rangle &= \begin{pmatrix} \cos\alpha & e^{-i\beta}\sin\alpha \\ e^{i\beta}\sin\alpha & -\cos\alpha \end{pmatrix} \begin{pmatrix} \cos(\alpha/2) \\ e^{i\beta}\sin(\alpha/2) \end{pmatrix} \\ &= \begin{pmatrix} \cos\alpha\cos(\alpha/2) + \sin\alpha\sin(\alpha/2) \\ e^{i\beta}\sin\alpha\cos(\alpha/2) - e^{i\beta}\cos\alpha\sin(\alpha/2) \end{pmatrix} \\ &= \begin{pmatrix} \cos(\alpha - \alpha/2) \\ e^{i\beta}\sin(\alpha - \alpha/2) \end{pmatrix} = \begin{pmatrix} \cos(\alpha/2) \\ e^{i\beta}\sin(\alpha/2) \end{pmatrix} = +1 \cdot |+\rangle. \end{aligned} \tag{524}$$

The other eigenstate, $|-\rangle$, is orthogonal to $|+\rangle$. From eq. (522) we infer that

$$|-\rangle = e^{-i\beta}\sin(\alpha/2)|0\rangle - \cos(\alpha/2)|1\rangle, \tag{525}$$

as is readily verified:

$$\begin{aligned}
\hat{\mathbf{v}}\cdot\boldsymbol{\sigma}|-\rangle &= \begin{pmatrix} \cos\alpha & e^{-i\beta}\sin\alpha \\ e^{i\beta}\sin\alpha & -\cos\alpha \end{pmatrix}\begin{pmatrix} e^{-i\beta}\sin(\alpha/2) \\ -\cos(\alpha/2) \end{pmatrix} \\
&= \begin{pmatrix} e^{-i\beta}\cos\alpha\sin(\alpha/2) - e^{-i\beta}\sin\alpha\cos(\alpha/2) \\ \sin\alpha\sin(\alpha/2) + \cos\alpha\cos(\alpha/2) \end{pmatrix} \\
&= \begin{pmatrix} -e^{-i\beta}\sin(\alpha-\alpha/2) \\ \cos(\alpha-\alpha/2) \end{pmatrix} = \begin{pmatrix} -e^{-i\beta}\sin(\alpha/2) \\ -[-\cos(\alpha/2)] \end{pmatrix} = -1\cdot|-\rangle. \qquad (526)
\end{aligned}$$

The projection operator for the $|+\rangle$ eigenstate is

$$\begin{aligned}
\mathsf{P}_+ &= |+\rangle\langle+| \\
&= [\cos(\alpha/2)|0\rangle + e^{i\beta}\sin(\alpha/2)|1\rangle][\cos(\alpha/2)\langle 0| + e^{-i\beta}\sin(\alpha/2)\langle 1|] \\
&= \cos^2(\alpha/2)|0\rangle\langle 0| + e^{-i\beta}\sin(\alpha/2)\cos(\alpha/2)|0\rangle|1\rangle] \\
&\quad + e^{i\beta}\sin(\alpha/2)\cos(\alpha/2)|1\rangle|0\rangle] + \sin^2(\alpha/2)|1\rangle\langle 1| \\
&= \begin{pmatrix} \cos^2(\alpha/2) & e^{-i\beta}\sin(\alpha/2)\cos(\alpha/2) \\ e^{i\beta}\sin(\alpha/2)\cos(\alpha/2) & \sin^2(\alpha/2) \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} 1+\cos\alpha & e^{-i\beta}\sin\alpha \\ e^{i\beta}\sin\alpha & 1-\cos\alpha \end{pmatrix} = \frac{\mathbf{I}+\hat{\mathbf{v}}\cdot\boldsymbol{\sigma}}{2}\,. \qquad (527)
\end{aligned}$$

Similarly,

$$\begin{aligned}
\mathsf{P}_- &= |-\rangle\langle-| = \begin{pmatrix} e^{-i\beta}\sin(\alpha/2) \\ -\cos(\alpha/2) \end{pmatrix}\begin{pmatrix} e^{i\beta}\sin(\alpha/2) & -\cos(\alpha/2) \end{pmatrix} \\
&= \begin{pmatrix} \sin^2(\alpha/2) & -e^{-i\beta}\sin(\alpha/2)\cos(\alpha/2) \\ -e^{i\beta}\sin(\alpha/2)\cos(\alpha/2) & \cos^2(\alpha/2) \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} 1-\cos\alpha & -e^{-i\beta}\sin\alpha \\ -e^{i\beta}\sin\alpha & 1+\cos\alpha \end{pmatrix} = \frac{\mathbf{I}-\hat{\mathbf{v}}\cdot\boldsymbol{\sigma}}{2}\,. \qquad (528)
\end{aligned}$$

The probability that a state $|\psi\rangle$ will be found in the $|+\rangle$ state is

$$P_+ = \langle\psi|\mathsf{P}_+|\psi\rangle. \qquad (529)$$

In particular, for the initial state $|0\rangle$

$$P_+ = \langle 0|\mathsf{P}_+|0\rangle = \mathsf{P}_{+,00} = \frac{1+\cos\alpha}{2} = \frac{1+v_z}{2}\,. \qquad (530)$$

If the initial state $|\psi\rangle$ is measured to be in the state $|+\rangle$, the final state is, of course, the state $|+\rangle$ described by eq. (522).

(b) The measurement operator to determine the value of the second bit of state $|\psi\rangle$ can be written as

$$\mathsf{M}_2 = 0 \cdot \mathsf{P}_{0,2} + 1 \cdot \mathsf{P}_{1,2} = 0 \cdot |0\rangle_2\langle 0|_2 + 1 \cdot |1\rangle_2\langle 1|_2. \tag{531}$$

Applying this to the state

$$|\psi\rangle = \psi_{00}|0\rangle|0\rangle + \psi_{01}|0\rangle|1\rangle + \psi_{10}|1\rangle|0\rangle + \psi_{11}|1\rangle|1\rangle, \tag{532}$$

we obtain

$$\mathsf{M}_2|\psi\rangle = 0 \cdot (\psi_{00}|0\rangle|0\rangle + \psi_{10}|1\rangle|0\rangle) + 1 \cdot (\psi_{01}|0\rangle|1\rangle + \psi_{11}|1\rangle|1\rangle). \tag{533}$$

This means that the probability that the second bit is found to be $|1\rangle$ is

$$P_2(1) = \langle\psi|\mathsf{P}^\dagger_{1,2}\mathsf{P}_{1,2}|\psi\rangle = |\psi_{01}|^2 + |\psi_{11}|^2, \tag{534}$$

and that if this is the result, the state after the measurement is

$$|\psi'\rangle = \frac{\mathsf{P}_{1,2}|\psi\rangle}{\sqrt{\langle\psi|\mathsf{P}^\dagger_{1,2}\mathsf{P}_{1,2}|\psi\rangle}} = \frac{\psi_{01}|0\rangle|1\rangle + \psi_{11}|1\rangle|1\rangle}{\sqrt{|\psi_{01}|^2 + |\psi_{11}|^2}}. \tag{535}$$

If the measurement does not yield the value 1 for the second bit, it yields the value 0. The probability of this outcome is

$$P_2(0) = \langle\psi|\mathsf{P}^\dagger_{1,2}\mathsf{P}_{1,2}|\psi\rangle = |\psi_{00}|^2 + |\psi_{10}|^2 = 1 - |\psi_{01}|^2 - |\psi_{11}|^2 = 1 - P_2(1), \tag{536}$$

and that if this is the result, the state after the measurement is

$$|\psi'\rangle = \frac{\mathsf{P}_{0,2}|\psi\rangle}{\sqrt{\langle\psi|\mathsf{P}^\dagger_{0,2}\mathsf{P}_{0,2}|\psi\rangle}} = \frac{\psi_{00}|0\rangle|0\rangle + \psi_{10}|1\rangle|0\rangle}{\sqrt{|\psi_{00}|^2 + |\psi_{10}|^2}}. \tag{537}$$

If we wish to determine the value of both bits of state (532), we note that the operators to determine if bits 1 and 2 have values $j = 0$ or 1 and $k = 0$ or 1 are P_{jk} where

$$\begin{aligned}
\mathsf{P}_{00} &= |0\rangle_1\langle 0|_1 \otimes |0\rangle_2\langle 0|_2, && (538)\\
\mathsf{P}_{01} &= |0\rangle_1\langle 0|_1 \otimes |1\rangle_2\langle 1|_2, && (539)\\
\mathsf{P}_{10} &= |1\rangle_1\langle 1|_1 \otimes |0\rangle_2\langle 0|_2, && (540)\\
\mathsf{P}_{11} &= |1\rangle_1\langle 1|_1 \otimes |1\rangle_2\langle 1|_2, && (541)
\end{aligned}$$

The eigenstates of these four operators are $|0\rangle|0\rangle$, $|0\rangle|1\rangle$, $|1\rangle|0\rangle$ and $|1\rangle|1\rangle$, and the corresponding eigenvalues are all unity. However, if we simply add the four operators with unit weights, we obtain the unit matrix. We construct a measurement operator M_{12} that can distinguish the four eigenstates by summing the operators (538)-(541) with weights ranging from binary numbers 00 to 11 according to

$$\mathsf{M}_{12} = 00 \cdot \mathsf{P}_{00} + 01 \cdot \mathsf{P}_{01} + 10 \cdot \mathsf{P}_{10} + 11 \cdot \mathsf{P}_{11}. \tag{542}$$

Applying this to the state (532) we obtain

$$\mathsf{M}_{12}|\psi\rangle = 00\cdot\psi_{00}|0\rangle|1\rangle + 01\cdot\psi_{01}|0\rangle|1\rangle + 10\cdot\psi_{10}|1\rangle|0\rangle + 11\cdot\psi_{11}|1\rangle|1\rangle. \quad (543)$$

We interpret this the with the aid of von Neumann's argument as meaning that the four terms of eq. (543) are correlated with four positions of a pointer. When the pointer is observed, it is found to be in one of the four positions. If the measurement could be repeated with many copies of state $|\psi\rangle$ the probabilities of observing the pointer would be $\left|\psi_{jk}\right|^2$ for the position with binary number jk. If the pointer is found to be at position jk, then we have measured bit 1 to have value j and bit 2 to have value k. The state of $|\psi\rangle$ after this measurement is $|j\rangle|k\rangle$.

(c) **Stern-Gerlach**

The property of the neutral spin-1/2 particle to be measured is the z component of its spin, for which the relevant operator is $\boldsymbol{\sigma}_z$.

The value of the pointer in the Stern-Gerlach apparatus is the angle of the trajectory of the particle in the x-z plane after is leaves the apparatus. This is equivalent to the pointer value being the z component of the particle's momentum. So, we take the "coordinate" to be measured as p_z. The "momentum" that is conjugate to p_z is just the spatial coordinate z.

Therefore, we need a Hamiltonian of the form,

$$\mathsf{h} = \lambda\boldsymbol{\sigma}_z\mathsf{z}, \quad (544)$$

where z is the hermitian operator associated with the observable coordinate z.

Such a Hamiltonian can be realized by passing the particle through a magnetic field $\mathbf{B}$. Then, its magnetic energy, $U = -\boldsymbol{\mu}\cdot\mathbf{B}$, where $\boldsymbol{\mu} = \gamma\mathbf{s} = \hbar\gamma\boldsymbol{\sigma}/2$ is the particle's magnetic moment and $\boldsymbol{\sigma} = (\boldsymbol{\sigma}_x, \boldsymbol{\sigma}_y, \boldsymbol{\sigma}_z)$ is the Pauli-matrix vector. The (reduced) interaction Hamiltonian $\mathsf{h} = \mathcal{H}/\hbar$ of the particle is,

$$\mathsf{h} = -\frac{\gamma}{2}\boldsymbol{\sigma}\cdot\mathbf{B}. \quad (545)$$

We make eq. (545) have the desired form (544) by choosing the magnetic field to be in the z direction and to vary linearly with position,

$$\mathbf{B} = -Az\hat{\mathbf{z}}, \quad (546)$$

$$\mathsf{h} = \frac{\gamma A}{2}\boldsymbol{\sigma}_z\mathsf{z}. \quad (547)$$

(d) **Quantum Nondemolition Measurement**

We desire the form of a unitary 4×4 matrix U that operates on Qbits $|\psi\rangle = a_0|0\rangle + a_1|1\rangle$ and $|0\rangle$ according to $\mathsf{U}|\psi\rangle|0\rangle = a_0|0\rangle|0\rangle + a_1|1\rangle|1\rangle$.

The initial state is

$$|\psi\rangle|0\rangle = a_0|0\rangle|0\rangle + a_1|1\rangle|0\rangle = \begin{pmatrix} a_0 \\ 0 \\ a_1 \\ 0 \end{pmatrix}, \quad (548)$$

and the desired final state is

$$\mathsf{U}|\psi\rangle|0\rangle = \mathsf{U}\begin{pmatrix} a_0 \\ 0 \\ a_1 \\ 0 \end{pmatrix} = \begin{pmatrix} a_0 \\ 0 \\ 0 \\ a_1 \end{pmatrix}. \tag{549}$$

A suitable 4×4 matrix representation of the unitary operator U is therefore

$$\mathsf{U} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & a & 0 & b \\ 0 & c & 0 & d \\ 0 & 0 & 1 & 0 \end{pmatrix}, \tag{550}$$

noting that each row and column of a unitary matrix, considered as a vector, is normalized to 1.

If U is symmetric, then $d = 1$. For the row and column containing d to be normalized, we have that $b = c = 0$, and finally, for the row and column continaing a to be normalized we must have $a = 1$. Thus,

$$\mathsf{U} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{551}$$

The matrix U of eq. (551) will soon become our favorite 4×4 matrix, and has been given the name Controlled-NOT or C-NOT.

6. **Quantum Cloning and Quantum Teleportation**

(a) **Quantum Cloning**

The Controlled-NOT operator C_{xy} does nothing to $|y\rangle$ when $|x=0\rangle$, so $|0\rangle|0\rangle$ goes to $|0\rangle|0\rangle$ and $|0\rangle|1\rangle$ goes to $|0\rangle|1\rangle$, but flips $|y\rangle$ when $|x=1\rangle$, so $|1\rangle|0\rangle$ goes to $|1\rangle|1\rangle$ and $|1\rangle|1\rangle$ goes to $|1\rangle|0\rangle$. These rules determine the matrix elements to be

$$\mathsf{C}_{xy} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \left(\begin{array}{c|c} \mathbf{I} & 0 \\ \hline 0 & \mathsf{X} \end{array}\right), \tag{552}$$

The two-bit state $(a|0\rangle + b|1\rangle)|0\rangle$ can be written as the column vector

$$\mathsf{v} = \begin{pmatrix} a \\ 0 \\ b \\ 0 \end{pmatrix}. \tag{553}$$

Applying transformation (552) to this we obtain

$$\mathsf{C}_{xy}\mathsf{v} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ 0 \\ b \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ 0 \\ 0 \\ b \end{pmatrix}. \tag{554}$$

However, if bit $|y\rangle$ were to be a copy of bit $|x\rangle$ we should have obtained the state

$$a|0\rangle + b|1\rangle)a|0\rangle + b|1\rangle) = \begin{pmatrix} a^2 \\ ab \\ ab \\ b^2 \end{pmatrix}. \tag{555}$$

The two states (554) and (555) are equal only if $ab = 0$, in which case we must have $(a, b) = (1, 0)$ or $(0, 1)$. That is, only cCbits can be copied successfully by the Controlled-NOT operator.

By the definition of the Controlled-NOT operation, we have that $\mathsf{C}_{xy}|0\rangle|0\rangle = |0\rangle|0\rangle$ and $\mathsf{C}_{xy}|1\rangle|1\rangle = |1\rangle|0\rangle$. Thus C_{xy} successfully deletes the second of a pair of identical Cbits. However, if we apply C_{xy} to a pair of identical Qbits, given by eq. (555), we find

$$\mathsf{C}_{xy}\mathsf{v} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a^2 \\ ab \\ ab \\ b^2 \end{pmatrix} = \begin{pmatrix} a^2 \\ ab \\ b^2 \\ ab \end{pmatrix}, \tag{556}$$

whereas we were hoping to delete the second bit which would have brought us back to the state (553). Again, this only works if $ab = 0$, *i.e.*, if the Qbits are actually Cbits.

(b) **Successful Cloning Would Imply Faster Than Light Communication**
Using the relations

$$|0\rangle = \frac{|+\rangle + |-\rangle}{\sqrt{2}} \qquad |1\rangle = \frac{|+\rangle - |-\rangle}{\sqrt{2}} \tag{557}$$

the entangled state (106) can be rewritten in various ways:

$$\begin{aligned}
|\psi\rangle &= \frac{|0\rangle_{\rm A}|0\rangle_{\rm B}}{\sqrt{2}} + \frac{|1\rangle_{\rm A}|1\rangle_{\rm B}}{\sqrt{2}} && (558)\\
&= \frac{(|0\rangle_{\rm A} + |1\rangle_{\rm A})(|0\rangle_{\rm B} + |1\rangle_{\rm B})}{2\sqrt{2}} + \frac{(|0\rangle_{\rm A} - |1\rangle_{\rm A})(|0\rangle_{\rm B} - |1\rangle_{\rm B})}{2\sqrt{2}} \\
&= \frac{|+\rangle_{\rm A}|+\rangle_{\rm B}}{\sqrt{2}} + \frac{|-\rangle_{\rm A}|-\rangle_{\rm B}}{\sqrt{2}} && (559)\\
&= \frac{|0\rangle_{\rm A}|+\rangle_{\rm B}}{2} + \frac{|0\rangle_{\rm A}|-\rangle_{\rm B}}{2} + \frac{|1\rangle_{\rm A}|+\rangle_{\rm B}}{2} - \frac{|1\rangle_{\rm A}|-\rangle_{\rm B}}{2} && (560)\\
&= \frac{|+\rangle_{\rm A}|0\rangle_{\rm B}}{2} + \frac{|+\rangle_{\rm A}|1\rangle_{\rm B}}{2} + \frac{|-\rangle_{\rm A}|0\rangle_{\rm B}}{2} - \frac{|-\rangle_{\rm A}|1\rangle_{\rm B}}{2}\,. && (561)
\end{aligned}$$

Thus, state $|\psi\rangle$ has the same type of entangled structure in both the $[|0\rangle, |1\rangle]$ and the $[|+\rangle, |-\rangle]$ bases, but if different bases are used to describe bits A and B the form of the state is more complicated.

When a measurement is made of one of the bits in the [0,1] basis, the appropriate measurement operator is

$$\mathsf{M}_{[0,1]} = 0 \cdot |0\rangle\langle 0| + 1 \cdot |1\rangle\langle 1|, \tag{562}$$

while when a measurement is made of one of the bits in the $[+,-]$ basis, the appropriate measurement operator is

$$\mathsf{M}_{[+,-]} = + \cdot |+\rangle\langle +| + - \cdot |-\rangle\langle -|, \tag{563}$$

If, as appears to be the case for quantum mechanics, Alice and Bob can make only a single measurement on bits A and B, respectively, per each physical example of state $|\psi\rangle$, the results of (a repeated set of) such measurements are

(a) Bob chooses to measure in the $[|0\rangle, |1\rangle]$ basis.
   i. Alice chooses to measure in the $[|0\rangle, |1\rangle]$ basis.
   We use eq. (558) to describe the state $|\psi\rangle$, and the appropriate measurement operator is

$$\begin{aligned}
\mathsf{M}_{[0,1]_A}\mathsf{M}_{[0,1]_B} &= (0_A \cdot |0\rangle_A\langle 0|_A + 1_A \cdot |1\rangle_A\langle 1|_A) \otimes (0_B \cdot |0\rangle_B\langle 0|_B + 1_B \cdot |1\rangle_B\langle 1|_B) \\
&= 0_A 0_B \cdot |0\rangle_A\langle 0|_A|0\rangle_B\langle 0|_B + 0_A 1_B \cdot |0\rangle_A\langle 0|_A|1\rangle_B\langle 1|_B \\
&\quad + 1_A 0_B \cdot |1\rangle_A\langle 1|_A|0\rangle_B\langle 0|_B + 1_A 1_B \cdot |1\rangle_A\langle 1|_A|1\rangle_B\langle 1|_B. \qquad (564)
\end{aligned}$$

Formally, the measurement yields

$$\mathsf{M}_{[0,1]_A}\mathsf{M}_{[0,1]_B}\frac{|0\rangle\mathrm{A}|0\rangle_\mathrm{B}+|1\rangle_\mathrm{A}|1\rangle_\mathrm{B}}{\sqrt{2}}=\frac{0_A0_B}{\sqrt{2}}\cdot|0\rangle_A|0\rangle_B+\frac{1_A1_B}{\sqrt{2}}\cdot|1\rangle_A|1\rangle_B. \tag{565}$$

This means that

A. Alice observes bit A to be in the $|0\rangle$ state (with 50% probability). Then, Bob observes bit B to be in the $|0\rangle$ state (with 50% probability).

B. Alice observes bit A to be in the $|1\rangle$ state (with 50% probability). Then, Bob observes bit B to be in the $|1\rangle$ state (with 50% probability).

Thus, Bob observes bit B to be $|0\rangle$ or $|1\rangle$ with 50% probability each, which is the same result that he would obtain if no measurement of bit A were made by Alice.

ii. Alice chooses to measure in the $[|+\rangle,|-\rangle]$ basis.
We use eq. (561) to describe the state $|\psi\rangle$, and the appropriate measurement operator is

$$\begin{aligned}\mathsf{M}_{[+,-]_A}\mathsf{M}_{[0,1]_B} &= (+_A\cdot|+\rangle_A\langle+|_A+-_A\cdot|-\rangle_A\langle-|_A)\otimes(0_B\cdot|0\rangle_B\langle0|_B+1_B\cdot|1\rangle_B\langle1|_B)\\ &= +_A0_B\cdot|+\rangle_A\langle+|_A|0\rangle_B\langle0|_B++_A1_B\cdot|+\rangle_A\langle+|_A|1\rangle_B\langle1|_B\\ &\quad+-_A\,0_B\cdot|-\rangle_A\langle-|_A|0\rangle_B\langle0|_B+-_A1_B\cdot|-\rangle_A\langle-|_A|1\rangle_B\langle1|_B.\end{aligned} \tag{566}$$

Formally, the measurement yields

$$\begin{aligned}&\mathsf{M}_{[+,-]_A}\mathsf{M}_{[0,1]_B}\frac{|+\rangle_\mathrm{A}|0\rangle_\mathrm{B}+|+\rangle_\mathrm{A}|1\rangle_\mathrm{B}+|-\rangle_\mathrm{A}|0\rangle_\mathrm{B}-|-\rangle_\mathrm{A}|1\rangle_\mathrm{B}}{2}\\ &\quad=\frac{+_A0_B}{2}\cdot|+\rangle_A|0\rangle_B+\frac{+_A1_B}{2}\cdot|+\rangle_A|1\rangle_B\\ &\qquad+\frac{-_A0_B}{2}\cdot|-\rangle_A|0\rangle_B+\frac{-_A1_B}{2}\cdot|-\rangle_A|1\rangle_B.\end{aligned} \tag{567}$$

This means that

A. Alice observes bit A to be in the $|+\rangle$ state (with 50% probability). Then, Bob observes bit B to be in the $|0\rangle$ state (with 25% probability), or in the $|-\rangle$ state with (with 25% probability).

B. Alice observes bit A to be in the $|-\rangle$ state (with 50% probability). Then, Bob observes bit B to be in the $|0\rangle$ state (with 25% probability), or in the $|-\rangle$ state (with 25% probability).

Again, Bob observes bit B to be $|0\rangle$ or $|1\rangle$ with 50% probability each, which is the same result that he would obtain if no measurement of bit A were made by Alice.

Indeed, the results of Bob's measurements of bit B in the [0,1] basis (in the absence of knowledge of Alice's actions) are the same whether Alice measures bit A in the [0,1] basis, or in the $[+,-]$ basis, or if Alice makes no measurement at all.

(b) Bob chooses to measure in the $[|+\rangle, |-\rangle]$ basis.

i. Alice chooses to measure in the $[|0\rangle, |1\rangle]$ basis.

A. Alice observes bit A to be in the $|0\rangle$ state (with 50% probability). Then, Bob observes bit B to be in the $|+\rangle$ state (with 25% probability), or in the $|-\rangle$ state with (with 25% probability).

B. Alice observes bit A to be in the $|1\rangle$ state (with 50% probability). Then, Bob observes bit B to be in the $|+\rangle$ state (with 25% probability), or in the $|-\rangle$ state (with 25% probability).

ii. Alice chooses to measure in the $[|+\rangle, |-\rangle]$ basis.

A. Alice observes bit A to be in the $|+\rangle$ state (with 50% probability). Then, Bob observes bit B to be in the $|+\rangle$ state (with 50% probability).

B. Alice observes bit A to be in the $|-\rangle$ state (with 50% probability). Then, Bob observes bit B to be in the $|-\rangle$ state (with 50% probability).

By himself, Bob observes bit B to be $|0\rangle$ or $|1\rangle$ with equal probability if he measures in the $[|0\rangle, |1\rangle]$ basis, or that bit B is $|+\rangle$ or $|-\rangle$ with equal probability if he measures in the $[|+\rangle, |-\rangle]$ basis. This gives him no clue as to what Alice has done; he doesn't know whether she made measurements in the $[|0\rangle, |1\rangle]$ basis or in the $[|+\rangle, |-\rangle]$ basis, or whether she made any measurements at all.

In my view, Bob has learned nothing at all about bit A from his measurements of bit B.[144] There is no "signal" from one part of an entangled state to the other.

Suppose, however, that Bob could make lots of copies of bit B, each having the entanglement with bit A given in eq. (558)-(561). Then he could measure half of the copies in the basis $[|0\rangle, |1\rangle]$ and the other half in the basis $[|+\rangle, |-\rangle]$. If he observes the various copies of bit B to be $|0\rangle$, $|1\rangle$, $|+\rangle$ and $|-\rangle$ with equal probability, he can conclude with good assurance that Alice made no measurement of bit A. But if he found no copies of bit B to be one of those four states (*e.g.*, no $|+\rangle$), while half of the copies to be its basis partner (*e.g.*, 50% $|-\rangle$), then he could conclude that Alice had made a measurement in that basis (*e.g.*, $[|+\rangle, |-\rangle]$), and that her result was that bit A is the same state as he found 50% of the time (*e.g.*, $|-\rangle$). He could interpret these results as a signal from Alice as to her choice of basis and of her result of a measurement, despite their measurements being spacelike-separated. This would imply faster-than-light communication!

So it is happily consistent with our trust in the theory of relativity that the needed cloning of the entangled state (106) cannot be done.[145]

[144] However, starting with Einstein, another kind of comment has been made. Namely, that if Bob measures bit B to be, say $|0\rangle$, then he can "predict with certainty" that IF Alice measured bit A in the $[|0\rangle, |1\rangle]$ basis, then she would find bit A to be $|0\rangle$ also. Similarly, Bob can "predict with certainty" that if he measures bit B to be $|+\rangle$, then IF Alice measured bit A in the $[|+\rangle, |-\rangle]$ basis she would find bit A to be $|+\rangle$ also. Such conditional predictions are obviously unsatisfactory because they do not contain useful knowledge about bit A. Nonetheless, in a very loose usage of words they imply "certain knowledge" simultaneously about spacelike-separated quantum states in two different bases, which led Einstein to remark that they implied some kind of "spooky" action at a distance, which in turn suggested to him that quantum theory was somehow "incomplete". I subscribe to the camp of Bohr that this logic does not warrant such conclusions. http://kirkmcd.princeton.edu/examples/QM/bohr_pr_48_696_35.pdf

[145] We don't need the full no-cloning theorem to conclude that we cannot make an "exact" copy of bit B if

Suppose Bob makes a "copy" of bit B using a Controlled-NOT gate with its second input, bit C, initially set to $|0\rangle$. He then measures bit B in the [0,1] basis and bit C in the $[+,-]$ basis. The 3-bit state $|\psi\rangle$ after the "copy" is made, but before the measurement, is

$$|\psi\rangle = \frac{|0\rangle_A|0\rangle_B|0\rangle_C}{\sqrt{2}} + \frac{|1\rangle_A|1\rangle_B|1\rangle_C}{\sqrt{2}} \tag{568}$$

$$= \frac{|0\rangle_A|0\rangle_B|+\rangle_C}{2} + \frac{|0\rangle_A|0\rangle_B|-\rangle_C}{2} + \frac{|1\rangle_A|1\rangle_B|+\rangle_C}{2} - \frac{|1\rangle_A|1\rangle_B|-\rangle_C}{2} \tag{569}$$

$$= \frac{|+\rangle_A|0\rangle_B|+\rangle_C}{2\sqrt{2}} + \frac{|-\rangle_A|0\rangle_B|+\rangle_C}{2\sqrt{2}} + \frac{|+\rangle_A|0\rangle_B|-\rangle_C}{2\sqrt{2}} + \frac{|-\rangle_A|0\rangle_B|-\rangle_C}{2\sqrt{2}}$$
$$+\frac{|+\rangle_A|1\rangle_B|+\rangle_C}{2\sqrt{2}} - \frac{|-\rangle_A|1\rangle_B|+\rangle_C}{2\sqrt{2}} - \frac{|+\rangle_A|1\rangle_B|-\rangle_C}{2\sqrt{2}} + \frac{|-\rangle_A|1\rangle_B|-\rangle_C}{2\sqrt{2}}. \tag{570}$$

If Alice makes no measurement of bit A, we read off from eq. (569) that Bob will find bits B and C in the four combinations $0_B{+}_C$, $0_B{-}_C$, $1_B{+}_C$, and $1_B{-}_C$ each with 25% probability.

Similarly, if Alice measures bit A in the [0,1] basis, she finds bit A to be 0 or 1 each with 50% probability, but Bob's measurements of bits B and C (in the absence of knowledge as to Alice's results) again yield the four combinations $0_B{+}_C$, $0_B{-}_C$, $1_B{+}_C$, and $1_B{-}_C$ each with 25% probability.

And if Alice measures bit A in the $[+,-]$ basis, she finds bit A to be $+$ or $-$ each with 50% probability, but Bob's measurements of bits B and C (in the absence of knowledge as to Alice's results) again yield the four combinations $0_B{+}_C$, $0_B{-}_C$, $1_B{+}_C$, and $1_B{-}_C$ each with 25% probability.

It appears that Bob has not increased his knowledge about bit A by making the Controlled-NOT "copy" of bit B.

One might argue that the state (568) actually includes in bit C as good a copy of bit B as possible (even if we had never heard of the no-cloning theorem). But, proper re-expression of this state in the appropriate bases for measurements by Bob (and Alice) shows that the observations claimed on the previous page could never occur.

(c) **Swap Two Bits**

The truth table of the two-bit SWAP operation is

S_{ab} :

a	b	$a' = b$	$b' = a$
0	0	0	0
0	1	1	0
1	0	0	1
1	1	1	1

(571)

it is entangled with bit A but we have no knowledge of bit A. An exact copy of part of a system could only be made without knowledge of the rest of the system if that system could be described as a direct product of the part with the rest of the system ($|\text{system}\rangle = |\text{part}\rangle|\text{rest}\rangle$). Hence, there can be no exact copying of part of an entangled system.

Hence, the 4×4 matrix for the SWAP operator is

$$\mathsf{S}_{ab} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{572}$$

This matrix is closely related to the matrix (552) for the Controlled-NOT operator C_{ab}.

The hint is to consider the Controlled-NOT operator C_{ba}, whose truth table is

$$\mathsf{C}_{ba}: \quad \begin{array}{cc|cc} a & b & a' & b' \\ \hline 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array} \tag{573}$$

Hence, the 4×4 matrix for the C_{ba} operator is

$$\mathsf{C}_{ba} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{574}$$

If we apply the two Controlled-NOT operators in sequence, we obtain the matrix

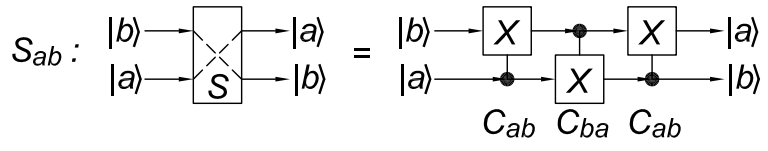
$$\mathsf{C}_{ba}\mathsf{C}_{ab} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{575}$$

This is not yet eq. (572), but we persevere and multiply eq. (575) by C_{ab},

$$\mathsf{C}_{ab}\mathsf{C}_{ba}\mathsf{C}_{ab} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \mathsf{S}_{ab}. \tag{576}$$

Likewise, $\mathsf{C}_{ba}\mathsf{C}_{ab}\mathsf{C}_{ba} = \mathsf{S}_{ab}$.

A diagram for a SWAP circuit is



As a formality, we verify the effect of S_{ab} on Qbits $|a\rangle = \alpha|0\rangle + \beta|1\rangle$ and $|b\rangle = \gamma|0\rangle + \delta|1\rangle$. Then,

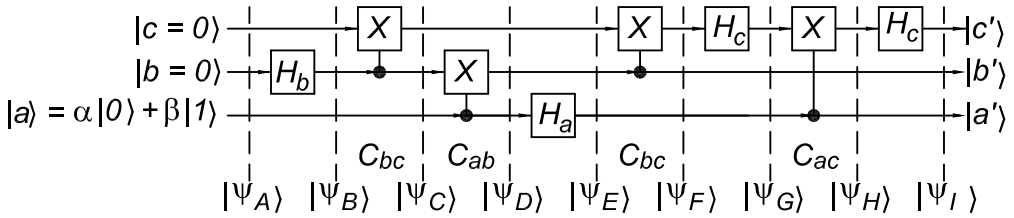
$$|ab\rangle = (\alpha|0\rangle + \beta|1\rangle)(\gamma|0\rangle + \delta|1\rangle) = \alpha\gamma|00\rangle + \alpha\delta|01\rangle + \beta\gamma|10\rangle + \beta\delta|11\rangle, \quad (577)$$

so that

$$\begin{aligned} \mathsf{S}_{ab}|ab\rangle &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha\gamma \\ \alpha\delta \\ \beta\gamma \\ \beta\delta \end{pmatrix} = \begin{pmatrix} \alpha\gamma \\ \beta\gamma \\ \alpha\delta \\ \beta\delta \end{pmatrix} \\ &= \alpha\gamma|00\rangle + \beta\gamma|01\rangle + \alpha\delta|10\rangle + \beta\delta|11\rangle \\ &= (\gamma|0\rangle + \delta|1\rangle)(\alpha|0\rangle + \beta|1\rangle) = |ba\rangle. \end{aligned} \quad (578)$$

(d) **Quantum Teleportation**

Some care is required to trace the effect of the 3-Qbit processor that contains 8 operations. We label the 3-Qbit state at various stages in the process as $|\psi_A\rangle$-$|\psi_I\rangle$, as shown in the figure below.



The initial state is

$$|\psi_A\rangle = (\alpha|0_a\rangle + \beta|1_a\rangle)|0_b\rangle|0_c\rangle \equiv \alpha|000\rangle + \beta|100\rangle. \quad (579)$$

We recall that the Hadamard transformation has the 2×2 matrix form

$$\mathsf{H} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad (580)$$

so that

$$\mathsf{H}|0\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \qquad \mathsf{H}|1\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \quad (581)$$

Then,

$$\begin{aligned} |\psi_A\rangle &= \alpha|000\rangle + \beta|100\rangle, \\ |\psi_B\rangle = \mathsf{H}_b|\psi_A\rangle &= \alpha\frac{|000\rangle + |010\rangle}{\sqrt{2}} + \beta\frac{|100\rangle + |110\rangle}{\sqrt{2}}, \\ |\psi_C\rangle = \mathsf{C}_{bc}|\psi_B\rangle &= \alpha\frac{|000\rangle + |011\rangle}{\sqrt{2}} + \beta\frac{|100\rangle + |111\rangle}{\sqrt{2}}, \\ |\psi_D\rangle = \mathsf{C}_{ab}|\psi_C\rangle &= \alpha\frac{|000\rangle + |011\rangle}{\sqrt{2}} + \beta\frac{|110\rangle + |101\rangle}{\sqrt{2}}, \end{aligned}$$

$$
\begin{aligned}
|\psi_E\rangle = \mathsf{H}_a|\psi_D\rangle &= \alpha\frac{|000\rangle+|100\rangle+|011\rangle+|111\rangle}{2}+\beta\frac{|010\rangle-|110\rangle+|001\rangle-|101\rangle}{2},\\
|\psi_F\rangle = \mathsf{C}_{bc}|\psi_E\rangle &= \alpha\frac{|000\rangle+|100\rangle+|010\rangle+|110\rangle}{2}+\beta\frac{|011\rangle-|111\rangle+|001\rangle-|101\rangle}{2},\\
|\psi_G\rangle = \mathsf{H}_c|\psi_F\rangle &= \alpha\frac{|000\rangle+|001\rangle+|100\rangle+|101\rangle+|010\rangle+|011\rangle+|110\rangle+|111\rangle}{2\sqrt{2}}\\
&\quad+\beta\frac{|010\rangle-|011\rangle-|110\rangle+|111\rangle+|000\rangle-|001\rangle-|100\rangle+|101\rangle}{2\sqrt{2}},\\
|\psi_H\rangle = \mathsf{C}_{ac}|\psi_G\rangle &= \alpha\frac{|000\rangle+|001\rangle+|101\rangle+|100\rangle+|010\rangle+|011\rangle+|111\rangle+|110\rangle}{2\sqrt{2}}\\
&\quad+\beta\frac{|010\rangle-|011\rangle-|111\rangle+|110\rangle+|000\rangle-|001\rangle-|101\rangle+|100\rangle}{2\sqrt{2}},\\
|\psi_I\rangle = \mathsf{H}_c|\psi_H\rangle &= \alpha\frac{|000\rangle+|001\rangle+|000\rangle-|001\rangle+|100\rangle-|101\rangle+|100\rangle+|101\rangle}{4}\\
&\quad+\alpha\frac{|010\rangle+|011\rangle+|010\rangle-|011\rangle+|110\rangle-|111\rangle+|110\rangle+|111\rangle}{4}\\
&\quad+\beta\frac{|010\rangle+|011\rangle-|010\rangle+|011\rangle-|110\rangle+|111\rangle+|110\rangle+|111\rangle}{4}\\
&\quad+\beta\frac{|000\rangle+|001\rangle-|000\rangle+|001\rangle-|100\rangle+|101\rangle+|100\rangle+|101\rangle}{4}\\
&= \alpha\frac{|000\rangle+|100\rangle+|010\rangle+|110\rangle}{2}+\beta\frac{|011\rangle+|111\rangle+|001\rangle+|101\rangle}{2}\\
&= \alpha\frac{|0\rangle+|1\rangle}{\sqrt{2}}\frac{|0\rangle+|1\rangle}{\sqrt{2}}|0\rangle+\beta\frac{|0\rangle+|1\rangle}{\sqrt{2}}\frac{|0\rangle+|1\rangle}{\sqrt{2}}|1\rangle\\
&= \frac{|0_a\rangle+|1_a\rangle}{\sqrt{2}}\frac{|0_b\rangle+|1_b\rangle}{\sqrt{2}}(\alpha|0_c\rangle+\beta|1_c\rangle). \qquad (582)
\end{aligned}
$$

Despite the entangling effects of the Hadamard operations, the final state $|\psi_I\rangle = |a'\rangle|b'\rangle|c'\rangle$ is a direct product state in which $|c'\rangle$ is the same as the initial state $|a\rangle$. The final states $|a'\rangle$ and $|b'\rangle$ could, if desired, be brought to $|0\rangle$ states with additional operations H_a and H_b.

Turning to the version of the above process in which Alice and Bob split things up, we see that Alice measures bits a and b when the system has state $|\psi_E\rangle$. There are 4 possible outcomes of her measurements, resulting in one of four states,

$$
\begin{aligned}
|\psi_{E_{00}}\rangle &= |00\rangle(\alpha|0\rangle+\beta|1\rangle), \qquad (583)\\
|\psi_{E_{01}}\rangle &= |01\rangle(\alpha|1\rangle+\beta|0\rangle), \qquad (584)\\
|\psi_{E_{10}}\rangle &= |10\rangle(\alpha|0\rangle-\beta|1\rangle), \qquad (585)\\
|\psi_{E_{11}}\rangle &= |11\rangle(\alpha|1\rangle-\beta|0\rangle). \qquad (586)
\end{aligned}
$$

We see that if Alice's measurements of bits a and b are $|00\rangle$, Bob doesn't actually have to do anything to bit c; it already is in the unknown state $\alpha|0\rangle+\beta|1\rangle$. This is perhaps the source of the description teleportation. It seems as if Alice's measurement transferred information from bit a to bit c without any obvious interaction

between them, and over the possibly large distance between the location of these bits. Since we can't identify a classical mechanism for this, we might declare this transfer to be "instantaneous", and hence a kind of "teleportation".

But, Bob can't really know that his bit c has taken on the initial state of bit a until the classical information about the measurements of bits a and c has arrived. So in practice, there is no faster-than-light transfer of knowledge. However, this quantum process remains somewhat surprising, even mysterious from a classical point of view.

For completeness, we should verify that the 4 transformations of Bob, C_{bc}, H_c, C_{ac} and finally H_c again, always result in bit c ending up as $\alpha|0\rangle + \beta|1\rangle$. Perhaps it suffices to do this for only the cases where the measurements of bits a and b are $|00\rangle$ and $|11\rangle$.

$$\begin{aligned}
|\psi_{F_{00}}\rangle = \mathsf{C}_{bc}|\psi_{E_{00}}\rangle &= |00\rangle(\alpha|0\rangle + \beta|1\rangle), \\
|\psi_{G_{00}}\rangle = \mathsf{H}_c|\psi_{F_{00}}\rangle &= |00\rangle\left(\alpha\frac{|0\rangle+|1\rangle}{\sqrt{2}} + \beta\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right), \\
|\psi_{H_{00}}\rangle = \mathsf{C}_{ac}|\psi_{G_{00}}\rangle &= |00\rangle\left(\alpha\frac{|0\rangle+|1\rangle}{\sqrt{2}} + \beta\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right), \\
|\psi_{I_{00}}\rangle = \mathsf{H}_c|\psi_{H_{00}}\rangle &= |00\rangle\left(\alpha\frac{|0\rangle+|1\rangle+|0\rangle-|1\rangle}{2} + \beta\frac{|0\rangle+|1\rangle-|0\rangle+|1\rangle}{\sqrt{2}}\right), \\
&= |00\rangle(\alpha|0\rangle + \beta|1\rangle). \qquad (587)
\end{aligned}$$

Similarly,

$$\begin{aligned}
|\psi_{F_{11}}\rangle = \mathsf{C}_{bc}|\psi_{E_{11}}\rangle &= |11\rangle(\alpha|0\rangle - \beta|1\rangle), \\
|\psi_{G_{11}}\rangle = \mathsf{H}_c|\psi_{F_{11}}\rangle &= |11\rangle\left(\alpha\frac{|0\rangle+|1\rangle}{\sqrt{2}} - \beta\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right), \\
|\psi_{H_{11}}\rangle = \mathsf{C}_{ac}|\psi_{G_{11}}\rangle &= |11\rangle\left(\alpha\frac{|0\rangle+|1\rangle}{\sqrt{2}} + \beta\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right), \\
|\psi_{I_{11}}\rangle = \mathsf{H}_c|\psi_{H_{11}}\rangle &= |11\rangle\left(\alpha\frac{-|0\rangle-|1\rangle+|0\rangle-|1\rangle}{2} + \beta\frac{|0\rangle+|1\rangle-|0\rangle+|1\rangle}{\sqrt{2}}\right), \\
&= |11\rangle(\alpha|0\rangle + \beta|1\rangle). \qquad (588)
\end{aligned}$$

The result that it makes no difference to the outcome of the final 4 transformations whether bits a and b are measured at the beginning or the end arises because those bits only serve as control bits in Controlled-NOT operations. Indeed, for any Controlled-U gate (where U is any 1-Qbit unitary transformation), if the control Qbit is to be measured in the [0,1] basis then the measurement may be performed either before or after the gate is executed, without making any difference to the final outcome.[146]

From the last line of eq. (582), we see that if Alice did not measure bits $|a\rangle$ and $\mathbf{b}$, and Bob performed Hadamard transformation H_a on the state $|\psi_I\rangle$, the result

[146] http://kirkmcd.princeton.edu/examples/QM/griffiths_prl_76_3228_96.pdf

would be

$$|\psi_J\rangle = \mathsf{H}_a|\psi_I\rangle = |0_a\rangle\frac{|0_b\rangle + |1_b\rangle}{\sqrt{2}}(\alpha|0_c\rangle + \beta|1_c\rangle), \tag{589}$$

since $\mathsf{H}_a|a_I\rangle = \mathsf{H}_a|+_a\rangle = |0_a\rangle$. A lesson is that the Controlled-NOT operations with $|a\rangle$ as the control bit entangle this Qbit with bits $|b\rangle$ and $|c\rangle$, such that the coefficients α and β in the initial state of $|a\rangle$ are no longer an exclusive property of that bit, but are shared with the other bits. Rather, in the final state $|\psi_I\rangle$ these coefficients are now associated with bit $|c'\rangle$, and not with $|a'\rangle$, which latter bit is in the state $|+_a\rangle$ independent of α and β.

The "delocalization" of the properties of a Qbit via entangling it with another Qbit will be exploited in other quantum computations discussed later in this course.

7. Quantum Optics

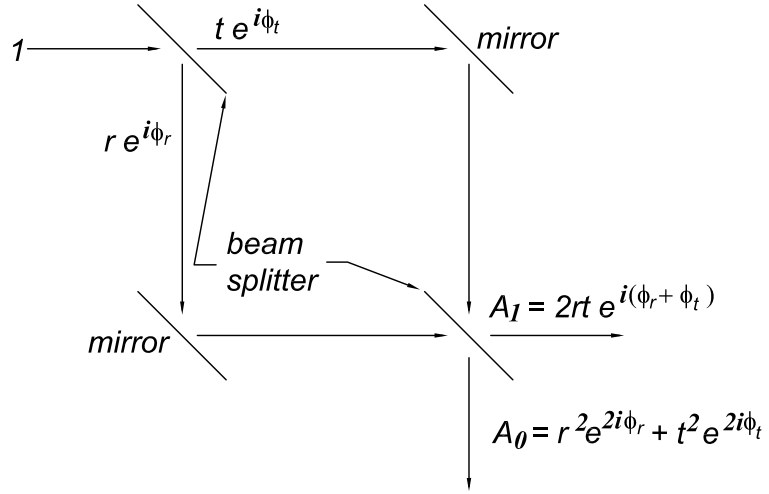
(a) Phase Shift in a Lossless Beam Splitter

Part (a) dates back to work of Airy (who used a different method) in 1833.

A beam of light of unit amplitude is incident on the interferometer from the upper left. The reflected and transmitted amplitudes are $re^{i\phi_r}$ and $te^{i\phi_t}$, where magnitudes r and t are real numbers. The condition of a lossless beam splitter is that

$$r^2 + t^2 = 1. \tag{590}$$

The reflected and transmitted beams are reflected off mirrors and recombined in a second lossless beam splitter, identical to the first.



Then, the amplitude for transmission at the first beam splitter, followed by reflection at the second, is $tre^{i(\phi_t+\phi_r)}$, *etc.* Hence, the recombined beam that moves to the right has amplitude

$$A_1 = 2rte^{i(\phi_r+\phi_t)}, \tag{591}$$

while the recombined beam that moves downwards has amplitude

$$A_0 = r^2e^{2i\phi_r} + t^2e^{2i\phi_t}. \tag{592}$$

The intensity of the output beam 1 is

$$I_1 = |A_1|^2 = 4r^2t^2, \tag{593}$$

and that of the output beam 0 is

$$I_0 = |A_2|^2 = r^4 + t^4 + 2r^2t^2\cos 2(\phi_t - \phi_r). \tag{594}$$

For lossless splitters, the total output intensity must be unity,

$$I_0 + I_1 = 1 = (r^2 + t^2)^2 + 2r^2t^2[1 + \cos 2(\phi_t - \phi_r)]. \tag{595}$$

Recalling eq. (590), we must have

$$\phi_t - \phi_r = \pm 90^\circ, \tag{596}$$

for any value of the splitting ratio $r^2 : t^2$.

Via additional arguments we can show that[147]

$$\phi_t - \phi_r = -90^\circ, \qquad i.e., \qquad \phi_r = \phi_t + \frac{\pi}{2}, \qquad \text{and} \qquad e^{i\phi_r} = ie^{i\phi_t}. \tag{597}$$

We will use this form in parts (c)-(e).

Note that eqs. (594) and (596) imply that the intensity of output beam 0 is

$$I_0 = r^4 + t^4 - 2r^2t^2 = (r^2 - t^2)^2, \tag{598}$$

so that for an interferometer with 50:50 beam splitters, no light would emerge on output path 0.

(b) **Bunching of Photons in a Beam Splitter**

Part (b) appears to have been first considered by Hong *et al.* in 1987.[148] See also.[149]

In a lossless beam splitter the amplitudes for a single photon to be reflected or transmitted have magnitudes $r = \sqrt{R}$ and $t = \sqrt{T}$, where R and T are the probability that a single photon is reflected or transmitted, respectively. For a 50:50 splitter, we have $r = t = 1/\sqrt{2}$. If two photons appear after the beam splitter on the same side (and none on the other), then one of these was reflected and one was transmitted. The amplitude for this is therefore

$$A_{2,0} = A_{0,2} \propto re^{i\phi_r}te^{i\phi_t}. \tag{599}$$

In this case both photons end up in the same final state. Since photons are bosons, the amplitude for having multiple particles in the same final state is enhanced by the square root of the number of ways the particles can be rearranged. Since the "first" output photon could be either the first or the second input photon, the enhancement factor is $\sqrt{2}$. That is,

$$A_{2,0} = A_{0,2} = \sqrt{2}re^{i\phi_r}te^{i\phi_t}. \tag{600}$$

so the probability that two output photons appear on the same side of the splitter is

$$P_{2,0} = P_{0,2} = |A_{2,0}|^2 = 2r^2t^2 = 2RT. \tag{601}$$

For a 50:50 splitter,

$$P_{2,0} = P_{0,2} = \frac{1}{2} \qquad \text{(50:50 splitter)}. \tag{602}$$

One output photon could appear on each side of beam splitter in either of two ways: both input photons are reflected, or both are transmitted. The amplitude for this is

$$A_{1,1} = re^{i\phi_r}re^{i\phi_r} + te^{i\phi_t}te^{i\phi_t}, \tag{603}$$

[147] See, for example, problem 4(b) of
http://kirkmcd.princeton.edu/examples/ph501set6.pdf

[148] http://kirkmcd.princeton.edu/examples/QM/hong_prl_59_2044_87.pdf

[149] http://kirkmcd.princeton.edu/examples/QM/campos_pra_40_1371_89.pdf

so the probability that one output photon appears on each side of the splitter is

$$P_{1,1} = |A_1|^2 = r^4 + t^4 + 2r^2t^2\cos[2(\phi_r - \phi_t)] = r^4 + t^4 - 2r^2t^2 = (R-T)^2, \quad (604)$$

using eq. (596). For a 50:50 splitter, we obtain the possibly surprising result that

$$P_{1,1} = 0 \qquad \text{(50:50 splitter)}; \quad (605)$$

i.e., a quantum analysis predicts that the case of one photon emerging from each side of the splitter does not occur.

Note that in general

$$P_{2,0} + P_{1,1} + P_{0,2} = (R-T)^2 + 2RT = (R+T)^2 = 1, \quad (606)$$

so that the total probability is indeed unity.

In contrast, a classical analysis for input beams of intensities i_1 and i_2 is that the amplitudes of the two (in-phase) input beams are $a_1 = \sqrt{\mathsf{i}_1}$ and $a_2 = \sqrt{\mathsf{i}_2}$. The output beam 1 (in the direction of the input beam 1) is due to transmission of input beam 1 and reflection of input beam 2, and so has amplitude

$$A_1 = a_1 t e^{i\phi_t} + a_2 r e^{i\phi_r}. \quad (607)$$

Similarly, the classical amplitude for output beam 2 is

$$A_2 = a_1 r e^{i\phi_r} + a_2 t e^{i\phi_t}. \quad (608)$$

The intensities of the output beams are therefore

$$I_1 = |A_1|^2 = \mathsf{i}_1 T + \mathsf{i}_2 R + 2\sqrt{\mathsf{i}_1\mathsf{i}_2}rt\cos(\phi_r - \phi_t) = \mathsf{i}_1 T + \mathsf{i}_2 R, \quad (609)$$

and

$$I_2 = |A_2|^2 = \mathsf{i}_1 R + \mathsf{i}_2 T + 2\sqrt{\mathsf{i}_1\mathsf{i}_2}rt\cos(\phi_r - \phi_t) = \mathsf{i}_1 R + \mathsf{i}_2 T. \quad (610)$$

When $R = T = 1/2$, we find that $I_1 = I_2 = (\mathsf{i}_1 + \mathsf{i}_2)/2$. Thus, the classical expectation is that a 50:50 beam splitter just splits the beams 50:50 no matter what direction they come from and what intensities they have.

But the quantum result for the case of only one photon in each input "beam" is that there is zero probability of $(1+1)/2 = 1$ photon in an output beam. Rather, the photons "bunch" into a single output beam on only one side of a lossless 50:50 splitter.[150]

(c) **A Beam Splitter as a Quantum Processor**

Part (c) was perhaps first considered by Yurke *et al.* in 1986.[151]

The transmission coefficient of the lossless beam splitter can be written as $T = 1 - R = \cos^2\frac{\beta}{2}$, since we have $R = \sin^2\frac{\beta}{2}$. The magnitudes of the reflected and transmitted amplitudes are then $r = \sin\frac{\beta}{2}$ and $t = \cos\frac{\beta}{2}$.

[150] The quantum result for input beams of arbitrary numbers of photons is discussed in http://kirkmcd.princeton.edu/examples/bunching.pdf

[151] http://kirkmcd.princeton.edu/examples/QM/yurke_pra_33_4033_86.pdf

We define, without loss of generality, the phase shift of a transmitted photon to be $\phi_t = 0$. Then from eq. (597) we have that the phase shift of a reflected photon is $\phi_r = \pi/2$. Hence, the beam splitter takes a photon $|0\rangle$ into $\cos\frac{\beta}{2}|0'\rangle + e^{i\pi/2}\sin\frac{\beta}{2}|1'\rangle$. Similarly, the beam splitter takes the input state $|1\rangle$ into $(e^{i\pi/2}\sin\frac{\beta}{2}|0'\rangle + \cos\frac{\beta}{2}|1'\rangle$. Therefore, we can write the action of the beam splitter as the 2×2 unitary matrix

$$\mathsf{M} = \begin{pmatrix} \cos\frac{\beta}{2} & e^{i\pi/2}\sin\frac{\beta}{2} \\ e^{i\pi/2}\sin\frac{\beta}{2} & \cos\frac{\beta}{2} \end{pmatrix}. \tag{611}$$

To understand the effect of a dielectric plate on the propagation of a photon, we recall that a photon of wavelength λ that travels in vacuum along the z direction has wave function

$$\psi_{\text{vacuum}} = e^{i(kz-\omega t)}, \tag{612}$$

where $k = 2\pi/\lambda$ is the (free-space) wave number and $\omega = kc$ is the angular frequency of the photon. Of course, c is the speed of light in vacuum. When this photon is inside a dielectric medium its speed is reduced to $v = c/n$, where n is the index of refraction. The angular frequency of the photon is unchanged, so the wave number becomes $k' = \omega/v = \omega n/c = kn$. The wave function of the photon inside the dielectric is therefore

$$\psi_{\text{dielectric}} = e^{i(k'z-\omega t)} = e^{i(knz-\omega t)}. \tag{613}$$

If the dielectric plate extends from $z = 0$ to d, then the wave function of the photon at the exit of the plate is

$$\psi(z=d) = e^{i(knd-\omega t)} = e^{i(kd-\omega t+k(n-1)d)} = e^{i(kd-\omega t+\phi)}, \tag{614}$$

where

$$\phi = k(n-1)d) = \frac{2\pi d(n-1)}{\lambda} \tag{615}$$

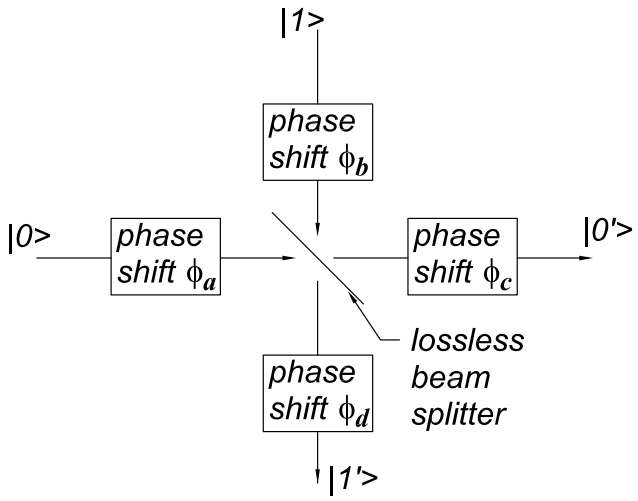
is the phase difference between the photon that has traversed the dielectric plate and a photon that traveled distance d in vacuum.

Hence, the addition of the four wave-shifting plates to the beam splitter, as shown in the figure below, has the effect that

$$\begin{aligned} |0\rangle &\rightarrow \cos\frac{\beta}{2}e^{i(\phi_a+\phi_c)}|0'\rangle + \sin\frac{\beta}{2}e^{i(\phi_a+\phi_d+\pi/2)}|1'\rangle, \\ |1\rangle &\rightarrow \sin\frac{\beta}{2}e^{i(\phi_b+\phi_c+\pi/2)}|0'\rangle + \cos\frac{\beta}{2}e^{i(\phi_b+\phi_d)}|1'\rangle. \end{aligned} \tag{616}$$

Expressing this transformation as a 2×2 unitary matrix U, we have

$$\mathsf{U} = \begin{pmatrix} \cos\frac{\beta}{2}e^{i(\phi_a+\phi_c)} & \sin\frac{\beta}{2}e^{i(\phi_b+\phi_c+\pi/2)} \\ \sin\frac{\beta}{2}e^{i(\phi_a+\phi_d+\pi/2)} & \cos\frac{\beta}{2}e^{i(\phi_b+\phi_d)} \end{pmatrix}. \tag{617}$$



The form (617) is similar to the general form of a 2×2 unitary matrix given by eqs. (44) and (56),

$$\mathsf{U} = e^{i\delta} \begin{pmatrix} \cos\frac{\beta}{2} e^{i(\alpha+\gamma)/2} & \sin\frac{\beta}{2} e^{i(-\alpha+\gamma)/2} \\ -\sin\frac{\beta}{2} e^{i(\alpha-\gamma)/2)} & \cos\frac{\beta}{2} e^{-i(\alpha+\gamma)/2} \end{pmatrix}. \tag{618}$$

Indeed, the two forms (617) and (618) will be identical when

$$\phi_a + \phi_c = \frac{\alpha}{2} + \frac{\gamma}{2} + \delta, \tag{619}$$

$$\phi_b + \phi_c = -\frac{\alpha}{2} + \frac{\gamma}{2} + \delta - \frac{\pi}{2}, \tag{620}$$

$$\phi_a + \phi_d = \frac{\alpha}{2} - \frac{\gamma}{2} + \delta + \frac{\pi}{2}, \tag{621}$$

$$\phi_b + \phi_d = -\frac{\alpha}{2} - \frac{\gamma}{2} + \delta. \tag{622}$$

Since, for example (622) = (620) + (621) - (619), we have only 3 relations among the 4 phases ϕ_a, ϕ_b, ϕ_c and ϕ_d. We can then, for instance, set any one of these 4 phases to 0 and still obtain a solution to eqs. (619)-(622). Thus,

$$\phi_a = 0, \qquad \phi_b = -\alpha - \frac{\pi}{2}, \qquad \phi_c = \frac{\alpha}{2} + \frac{\gamma}{2} + \delta, \qquad \phi_d = \frac{\alpha}{2} - \frac{\gamma}{2} + \delta + \frac{\pi}{2}. \tag{623}$$

permits a beam splitter (with $R = \sin^2\frac{\beta}{2}$) + 3 wave-shifter plates to represent an arbitrary 2×2 unitary transformation.

The gate

$$\mathsf{Z} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{624}$$

does nothing to $|0\rangle$ and changes the phase of $|1\rangle$ by π. Hence, we expect that we can implement an optical Z gate with no beam splitter, phase shifts $\phi_a = \phi_c = 0$ on the path of $|0\rangle$, and phase shifts $\phi_b + \phi_d = \pi$ on the path of $|1\rangle$.
Comparing eq. (624) to the general form (618), we quickly see that

$$\alpha = -\pi, \qquad \beta = 0, \qquad \gamma = 0, \qquad \delta = \frac{\pi}{2}. \tag{625}$$

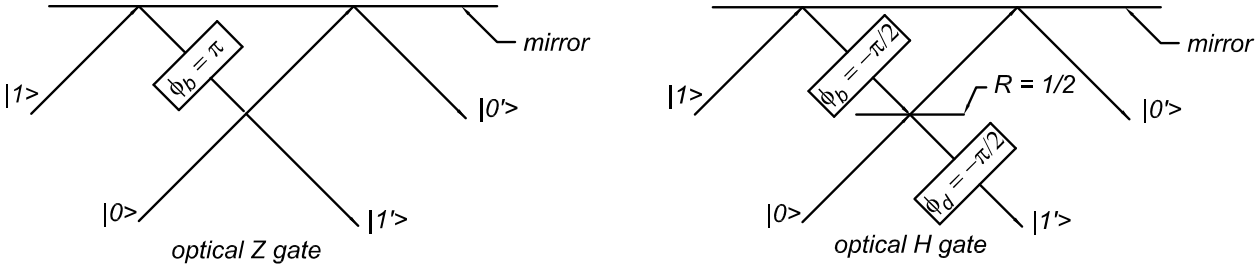
Using these values in eq. (623), we find the phase shifts to be

$$\phi_a = 0, \qquad \phi_b = -\frac{3\pi}{2}, \qquad \phi_c = 0, \qquad \phi_d = \frac{\pi}{2}\,. \tag{626}$$

The reflection coefficient of the beam splitter is $R = \sin^2\frac{\beta}{2} = 0$, so we don't need a beam splitter to implement an optical Z gate. Whenever there is no beam splitter, the phase shifts ϕ_b and ϕ_d can be combined into a single phase shift of $\phi_b + \phi_d = -pi$, which is equivalent to a phase shift of $+\pi$. Thus, we can set the phase shifts to be

$$\phi_a = 0, \qquad \phi_b = \pi, \qquad \phi_c = 0, \qquad \phi_d = 0. \tag{627}$$

An arrangement of phase-shifting plates to implement an optical Z gate is shown in the figure on the left below.



Clearly, an optical Z^p gate could be implemented by changing the phase shift ϕ_b to $p\pi$.

The Hadamard gate

$$\mathsf{H} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{628}$$

splits both $|0\rangle$ and $|1\rangle$ evenly, and then shifts the phase of the final $|1\rangle$ by π when the initial state is $|1\rangle$. Hence, we expect that we can implement an optical H gate with a 50:50 beam splitter. In an arrangement of the type shown above, part of the initial $|0\rangle$ is transmitted to the final state $|0'\rangle$ with no phase shift, but the part of $|0\rangle$ that is reflected into $|1'\rangle$ receives a phase shift of $\pi/2$ by the splitter, which must be restored to 0 by the phase shift $\phi_d = -\pi/2$ (or $3\pi/2$ if you consider that the phase shift of a physical wave plate must be positive). The part of $|1\rangle$ that is transmitted to $|1'\rangle$ needs an overall phase shift of π (or $-\pi$), so we must add a phase shift $\phi_b = -\pi/2$ to accomplish this. Then, the part of $|1\rangle$ that is reflected into $|0'\rangle$ experiences phase shift $-\pi/2$ from plate b and shift $\pi/2$ from the beam splitter, for a total phase shift of 0 as desired.

The implementation of an optical H gate is shown on the right of the above figure.

We confirm the preceding analysis by comparing eq. (628) to the general form (618), which leads to

$$\alpha = 0, \qquad \beta = \frac{\pi}{4}\,, \qquad \gamma = -\pi, \qquad \delta = \frac{\pi}{2}\,. \tag{629}$$

Using these values in eq. (623), we find the phase shifts to be

$$\phi_a = 0, \qquad \phi_b = -\frac{\pi}{2}, \qquad \phi_c = 0, \qquad \phi_d = \frac{3\pi}{2}\,, \tag{630}$$

as previously deduced.

(d) **A Mach-Zehnder Interferometer as a Quantum Processor**

Part (d) was noted by Reck *et al.* in 1994.[152]

A lossless 50:50 beam splitter is described by eq. (611) with $\beta = \pi$. Thus, the desired unitary matrix M is given by

$$\mathsf{M} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} = \frac{\mathbf{I} + i\mathsf{X}}{\sqrt{2}} \,. \tag{631}$$

Similarly, a matrix description of the effect of the phase shifting plates is

$$\phi = \begin{pmatrix} e^{i\phi_0} & 0 \\ 0 & e^{i\phi_1} \end{pmatrix} . \tag{632}$$

The effect of the entire Mach-Zehnder interferometer is then

$$\begin{aligned}
\mathsf{U} &= \phi_c \mathsf{M}\, \phi_b \mathsf{M}\, \phi_a \\
&= \frac{1}{\sqrt{2}} \phi_c \mathsf{M}\, \phi_b \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} \begin{pmatrix} e^{i\phi_{a0}} & 0 \\ 0 & e^{i\phi_{a1}} \end{pmatrix} \\
&= \frac{1}{\sqrt{2}} \phi_c \mathsf{M} \begin{pmatrix} e^{i\phi_{b0}} & 0 \\ 0 & e^{i\phi_{b1}} \end{pmatrix} \begin{pmatrix} e^{i\phi_{a0}} & ie^{i\phi_{a1}} \\ ie^{i\phi_{a0}} & e^{i\phi_{a1}} \end{pmatrix} \\
&= \frac{1}{2} \phi_c \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} \begin{pmatrix} e^{i(\phi_{a0}+\phi_{b0})} & ie^{i(\phi_{a1}+\phi_{b0})} \\ ie^{i(\phi_{a0}+\phi_{b1})} & e^{i(\phi_{a1}+\phi_{b1})} \end{pmatrix} \\
&= \frac{1}{2} \begin{pmatrix} e^{i\phi_{c0}} & 0 \\ 0 & e^{i\phi_{c1}} \end{pmatrix} \begin{pmatrix} e^{i\phi_{a0}}[e^{i\phi_{b0}} - e^{i\phi_{b1}}] & ie^{i\phi_{a1}}[e^{i\phi_{b0}} + e^{i\phi_{b1}}] \\ ie^{i\phi_{a0}}[e^{i\phi_{b0}} + e^{i\phi_{b1}}] & -e^{i\phi_{a1}}[e^{i\phi_{b0}} - e^{i\phi_{b1}}] \end{pmatrix} \\
&= \frac{1}{2} \begin{pmatrix} e^{i(\phi_{a0}+\phi_{c0})}[e^{i\phi_{b0}} - e^{i\phi_{b1}}] & ie^{i(\phi_{a1}+\phi_{c0})}[e^{i\phi_{b0}} + e^{i\phi_{b1}}] \\ ie^{i(\phi_{a0}+\phi_{c1})}[e^{i\phi_{b0}} + e^{i\phi_{b1}}] & -e^{i(\phi_{a1}+\phi_{c1})}[e^{i\phi_{b0}} - e^{i\phi_{b1}}] \end{pmatrix} .
\end{aligned} \tag{633}$$

Comparing eq. (633) with the general form (618) for a 2×2 unitary matrix, the two will be the same if we take

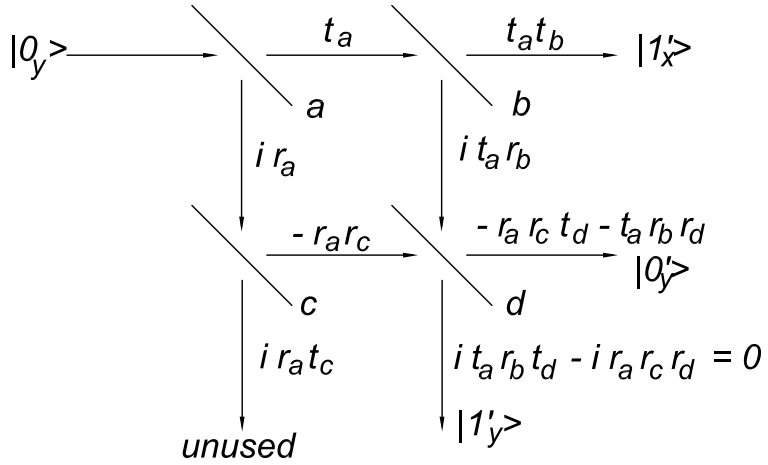
$$\begin{aligned}
\phi_{a0} &= \frac{\gamma}{2} \,, & \phi_{a1} &= -\frac{\gamma}{2} \,, \\
\phi_{b0} &= \frac{\beta}{2} + \delta, & \phi_{b1} &= -\frac{\beta}{2} + \delta + \pi, \\
\phi_{c0} &= \frac{\alpha}{2} \,, & \phi_{c1} &= -\frac{\alpha}{2} + \pi .
\end{aligned} \tag{634}$$

[152] http://kirkmcd.princeton.edu/examples/QM/reck_prl_73_58_94.pdf

(e) **An Optical Controlled-NOT Gate**

N-Qbit gates based on only a single photon were discussed by Reck *et al.*[153] Two-Qbit gates based on 2 photons have been discussed by Koashi *et al.*,[154] by Knill *et al.*,[155] and by O'Brien *et al.*[156] A variant on this part using polarized photons has been discussed by Ralph *et al.*[157]

When the only photon incident on the interferometer is $|0_y\rangle$, then the amplitudes for the various splittings are shown in the figure below.



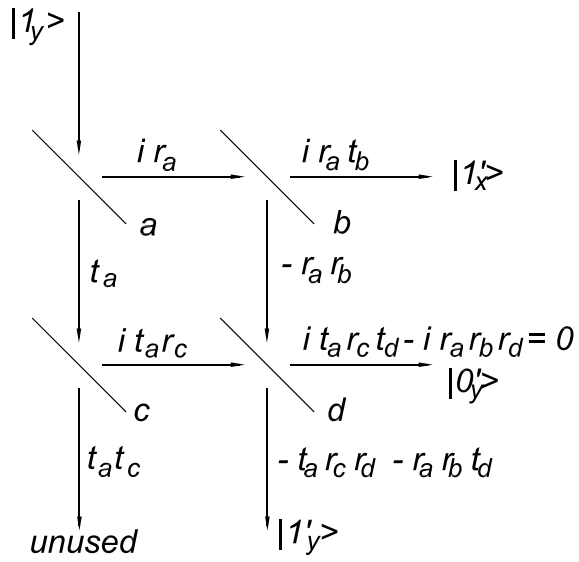
In this case we never want a photon to emerge as $|1'_y\rangle$, which requires that $ir_ar_cr_d = it_ar_bt_d$. The simplest solution to this is, recalling that $r^2 + t^2 = 1$,

$$r_a = t_a = \frac{1}{\sqrt{2}}\,, \qquad r_b = r_c, \qquad r_d = t_b = \frac{1}{\sqrt{2}}\,. \tag{635}$$

The probability that a photon in initial state $|0_y\rangle$ emerges as $|0'_y\rangle$ is then $r_b^2 = R_b$. The probability that the photon emerges at the unused output port is $r_a^2 t_c^2 = t_b^2/2$, and the probability that the photon emerges at the output port labeled $|1'_x\rangle$ is $t_a^2 t_b^2 = t_b^2/2$. Thus, the total probability for the final states is $r_b^2 + 2(t_b^2/2) = 1$, as expected.

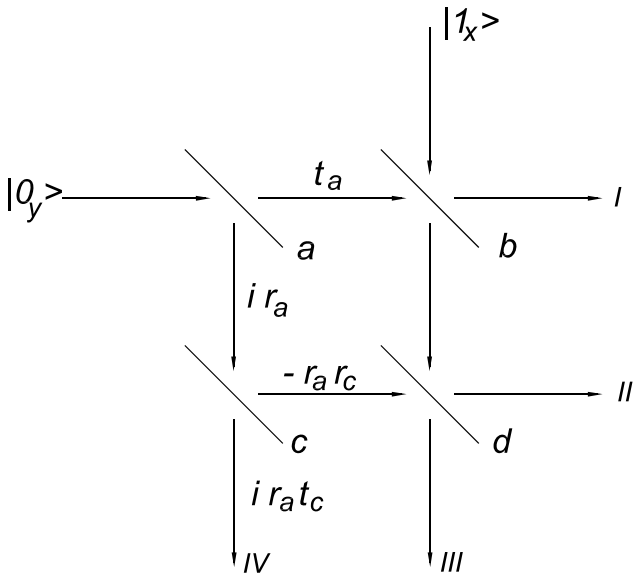
When the only photon incident on the interferometer is $|1_y\rangle$, then the amplitudes for the various splittings are shown in the figure on the next page.

[153] http://kirkmcd.princeton.edu/examples/QM/reck_prl_73_58_94.pdf
[154] http://kirkmcd.princeton.edu/examples/QM/koashi_pra_63_030301_01.pdf
[155] http://kirkmcd.princeton.edu/examples/QM/knill_nature_409_46_01.pdf
[156] http://kirkmcd.princeton.edu/examples/QM/obrien_nature_426_264_03.pdf
[157] http://kirkmcd.princeton.edu/examples/QM/ralph_pra_65_024308_02.pdf



In this case we never want a photon to emerge as $|0'_y\rangle$, which requires that $r_a r_b r_d = t_a r_c t_d$. A solution to this is again eq. (635). The probability that $|1_y\rangle$ emerges as $|1'_y\rangle$ is $r_b^2 = R_b$.

Turning now to the case when the "control" photon is in the state $|1_x\rangle$ and the "target" photon is in the state $|0_y\rangle$, the splittings are shown in the figure below. To organize the bookkeeping, we label the 4 output ports as I, II, III and IV.



One photon is incident on each side of splitter b, and we desire that one photon emerge from each side. Recalling eq. (603), with $\phi_t = 0$ and $\phi_r = \pi/2$, the amplitude that one photon emerges from each side of splitter b is $t_b^2 - r_b^2$ times the product of the incident amplitudes. Hence, splitter b cannot be a 50:50 device.

We desire that one photon emerges from port I (in state $|1'_x\rangle$) and the other photon emerges from port III (in state $|1'_y\rangle$) and not from port II (in state $|0'_y\rangle$).

There are two ways that one photon could emerge from port I and one from port II:

i. One travels on path $|0_y\rangle$-*a*-*c*-*d*-II and the other on path $|1_x\rangle$-*b*-I. The combined amplitude for this option is $(-r_a r_c t_d)(ir_b) = -ir_b^2/2$, recalling that $r_a = t_a = r_d = t_d = 1/\sqrt{2}$ and $r_b = r_c$.

ii. The two photons impinge on splitter b, then one goes to port I and the other goes down to splitter d and into port II. The combined amplitude for this option is $(t_a)(t_b^2 - r_b^2)(ir_d) = i(1 - 2r_b^2)/\sqrt{2}$.

We desire that the sum of these two amplitudes be 0,

$$r_b^2 - (1 - 2r_b^2) = 0, \tag{636}$$

which requires that $r_b^2 = 1/3$.

The desired reflection coefficients are therefore

$$R_a = \frac{1}{2}, \qquad R_b = \frac{1}{3}, \qquad R_c = \frac{1}{3}, \qquad R_d = \frac{1}{2}. \tag{637}$$

As a check, we verify that the sum of all possible outcomes is unity.

Among the possible fates of the 2 input photons is the case that both photons travel on the path from splitter b to splitter d. We then need to know the amplitudes that these 2 photons both emerge in port II, or both in port III, or 1 in port II and 1 in port III. If the photons were distinguishable, the amplitude that both are reflected to port II would be $(ir_d)^2$, the amplitude that both are transmitted to port III would be t_d^2, and the amplitude that one is transmitted to port II and one is reflected to port III would be $ir_d t_d$. However, photons are bosons, so the fact that there are two ways to arrange that one is transmitted and one is reflected means that the probability for this is twice that mentioned above, and the amplitude is $\sqrt{2}$ times as large: the amplitude that one is transmitted to port II and one is reflected to port III is actually $i\sqrt{2}r_d t_d$.

We now list the amplitudes for all possible trajectories for the two photons:

$$\begin{aligned}
A(1 \text{ in I, } 1 \text{ in II}) &= (-r_a r_c t_d)(ir_b) + (t_a)(t_b^2 - r_b^2)(ir_d) = 0, \\
A(1 \text{ in I, } 1 \text{ in III}) &= (ir_a r_c r_d)(ir_b) + (t_a)(t_b^2 - r_b^2)(t_d), \\
A(1 \text{ in I, } 1 \text{ in IV}) &= (ir_a t_c)(ir_b), \\
A(1 \text{ in II, } 1 \text{ in III}) &= (-r_a r_c)(t_b)(t_d^2 - r_d^2) + (t_a)(i\sqrt{2}r_b t_b)(i\sqrt{2}r_d t_d), \\
A(1 \text{ in II, } 1 \text{ in IV}) &= (ir_a t_c)(it_b r_d), \\
A(1 \text{ in III, } 1 \text{ in IV}) &= (ir_a t_c)(t_b t_d), \\
A(2 \text{ in I}) &= (t_a)(i\sqrt{2}r_b t_b). \\
A(2 \text{ in II}) &= (-r_a r_c)(t_b)(i\sqrt{2}r_d t_d) + (t_a)(i\sqrt{2}r_b t_b)(ir_d)^2, \\
A(2 \text{ in III}) &= (-r_a r_c)(t_b)(i\sqrt{2}r_d t_d) + (t_a)(i\sqrt{2}r_b t_b)(t_d)^2 = 0.
\end{aligned} \tag{638}$$

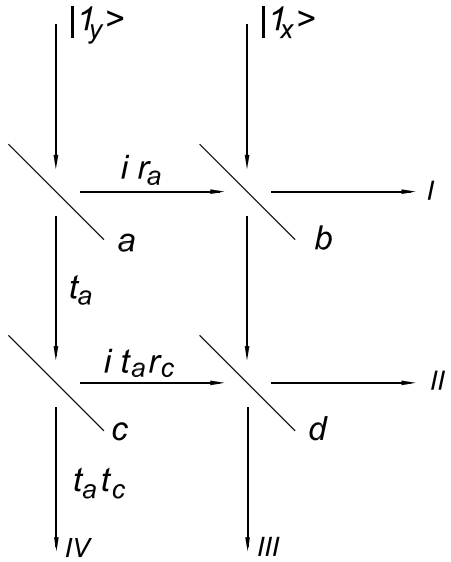
The corresponding probabilities are, recalling that $r_a^2 = t_a^2 = r_d^2 = t_d^2 = 1/2$, $r_b^2 = r_c^2 = 1/3$, $t_b^2 = t_c^2 = 2/3$:

$$\begin{aligned}
P(1 \text{ in I, } 1 \text{ in II}) &= 0, \\
P(1 \text{ in I, } 1 \text{ in III}) &= r_a^2 t_b^4 t_d^2 = \frac{4}{36}, \\
P(1 \text{ in I, } 1 \text{ in IV}) &= r_a^2 r_b^2 t_c^2 = \frac{4}{36}, \\
P(1 \text{ in II, } 1 \text{ in III}) &= 4 t_a^2 r_b^2 t_b^2 r_d^2 t_d^2 = \frac{4}{36}, \\
P(1 \text{ in II, } 1 \text{ in IV}) &= r_a^2 t_b^2 t_c^2 r_d^2 = \frac{4}{36}, \\
P(1 \text{ in III, } 1 \text{ in IV}) &= r_a^2 t_b^2 t_c^2 t_d^2 = \frac{4}{36}, \\
P(2 \text{ in I}) &= 2 t_a^2 r_b^2 t_b^2 = \frac{8}{36}. \\
P(2 \text{ in II}) &= 8 r_a^6 r_b^2 t_b^2 = \frac{8}{36}, \\
P(2 \text{ in III}) &= 0.
\end{aligned} \tag{639}$$

The sum of the probabilities is 1, as expected.

The probability that this configuration behaved like the Controlled-NOT operation $|1_x 0_y\rangle \rightarrow |1_x 1_y\rangle$ is only $4/36 = 1/9$.

Finally, we examine the case when the "control" photon is in the state $|1_x\rangle$ and the "target" photon is in the state $|1_y\rangle$ with the aid of the figure below.



$$\begin{aligned}
A(1 \text{ in I, } 1 \text{ in II}) &= (i t_a r_c t_d)(i r_b) + (i r_a)(t_b^2 - r_b^2)(i r_d) \\
&= -r_a^2 t_b^2 = -\frac{1}{3}, \\
A(1 \text{ in I, } 1 \text{ in III}) &= (-t_a r_c r_d)(i r_b) + (i r_a)(t_b^2 - r_b^2)(t_d) \\
&= i r_a t_a (1 - 3 r_b^2) = 0.
\end{aligned} \tag{640}$$

Thus, $P(|1_x 1_y\rangle \rightarrow |1'_x 1'_y\rangle) = 0$ as desired, and $P(|1_x 1_y\rangle \rightarrow |1'_x 0'_y\rangle) = P(|1_x 0_y\rangle \rightarrow |1'_x 1'_y\rangle) = 1/9$.

It appears that an optical Controlled-NOT gate that succeeds 1/6 of the time has now been demonstrated, although the authors don't actually mention this:
http://kirkmcd.princeton.edu/examples/QM/pryde_prl_92_190402_04.pdf

8. **A Programmable Quantum Computer?**

This problem is based on a paper by Nielsen and Chuang,[158] who do not require that the program register be unchanged by operation V.[159]

Given two $2^n \times 2^n$ unitary matrices U_p and U_q, their associated m-Qbit program register states $|p\rangle$ and $|q\rangle$, and the $(m+n) \times (m+n)$ unitary matrix V which implements a programmable quantum gate array such that

$$\mathsf{V}[|p\rangle \otimes |d\rangle] = |p\rangle \otimes \mathsf{U}_p|d\rangle, \tag{111}$$

$$\mathsf{V}[|q\rangle \otimes |d\rangle] = |q\rangle \otimes \mathsf{U}_q|d\rangle, \tag{112}$$

we take the scalar product of eqs. (111) and (112) to find

$$\begin{aligned} [\langle p| \otimes \langle d|]\mathsf{V}^\dagger\mathsf{V}[|q\rangle \otimes |d\rangle] &= [\langle p| \otimes \langle d|][|q\rangle \otimes |d\rangle] = \langle p|q\rangle \\ &= [\langle p| \otimes \langle d|\mathsf{U}_p^\dagger][|q\rangle \otimes \mathsf{U}_q|d\rangle] = \langle p|q\rangle\langle d|\mathsf{U}_p^\dagger\mathsf{U}_q|d\rangle, \end{aligned} \tag{643}$$

for all n-bit words $|d\rangle$.

If U_p is distinct from U_q then $\langle d|\mathsf{U}_p^\dagger\mathsf{U}_q|d\rangle$ cannot be 1 for all Qbits $|d\rangle$, so in this case we must have $\langle p|q\rangle = 0$; *i.e.*, the program-register words $|p\rangle$ and $|q\rangle$ are orthogonal whenever U_p and U_q are distinct.

---

[158] http://kirkmcd.princeton.edu/examples/QM/nielsen_prl_79_321_97.pdf

[159] This requirement simplifies the analysis, and could always be enforced since the action of V on the program register $|p\rangle$ is a direct product with the action of V on the data register $|d\rangle$. That is, we can write

$$\mathsf{V}[|p\rangle \otimes |d\rangle] = \mathsf{V}_p|p\rangle \otimes \mathsf{U}_p|d\rangle, \tag{641}$$

where V_p is a unitary transformation. If V_p is not the unit matrix, we could define $\mathsf{V}' = \mathsf{V}_p^{-1} \otimes \mathsf{V}$. Then,

$$\mathsf{V}'[|p\rangle \otimes |d\rangle] = |p\rangle \otimes \mathsf{U}_p|d\rangle. \tag{642}$$

9. **Designer Hamiltonians**

(a) The Controlled-NOT operator C_{ab} has the matrix form

$$\mathsf{C}_{ab} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \left(\begin{array}{c|c} \mathbf{I} & \mathbf{0} \\ \hline \mathbf{0} & \mathsf{X} \end{array}\right) \tag{644}$$

If matrix A acts on bit $|a\rangle = a_0|0\rangle_a + a_1|1\rangle_a$, and matrix B acts on bit $|b\rangle = b_0|0\rangle_b + b_1|1\rangle_b$, where we write

$$\mathsf{A} = \begin{pmatrix} A_{00} & A_{01} \\ A_{10} & A_{11} \end{pmatrix}, \qquad \text{and} \qquad \mathsf{B} = \begin{pmatrix} B_{00} & B_{01} \\ B_{10} & B_{11} \end{pmatrix}, \tag{645}$$

The the operator tensor product $\mathsf{A} \otimes \mathsf{B}$ acts on the bit tensor product $|a\rangle \otimes |b\rangle$ according to

$$\begin{aligned}
(\mathsf{A} \otimes \mathsf{B})(|a\rangle \otimes |b\rangle) &= (\mathsf{A}|a\rangle) \otimes (\mathsf{B}|b\rangle) \\
&= \begin{pmatrix} A_{00} & A_{01} \\ A_{10} & A_{11} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} \otimes \begin{pmatrix} B_{00} & B_{01} \\ B_{10} & B_{11} \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \end{pmatrix} \\
&= \begin{pmatrix} A_{00}a_0 + A_{01}a_1 \\ A_{10}a_0 + A_{11}a_1 \end{pmatrix} \otimes \begin{pmatrix} B_{00}b_0 + B_{01}b_1 \\ B_{10}b_0 + B_{11}b_1 \end{pmatrix} \\
&= \begin{pmatrix} A_{00}B_{00}a_0b_0 + A_{00}B_{01}a_0b_1 + A_{10}B_{00}a_1b_0 + A_{10}B_{10}a_1b_1 \\ A_{00}B_{10}a_0b_0 + A_{00}B_{11}a_0b_1 + A_{10}B_{10}a_1b_0 + A_{10}B_{11}a_1b_1 \\ A_{10}B_{00}a_0b_0 + A_{10}B_{01}a_0b_1 + A_{11}B_{00}a_1b_0 + A_{11}B_{10}a_1b_1 \\ A_{10}B_{10}a_0b_0 + A_{10}B_{11}a_0b_1 + A_{11}B_{10}a_1b_0 + A_{11}B_{11}a_1b_1 \end{pmatrix} \\
&= \begin{pmatrix} A_{00}B_{00} & A_{00}B_{01} & A_{10}B_{00} & A_{10}B_{10} \\ A_{00}B_{10} & A_{00}B_{11} & A_{10}B_{10} & A_{10}B_{11} \\ A_{10}B_{00} & A_{10}B_{01} & A_{11}B_{00} & A_{11}B_{10} \\ A_{10}B_{10} & A_{10}B_{11} & A_{11}B_{10} & A_{11}B_{11} \end{pmatrix} \begin{pmatrix} a_0b_0 \\ a_0b_1 \\ a_1b_0 \\ a_1b_1 \end{pmatrix}.
\end{aligned} \tag{646}$$

Thus,

$$\mathsf{A} \otimes \mathsf{B} = \left(\begin{array}{c|c} A_{00}\mathsf{B} & A_{01}\mathsf{B} \\ \hline A_{10}\mathsf{B} & A_{11}\mathsf{B} \end{array}\right). \tag{647}$$

The result (647) suggests that we write

$$\begin{aligned}
\mathsf{C}_{ab} &= \left(\begin{array}{c|c} \mathbf{I} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right) + \left(\begin{array}{c|c} \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathsf{X} \end{array}\right) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}_a \otimes \mathbf{I}_b + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}_a \otimes \mathsf{X}_b \\
&= \bar{\mathsf{n}}_a \otimes \mathbf{I}_b + \mathsf{n}_a \otimes \mathsf{X}_b = \mathsf{a}\mathsf{a}^\dagger \otimes (\mathsf{b}\mathsf{b}^\dagger + \mathsf{b}^\dagger\mathsf{b}) + \mathsf{a}^\dagger\mathsf{a} \otimes (\mathsf{b} + \mathsf{b}^\dagger).
\end{aligned} \tag{648}$$

Note that $\bar{\mathsf{n}}_a \otimes \mathbf{I}_b$ means "do nothing to b if a is 0", and that $\mathsf{n}_a \otimes \mathsf{X}_b$ means "flip b if a is 1".

(b) The Controlled-Controlled-NOT operation C_{abc} can be expressed as "do nothing if bit $a = 0$", "do nothing if bit $a = 1$ and bit $b = 0$", and "flip bit c if both bits a and b are 1". In the spirit of the last line of part (a), we anticipate that C_{abc} can be written as $\bar{\mathsf{n}}_a \otimes \mathbf{I}_b \otimes \mathbf{I}_c + \mathsf{n}_a \otimes \bar{\mathsf{n}}_b \otimes \mathbf{I}_c + \mathsf{n}_a \otimes \mathsf{n}_b \otimes \mathsf{X}_c$.

To verify this, we first note that the truth table for the Controlled-Controlled-NOT operator is

C_{abc} :

a	b	c	a'	b'	c'
0	0	0	0	0	0
0	0	1	0	0	1
0	1	0	0	1	0
0	1	1	0	1	1
1	0	0	1	0	0
1	0	1	1	0	1
1	1	0	1	1	1
1	1	1	1	1	0

(649)

so its matrix representation is

$$\mathsf{C}_{abc} = \begin{pmatrix} 1&0&0&0&0&0&0&0\\ 0&1&0&0&0&0&0&0\\ 0&0&1&0&0&0&0&0\\ 0&0&0&1&0&0&0&0\\ 0&0&0&0&1&0&0&0\\ 0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&0&1\\ 0&0&0&0&0&0&1&0 \end{pmatrix} = \left(\begin{array}{c|c|c|c} \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathsf{X} \end{array}\right). \tag{650}$$

The relation (647) suggests that we write

$$\mathsf{A}\otimes\mathsf{B}\otimes\mathsf{C} = \left(\begin{array}{c|c} A_{00}(\mathsf{B}\otimes\mathsf{C}) & A_{01}(\mathsf{B}\otimes\mathsf{C}) \\ \hline A_{10}(\mathsf{B}\otimes\mathsf{C}) & A_{11}(\mathsf{B}\otimes\mathsf{C}) \end{array}\right) = \left(\begin{array}{c|c|c|c} A_{00}B_{00}\mathsf{C} & A_{00}B_{01}\mathsf{C} & A_{01}B_{00}\mathsf{C} & A_{01}B_{01}\mathsf{C} \\ \hline A_{00}B_{10}\mathsf{C} & A_{00}B_{11}\mathsf{C} & A_{01}B_{10}\mathsf{C} & A_{01}B_{11}\mathsf{C} \\ \hline A_{10}B_{00}\mathsf{C} & A_{10}B_{01}\mathsf{C} & A_{11}B_{00}\mathsf{C} & A_{11}B_{01}\mathsf{C} \\ \hline A_{10}B_{10}\mathsf{C} & A_{10}B_{11}\mathsf{C} & A_{11}B_{10}\mathsf{C} & A_{11}B_{11}\mathsf{C} \end{array}\right). \tag{651}$$

We therefore expand the matrix C_{abc} further as

$$\mathsf{C}_{abc} \;=\; \left(\begin{array}{c|c|c|c} \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{array}\right) + \left(\begin{array}{c|c|c|c} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathsf{X} \end{array}\right)$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}_a \otimes \mathbf{I}_b \otimes \mathbf{I}_c + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}_a \otimes \left(\begin{array}{c|c} \mathbf{I} & \mathbf{0} \\ \hline \mathbf{0} & \mathsf{X} \end{array}\right)_{bc}$$
$$= \bar{\mathsf{n}}_a \otimes \mathbf{I}_b \otimes \mathbf{I}_c + \mathsf{n}_a \otimes (\bar{\mathsf{n}}_b \otimes \mathbf{I}_c + \mathsf{n}_b \otimes \mathsf{X}_c) \quad (652)$$

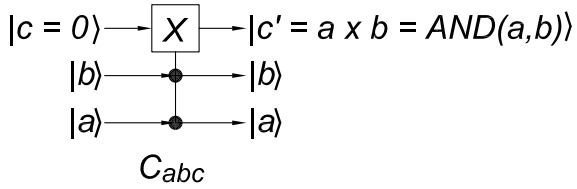
The operator C_{abc} is symmetric in a and b, but the form (652) does not display this symmetry well. We can make this symmetry manifest if we replace $\bar{\mathsf{n}}_a$ by $\mathbf{I}_a - \mathsf{n}_a$, and likewise for $\bar{\mathsf{n}}_b$. This gives

$$\begin{aligned} \mathsf{C}_{abc} &= \mathbf{I}_a \otimes \mathbf{I}_b \otimes \mathbf{I}_c + \mathsf{n}_a \otimes \mathsf{n}_b \otimes (\mathsf{X}_c - \mathbf{I}_c) \\ &= \mathbf{I}_a \otimes \mathbf{I}_b \otimes \mathbf{I}_c + \mathsf{a}^\dagger\mathsf{a} \otimes \mathsf{b}^\dagger\mathsf{b} \otimes (\mathsf{c} + \mathsf{c}^\dagger - \mathbf{I}_c), \end{aligned} \quad (653)$$

where we suppress the expansion of the unit matrices in term of annihilation and creation operators.

(c) $\mathsf{AND}(\mathbf{a}, \mathbf{b}) =$ **Multiply Bits** a **and** b**.**

The Controlled-Controlled-NOT operator C_{abc} also performs the role of multiplying bits a and b with result c, provided that the initial state of bit c is $|0\rangle$. Hence, this operator also implements the AND gate: $c = \mathsf{AND}(a, b)$.



We verify this with the Mathematica command
`TruthTable[CCNGate[1, 2, 3, 3]]`, whose results for bit $c = |0\rangle$ are

```
ket[0, 0, 0] -> ket[0, 0, 0],
ket[0, 1, 0] -> ket[0, 1, 0],
ket[1, 0, 0] -> ket[1, 0, 0],
ket[1, 1, 0] -> ket[1, 1, 1].
```

This TruthTable is the same as for the operation of multiplying two Cbits.

To observe the result of the two-bit AND/Multiply, we need only to observe the final state of bit c. A suitable measurement operator is therefore

$$\mathsf{M} = 0 \cdot |0\rangle_c\langle 0|_c + 1 \cdot |1\rangle_c\langle 1|_c. \quad (654)$$

If the input bits are general Qbits of the form $|a\rangle = a_0|0\rangle + a_1|1\rangle$ and $|b\rangle = b_0|0\rangle + b_1|1\rangle$, then the direct product input state of the circuit is

$$|\text{in}\rangle = \mathrm{a_0b_0}|0\rangle|0\rangle|0\rangle + \mathrm{a_0b_1}|0\rangle|1\rangle|0\rangle + \mathrm{a_1b_0}|1\rangle|0\rangle|0\rangle + \mathrm{a_1b_1}|1\rangle|1\rangle|0\rangle. \quad (655)$$

Referring to the above TruthTable, we see that the output state is then

$$|\text{out}\rangle = \mathrm{a_0b_0}|0\rangle|0\rangle|0\rangle + \mathrm{a_0b_1}|0\rangle|1\rangle|0\rangle + \mathrm{a_1b_0}|1\rangle|0\rangle|0\rangle + \mathrm{a_1b_1}|1\rangle|1\rangle|1\rangle. \quad (656)$$

Applying the measurement operator (654) to the output state, we obtain

$$\mathsf{M}|\text{out}\rangle = 0\cdot(\mathrm{a_0b_0}|0\rangle|0\rangle|0\rangle + \mathrm{a_0b_1}|0\rangle|1\rangle|0\rangle + \mathrm{a_1b_0}|1\rangle|0\rangle|0\rangle) + 1\cdot\mathrm{a_1b_1}|1\rangle|1\rangle|1\rangle. \quad (657)$$

The probabilities of observing the result to be 0 or 1 are

$$P_0 = |a_0 b_0|^2 + |a_0 b_1|^2 + |a_0 b_1|^2 = |a_0|^2 |b_0|^2 + |a_0|^2 |b_1|^2 + |a_1|^2 |b_0|^2 , \qquad (658)$$

and

$$P_1 = |a_1 b_1|^2 = |a_1|^2 |b_1|^2 . \qquad (659)$$

The total probability of the result is

$$P_0 + P_1 = |a_0 b_0|^2 + |a_0 b_1|^2 + |a_0 b_1|^2 + |a_1 b_1|^2 = (|a_0|^2 + |a_1|^2)(|b_0|^2 + |b_1|^2) = 1. \qquad (660)$$

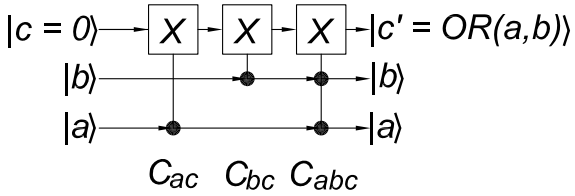
Thus, the effect of the Multiply circuit on Qbits is to produce a 1 with a probability equal to the product of the probabilities that the input bits are both 1, and to produce a 0 with a probability equal to the sum of the products of the probabilities that the input bits are 0 and 0, 0 and 1, or 1 and 0. That is, the quantum Multiply circuit for Qbits is a kind of probabilistic multiply operation.

(d) $\mathsf{OR}(\mathbf{a}, \mathbf{b})$.

One way to convert bit c from $|0\rangle$ to the OR of bits a and b, while leaving bits a and b in their initial states, is as follows. Use `CNGate[1,3,3]` to flip bit c to $|1\rangle$ if bit $a = |1\rangle$. Then apply `CNGate[2,3,3]` to flip bit c to $|1\rangle$ if bit $b = |1\rangle$. However, if both bits a and b are $|1\rangle$, the second `CNGate` flips bit c back to $|0\rangle$. This can be fixed by applying `CCNGate[1,2,3,3]`.

The order of application of the three gates does not matter.

The diagram of our OR gate is



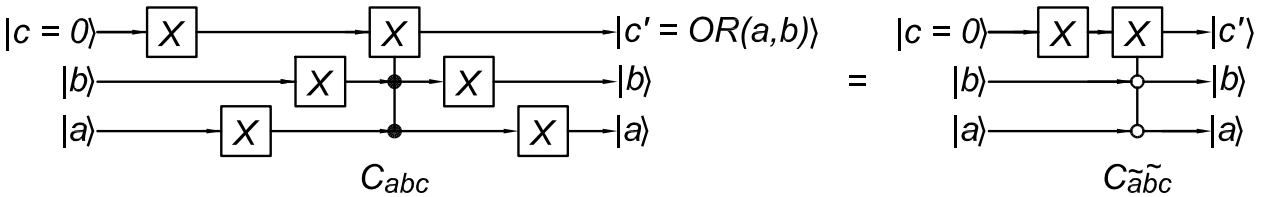
We verify its functionality via the Mathematica command

```
TruthTable[CCNGate[1, 2, 3, 3] . CNGate[2, 3, 3] . CNGate[1, 3, 3]]
```

which yields

```
ket[0, 0, 0] -> ket[0, 0, 0],
ket[0, 1, 0] -> ket[0, 1, 1],
ket[1, 0, 0] -> ket[1, 0, 1],
ket[1, 1, 0] -> ket[1, 1, 1].
```

Another possible implementation of the OR gate is



Indeed, the combination of the 2nd through the 5th gates in the above circuit flips bit $|c\rangle$ only if both bits $|a\rangle$ and $|b\rangle$ are $|0\rangle$. Later in the course we will call this operation $\mathsf{C}_{\tilde{a}\tilde{b}c}$, and symbolize it as shown on the right above.

(e) **Add Two 1-Bit Numbers.**

The Controlled-NOT operator C_{ab} also performs the role of addition modulo 2, with the result appearing as bit b'. This can be seen via

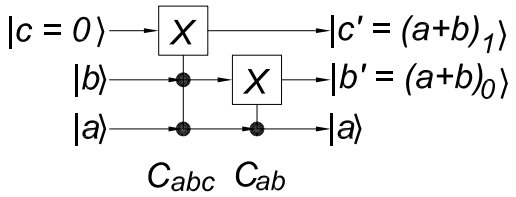
```
TruthTable[CNGate[1, 2, 2]]
```

which returns

```
ket[0, 0] -> ket[0, 0],
ket[0, 1] -> ket[0, 1],
ket[1, 0] -> ket[1, 1],
ket[1, 1] -> ket[1, 0].
```

The remaining task is to "carry" the 1 that results from adding 1 to 1. For this we need an operator that produces a $|1\rangle$ only when both bits a and b are $|1\rangle$, and zero otherwise. Further, since we can't copy the initial Qbits a and b, the desired operator should simply pass these bits on unaltered. On reflection, we see that the Controlled-Controlled-NOT operator is what we want.

A circuit for a quantum adder that adds bits a and b and presents the answer as $c'b'$ (*i.e.*, the high-order bit is $c' = (a+b)_1$) is



Our adder gate is the product $\mathsf{C}_{ij}\mathsf{C}_{ijk}$, meaning that the C_{ijk} is applied first. To verify that we should use `CNGate[1,2,3] . CCNGate[1,2,3,3]` and not `CCNGate[1,2,3,3] . CNGate[1,3,3]`, we execute

```
TruthTable[CNGate[1, 2, 3] . CCNGate[1, 2, 3, 3]]
```

which returns

```
ket[0, 0, 0] -> ket[0, 0, 0],
ket[0, 1, 0] -> ket[0, 1, 0],
ket[1, 0, 0] -> ket[1, 1, 0],
ket[1, 1, 0] -> ket[1, 0, 1].
```

The addends are the initial bits 1 and 3. The low-order bit of the answer is bit 2, and the high-order bit is bit 3. Thus, `ket[1, 1, 0] -> ket[1, 0, 1]` corresponds to $1 + 1 = 10$, as desired.

To observe the result of the two-bit ADD circuit, we need to observe the final state of bits b and c. A suitable measurement operator is therefore

$$\mathsf{A} = 0 \cdot |0\rangle_b\langle 0|_b \cdot |0\rangle_c\langle 0|_c + 1 \cdot |0\rangle_b\langle 0|_b \cdot |1\rangle_c\langle 1|_c + 10 \cdot |1\rangle_b\langle 1|_b \cdot |0\rangle_c\langle 0|_c \qquad (661)$$

If the input bits are general Qbits of the form $|a\rangle = a_0|0\rangle + a_1|1\rangle$ and $|b\rangle = b_0|0\rangle + b_1|1\rangle$, then the input state of the circuit is

$$|\text{in}\rangle = \text{a}_0\text{b}_0|0\rangle|0\rangle|0\rangle + \text{a}_0\text{b}_1|0\rangle|1\rangle|0\rangle + \text{a}_1\text{b}_0|1\rangle|0\rangle|0\rangle + \text{a}_1\text{b}_1|1\rangle|1\rangle|0\rangle. \qquad (662)$$

Referring to the above TruthTable, we see that the output state is then

$$|\text{out}\rangle = \text{a}_0\text{b}_0|0\rangle|0\rangle|0\rangle + \text{a}_0\text{b}_1|0\rangle|0\rangle|1\rangle + \text{a}_1\text{b}_0|1\rangle|0\rangle|1\rangle + \text{a}_1\text{b}_1|1\rangle|1\rangle|0\rangle. \qquad (663)$$

Applying the measurement operator (654) to the output state, we obtain

$$\mathsf{A}|\text{out}\rangle = 0 \cdot \mathrm{a_0 b_0}|0\rangle|0\rangle|0\rangle + 1 \cdot (\mathrm{a_0 b_1}|0\rangle|0\rangle|1\rangle + \mathrm{a_1 b_0}|1\rangle|0\rangle|1\rangle) + 10 \cdot \mathrm{a_1 b_1}|1\rangle|1\rangle|0\rangle. \tag{664}$$

The probabilities of observing the result to be 0, 1 or 10 are

$$P_0 = |a_0 b_0|^2 , \tag{665}$$

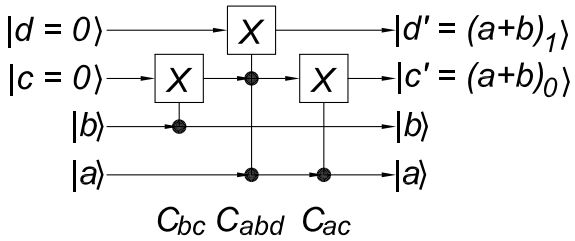
$$P_1 = |a_0 b_1|^2 + |a_1 b_0|^2 , \tag{666}$$

and

$$P_{10} = |a_1 b_1|^2 . \tag{667}$$

Thus, the effect of the Add circuit on Qbits is to produce a sum with a probability equal to the probability of observing the corresponding addends. Like the quantum Multiply circuit, the quantum Add circuit for Qbits functions probabilistically.

A possible defect of the above circuit is that the initial value of the bit b is overwritten. We could make a classical copy of the second addend bit using a Controlled-NOT gate that acts on that bit and a fourth bit whose initial value is $|0\rangle$, as shown in the diagram below.



The action of this circuit on classical bits is the same as that of the previous adder. Thus,

```
TruthTable[CNGate[1, 3, 4] . CCNGate[1, 3, 4, 4] . CNGate[2, 3, 4]]
```

returns

```
ket[0, 0, 0, 0] -> ket[0, 0, 0, 0],
ket[0, 1, 0, 0] -> ket[0, 1, 1, 0],
ket[1, 0, 0, 0] -> ket[1, 0, 1, 0],
ket[1, 1, 0, 0] -> ket[1, 1, 0, 1].
```

To observe the result of the revised two-bit Add circuit, we need to observe the final state of bits c and d. A suitable measurement operator is therefore

$$\mathsf{A} = 0 \cdot |0\rangle_c\langle 0|_c \cdot |0\rangle_d\langle 0|_d + 1 \cdot |0\rangle_c\langle 0|_c \cdot |1\rangle_d\langle 1|_d + 10 \cdot |1\rangle_c\langle 1|_c \cdot |0\rangle_d\langle 0|_d \tag{668}$$

If the input bits are general Qbits of the form $|a\rangle = a_0|0\rangle + a_1|1\rangle$ and $|b\rangle = b_0|0\rangle + b_1|1\rangle$, then the direct product input state of the circuit is

$$|\text{in}\rangle = \mathrm{a_0 b_0}|0\rangle|0\rangle|0\rangle|0\rangle + \mathrm{a_0 b_1}|0\rangle|1\rangle|0\rangle|0\rangle + \mathrm{a_1 b_0}|1\rangle|0\rangle|0\rangle|0\rangle + \mathrm{a_1 b_1}|1\rangle|1\rangle|0\rangle|0\rangle. \tag{669}$$

Referring to the above TruthTable, we see that the output state is then

$$|\text{out}\rangle = \mathrm{a_0 b_0}|0\rangle|0\rangle|0\rangle|0\rangle + \mathrm{a_0 b_1}|0\rangle|1\rangle|0\rangle|1\rangle + \mathrm{a_1 b_0}|1\rangle|0\rangle|0\rangle|1\rangle + \mathrm{a_1 b_1}|1\rangle|1\rangle|1\rangle|0\rangle. \tag{670}$$

Applying the measurement operator (654) to the output state, we obtain

$$\begin{aligned} \mathsf{A}|\text{out}\rangle &= 0 \cdot a_0 b_0 |0\rangle|0\rangle|0\rangle|0\rangle + 1 \cdot (a_0 b_1 |0\rangle|1\rangle|0\rangle|1\rangle + a_1 b_0 |1\rangle|0\rangle|0\rangle|1\rangle) \\ &\quad +10 \cdot a_1 b_1 |1\rangle|1\rangle|1\rangle|0\rangle. \end{aligned} \tag{671}$$

The probabilities of observing the result to be 0, 1 or 10 are again

$$P_0 = |a_0 b_0|^2 , \tag{672}$$

$$P_1 = |a_0 b_1|^2 + |a_1 b_0|^2 , \tag{673}$$

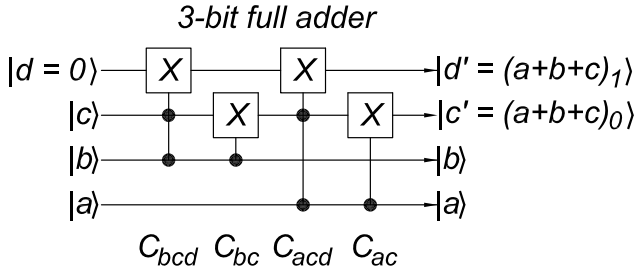
and

$$P_{10} = |a_1 b_1|^2 . \tag{674}$$

Although the revised adder uses an entangled "copy" of bit b, the results of the circuit are in effect that same as those of the previous adder, in that the output bits would be observed with the same probabilities.[160] Thus, while the meaning of the adder circuits as applied to Qbits is only probabilistic (rather than deterministic), this meaning is not altered when we use the C-NOT gate to produce a "copy" of one of the input bits.

(f) **Add Three Bits.**

To add bits a, b and c, overwriting bit c in the process, we first add bits b and c using the circuit of part (e). If we then add bits a and c with a similar circuit, the result is



It is pleasing, but perhaps surprising that we are done!

The command `TruthTable[CNGate[1, 3, 4] . CCNGate[1, 3, 4, 4] . CNGate[2, 3, 4] . CCNGate[2, 3, 4, 4]]`

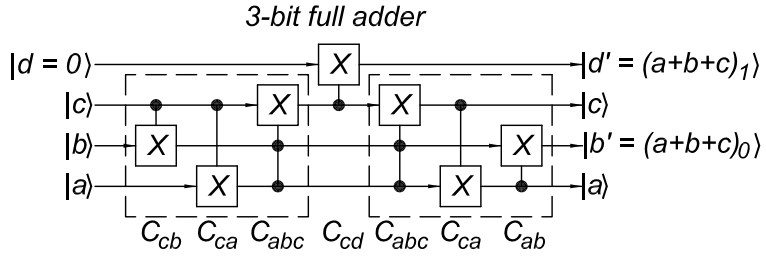
yields

```
ket[0, 0, 0, 0] -> ket[0, 0, 0, 0],
ket[0, 0, 1, 0] -> ket[0, 0, 1, 1],
ket[0, 1, 0, 0] -> ket[0, 1, 1, 1],
ket[0, 1, 1, 0] -> ket[0, 1, 0, 1],
ket[1, 0, 0, 0] -> ket[1, 1, 1, 1],
ket[1, 0, 1, 0] -> ket[1, 1, 0, 1],
ket[1, 1, 0, 0] -> ket[1, 0, 0, 1],
ket[1, 1, 1, 0] -> ket[1, 0, 1, 1]
```

as desired, with the 2-bit sum appearing in the 3rd and 4th output bits.

[160] This remark remains true even if input bits a and b are in an entangled state, such as $(|0\rangle|0\rangle + |1\rangle|1\rangle)/\sqrt{2}$.

Another way to add 3 bits is shown below.[161] This circuit uses more gates, so its advantages are not immediately apparent. However, it turns out that this circuit can be extended to add two n-bit numbers more compactly than our previous circuit.

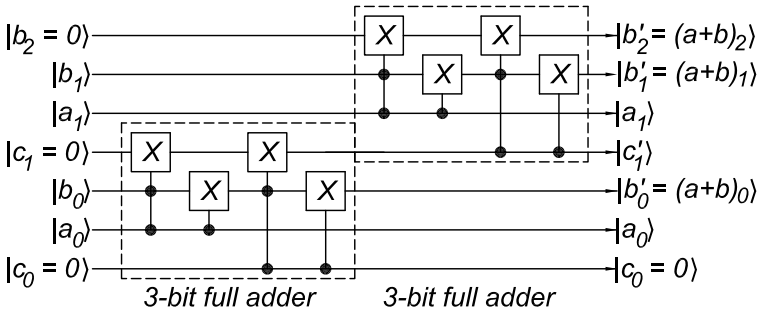


Note that the subcircuit in the left box is a 3-bit majority, meaning that bit $|c\rangle$ will be set to $|1\rangle$ whenever 2 or more of the 3 input bits are $|1\rangle$. Thus, at the output of the left circuit, bit $|c\rangle$ is the high-order bit of the sum $a+b+c$. The Controlled-NOT C_{cd} then transfers this bit to line d.

The subcircuit in the right box almost undoes the circuit in the left box, but by changing the last gate slightly, the combined effect of the two circuits in boxes is to calculate the low-order bit of the sum $a+b+c$.

(g) **Add Two 2-Bit Numbers.**

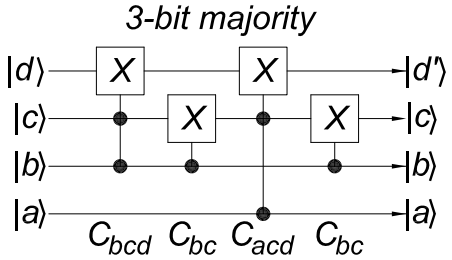
To add two 2-bit numbers $a = \sum_{j=0} a_j 2^j$ and $b = \sum_{j=0} b_j 2^j$ using the 3-bit full adder of part (f), we use one copy of the 3-bit adder to add bits $|a_0\rangle$ and $|b_0\rangle$. We introduce a dummy bit $|c_0\rangle = |0\rangle$ as the first input bit of this adder. The 4th output bit is the "carry" bit $|c_1\rangle$ from the addition $a_0 + b_0$. Then, we use a 2nd copy of the 3-bit adder, with inputs $|c_1\rangle$, $|a_1\rangle$ and $|b_1\rangle$, to complete the calculation.



The classical version of this algorithm is called a "ripple adder".[162]

If we want to restore the ancillary bit $|c_1\rangle$ to its initial value of $|0\rangle$, we need an additional group of gates that operate on the upper four bits, without changing the first three of those bits. The desired circuit is no doubt closely related to that of part (f) which sums the first three bits. In particular, we note that if the 4rth gate of the adder circuit of part (f) is changed from C_{ac} to C_{bc}, then bit c would be unaffected, while the calculation of bit d is as before.

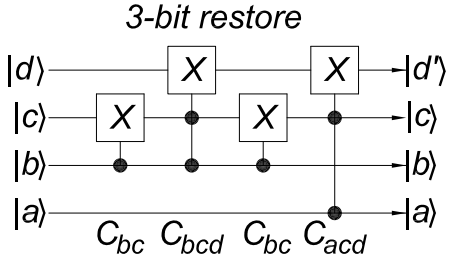
[161] http://kirkmcd.princeton.edu/examples/QM/cuccaro_quant-ph-0410184.pdf
[162] http://kirkmcd.princeton.edu/examples/QM/gossett_quant-ph-9808061.pdf



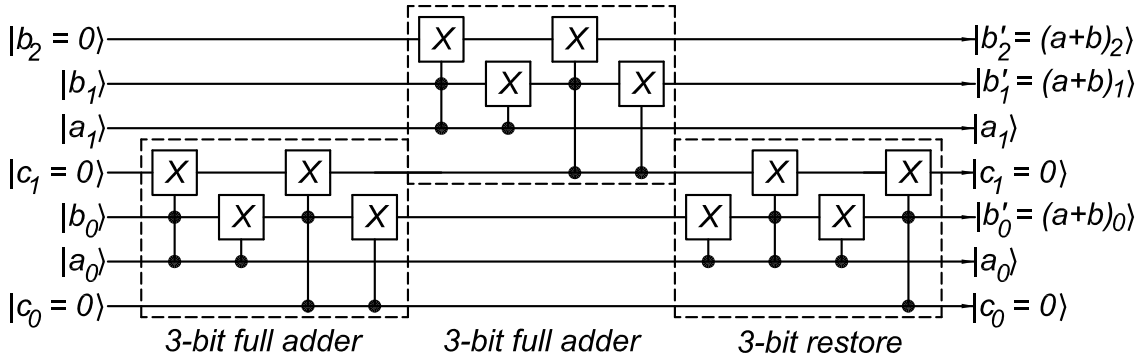
In this circuit, bit $|d\rangle$ is flipped whenever two or more of the inputs $|a\rangle$, $|b\rangle$ and $|c\rangle$ are $|1\rangle$. Therefore, this is also a 3-bit majority circuit.

However, this circuit does not quite performs the desired restoration, since if $|a_0\rangle = |b_0\rangle = |1\rangle$ are input to the ripple adder, then $|b_0'\rangle = |0\rangle$ and $|a_0\rangle = |c_1'\rangle = |1\rangle$ would be input to the majority circuit, which leaves bit $|c_1'\rangle$ unchanged.

However, we can fix things if we move the two Controlled-NOT gates forward, as shown below.



To see if the restore circuit indeed restores bit $|c_1\rangle$, we consider the arrangement



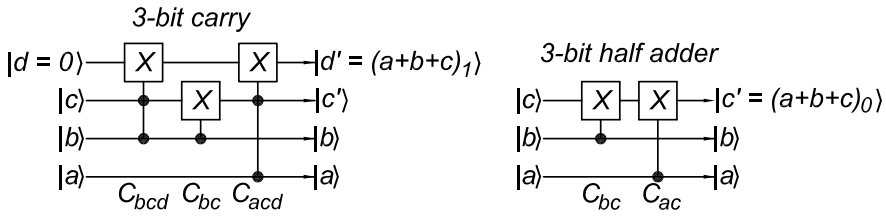
Because every Controlled-NOT gate that affects bit $|c_1\rangle$ appears twice, we expect that this bit is unchanged. To verify this, we calculate

```
TruthTable[ CCNGate[1, 3, 4, 7] . CNGate[2, 3, 7] .
            CCNGate[2, 3, 4, 7] . CNGate[2, 3, 7] .
            CNGate[4, 6, 7] . CCNGate[4, 6, 7, 7] .
            CNGate[5, 6, 7] . CCNGate[5, 6, 7, 7] .
            CNGate[1, 3, 7] . CCNGate[1, 3, 4, 7] .
            CNGate[2, 3, 7] . CCNGate[2, 3, 4, 7]],
```

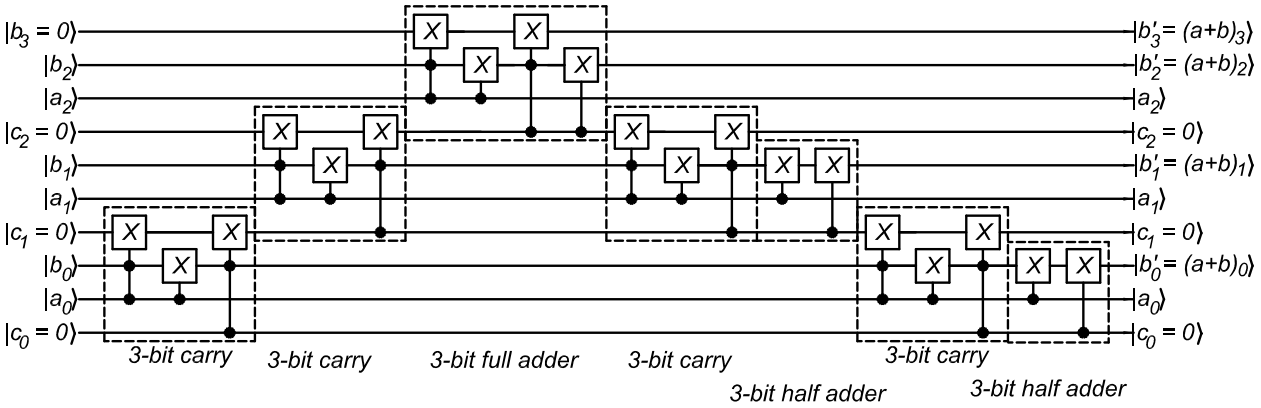
which yields

```
ket[0, 0, 0, 0, 0, 0, 0] -> ket[0, 0, 0, 0, 0, 0, 0],
ket[0, 0, 1, 0, 0, 0, 0] -> ket[0, 0, 1, 0, 0, 0, 0],
ket[0, 1, 0, 0, 0, 0, 0] -> ket[0, 1, 1, 0, 0, 0, 0],
ket[0, 1, 1, 0, 0, 0, 0] -> ket[0, 1, 1, 0, 0, 1, 0],
ket[0, 0, 0, 0, 0, 1, 0] -> ket[0, 0, 0, 0, 0, 1, 0],
ket[0, 0, 1, 0, 0, 1, 0] -> ket[0, 0, 1, 0, 0, 1, 0],
ket[0, 1, 0, 0, 0, 1, 0] -> ket[0, 1, 1, 0, 0, 1, 0],
ket[0, 1, 1, 0, 0, 1, 0] -> ket[0, 1, 0, 0, 0, 0, 1],
ket[0, 0, 0, 0, 1, 0, 0] -> ket[0, 0, 0, 0, 1, 1, 0],
ket[0, 0, 1, 0, 1, 0, 0] -> ket[0, 0, 1, 0, 1, 1, 0],
ket[0, 1, 0, 0, 1, 0, 0] -> ket[0, 1, 1, 0, 1, 1, 0],
ket[0, 1, 1, 0, 1, 0, 0] -> ket[0, 1, 0, 0, 1, 0, 1],
ket[0, 0, 0, 0, 1, 1, 0] -> ket[0, 0, 0, 0, 1, 0, 1],
ket[0, 0, 1, 0, 1, 1, 0] -> ket[0, 0, 1, 0, 1, 0, 1],
ket[0, 1, 0, 0, 1, 1, 0] -> ket[0, 1, 1, 0, 1, 0, 1],
ket[0, 1, 1, 0, 1, 1, 0] -> ket[0, 1, 0, 0, 1, 1, 1].
```

The quantum computation literature[163] shows a variant on the above circuit in which the 3-bit full adder of part (f) is split into a 3-bit carry and a 3-bit half adder, that calculate the higher- and lower-order bits of the sum, respectively:



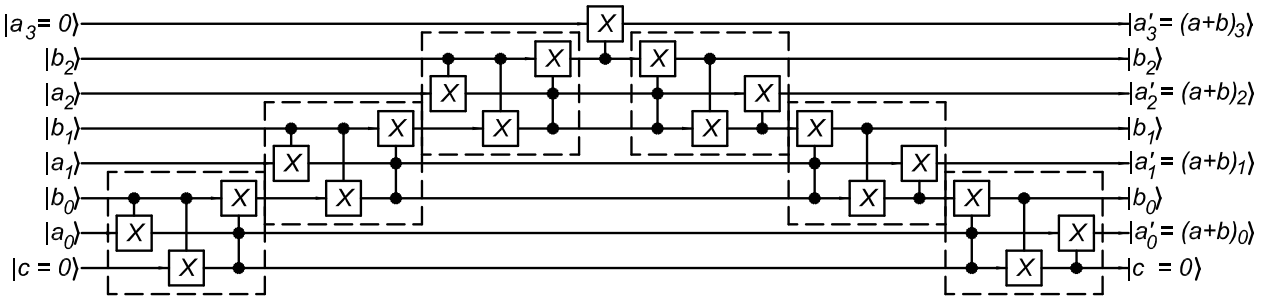
Note that a second 3-bit carry operation undoes the effect of the first. Hence, a ripple adder could also be constructed with a cascade of 3-bit carry operations that terminates in a 3-bit full adder for the highest-order bits, followed by an inverse cascade of carry operations to reset the ancillary bits on all but the highest-order bits of a and b to reset the ancillary bits. Also, 3-bit half adders must be inserted into the inverse cascade so as to utilize the carry bits just before they are reset:



The two versions of the quantum ripple adders involve the same number of gates.

[163] http://kirkmcd.princeton.edu/examples/QM/vedral_pra_54_147_96.pdf

Another ripple adder can be built from the circuit introduced at the end of part (f), as shown below. This circuit uses only one auxiliary bit, and has fewer gates $(6n+1)$ that our two previous adders (which have $8n-4$ gates to add two n-bit numbers).



(h) **Multiply Two 2-Bit Numbers.**

The product of two n-bit numbers $a = \sum_{j=0}^{n-1} a_j 2^j$ and $b = \sum_{j=0}^{n-1} b_j 2^j$ can have up to $2n$ bits,

$$\begin{aligned} a \times b &= \sum_{j=0}^{n-1} a_j 2^j \times \sum_{k=0}^{n-1} b_k 2^k \\ &= \sum_{j=0}^{n-1} a_j 2^j \sum_{k=0}^{n-1} b_k 2^{j+k} \qquad (675) \\ &= \sum_{l=0}^{2n-1} \sum_{m=0}^{l} a_l b_{l-m} 2^l. \qquad (676) \end{aligned}$$

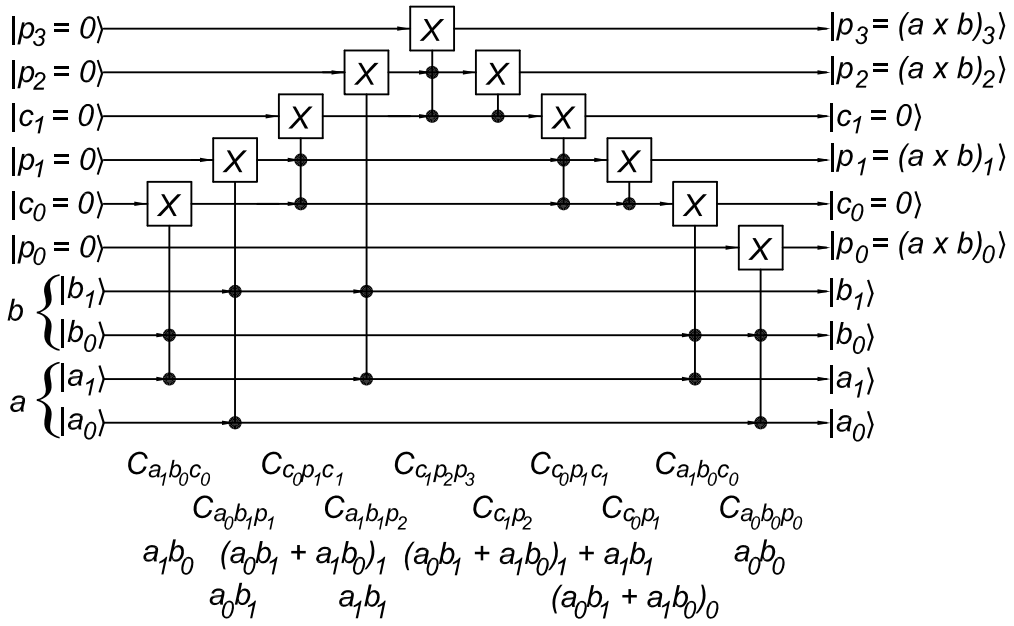
Equation (675) can be implemented via shifts and (controlled) adds, requiring a total of about $8n$ bits. Equation (676) involves n^2 1×1 bit multiplications and additions, requiring about $n(n+3)$ bits. For $n > 5$, multiplication based on shifting and adding has a lower bit count, and would likely be prefered for quantum computation.

For the present example of multiplication of two 2-bit number, we use eq. (676) as it is simpler in this case. Writing out the sum explicity,

$$a \times b = a_0 b_0 + 2(a_0 b_1 + a_1 b_0) + 4 a_2 b_2. \qquad (677)$$

If both $a_0 b_1$ and $a_1 b_0$ are 1, a 1 carries over into the 4's column of the product. If $a_2 b_2 = 1$ also, then a 1 carries over into the 8's column.

A circuit that performs the multiplication is shown in the figure on the next page.



The command

```
TruthTable[CCNGate[1, 3, 5, 10] . CCNGate[2, 3, 6, 10] .
           CNGate[6, 7, 10] . CCNGate[6, 7, 8, 10] .
           CNGate[8, 9, 10] . CCNGate[8, 9, 10, 10] .
           CCNGate[2, 4, 9, 10] . CCNGate[6, 7, 8, 10] .
           CCNGate[1, 4, 7, 10] . CCNGate[2, 3, 6, 10]]
```

yields

```
ket[0, 0, 0, 0, 0, 0, 0, 0, 0, 0] -> ket[0, 0, 0, 0, 0, 0, 0, 0, 0, 0],
ket[0, 0, 0, 1, 0, 0, 0, 0, 0, 0] -> ket[0, 0, 0, 1, 0, 0, 0, 0, 0, 0],
ket[0, 0, 1, 0, 0, 0, 0, 0, 0, 0] -> ket[0, 0, 1, 0, 0, 0, 0, 0, 0, 0],
ket[0, 0, 1, 1, 0, 0, 0, 0, 0, 0] -> ket[0, 0, 1, 1, 0, 0, 0, 0, 0, 0],
ket[0, 1, 0, 0, 0, 0, 0, 0, 0, 0] -> ket[0, 1, 0, 0, 0, 0, 0, 0, 0, 0],
ket[0, 1, 0, 1, 0, 0, 0, 0, 0, 0] -> ket[0, 1, 0, 1, 0, 0, 0, 0, 1, 0],
ket[0, 1, 1, 0, 0, 0, 0, 0, 0, 0] -> ket[0, 1, 1, 0, 0, 0, 1, 0, 0, 0],
ket[0, 1, 1, 1, 0, 0, 0, 0, 0, 0] -> ket[0, 1, 1, 1, 0, 0, 1, 0, 1, 0],
ket[1, 0, 0, 0, 0, 0, 0, 0, 0, 0] -> ket[1, 0, 0, 0, 0, 0, 0, 0, 0, 0],
ket[1, 0, 0, 1, 0, 0, 0, 0, 0, 0] -> ket[1, 0, 0, 1, 0, 0, 1, 0, 0, 0],
ket[1, 0, 1, 0, 0, 0, 0, 0, 0, 0] -> ket[1, 0, 1, 0, 1, 0, 0, 0, 0, 0],
ket[1, 0, 1, 1, 0, 0, 0, 0, 0, 0] -> ket[1, 0, 1, 1, 1, 0, 1, 0, 0, 0],
ket[1, 1, 0, 0, 0, 0, 0, 0, 0, 0] -> ket[1, 1, 0, 0, 0, 0, 0, 0, 0, 0],
ket[1, 1, 0, 1, 0, 0, 0, 0, 0, 0] -> ket[1, 1, 0, 1, 0, 0, 1, 0, 1, 0],
ket[1, 1, 1, 0, 0, 0, 0, 0, 0, 0] -> ket[1, 1, 1, 0, 1, 0, 1, 0, 0, 0],
ket[1, 1, 1, 1, 0, 0, 0, 0, 0, 0] -> ket[1, 1, 1, 1, 1, 0, 0, 0, 0, 1],
```

as desired.

10. **Deutsch's Algorithm**

(a) For function f_0, eq. (131), we have $y \oplus f_0(x) = y$, and the transformation U_{f_0} is the 4×4 unit matrix,

$$\mathsf{U}_{f_0} = \mathbf{I}. \tag{678}$$

For function f_1, eq. (132), we have

$$\begin{aligned} |0\rangle|0 \oplus f_1(0)\rangle &= |0\rangle|0\rangle, \\ |0\rangle|1 \oplus f_1(0)\rangle &= |0\rangle|1\rangle, \\ |1\rangle|0 \oplus f_1(1)\rangle &= |1\rangle|1\rangle, \\ |1\rangle|1 \oplus f_1(1)\rangle &= |1\rangle|0\rangle, \end{aligned} \tag{679}$$

and the transformation U_{f_1} is the 4×4 matrix

$$\mathsf{U}_{f_1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \mathsf{C}_{xy}, \tag{680}$$

recalling the definition (137).

For function f_2, eq. (133), we have

$$\begin{aligned} |0\rangle|0 \oplus f_2(0)\rangle &= |0\rangle|1\rangle, \\ |0\rangle|1 \oplus f_2(0)\rangle &= |1\rangle|0\rangle, \\ |1\rangle|0 \oplus f_2(1)\rangle &= |1\rangle|0\rangle, \\ |1\rangle|1 \oplus f_2(1)\rangle &= |1\rangle|1\rangle, \end{aligned} \tag{681}$$

which flips bit $|y\rangle$ only if bit $|x\rangle = |0\rangle$. This is a variant on the Controlled-NOT operation C_{xy} that we will call $\mathsf{C}_{\tilde{x}y}$ in anticipation of prob. 12(b). The transformation U_{f_2} is the 4×4 matrix

$$\mathsf{U}_{f_2} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \left(\begin{array}{c|c} \mathsf{X} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{I} \end{array}\right). \tag{682}$$

For function f_3, eq. (134), we have

$$\begin{aligned} |0\rangle|0 \oplus f_3(0)\rangle &= |0\rangle|0\rangle, \\ |0\rangle|1 \oplus f_3(0)\rangle &= |0\rangle|1\rangle, \\ |1\rangle|0 \oplus f_3(1)\rangle &= |1\rangle|1\rangle, \\ |1\rangle|1 \oplus f_3(1)\rangle &= |1\rangle|0\rangle, \end{aligned} \tag{683}$$

and the transformation U_{f_3} is the 4×4 matrix

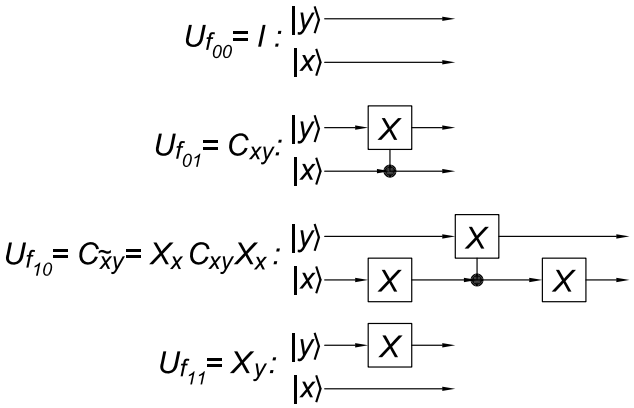
$$\mathsf{U}_{f_3} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \mathsf{X}_y, \tag{684}$$

recalling eq. (141).

(b) Only the transformation $\mathsf{U}_{f_2} = \mathsf{C}_{\tilde{x}y}$ remains to be re-expressed in terms of the Controlled-NOT operation C_{xy}. If we flip bit $|x\rangle$ prior to applying a C_{xy} operation, then bit $|y\rangle$ will be flipped when the initial $|x\rangle$ was $|0\rangle$ as desired. To restore state $|x\rangle$ to its initial value, we flip it a second time, but after the operation C_{xy}. Thus,

$$\begin{aligned} \mathsf{C}_{\tilde{x}y} &= \mathsf{X}_x \mathsf{C}_{xy} \mathsf{X}_x = \left(\begin{array}{c|c} \mathbf{0} & \mathsf{X} \\ \hline \mathsf{X} & \mathbf{0} \end{array}\right) \left(\begin{array}{c|c} \mathbf{I} & \mathbf{0} \\ \hline \mathbf{0} & \mathsf{X} \end{array}\right) \left(\begin{array}{c|c} \mathbf{0} & \mathsf{X} \\ \hline \mathsf{X} & \mathbf{0} \end{array}\right) \\ &= \left(\begin{array}{c|c} \mathbf{0} & \mathsf{X} \\ \hline \mathsf{X} & \mathbf{0} \end{array}\right) \left(\begin{array}{c|c} \mathbf{0} & \mathsf{X} \\ \hline \mathbf{I} & \mathbf{0} \end{array}\right) = \left(\begin{array}{c|c} \mathsf{X} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{I} \end{array}\right) = \mathsf{U}_{f_2}. \end{aligned} \tag{685}$$

(c) Diagrams for the four transformations U_{f_j} found in part (b) are



(d) Note that the unitary transformation U_f is represented by real, symmetric matrices, and hence $\mathsf{U}^{-1} = \mathsf{U}^\dagger = \mathsf{U}$. This suggests that a second application of Deutsch's transformation (135) may undo the effect of the first application. But since a measurement was made after the first application, we must look into the details.

Following the application of transformation (135) to the input bits $|x\rangle = |+\rangle$ and $|y\rangle = |-\rangle$, the measurement of the first bit yields $(-1)^{f(0)}|+\rangle$ if $f(0) = f(1)$ and $(-1)^{f(0)}|-\rangle$ if $f(0) \neq f(1)$. Hence, if $f(0) = f(1)$, a second application of the transformation (135), now using the measured bit as the input bit $|x\rangle$, will bring that bit to $[(-1)^{f(0)}]^2|+\rangle = |+\rangle$, which restores it to its initial value.

This suggests that we try the transformation (135) for the case where $|x\rangle = (-1)^{f(0)}|-\rangle$ and $|y\rangle = |-\rangle$. Recalling eq. (142) we obtain

$$|x\rangle|y \oplus f(x)\rangle = (-1)^{f(0)} \frac{(-1)^{f(0)}|0\rangle - (-1)^{f(1)}|1\rangle}{\sqrt{2}} \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \tag{686}$$

Thus, if we apply the transformation (135) to the results of the measurement, we get the $|x\rangle = |+\rangle$ and $|y\rangle = |-\rangle$ when $f(0) = f(1)$, and we get the result (686) when $f(0) \neq f(1)$. In the latter case this also gives $|x\rangle = [(-1)^{f(0)}]^2|+\rangle = |+\rangle$ and $|y\rangle = |-\rangle$. Thus, simply applying transformation (135) to the results of the measurement restores the input bits to their initial states.

If the effect of the measurement is that the phase factor $(-1)^{f(0)}$ is lost, then to restore the bits precisely a measurement of bit $|x\rangle$ should be made after the second application of transformation (135) to discard the phase factor introduced by that transformation.

(e) The output of Deutsch's algorithm for input bits $|x\rangle = |+\rangle$ and $|y\rangle = |-\rangle$ is the state (143), which has useful properties in the $[|+\rangle, |-\rangle]$ basis. To bring this state to the $[|0\rangle, |1\rangle]$ basis we apply the Hadamard transformation,

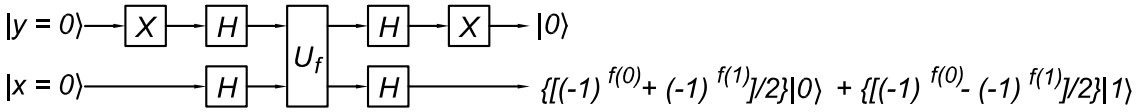
$$\mathsf{H} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{687}$$

separately to each of the bits. Then,

$$\begin{aligned} (\mathsf{H}|x\rangle)(\mathsf{H}|y \oplus f(x)\rangle) &= \frac{(-1)^{f(0)}|0\rangle + (-1)^{f(1)}|1\rangle}{\sqrt{2}} \frac{|0\rangle - |1\rangle}{\sqrt{2}} \\ &= \left(\frac{(-1)^{f(0)} + (-1)^{f(1)}}{2} |0\rangle + \frac{(-1)^{f(0)} - (-1)^{f(1)}}{2} |1\rangle \right) |1\rangle. \end{aligned} \tag{688}$$

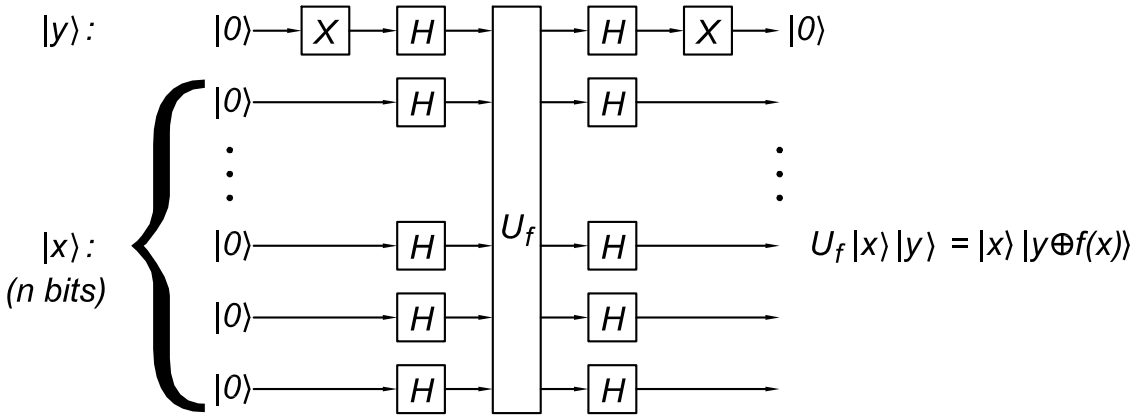
This has the desired result that a measurement of the first bit in the $[|0\rangle, |1\rangle]$ basis yields $|0\rangle$ if $f(0) = f(1)$ and $|1\rangle$ if $f(0) \neq f(1)$.

If we wish to bring the second bit back to its initial state of $|0\rangle$, we could apply an additional NOT transformation to this. A diagram for this algorithm is therefore,



Since this diagram is symmetric and reversible, it is clear that a second application of it would bring the final state of the first application back to the initial state.

(f) An "obvious" generalization of the circuit of part (e) to n input bits is the circuit shown below, which is called the Deutsch-Jozsa algorithm.



The input state consists of the n-bit word $|x\rangle_n$ and the 1-bit state $|y\rangle$. The unitary transformation U_f is

$$\mathsf{U}_f|x\rangle_n|y\rangle = |x\rangle_n|y \oplus f(x)\rangle, \tag{689}$$

where $f(x)$ is a function that maps n bits onto 1 bit.

Based on our experience with Deutsch's algorithm for the case that input state $|x\rangle$ has only 1 bit, we expect that by observing the n-bit output state of $|x\rangle_n$ we may learn something useful about the n-to-1 function f.

We again prepare the state $|y\rangle$ that is input to U_f to be

$$|y\rangle = \mathsf{HX}|0\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}\,. \tag{690}$$

Then, the output state of bit y, after applying the transformation U_f, is again given by eq. (142),

$$|y \oplus f(x)\rangle = \frac{|0 \oplus f(x)\rangle - |1 \oplus f(x)\rangle}{\sqrt{2}} = (-1)^{f(x)}\frac{|0\rangle - |1\rangle}{\sqrt{2}}\,. \tag{691}$$

The state $|x\rangle_n$ that is input to U_f is given by

$$|x\rangle_n = \mathsf{H}^{\otimes n}|0\rangle_n = \frac{1}{2^{n/2}}\sum_{j=0}^{2^n-1}|j\rangle_n, \tag{692}$$

recalling eq. (130). This is a linear combination of all possible n-bit basis states $|j\rangle_n$. Hence, a single operation of U_f on this input state can tell us something about $f(x)$ for all possible values of x.

Using the states (692) and (690) as the input states, we have

$$\mathsf{U}_f(\mathsf{H}^{\otimes n} \otimes \mathsf{HX})|0\rangle_n|0\rangle = \frac{1}{2^{n/2}}\sum_{j=0}^{2^n-1}(-1)^{f(j)}|j\rangle_n\frac{|0\rangle - |1\rangle}{\sqrt{2}}\,. \tag{693}$$

The hint is that this result may be particularly simple if function f is constant, *i.e.*, if $f(x) = f(0)$ for all values of x. In this case all n terms of the output state are multiplied by the same phase factor $(-1)^{f(0)}$,

$$\mathsf{U}_f(\mathsf{H}^{\otimes n} \otimes \mathsf{HX})|0\rangle_n|0\rangle = \frac{(-1)^{f(0)}}{2^{n/2}}\sum_{j=0}^{2^n-1}|j\rangle_n\frac{|0\rangle - |1\rangle}{\sqrt{2}} \qquad (f \text{ is constant}). \tag{694}$$

As shown in the figure on the previous page, we now apply a Hadamard transformation to each of the final-state bits (plus a NOT transformation to the final bit y). Since $\mathsf{H}^2 = \mathbf{I}$, each of the n bits of the state $|x\rangle$ is restored to $|0\rangle$ (and the y bit is also restored to $|0\rangle$). That is,

$$(\mathsf{H}^{\otimes n} \otimes \mathsf{XH})\mathsf{U}_f(\mathsf{H}^{\otimes n} \otimes \mathsf{HX})|0\rangle_n|0\rangle = (-1)^{f(0)}|0\rangle_n|0\rangle \qquad (f \text{ is constant}). \tag{695}$$

If we observe all n final-state bits of $|x\rangle_n$ in the circuit shown on the preceding page, we will find them all to be $|0\rangle$ when function f is constant.

Of course, we might also find that all n final-state bits of $|x\rangle_n$ are $|0\rangle$ for other versions of the function f.

Our task now is to find a restriction on function f so that the circuit can never produce final states $|x\rangle_n$ with all bits $|0\rangle$. Then, a single application of the circuit can distinguish between a constant function f and this restricted class of f.

When we apply the final Hadamard transformations to the general case (693) we obtain

$$(\mathsf{H}^{\otimes n} \otimes \mathsf{XH})\mathsf{U}_f(\mathsf{H}^{\otimes n} \otimes \mathsf{HX})|0\rangle_n|0\rangle = \frac{1}{2^{n/2}} \sum_{j=0}^{2^n-1} (-1)^{f(j)} \mathsf{H}^{\otimes n}|j\rangle_n|0\rangle. \tag{696}$$

Each n-bit basis state $|j\rangle_n$ can be written as a direct product,

$$|j\rangle_n = \prod_{l=0}^{n-1} |j_l\rangle, \tag{16}$$

where j_l is either 0 or 1. Then,

$$\begin{aligned} \mathsf{H}^{\otimes n}|j\rangle_n &= \prod_{l=0}^{n-1} \mathsf{H}|j_l\rangle = \frac{1}{2^{n/2}} \prod_{l=0}^{n-1} (|0\rangle + (-1)^{j_l}|1\rangle) = \frac{1}{2^{n/2}} \prod_{l=0}^{n-1} \sum_{k_l=0}^{1} (-1)^{j_l k_l}|k_l\rangle \\ &= \frac{1}{2^{n/2}} \sum_{k=0}^{2^n-1} (-1)^{\sum_l j_l k_l}|k\rangle_n = \frac{1}{2^{n/2}} \sum_{k=0}^{2^n-1} (-1)^{j\cdot k}|k\rangle_n \\ &= \frac{1}{2^{n/2}} \sum_{k=0}^{2^n-1} (-1)^{j\odot k}|k\rangle_n, \end{aligned} \tag{697}$$

noting that the scalar product $j \cdot k = \sum_l j_l k_l$ has the same effect as the product $j \odot k = \sum_l j_l k_l \pmod 2$ when used as the exponent of -1. Thus,

$$(\mathsf{H}^{\otimes n} \otimes \mathsf{XH})\mathsf{U}_f(\mathsf{H}^{\otimes n} \otimes \mathsf{HX})|0\rangle_n|0\rangle = \frac{1}{2^n} \sum_{j=0}^{2^n-1} (-1)^{f(j)} \sum_{k=0}^{2^n-1} (-1)^{j\odot k}|k\rangle_n|0\rangle. \tag{698}$$

We desire that the result (698) does not contain a term with the state $|k\rangle_n = |0\rangle_n$. For this state the scalar product $j \odot k$ vanishes for all j. Hence, if

$$\sum_{j=0}^{2^n-1} (-1)^{f(j)} = 0, \tag{699}$$

i.e., if $f(j) = 0$ for exactly half of the basis states $|j\rangle$ (and $f(j) = 1$ for the other half), then the function f is balanced and the result (698) does not contain the state $|0\rangle_n$.

Hence, the Deutsch-Jozsa algorithm solves the problem of determining whether the n-to-1 function f is either balanced or constant in a single quantum evaluation of f (provided that we know in advance that f is one or the other of these types).

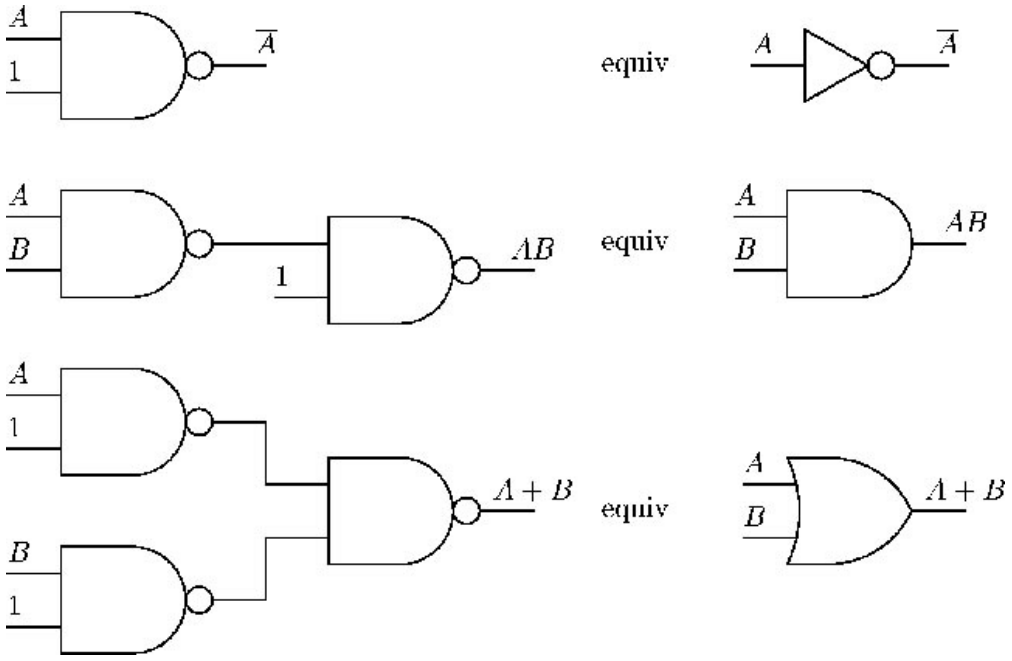
If we are to make this determination by classical evaluations of the function f, we must make $2^{n-1} + 1$ such evaluations. So it is often said that the Deutsch-Jozsa

algorithm shows how a quantum computer could be $2^{n-1}+1$ times faster than a classical computer.

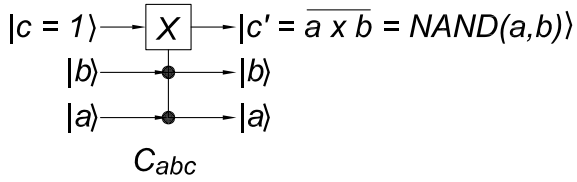
Note however, that if we know that f is either balanced or constant, we know that it is balanced as soon as we find two different values of x that give different values of f. If indeed $f(x)$ is balanced and we make m evaluations of it with different, random values of x, then the probability that all m evaluations yield the same value of f is $2/2^m$. If, say, we are satisfied to know whether f is balanced or constant to 99.9% probability, this can be determined by only 10 classical evaluations no matter how large the word size n is.

11. **Universal Gates for Classical Computation**

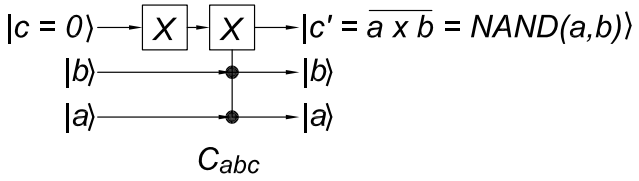
(a) **The 2-Bit NAND Gate is Universal for Classical Computation.**
One way to implement the NOT, AND and OR gates with NAND gates is



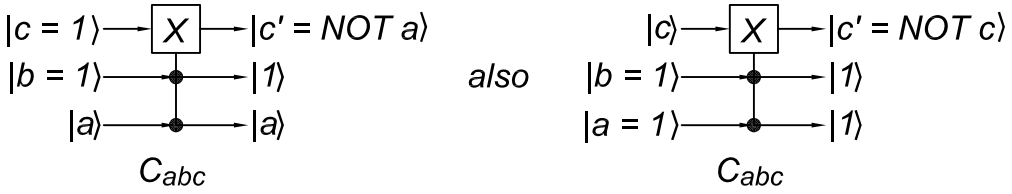
(b) **The Controlled-Controlled-NOT Gate is Universal for Classical Computation.**
A single Controlled-Controlled-NOT gate serves as a NAND gate if its third input is initially $|1\rangle$.



If the third bit is initially $|0\rangle$, then we must be able to add a 1-bit NOT gate before the third input to the Controlled-Controlled-NOT gate.

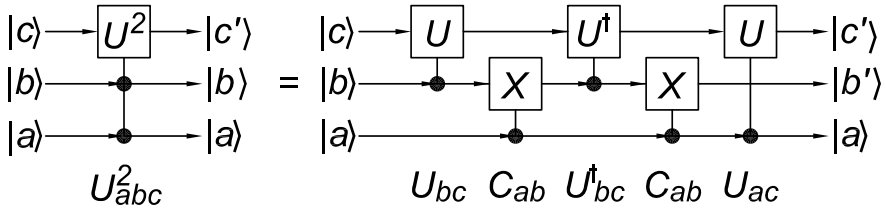


We note that a Controlled-Controlled-NOT gate functions as a NOT gate for its first input if its second and third inputs are $|1\rangle$, but not if they are $|0\rangle$.



(c) **The Controlled-NOT Gate is Universal for Classical Computation.**[164]

i. In the circuit on the right below,



bit b flips twice if bit $a = |1\rangle$ and not at all if bit $a = |0\rangle$; either way bit b ends up unchanged. The effect of the circuit on bit c is

A. If $|a\rangle$ and $|b\rangle$ are both $|0\rangle$, the Controlled-U gates U_{bc}, $\mathsf{U}_{bc}^\dagger$ and U_{ac} are all inactive, so $|c'\rangle = |c\rangle$.

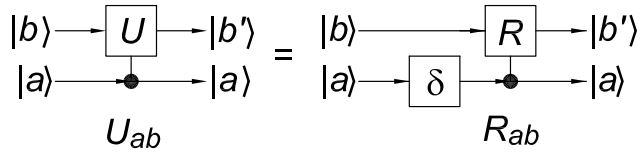
B. If $|a\rangle = |0\rangle$ and $|b\rangle = |1\rangle$, the Controlled-U gates U_{bc} and $\mathsf{U}_{bc}^\dagger$ are active and U_{ac} is inactive, so $|c'\rangle = \mathsf{U}^\dagger\mathsf{U}|c\rangle = |c\rangle$.

C. If $|a\rangle = |1\rangle$ and $|b\rangle = |0\rangle$, the Controlled-U gates $\mathsf{U}_{bc}^\dagger$ and U_{ac} are active and U_{bc} is inactive, so $|c'\rangle = \mathsf{U}\mathsf{U}^\dagger|c\rangle = |c\rangle$.

D. If $|a\rangle = |1\rangle$ and $|b\rangle = |1\rangle$, the Controlled-U gates U_{bc} and U_{ac} are active and $\mathsf{U}_{bc}^\dagger$ is inactive, so $|c'\rangle = \mathsf{U}^2|c\rangle$.

This establishes that the two circuits shown above are the same.

ii. To construct a Controlled-U gate for a general 1-bit unitary operator, $\mathsf{U} = e^{i\delta}\mathsf{R}$, we can use the circuit



where the $\boldsymbol{\delta}$ gate is given by

$$\boldsymbol{\delta} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\delta} \end{pmatrix}. \tag{700}$$

To see this, we note that the Controlled-R gate behaves as

$$\begin{aligned} |0\rangle|b\rangle &\rightarrow |0\rangle|b\rangle, \\ |1\rangle|b\rangle &\rightarrow |1\rangle\mathsf{R}|b\rangle. \end{aligned} \tag{701}$$

When the $\boldsymbol{\delta}$ operator is applied to bit a, $|0\rangle$ remains $|0\rangle$, while $|1\rangle$ becomes $e^{i\delta}|1\rangle$. When the phased-shifted bit a is presented to the Controlled-R operator, the output is

$$\begin{aligned} |0\rangle|b\rangle &\rightarrow |0\rangle|b\rangle, \\ e^{i\delta}|1\rangle|b\rangle &\rightarrow e^{i\delta}|1\rangle\mathsf{R}|b\rangle = |1\rangle e^{i\delta}\mathsf{R}|b\rangle = |1\rangle\mathsf{U}|b\rangle. \end{aligned} \tag{702}$$

Thus, the output bit b is exactly as expected for a Controlled-U operator, as desired. (This conclusion is clearer if we measure the final bit a in the [0,1] basis.)

[164] http://kirkmcd.princeton.edu/examples/QM/barenco_pra_52_3457_95.pdf

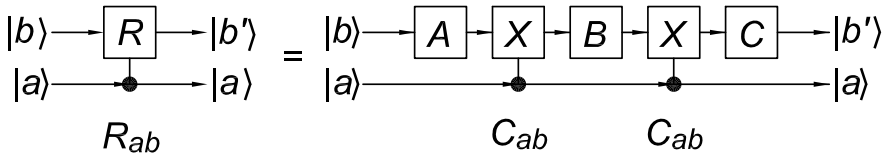
Recalling that the number operator n has the matrix form

$$\mathsf{n} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \tag{703}$$

and that $\mathsf{n}^2 = \mathsf{n}$, we see that

$$\begin{aligned} e^{i\delta\mathsf{n}} &= \mathbf{I} + \sum_{k=1}^{\infty} \frac{(i\delta\mathsf{n})^k}{k!} = \mathbf{I} + \mathsf{n}\sum_{k=1}^{\infty} \frac{(i\delta)^k}{k!} = \mathbf{I} + \mathsf{n}(e^{i\delta} - 1) \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 + (e^{i\delta} - 1) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\delta} \end{pmatrix} = \boldsymbol{\delta}. \end{aligned} \tag{704}$$

iii. In the circuit on the right below,



we obtain $|b'\rangle = \mathsf{CBA}|b\rangle$ if $|a\rangle = |0\rangle$, and $|b'\rangle = \mathsf{CXBXA}|b\rangle$ if $|a\rangle = |1\rangle$. For this circuit to function as Controlled-R, we need

$$\mathsf{CBA} = \mathbf{I}, \tag{705}$$

$$\mathsf{CXBXA} = \mathsf{R}. \tag{706}$$

The rotation operator R can be decomposed as the product of 3 rotations according to eq. (44) as

$$\mathsf{R} = \mathsf{R}_z(\gamma)\mathsf{R}_y(\beta)\mathsf{R}_z(\alpha). \tag{707}$$

Likewise, we take the operators A, B and C to be special unitary operators with similar decompositions into products of rotations. Thus writing

$$\mathsf{B} = \mathsf{R}_z(\gamma_B)\mathsf{R}_y(\beta_B)\mathsf{R}_z(\alpha_B), \tag{708}$$

eq. (57) tells us that

$$\mathsf{XBX} = \mathsf{R}_z(-\gamma_B)\mathsf{R}_y(-\beta_B)\mathsf{R}_z(-\alpha_B). \tag{709}$$

The conditions (705)-(706) can now be expressed as

$$\begin{aligned} \mathbf{I} &= \mathsf{R}_z(\gamma_C)\mathsf{R}_y(\beta_C)\mathsf{R}_z(\alpha_C)\mathsf{R}_z(\gamma_B)\mathsf{R}_y(\beta_B)\mathsf{R}_z(\alpha_B)\mathsf{R}_z(\gamma_A)\mathsf{R}_y(\beta_A)\mathsf{R}_z(\alpha_A) \\ &= \mathsf{R}_z(\gamma_C)\mathsf{R}_y(\beta_C)\mathsf{R}_z(\alpha_C + \gamma_B)\mathsf{R}_y(\beta_B)\mathsf{R}_z(\alpha_B + \gamma_A)\mathsf{R}_y(\beta_A)\mathsf{R}_z(\alpha_A) \qquad (710) \\ \mathsf{CXBXA} &= \mathsf{R} = \mathsf{R}_z(\gamma)\mathsf{R}_y(\beta)\mathsf{R}_z(\alpha) \qquad (711) \\ &= \mathsf{R}_z(\gamma_C)\mathsf{R}_y(\beta_C)\mathsf{R}_z(\alpha_C)\mathsf{R}_z(-\gamma_B)\mathsf{R}_y(-\beta_B)\mathsf{R}_z(-\alpha_B)\mathsf{R}_z(\gamma_A)\mathsf{R}_y(\beta_A)\mathsf{R}_z(\alpha_A) \\ &= \mathsf{R}_z(\gamma_C)\mathsf{R}_y(\beta_C)\mathsf{R}_z(\alpha_C - \gamma_B)\mathsf{R}_y(-\beta_B)\mathsf{R}_z(-\alpha_B + \gamma_A)\mathsf{R}_y(\beta_A)\mathsf{R}_z(\alpha_A). \end{aligned}$$

We first consider eq. (711). While we can simply combine two rotations about y, or two rotations about z, we cannot simply combine a rotation about y

with a rotation about z. Thus, to collapse the last line of eq. (711) onto the first, we will have to take some rotations to be trivial. For example, we set

$$\beta_A = 0, \qquad \text{and} \qquad \alpha_C - \gamma_B = 0. \tag{712}$$

Then, eq. (711) reduces to

$$\mathsf{R}_z(\gamma)\mathsf{R}_y(\beta)\mathsf{R}_z(\alpha) = \mathsf{R}_z(\gamma_C)\mathsf{R}_y(\beta_C - \beta_B)\mathsf{R}_z(-\alpha_B + \gamma_A + \alpha_A). \tag{713}$$

This equation is satisfied provided that

$$\gamma_C = \gamma, \qquad \beta_C - \beta_B = \beta, \qquad \text{and} \qquad -\alpha_B + \gamma_A + \alpha_A = \alpha. \tag{714}$$

Our knowledge thus far can be summarized as

$$\begin{array}{lll} \alpha_C = \gamma_B, & \alpha_B = \alpha_A + \gamma_A - \alpha, & \alpha_A = ? \\ \beta_C = \beta_B + \beta, & \beta_B = ? & \beta_A = 0, \\ \gamma_C = \gamma, & \gamma_B = ? & \gamma_A = ? \end{array} \tag{715}$$

With this, eq. (710) becomes

$$\begin{aligned} \mathbf{I} &= \mathsf{R}_z(\gamma)\mathsf{R}_y(\beta_B + \beta)\mathsf{R}_z(2\gamma_B)\mathsf{R}_y(\beta_B)\mathsf{R}_z(\alpha_A + \gamma_A - \alpha)\mathsf{R}_z(\gamma_A)\mathsf{R}_z(\alpha_A) \\ &= \mathsf{R}_z(\gamma)\mathsf{R}_y(\beta_B + \beta)\mathsf{R}_z(2\gamma_B)\mathsf{R}_y(\beta_B)\mathsf{R}_z(2\alpha_A + 2\gamma_A - \alpha). \end{aligned} \tag{716}$$

To collapse this further, we set

$$\gamma_B = 0 = \alpha_C. \tag{717}$$

Equation (716) is now

$$\mathbf{I} = \mathsf{R}_z(\gamma)\mathsf{R}_y(2\beta_B + \beta)\mathsf{R}_z(2\alpha_A + 2\gamma_A - \alpha). \tag{718}$$

To collapse this to the identity, we set

$$\beta_B = -\frac{\beta}{2}, \qquad \alpha_A + \gamma_A = \frac{\alpha - \gamma}{2}. \tag{719}$$

The second equation of this can be satisfied by taking

$$\alpha_A = \frac{\alpha}{2}, \qquad \gamma_A = -\frac{\gamma}{2}. \tag{720}$$

The solution is

$$\begin{array}{lll} \alpha_C = 0, & \alpha_B = -\frac{\alpha+\gamma}{2}, & \alpha_A = \frac{\alpha}{2}, \\ \beta_C = \frac{\beta}{2}, & \beta_B = -\frac{\beta}{2}, & \beta_A = 0, \\ \gamma_C = \gamma, & \gamma_B = 0, & \gamma_A = -\frac{\gamma}{2}, \end{array} \tag{721}$$

$$\mathsf{A} = \mathsf{R}_z\left(\frac{\alpha-\gamma}{2}\right), \tag{722}$$

$$\mathsf{B} = \mathsf{R}_y\left(-\frac{\beta}{2}\right)\mathsf{R}_z\left(-\frac{\alpha+\gamma}{2}\right), \tag{723}$$

$$\mathsf{C} = \mathsf{R}_z(\gamma)\mathsf{R}_y\left(\frac{\beta}{2}\right). \tag{724}$$

As a check,

$$\begin{aligned}\mathsf{CBA} &= \mathsf{R}_z(\gamma)\mathsf{R}_y\left(\frac{\beta}{2}\right)\cdot\mathsf{R}_y\left(-\frac{\beta}{2}\right)\mathsf{R}_z\left(-\frac{\alpha+\gamma}{2}\right)\cdot\mathsf{R}_z\left(\frac{\alpha-\gamma}{2}\right)\\ &= \mathsf{R}_z(\gamma)\mathsf{R}_z(-\gamma) = \mathbf{I}, \qquad (725)\\ \mathsf{CXBXA} &= \mathsf{R}_z(\gamma)\mathsf{R}_y\left(\frac{\beta}{2}\right)\cdot\mathsf{R}_y\left(\frac{\beta}{2}\right)\mathsf{R}_z\left(\frac{\alpha+\gamma}{2}\right)\cdot\mathsf{R}_z\left(\frac{\alpha-\gamma}{2}\right)\\ &= \mathsf{R}_z(\gamma)\mathsf{R}_y(\beta)\mathsf{R}_z(\alpha) = \mathsf{R}. \qquad (726)\end{aligned}$$

This establishes that the two circuits in the figure 2 pages ago are identical.

iv. Using eq. (449), the NOT operator can be written.

$$\mathsf{X} = \boldsymbol{\sigma}_x = -i\mathsf{R}_x(\pi) = e^{-i\pi/2}\mathsf{R}_x(\pi). \tag{727}$$

Taking the square root, we obtain

$$\begin{aligned}\sqrt{\mathsf{X}} &= e^{-i\pi/4}\mathsf{R}_x(\pi/2) = \frac{1-i}{\sqrt{2}}\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}}\\ \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1-i & 1+i\\ 1+i & 1-i\end{pmatrix}\\ &= e^{-i\pi/4}R_z(\pi/2)R_y(\pi/2)R_z(-\pi/2), \qquad (728)\end{aligned}$$

recalling eq. (453). To represent $\sqrt{\mathsf{X}}$ in the general form $e^{-i\delta}R_z(\gamma)R_y(\beta)R_z(\alpha)$ we use

$$\alpha = -\frac{\pi}{2}, \qquad \beta = \frac{\pi}{2}, \qquad \gamma = \frac{\pi}{2}, \qquad \delta = -\frac{\pi}{4}. \tag{729}$$

The 1-bit operators needed for the construction of the Controlled-$\sqrt{\mathsf{X}}$ operator as in sec. ii-iii are

$$\mathsf{A} = \mathsf{R}_z\left(\frac{\alpha-\gamma}{2}\right) = \mathsf{R}_z(-\pi), \tag{730}$$

$$\mathsf{B} = \mathsf{R}_y\left(-\frac{\beta}{2}\right)\mathsf{R}_z\left(-\frac{\alpha+\gamma}{2}\right) = \mathsf{R}_y\left(-\frac{\pi}{4}\right), \tag{731}$$

$$\mathsf{C} = \mathsf{R}_z(\gamma)\mathsf{R}_y\left(\frac{\beta}{2}\right) = \mathsf{R}_z(\pi)\mathsf{R}_y\left(\frac{\pi}{4}\right), \tag{732}$$

$$\boldsymbol{\delta} = \begin{pmatrix}1 & 0\\ 0 & e^{i\delta}\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & e^{-i\pi/4}\end{pmatrix}. \tag{733}$$

12. Universal Gates for Quantum Computation

(a) All 1-Bit Quantum Gates Can Be Built from the H and $\boldsymbol{\sigma}_z^{1/4}$ Gates.[165]

$$\begin{aligned}
\boldsymbol{\sigma}_z^{-1/4}\boldsymbol{\sigma}_x^{1/4} &= \boldsymbol{\sigma}_z^{-1/4}\mathsf{H}\boldsymbol{\sigma}_z^{1/4}\mathsf{H} \\
&= \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & e^{-i\pi/4} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & e^{-i\pi/4} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ e^{i\pi/4} & -e^{i\pi/4} \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & e^{-i\pi/4} \end{pmatrix}\begin{pmatrix} 1+e^{i\pi/4} & 1-e^{i\pi/4} \\ 1-e^{i\pi/4} & 1+e^{i\pi/4} \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} 1+e^{i\pi/4} & 1-e^{i\pi/4} \\ e^{-i\pi/4}-1 & e^{-i\pi/4}+1 \end{pmatrix} = \begin{pmatrix} e^{i\pi/8}\cos\frac{\pi}{8} & -ie^{i\pi/8}\sin\frac{\pi}{8} \\ -ie^{-i\pi/8}\sin\frac{\pi}{8} & e^{-i\pi/8}\cos\frac{\pi}{8} \end{pmatrix} \\
&= \begin{pmatrix} \cos^2\frac{\pi}{8}+i\sin\frac{\pi}{8}\cos\frac{\pi}{8} & -i\sin\frac{\pi}{8}(\cos\frac{\pi}{8}+i\sin\frac{\pi}{8}) \\ -i\sin\frac{\pi}{8}(\cos\frac{\pi}{8}-i\sin\frac{\pi}{8}) & \cos^2\frac{\pi}{8}-i\sin\frac{\pi}{8}\cos\frac{\pi}{8} \end{pmatrix}. \qquad (734)
\end{aligned}$$

We desire this to be of the form

$$e^{i\frac{\theta}{2}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}} = \cos\frac{\theta}{2}\,\mathbf{I} + i\sin\frac{\theta}{2}(u_x\boldsymbol{\sigma}_x + u_y\boldsymbol{\sigma}_y + u_z\boldsymbol{\sigma}_z). \qquad (735)$$

Clearly we set

$$\cos\frac{\theta}{2} = \cos^2\frac{\pi}{8} = \frac{1+\cos\frac{\pi}{4}}{2} = \frac{2+\sqrt{2}}{4}. \qquad (736)$$

Corresponding to this, we have

$$\sin\frac{\theta}{2} = \sqrt{1-\cos^2\frac{\theta}{2}} = \sqrt{1-\cos^4\frac{\pi}{8}} = \sin\frac{\pi}{8}\sqrt{1+\cos^2\frac{\pi}{8}}. \qquad (737)$$

Using eqs. (736)-(737) in eq. (734), we can write

$$\begin{aligned}
\boldsymbol{\sigma}_z^{-1/4}\boldsymbol{\sigma}_x^{1/4} &= \begin{pmatrix} \cos\frac{\theta}{2}+i\sin\frac{\theta}{2}\frac{\cos\frac{\pi}{8}}{\sqrt{1+\cos^2\frac{\pi}{8}}} & i\sin\frac{\theta}{2}\frac{-\cos\frac{\pi}{8}-i\sin\frac{\pi}{8}}{\sqrt{1+\cos^2\frac{\pi}{8}}} \\ i\sin\frac{\theta}{2}\frac{-\cos\frac{\pi}{8}+i\sin\frac{\pi}{8}}{\sqrt{1+\cos^2\frac{\pi}{8}}} & \cos\frac{\theta}{2}+i\sin\frac{\theta}{2}\frac{-\cos\frac{\pi}{8}}{\sqrt{1+\cos^2\frac{\pi}{8}}} \end{pmatrix} \\
&= \cos\frac{\theta}{2}\,\mathbf{I} + i\sin\frac{\theta}{2}\left(\frac{-\cos\frac{\pi}{8}\,\boldsymbol{\sigma}_x + \sin\frac{\pi}{8}\,\boldsymbol{\sigma}_y + \cos\frac{\pi}{8}\,\boldsymbol{\sigma}_z}{\sqrt{1+\cos^2\frac{\pi}{8}}}\right). \qquad (738)
\end{aligned}$$

Comparing with eq. (735), we see that the axis of rotation has unit vector

$$\hat{\mathbf{u}} = \frac{(-\cos\frac{\pi}{8},\ \sin\frac{\pi}{8},\ \cos\frac{\pi}{8})}{\sqrt{1+\cos^2\frac{\pi}{8}}}, \qquad (739)$$

[165] http://kirkmcd.princeton.edu/examples/QM/boykin_quant-ph-9906054.pdf

which obeys the desired condition that $u_z = -u_x$.
Then from eq. (506) of Prob. 4(f), we can write

$$\mathsf{H}^{-1/2}\ \boldsymbol{\sigma}_z^{-1/4}\boldsymbol{\sigma}_x^{1/4}\ \mathsf{H}^{1/2} = e^{i\frac{\theta}{2}\hat{\mathbf{v}}\cdot\boldsymbol{\sigma}}, \tag{740}$$

where the unit vector $\hat{\mathbf{v}}$ is given by

$$\begin{aligned}\hat{\mathbf{v}} &= \frac{(u_x+\sqrt{2}u_y+u_z, -\sqrt{2}(u_x-\ u_z), u_x-\sqrt{2}u_y+u_z)}{2} \\ &= \frac{(\sin\frac{\pi}{8},\ 2\cos\frac{\pi}{8},\ -\sin\frac{\pi}{8})}{\sqrt{2\left(1+\cos^2\frac{\pi}{8}\right)}},\end{aligned} \tag{741}$$

which is seen to be orthogonal to $\hat{\mathbf{u}}$ of eq. (739).
Number theory experts may be amused to note that $e^{i\theta}$, for θ given by eq. (736), is a root of the quartic equation

$$4x^4+4x^3+x^2+4x+4=0, \tag{742}$$

which apparently tells some of us that θ/π is irrational.[166]

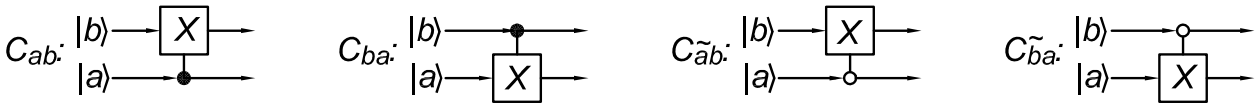
(b) **Generalized Controlled-NOT Operations.**[167]
The truth table of the 2-bit Controlled-NOT operator $\mathsf{C}_{\tilde{a}b}$ is

$\mathsf{C}_{\tilde{a}b}$:

a	b	a'	b'
0	0	0	1
0	1	0	0
1	0	0	1
1	1	1	1

(743)

From this we read off the elements of the 4×4 matrix representation as shown below, and similarly for the other three 2-bit Controlled-NOT operators.

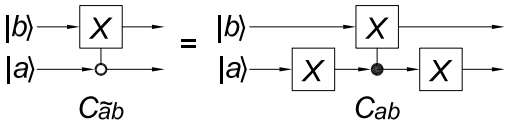


$$\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix},\quad \begin{pmatrix}1&0&0&0\\0&0&0&1\\0&0&1&0\\0&1&0&0\end{pmatrix},\quad \begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix},\quad \begin{pmatrix}0&0&1&0\\0&1&0&0\\1&0&0&0\\0&0&0&1\end{pmatrix}. \tag{744}$$

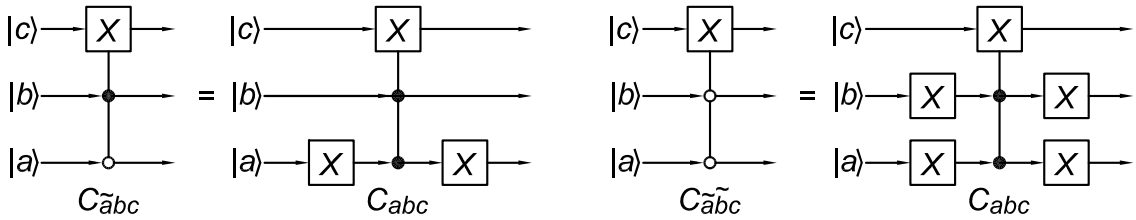
The four Controlled-NOT operators are all permutations of the identity matrix in which a pair of rows (or columns) are swapped.

[166] See also, prob. 7.4 of
http://kirkmcd.princeton.edu/examples/QM/Preskill/prob7_01.ps
[167] http://kirkmcd.princeton.edu/examples/QM/barenco_pra_52_3457_95.pdf

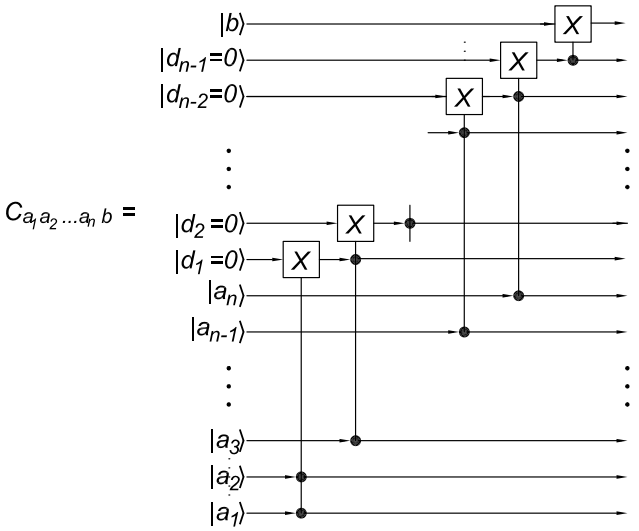
The 2-bit operator $\mathsf{C}_{\tilde{a}b}$ can be built from the regular Controlled-NOT operator C_{ab} if we flip bit a first, using a 1-bit NOT operation. Then, at the end we flip bit a again to restore its initial value.



Clearly this trick can be used for Controlled-NOT operators with any number of control bits. For example,



A Controlled-NOT operation with n control bits can be built up out of a sequence of $n-1$ Controlled-Controlled-NOT operations, together with $n-1$ auxiliary bits that are initially $|0\rangle$ as shown below. In this construction, the auxiliary bit d_j is flipped to $|1\rangle$ only when the control bits a_1, a_2, ... a_{j+1} are all $|1\rangle$. Thus the last auxiliary bit d_{n-1} serves as the desired control bit to convert the 2-bit Controlled-NOT gate into an n-control-bit Controlled-NOT gate.

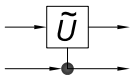


(c) **Two-Level Unitary Matrices.**

We first construct the two-level matrix

$$\begin{pmatrix} a & 0 & 0 & b \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ c & 0 & 0 & d \end{pmatrix} \tag{745}$$

from the Controlled-$\tilde{\mathsf{U}}$ operator



$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & a & b \\ 0 & 0 & c & d \end{pmatrix}. \tag{746}$$

Note that multiplying matrix (745) on the left by another matrix will redistribute the elements of matrix (745) within columns, but does not mix matrix elements within rows. Similarly, multiplication of matrix (745) by a matrix on the right mixes its elements within rows but not within columns. Since to get from matrix (745) to matrix (746) we need to move elements in both rows and columns, we will need to multiply matrix (745) on both the left and on the right.

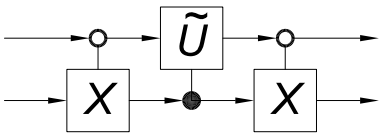
The realization of matrix multiplication by a sequence of operation corresponds to performing the rightmost matrix multiplication first, so we begin with a right multiplication with the goal of moving elements a and c from the 3rd column into the first by swapping the 1st of 3rd columns of matrix (745). After a little reflection we see that this can be accomplished via right multiplication of matrix (745) by a permutation of the unit matrix in which its 1st and 3rd columns have been swapped.

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & a & b \\ 0 & 0 & c & d \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ a & 0 & 0 & b \\ c & 0 & 0 & d \end{pmatrix}. \tag{747}$$

Next, we perform a left multiplication of eq. (747) with the goal of swapping the 1st and 3rd rows. The appropriate matrix of this is a permutation of the unit matrix in which the 1st and 3rd rows have been swapped (which is the same as if the 1st and 3rd columns had been swapped). Thus,

$$\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ a & 0 & 0 & b \\ c & 0 & 0 & d \end{pmatrix} = \begin{pmatrix} a & 0 & 0 & b \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ c & 0 & 0 & d \end{pmatrix}. \tag{748}$$

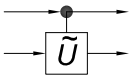
From eq. (744) we recognized our permutation of the unit matrix as $\mathsf{C}_{\tilde{b}a}$, so that our construction of matrix (745) can be represented by the bit-flow diagram,



Similarly, to construct the two-level matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & a & b & 0 \\ 0 & c & d & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{749}$$

from the Controlled-$\tilde{\mathsf{U}}$ operator

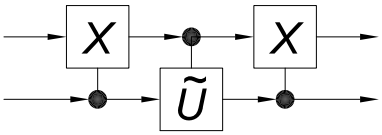


$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & a & 0 & b \\ 0 & 0 & 1 & 0 \\ 0 & c & 0 & d \end{pmatrix}, \tag{750}$$

we need to multiply matrix (749) on the left and right by the permutation of the unit matrix in which the 3rd and 4th columns (or rows) are swapped,

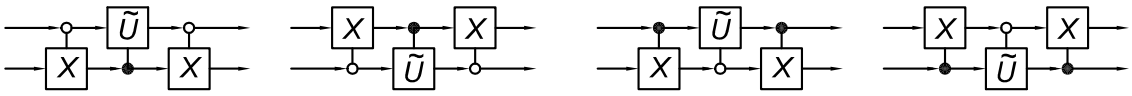
$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \tag{751}$$

which is, of course, just the basic Controlled-NOT operator C_{ab}. Thus, the bit-flow diagram to produce matrix (749) is

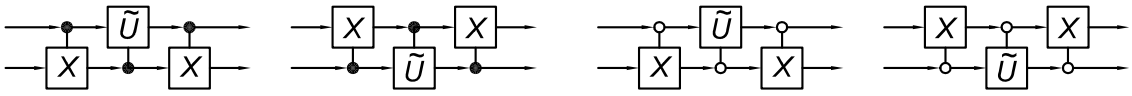


When $\tilde{\mathsf{U}} = \mathsf{X}$, we see that U of eq. (749) is the SWAP operation S_{ab} introduced in prob. 9(c) and written as a 4×4 matrix in eq. (572). The bit-flow diagram following eq. (576) is a special case of the diagram immediately above.

With the spirit of the above constructions in mind, we readily see that the two-level matrix (745) can be built using any of the four Controlled-$\tilde{\mathsf{U}}$ operators together with appropriate Controlled-NOT operators,



and likewise the two-level matrix (749) can be built four ways,



We end this part with a brief summary of the procedure to build up any $2^n \times 2^n$ two-level matrix U from the 1-bit unitary matrix $\tilde{\mathsf{U}}$.

First, we note that a generalized Controlled-NOT operator is a two-level operator that flips a single bit of an n-bit state if the other bits match the control-bit pattern. These operators are permutations of the identity matrix in which one pair of rows (or columns), say rows j and k, have been swapped. The effect of applying this Controlled-NOT operator on the left of an arbitrary matrix U is to swap its rows j and k. Or, if the Controlled-NOT operator is applied on the right of U, its columns j and k will be swapped. Then, applying the Controlled-NOT operator to both the left and right of U will swap both its rows and columns j and k.

Similarly, a generalized Controlled-$\tilde{\mathsf{U}}$ operator is a two-level operator that applies an arbitrary 1-bit operator $\tilde{\mathsf{U}}$ to a specified bit of an n-bit state if the other bits match the control-bit pattern.

Turning to a general $2^n \times 2^n$ two-level matrix U such as shown in eq. (752), we consider the nontrivial row indices j and k as binary numbers. If these numbers are the same except for one bit being a 0 in one case and a 1 in the other, then U is the same as a Controlled-$\tilde{\mathsf{U}}$ operator where the control bits are equal to the common bits in the binary representation of j and k.

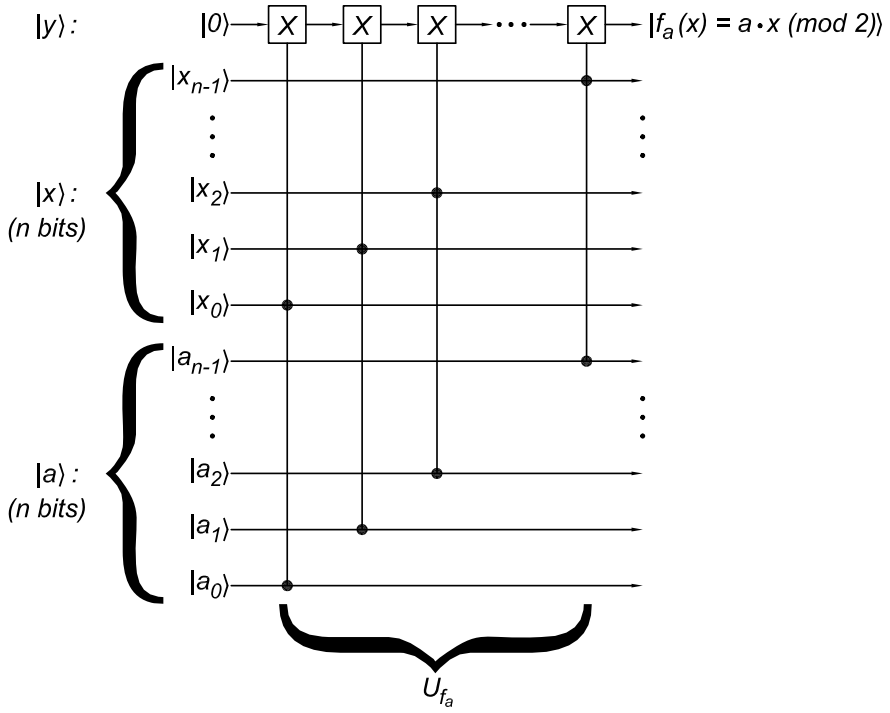
$$\mathsf{U}_{\text{two-level}} = \begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & a & \dots & 0 & \dots & b & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & c & \dots & 0 & \dots & d & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix}, \qquad \tilde{\mathsf{U}} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \quad (752)$$

If the binary representations of j and k differ in more than one bit, then we need to perform a sequence of bit flips on, say, index j, using appropriate generalized Controlled-NOT operators until the flipped j differs from k in only one digit. We can then construct U using the Controlled-$\tilde{\mathsf{U}}$ operator associated with the flipped j, together with the sequence of generalized Controlled-NOT operators needed to produce the flipped j. This sequence of operators must be applied symmetrically to both the left and right of the Controlled-$\tilde{\mathsf{U}}$ operator so as to move both the rows and columns from their location in the Controlled-$\tilde{\mathsf{U}}$ operator to their desired location in U.

When I wrote the above, I believed that I understood it. But can anyone else follow it?

13. **The Bernstein-Vazirani Problem**

(a) To implement $f_a(x) = a \cdot x \pmod 2$ we recall that a Controlled-Controlled-NOT operation flips the target bit if the product of the control bits is 1. Also, addition modulo 2 on a set of bits can be performed by flipping the sum bit once for each 1 bit in the set. Hence, a sequence of Controlled-Controlled-NOT operations whose control bits are corresponding bits of $|a\rangle$ and $|x\rangle$, and whose target bit is $|y\rangle$, performs the function $f_a(x)$.

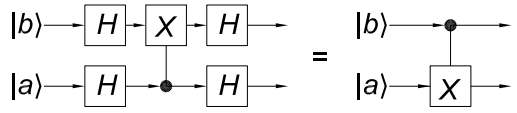


If we regard a as fixed, the above circuit could be simplified by replacing the Controlled-Controlled-NOT gates by Controlled-NOT gates, which are only needed for a bit $|x_j\rangle$ if $a_j = 1$.

If the initial state of bit $|y\rangle$ had been $|1\rangle$, the final state of bit $|y\rangle$ would be flipped compared to the case that $|y\rangle = |0\rangle$ initially. That is, in general, $|y\rangle \to |y \oplus f_a(x)\rangle$. Summarizing the entire process as a unitary operator U_{f_a}, we have

$$\mathsf{U}_{f_a}|a\rangle|x\rangle|y\rangle = |a\rangle|x\rangle|y \oplus f_a(x)\rangle. \tag{166}$$

(b) To demonstrate the identity



we note that applying the Hadamard operation,

$$\mathsf{H} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{753}$$

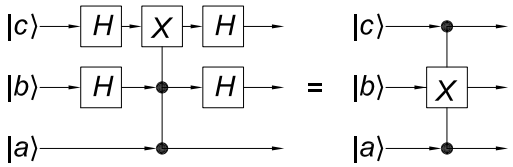
to both input lines is described by the 4×4 matrix (recall eq. (647))

$$\mathsf{H}^{\otimes 2} = \mathsf{H} \otimes \mathsf{H} = \frac{1}{2} \left(\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ \hline 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{array} \right). \tag{754}$$

Then, recalling eqs. (552) and (574), we have

$$\begin{aligned} \mathsf{H}^{\otimes 2}\mathsf{C}_{ab}\mathsf{H}^{\otimes 2} &= \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \\ &= \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = \mathsf{C}_{ba}. \end{aligned} \tag{755}$$

It is perhaps now obvious that



but we can verify this explicitly using

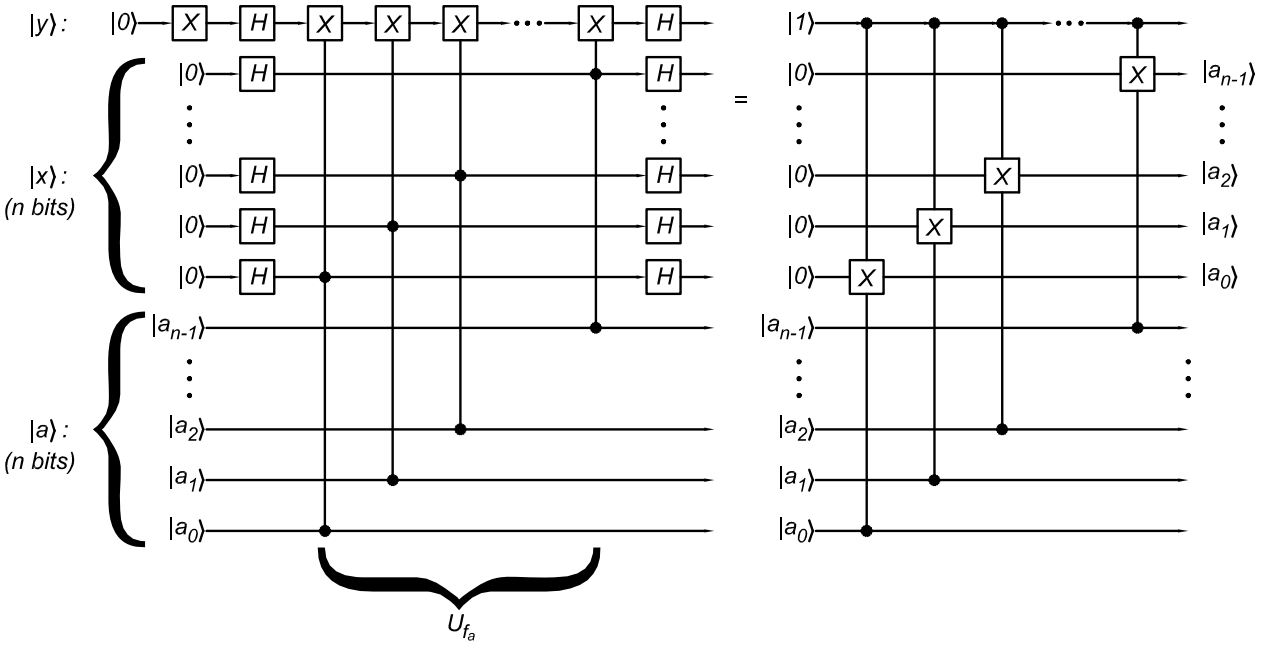
$$\mathbf{I}_a \otimes \mathsf{H}_b \otimes \mathsf{H}_c = \frac{1}{2} \left(\begin{array}{cccc|cccc} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 1 & -1 & 0 & 0 & 0 & 0 \\ 1 & 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 1 & -1 & -1 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 & -1 & 1 \end{array} \right). \tag{756}$$

Recalling eqs. (649)-(650) for the Controlled-Controlled-NOT operation C_{abc}, we have

$$(\mathbf{I}_a \otimes \mathsf{H}_b \otimes \mathsf{H}_c)\mathsf{C}_{abc}(\mathbf{I}_a \otimes \mathsf{H}_b \otimes \mathsf{H}_c) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} = \mathsf{C}_{acb}. \quad (757)$$

We now readily see that applying Hadamard gates at the beginning and end of all the x bit lines in our construction for the Bernstein-Vazirani function $f_a(x)$ leads to the appearance of a "copy" of a at the output of the x lines, as shown in the figure below. We must also add a NOT gate on the input of the bottom line, to provide a control bit with the value 1.

This procedure cannot, of course, make an exact copy of an arbitrary quantum state $|a\rangle$, but it does copy a Cbit a correctly.



We see that the Bernstein-Vazirani prodedure is an elaborate disguise of an n-bit copy operation as a function evaluation.

(c) Using the general form

$$\mathsf{H}^{\otimes n}|j\rangle_n = \frac{1}{2^{n/2}} \sum_{k=0}^{2^n-1} (-1)^{j\odot k}|k\rangle_n, \quad (145)$$

which in particular tells us that

$$\mathsf{H}^{\otimes n}|0\rangle_n = \frac{1}{2^{n/2}} \sum_{j=0}^{2^n-1} |j\rangle_n, \quad (130)$$

along with the relations

$$U_{f_a}|a\rangle|x\rangle|y\rangle = |a\rangle|x\rangle|y \oplus f_a(x)\rangle = |a\rangle|x\rangle|y \oplus (a \odot x)\rangle, \tag{166}$$

and

$$\frac{|0 \oplus f(x)\rangle - |1 \oplus f(x)\rangle}{\sqrt{2}} = (-1)^{f(x)} \frac{|0\rangle - |1\rangle}{\sqrt{2}}, \tag{142}$$

the above bit-flow diagram can be represented algebraically as the sequence of operations

$$\begin{aligned}
(\mathsf{H}_x^{\otimes n} \otimes \mathsf{H}_y)\mathsf{U}_{f_a}(\mathsf{H}_x^{\otimes n} \otimes \mathsf{H}_y\mathsf{X}_y)|0\rangle_x|0\rangle_y &= (\mathsf{H}_x^{\otimes n} \otimes \mathsf{H}_y)\mathsf{U}_{f_a} \frac{1}{2^{n/2}} \sum_{j=0}^{2^n-1} |j\rangle_x \left(\frac{|0\rangle_y - |1\rangle_y}{\sqrt{2}} \right) \\
&= (\mathsf{H}_x^{\otimes n} \otimes \mathsf{H}_y) \frac{1}{2^{n/2}} \sum_{j=0}^{2^n-1} (-1)^{f_a(j)} |j\rangle_x \left(\frac{|0\rangle_y - |1\rangle_y}{\sqrt{2}} \right) \\
&= \mathsf{H}_x^{\otimes n} \frac{1}{2^{n/2}} \sum_{j=0}^{2^n-1} (-1)^{a \odot j} |j\rangle_x |1\rangle_y \\
&= \mathsf{H}_x^{\otimes n} \mathsf{H}_x^{\otimes n} |a\rangle_x |1\rangle_y \\
&= |a\rangle_x |1\rangle_y,
\end{aligned} \tag{758}$$

noting that $\mathsf{H}^2 = \mathbf{I}$, so that $(\mathsf{H}_x^{\otimes n})^2 = \mathbf{I}_n$. Hence, the output lines of $|x\rangle$ are in the desired state $|a\rangle$ at the end of the operation, as we also found via the bit-flow diagram of part (b).

As a side remark, we note that the 3rd and 5th lines of eq. (758) also tell us that

$$|a\rangle = \mathsf{H}_x^{\otimes n} \frac{1}{2^{n/2}} \sum_{j=0}^{2^n-1} (-1)^{j \odot a} |j\rangle_n = \sum_{k=0}^{2^n-1} \frac{1}{2^n} \sum_{j=0}^{2^n-1} (-1)^{j \odot k} (-1)^{j \odot a} |k\rangle_n. \tag{759}$$

Hence, we must have the identity

$$\frac{1}{2^n} \sum_{j=0}^{2^n-1} (-1)^{j \odot k} (-1)^{a \odot j} = \delta_{ak}. \tag{168}$$

To verify this explicitly, we write $j = \sum_{l=0}^{n-1} j_l 2^l$ where where j_l is either 0 or 1. Thus,

$$\begin{aligned}
\frac{1}{2^n} \sum_{j=0}^{2^n-1} (-1)^{j \odot k} (-1)^{a \odot j} &= \frac{1}{2^n} \sum_{j=0}^{2^n-1} (-1)^{j \odot (a \oplus k)} = \frac{1}{2^n} \sum_{j=0}^{2^n-1} (-1)^{\sum_l j_l (a_l \oplus k_l)} \\
&= \frac{1}{2^n} \sum_{j=0}^{2^n-1} \prod_{l=0}^{n-1} (-1)^{j_l (a_l \oplus k_l)} = \frac{1}{2^n} \prod_{l=0}^{n-1} \sum_{j_l=0}^{1} (-1)^{j_l (a_l \oplus k_l)}.
\end{aligned} \tag{760}$$

If k does not equal a then $a_l \oplus k_l = 1$ for a nonzero number of indices l. Whenever number j has $j_l = 1$ for exactly one of those indices, the sum in the righthand side of eq. (760) vanishes. Hence, eq. (760) is zero unless $a = k$, in which case it is equal to 1.

14. **Simon's Problem**

(a) We can obtain the value of a from the addition $a = x \oplus y$ once we have found two numbers x and y such that $f_a(x) = f_a(y)$. If we pick x and y at random among the 2^n n-bit numbers, we will need to sample roughly a such pairs to have a good probability of finding one pair for which $f_a(x) = f_a(y)$. Since a is a number of order 2^n, we must make of order 2^n evaluations of the function f_a to solve the problem classically.

Therefore, we can say that Simon's problem is exponentially hard to solve via classical computation.

(b) To analyze the operation

$$\mathsf{H}_x^{\otimes n}\mathsf{U}_{f_a}\mathsf{H}_x^{\otimes n}|0\rangle_x|0\rangle_y \tag{174}$$

where

$$\mathsf{U}_{f_a}|x\rangle_n|y\rangle_n = |x\rangle_n|y \oplus f_a(x)\rangle_n, \tag{171}$$

and

$$f_a(x) = f_a(y) \qquad \text{iff} \qquad y = x \oplus a, \tag{170}$$

we again use the general form

$$\mathsf{H}^{\otimes n}|j\rangle_n = \frac{1}{2^{n/2}}\sum_{k=0}^{2^n-1}(-1)^{j\odot k}|k\rangle_n, \tag{145}$$

and its particular case

$$\mathsf{H}^{\otimes n}|0\rangle_n = \frac{1}{2^{n/2}}\sum_{j=0}^{2^n-1}|j\rangle_n, \tag{130}$$

Thus,

$$\begin{aligned}
\mathsf{H}_x^{\otimes n}\mathsf{U}_{f_a}\mathsf{H}_x^{\otimes n}|0\rangle_x|0\rangle_y &= \mathsf{H}_x^{\otimes n}\mathsf{U}_f\frac{1}{2^{n/2}}\sum_{j=0}^{2^n-1}|j\rangle_x|0\rangle_y \\
&= \mathsf{H}_x^{\otimes n}\frac{1}{2^{n/2}}\sum_{j=0}^{2^n-1}|j\rangle_x|f_a(j)\rangle_y \\
&= \mathsf{H}_x^{\otimes n}\frac{1}{2^{(n+1)/2}}\sum_{j=0}^{2^n-1}(|j\rangle_x + |j\oplus a\rangle_x)|f_a(j)\rangle_y \qquad (761)\\
&= \frac{1}{2^{n+1/2}}\sum_{k=0}^{2^n-1}\sum_{j=0}^{2^n-1}[(-1)^{j\odot k} + (-1)^{(j\oplus a)\odot k}]|k\rangle_x|f_a(j)\rangle_y \\
&= \frac{1}{2^{n+1/2}}\sum_{k=0}^{2^n-1}[1 + (-1)^{a\odot k}]|k\rangle_x\sum_{j=0}^{2^n-1}(-1)^{j\odot k}|f_a(j)\rangle_y,
\end{aligned}$$

where we have used the fact that

$$\mathsf{U}_{f_a}\sum_{j=0}^{2^n-1}|j\rangle_n|0\rangle_n = \frac{1}{\sqrt{2}}\sum_{j=0}^{2^n-1}(|j\rangle_n + |j\oplus a\rangle_n)|f_a(j)\rangle_n. \tag{173}$$

in going from the 2nd to the 3rd lines.

We see that the amplitude of a state $|k\rangle_n$ is zero if $a \odot k = 1$. Hence, if we measure the x output lines we find them to have value k only if $a \odot k = 0$, and any such k are observed with equal probability. (A measurement of the y output lines is not very informative.)

If we repeat the process (174) m times, each followed by measurement of the x lines, we build up a set of linear equation

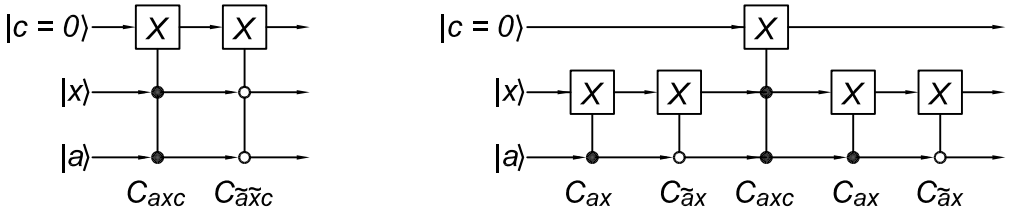
$$a \odot k_l = 0, \qquad l = 1, 2, ..., m, \tag{762}$$

for the n binary digits of the unknown number a. Not all of these equations are independent, so in general we need to make $m > n$ repetitions to obtain n independent (classical) linear equations,[168] which we can invert to find a. But the required number m of repetitions remains of order n, in contrast to the need for of order 2^n evaluations of f_a to determine a by a classical algorithm.

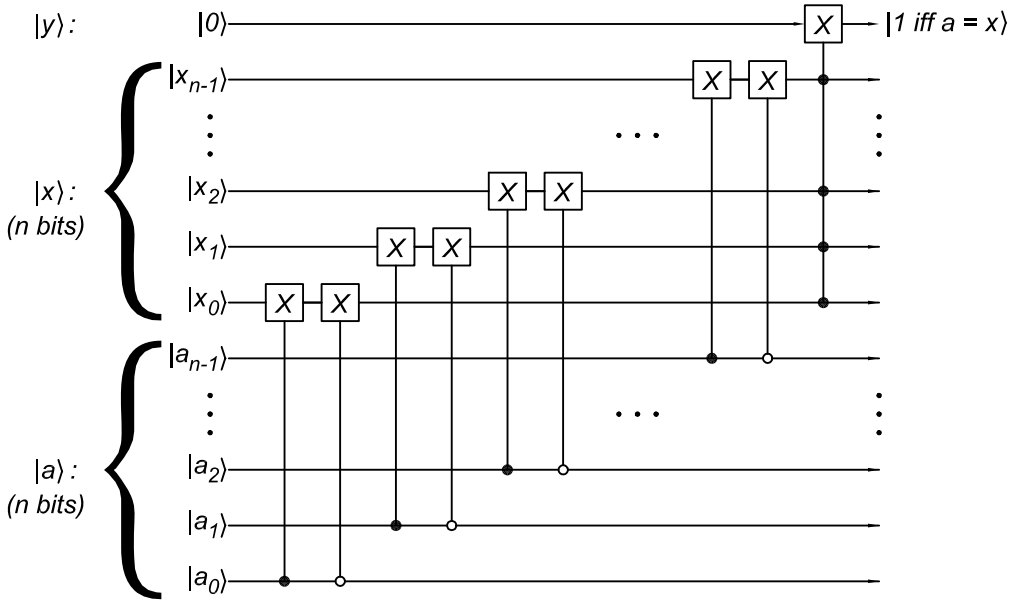
[168] For a discussion of the accuracy to which a can be determined as a function of the number m of repetitions, see, for example, sec. IID2 of
http://people.ccmr.cornell.edu/~mermin/qcomp/chap2.pdf

15. **Grover's Search Algorithm**

(a) We first construct a circuit that sets bit $|c\rangle$ to 1 only if bits $|a\rangle$ and $|x\rangle$ are equal. This can be accomplished, for example, using the two Controlled-Controlled-NOT operators C_{axc} and $\mathsf{C}_{\tilde{a}\tilde{x}c}$ for bit $|c\rangle$ initially set to $|0\rangle$, as shown on the left below,



A slight rearrangement of this circuit, as shown on the right (D. Peng, 3/31/05), permits generalization to the case that $|a\rangle$ and $|x\rangle$ and n-bit states to obtain the operation $\mathsf{U}_{f_a}|a\rangle_n|x\rangle_n|0\rangle = |a\rangle_n|x\rangle_n|1 \text{ iff } a = x\rangle$, without the need for any ancillary bits,



*To restore the initial state of $|x\rangle_n$, we should symmetrize the above diagram about the generalized Controlled-*NOT *operator at the right.*

(b) As in Deutsch's algorithm (prob. 10(c)), it is useful to initialize the auxiliary bit $|y\rangle$ to $|-\rangle_y = (|0\rangle - |1\rangle)/\sqrt{2} = \mathsf{H}|1\rangle = \mathsf{HX}|0\rangle$. Then, as before, we have

$$|y \oplus f(x)\rangle = \frac{|0 \oplus f(x)\rangle - |1 \oplus f(x)\rangle}{\sqrt{2}} = (-1)^{f(x)}\frac{|0\rangle - |1\rangle}{\sqrt{2}} = (-1)^{f(x)}|-\rangle_y. \quad (142)$$

Using state $|x\rangle_n$ of the form

$$|\phi\rangle_n = \mathsf{H}^{\otimes n}|0\rangle_n = \frac{1}{2^{n/2}}\sum_{j=0}^{2^n-1}|j\rangle_n \quad (130)$$

leads to

$$\mathsf{U}_{f_a}|x\rangle_n|y\rangle = \mathsf{U}_{f_a}(\mathsf{H}_x^{\otimes n} \otimes \mathsf{H}_y\mathsf{X}_y)|0\rangle_x|0\rangle_y = \frac{1}{2^{n/2}}\sum_{j=0}^{2^n-1}(-1)^{f_a(j)}|j\rangle_n|-\rangle_y. \quad (763)$$

Since the function f_a is

$$f_a(x) = \begin{cases} 0, & x \neq a, \\ 1, & x = a. \end{cases}, \tag{175}$$

the amplitude of state $|j\rangle_n$ in eq. (763) is $1/2^{n/2}$ unless $j = a$, in which case the amplitude is $-1/2^{n/2}$.

Thus we have "marked" the amplitude of the desired state $|a\rangle_n$ differently than all other (basis) states $|j\rangle_n$. It remains to find a procedure that can identify the "marked" state efficiently.

(c) The projection of state

$$|\psi\rangle_n = \sum_{j=0}^{N-1} \psi_j |j\rangle_n, \tag{182}$$

onto state

$$|\phi\rangle_n = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} |j\rangle_n \tag{130}$$

is

$$\langle\phi|\psi\rangle_n|\phi\rangle_n = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} \psi_j |\phi\rangle_n = \sqrt{N} \left\langle \psi_j \right\rangle |\phi\rangle_n = \sum_{j=0}^{N-1} \left\langle \psi_j \right\rangle |j\rangle_n, \tag{764}$$

where

$$\left\langle \psi_j \right\rangle = \frac{1}{N} \sum_{j=0}^{N-1} \psi_j = \frac{1}{\sqrt{N}} \langle\phi|\psi\rangle_n. \tag{183}$$

Hence, we can re-express $|\psi\rangle_n$ as

$$\begin{aligned} |\psi\rangle_n &= \langle\phi|\psi\rangle_n|\phi\rangle_n + |\psi\rangle_n - \langle\phi|\psi\rangle_n|\phi\rangle_n \\ &= \sqrt{N} \left\langle \psi_j \right\rangle |\phi\rangle_n + \sum_{j=0}^{N-1} \left(\psi_j - \left\langle \psi_j \right\rangle\right) |j\rangle_n. \end{aligned} \tag{765}$$

The reflection of $|\psi\rangle_n$ about $|\phi\rangle_n$ by operator $\mathsf{U}_{|\phi\rangle_n}$ can now be written as[169]

$$\begin{aligned} \mathsf{U}_{|\phi\rangle_n}|\psi\rangle_n &= \langle\phi|\psi\rangle_n|\phi\rangle_n - |\psi\rangle_n + \langle\phi|\psi\rangle_n|\phi\rangle_n = (2|\phi\rangle_n\langle\phi|_n - \mathbf{I}_n)\,|\psi\rangle_n \\ &= \sqrt{N} \left\langle \psi_j \right\rangle |\phi\rangle_n + \sum_{j=0}^{N-1} \left(\left\langle \psi_j \right\rangle - \psi_j\right) |j\rangle_n. \end{aligned} \tag{766}$$

Thus, we identify $\mathsf{U}_{|\phi\rangle_n}$ with the projection operator

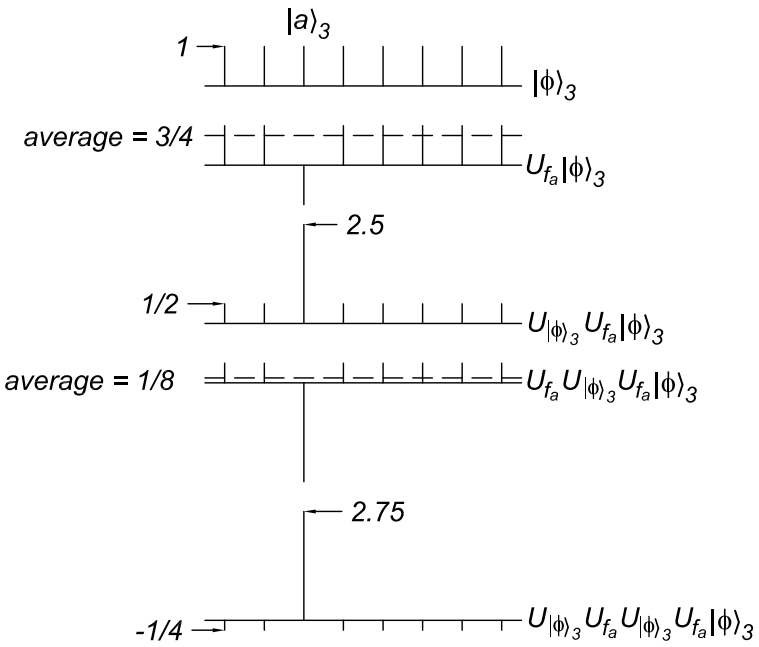
$$\mathsf{U}_{|\phi\rangle_n} = 2|\phi\rangle_n\langle\phi|_n - \mathbf{I}_n, \tag{767}$$

and we see that its effect is to change the sign of the amplitudes $\psi_j - \left\langle \psi_j \right\rangle$.

We also get another perspective on Grover's procedure, in which each interaction consists of reflecting a state about a plane orthogonal to the desired state $|a\rangle_n$ followed by a reflection of all amplitudes about their mean.

[169] We tacitly assume in eq. (766) that the $|y\rangle$ state is $|-\rangle_y$ and remains so under the action of $\mathsf{U}_{|\phi\rangle_n}$.

The figure below illustrates Grover's search algorithm for the case that $n = 8$. The probability of success of a single classical enquiry is $1/8 = 12.5\%$.



After the first iteration, the amplitude of the desired state $|a\rangle_3$ has been raised from 1 (in relative units) to 2.5, while that of the other 7 states has been reduced to 0.5. If a measurement were made now, the probability of success would be about 78%. For $n = 3$, $\sin\theta = 1/\sqrt{8}$, so $\theta = 20.7°$. Therefore, 2 iterations will bring the initial state $|\phi\rangle_3$ closer to $|a\rangle_3$ than 3 iterations. Hence, Grover's prescription for a list of 8 items involves 2 iterations, after which a measurement succeeds in finding $|a\rangle_3$ with probability 94.5%. If the procedure is repeated once, the combined probability of success is 99.7%.

We note that the transformation U_{f_a} which reflects a state about the plane perpendicular to $|a\rangle_n$ can also be written in a form involving a projection operator. Indeed, similarly to eq. (765), we can re-express a general state $|\psi\rangle_n$ as

$$|\psi\rangle_n = \langle\phi|\psi\rangle_n|a\rangle_n + |\psi\rangle_n - \langle a|\psi\rangle_n|a\rangle_n. \tag{768}$$

The reflection of $|\psi\rangle_n$ about the plane perpendicular to $|a\rangle_n$ by operator U_{f_a} changes the sign of the projection $\langle\phi|\psi\rangle_n|a\rangle_n$ of $|\psi\rangle_n$ onto $|a\rangle_n$, and so

$$\mathsf{U}_{f_a}|\psi\rangle_n = -\langle a|\psi\rangle_n|a\rangle_n + |\psi\rangle_n - \langle a|\psi\rangle_n|a\rangle_n = (\mathbf{I}_n - 2|a\rangle_n\langle a|_n)\,|\psi\rangle_n \tag{769}$$

Thus, we identify U_{f_a} with the operator

$$-\mathsf{U}_{|a\rangle_n} = \mathbf{I}_n - 2|a\rangle_n\langle a|_n = \mathsf{U}_{f_a}. \tag{770}$$

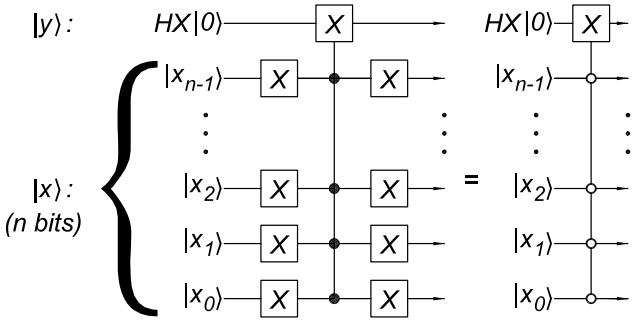
(d) Using the relation

$$|\phi\rangle_n = \mathsf{H}^{\otimes n}|0\rangle_n \tag{130}$$

and the fact that $\mathsf{H} = \mathsf{H}^\dagger = \mathsf{H}^{-1}$, eq. (767) can be written as[170]

$$\begin{aligned} -\mathsf{U}_{|\phi\rangle_n} &= \mathbf{I}_n - 2|\phi\rangle_n\langle\phi|_n = \mathbf{I}_n - 2\mathsf{H}^{\otimes n}|0\rangle_n\langle 0|_n\mathsf{H}^{\otimes n} \\ &= \mathsf{H}^{\otimes n}(\mathbf{I}_n - 2|0\rangle_n\langle 0|_n)\mathsf{H}^{\otimes n} = -\mathsf{H}^{\otimes n}\mathsf{U}_{|0\rangle_n}\mathsf{H}^{\otimes n} \\ &= \mathsf{H}^{\otimes n}\mathsf{U}_{f_0}\mathsf{H}^{\otimes n}. \end{aligned} \tag{771}$$

Recalling our construction of U_{f_a} in part (a), we can implement $\mathsf{U}_{f_0} = -\mathsf{U}_{|0\rangle_n}$ according to eq. (771) as shown below, deleting the unnecessary lines for $|a\rangle_n = |0\rangle_n$,



Note that for the above circuit to change the sign of the amplitude of $|0\rangle_x$ (and that of no other basis state $|j\rangle_x$), the auxiliary bit $|y\rangle$ must be prepared in the state $\mathsf{HX}|0\rangle_y = |-\rangle_y$. If we use the same bit $|y\rangle$ for this purpose as was used in the operation U_{f_a}, that bit is already in the needed state, so we don't add any new gates for this purpose to our construction of U_{f_0}.[171]

Our prescription for the m iterations of Grover's search algorithm can now be summarized as

$$[(\mathsf{H}_x^{\otimes n} \otimes \mathbf{I}_y)\mathsf{U}_{|0\rangle_n}(\mathsf{H}_x^{\otimes n} \otimes \mathbf{I}_y)\mathsf{U}_{f_a}]^m(\mathsf{H}_x^{\otimes n} \otimes \mathsf{H}_y\mathsf{X}_y)|0\rangle_x|0\rangle_y. \tag{772}$$

To use the same n auxiliary lines in all operators U_{f_a}, the construction given in part (a) must be symmetrized. If bits are cheaper than gates, it might be simpler to introduce a new set of n auxiliary bits for each U_{f_a} to reduce the overall number of gates.

[170] Again, the $|y\rangle$ state is to be $|-\rangle_y$ before and after the operation $\mathsf{U}_{|\phi\rangle_n}$.

[171] Most texts seem to consider the bit $|y\rangle$ needed for U_{f_0} as different from the bit $|y\rangle$ needed for U_{f_a}. Then, they note the appearance of the product $\mathsf{H}_y\mathsf{X}_y\mathsf{H}_y$ on the $|y\rangle$ line in U_{f_0}. Since this product equals $\mathsf{Z} = \boldsymbol{\sigma}_z$, one often sees diagrams with Controlled-Z operations describing Grover's algorithm.

16. **Parity of a Function**

A single application of the circuit yields

$$(\mathsf{H}_x^{\otimes n} \otimes \mathsf{X}_y\mathsf{H}_y)\mathsf{U}_f(\mathsf{H}_x^{\otimes n} \otimes \mathsf{H}_y\mathsf{X}_y)|0\rangle_x|0\rangle_y = \frac{1}{2^n}\sum_{j=0}^{2^n-1}\sum_{k=0}^{2^n-1}(-1)^{f(j)}(-1)^{j\odot k}|k\rangle_n|0\rangle_y, \quad (773)$$

using eqs. (130), (142), (145) and (186).

For the case that $n = 1$, the righthand side of eq. (773) is

$$\begin{aligned}
&\frac{1}{2}\Big\{(-1)^{f(0)}[(-1)^{0\odot 0}|0\rangle + (-1)^{0\odot 1}|1\rangle] + (-1)^{f(1)}[(-1)^{1\odot 0}|0\rangle + (-1)^{1\odot 1}|1\rangle]\Big\}\\
&= \frac{(-1)^{f(0)}}{2}\Big\{[(-1)^0|0\rangle + (-1)^0|1\rangle] + \Pi_f[(-1)^0|0\rangle + (-1)^1|1\rangle]\Big\}\\
&= \frac{(-1)^{f(0)}}{2}[(1+\Pi_f)|0\rangle + (1-\Pi_f)|1\rangle], \qquad (774)
\end{aligned}$$

using the fact that

$$(-1)^{f(1)} = (-1)^{f(0)}(-1)^{f(1)-f(0)} = (-1)^{f(0)}(-1)^{f(1)+f(0)} = (-1)^{f(0)}\Pi_f, \qquad (187)$$

since $f = 0$ or 1 only.

Thus, a measurement of the righthand side of eq. (773) yields the state $|0\rangle$ if $\Pi_f = 1$ and $|1\rangle$ if $\Pi_f = -1$.

17. **Quantum Fourier Transform,[172] Shor's Period-Finding Algorithm**

(a) We recall that the (binary) number k can be written $k = \sum_{l=0}^{n-1} k_l 2^l$, where $k_l = 0$ or 1, so that

$$\sum_{k=0}^{2^n-1} |k\rangle_n = \sum_{k=0}^{2^n-1} \prod_{l=0}^{n-1} |k_l\rangle = \prod_{l=0}^{n-1} \sum_{k_l=0}^{1} |k_l\rangle. \tag{17}$$

Then, the quantum Fourier transform (197) can be expanded as

$$\begin{aligned}
\mathbf{\Phi}|j\rangle_n &= \frac{1}{2^{n/2}} \sum_{k=0}^{2^n-1} e^{2\pi ijk/2^n} |k\rangle_n = \frac{1}{2^{n/2}} \sum_{k=0}^{2^n-1} \prod_{l=0}^{n-1} e^{2\pi ijk_l 2^l/2^n} \prod_{m=0}^{n-1} |k_m\rangle \\
&= \frac{1}{2^{n/2}} \sum_{k=0}^{2^n-1} \prod_{l=0}^{n-1} e^{2\pi ijk_l 2^l/2^n} |k_l\rangle = \frac{1}{2^{n/2}} \prod_{l=0}^{n-1} \sum_{k_l=0}^{1} e^{2\pi ijk_l 2^l/2^n} |k_l\rangle \\
&= \prod_{l=0}^{n-1} \frac{|0\rangle + e^{2\pi ij2^l/2^n}|1\rangle}{\sqrt{2}} = \prod_{l=0}^{n-1} \frac{|0\rangle + e^{2\pi i \sum_{m=0}^{n-1} j_m 2^m 2^l/2^n}|1\rangle}{\sqrt{2}} \\
&= \prod_{l=0}^{n-1} \frac{|0\rangle + e^{2\pi i \sum_{m=0}^{n-l-1} j_m 2^{m+l-n}}|1\rangle}{\sqrt{2}} \\
&= \prod_{l=0}^{n-1} \frac{|0\rangle + \prod_{m=0}^{n-l-1} e^{\pi i j_m 2^{m+l-(n-1)}}|1\rangle}{\sqrt{2}},
\end{aligned} \tag{775}$$

where we have expanded the number j as $\sum_{m=0}^{n-1} j_m 2^m$ where $j_m = 0$ or 1, and we note that once $m + l \geq n$ in the exponent $2\pi ij2^l/2^n = 2\pi i \sum_{m=0}^{n-1} j_m 2^{m+l-n}$ the phase angle of the mth term is an integer multiple of 2π and can be ignored.

The product over m in eq. (775) can be written out as

$$\prod_{m=0}^{n-l-1} e^{\pi i j_m 2^{m+l-(n-1)}} = \begin{cases} e^{\pi i j_0}, & l = n-1, \\ e^{\pi i j_1} e^{\pi i j_0/2}, & l = n-2, \\ e^{\pi i j_2} e^{\pi i j_1/2} e^{\pi i j_0/2^2}, & l = n-3, \\ \vdots & \\ e^{\pi i j_{n-1}} e^{\pi i j_{n-2}/2} \cdots e^{\pi i j_0/2^{n-1}}, & l = 0. \end{cases} \tag{776}$$

The highest-order output bit corresponds to the factor $l = n - 1$ in eq. (775), which is simply related to the lowest-order input bit j_0 by eq. (776).

(b) In the simplest circuit for the operation $\mathbf{\Phi}|j\rangle$, the lth input line corresponds to bit $|j_l\rangle$, and this line leads to the $n - 1 - l$th-order bit of $\mathbf{\Phi}|j\rangle$.

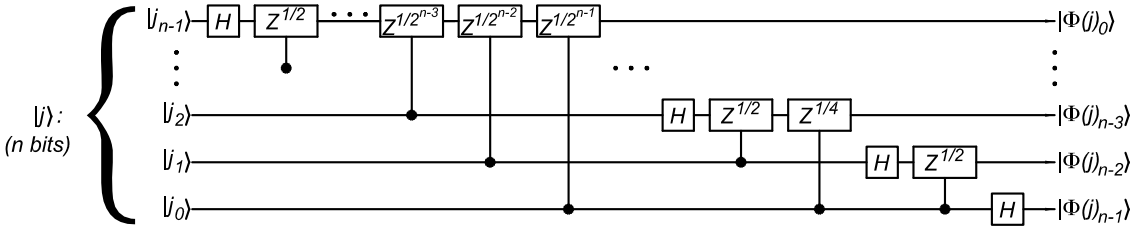
The leading terms of the $n - 1 - l$th-order bit of $\mathbf{\Phi}|j\rangle$) of eq. (775) have the form

$$\frac{|0\rangle + e^{\pi i j_l}|1\rangle}{\sqrt{2}} = \frac{|0\rangle + (-1)^{j_l}|1\rangle}{\sqrt{2}} = \mathsf{H}|j_l\rangle. \tag{61}$$

[172] For an early discussion of the quantum Fourier transform, see
http://kirkmcd.princeton.edu/examples/QM/coppersmith_ibm_rc19642_94.pdf

An additional factor of $e^{\pi i/2^{n-1-l-m}} = (-1)^{1/2^{n-1-l-m}}$ multiples the $|1\rangle$ in the lth factor whenever $|j_m\rangle = |1\rangle$ for $0 \le m < n-1-l$. This suggests use of an operation that affects only the $|1\rangle$ state of bit l whenever bit m is 1, *i.e.*, some kind of Controlled operation U_{ml}. We readily see that the desired operation is $\mathsf{U} = \mathsf{Z}^{1/2^{n-1-l-m}} = (\boldsymbol{\sigma}_z^{1/4})^{1/2^{n-l-m-3}}$.

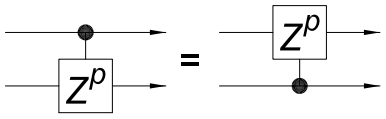
Putting together these pieces, we apply the Hadamard transformation to input line l only after the state of that line has been used to generate the needed phase factors for lines $m > l$. Hence, we can represent the quantum Fourier transform $\boldsymbol{\Phi}|j\rangle_n$ as



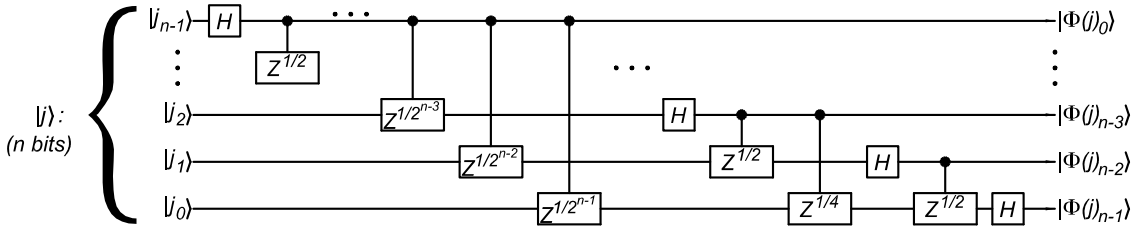
The above diagram can be drawn slightly differently by noting a property of Controlled-Z^p that arises because the only effect of Z^p is to change the phase of the $|1\rangle$ state. Since

$$\mathsf{Z}^p = \begin{pmatrix} 1 & 0 \\ 0 & e^{\pi i p} \end{pmatrix}, \tag{777}$$

the 2-Qbit Controlled-Z^p matrix has the form

$$\mathsf{Z}^p_{ab} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{\pi i p} \end{pmatrix} = \mathsf{Z}^p_{ba}. \qquad \text{  } \tag{778}$$

Thus, there is actually no distinction between the control and target bits of a Controlled-Z^p gate. Hence, the quantum Fourier transform $\boldsymbol{\Phi}$ can also be represented as



There is no hardware difference between this diagram and the previous; the difference is only one of notation.

(c) Recalling the definition (194) of the discrete Fourier transform, we see that the inverse Fourier transform is the same as the Fourier transform except for the reversal of the signs of all the phases in the exponential phase factors. Hence, the

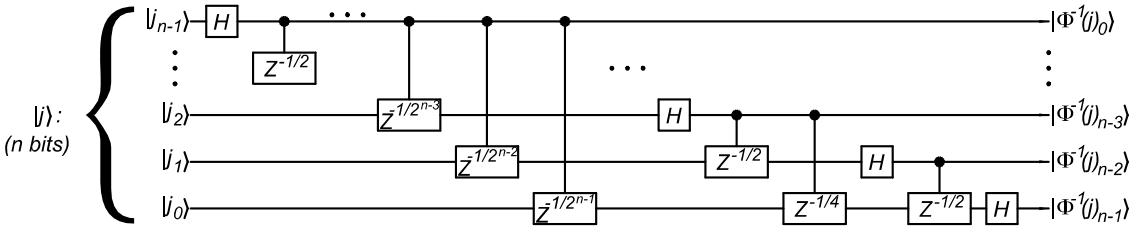
inverse form of eq. (197) is

$$\mathbf{\Phi}^{-1}|j\rangle_n = \frac{1}{2^{n/2}} \sum_{k=0}^{2^n-1} e^{-2\pi ijk/2^n}|k\rangle_n. \tag{779}$$

Rearranging this in the fashion of eq. (775), we have

$$\mathbf{\Phi}^{-1}|j\rangle_n = \prod_{l=0}^{n-1} \frac{|0\rangle + \prod_{m=0}^{n-l-1} e^{-\pi i j_m 2^{m+l-(n-1)}}|1\rangle}{\sqrt{2}}. \tag{780}$$

A bit-flow diagram corresponding to eq. (780) is



Although $\mathsf{Z}^{-1} = \mathsf{Z}$, in general $\mathsf{Z}^{-p} \neq \mathsf{Z}^p$. Hence, the circuit for the inverse Fourier transform $\mathbf{\Phi}^{-1}$ is distinct from that for the Fourier transform $\mathbf{\Phi}$.

(d) **Period Finding.**

Applying the quantum Fourier transform,

$$\mathbf{\Phi}|j\rangle_n = \frac{1}{2^{n/2}} \sum_{k=0}^{2^n-1} e^{2\pi ijk/2^n}|k\rangle_n, \tag{197}$$

to the state (202) and noting that $ab \approx 2^n$, we have

$$\begin{aligned}
\mathbf{\Phi}_x \mathsf{U}_f \mathsf{H}_x^{\otimes n}|0\rangle_x|0\rangle_y &= \mathbf{\Phi}_x \frac{1}{\sqrt{a}} \sum_{j=0}^{a-1} \frac{1}{\sqrt{b}} \sum_{m=0}^{b-1} |j+ma\rangle_x|f(j)\rangle_y \\
&= \frac{1}{\sqrt{a}} \sum_{j=0}^{a-1} \frac{1}{\sqrt{b}} \sum_{m=0}^{b-1} \frac{1}{2^{n/2}} \sum_{k=0}^{2^n-1} e^{2\pi i(j+ma)k/2^n}|k\rangle_x|f(j)\rangle_y \\
&= \frac{1}{\sqrt{ab2^n}} \sum_{k=0}^{2^n-1} \sum_{j=0}^{a-1} e^{2\pi ijk/2^n} \sum_{m=0}^{b-1} e^{2\pi imak/2^n}|k\rangle_x|f(j)\rangle_y \\
&= \frac{1}{2^n} \sum_{k=0}^{2^n-1} \sum_{j=0}^{a-1} e^{2\pi ijk/2^n} \sum_{m=0}^{b-1} (e^{2\pi iak/2^n})^m|k\rangle_x|f(j)\rangle_y \\
&= \frac{1}{2^n} \sum_{k=0}^{2^n-1} \sum_{j=0}^{a-1} e^{2\pi ijk/2^n} \frac{1-e^{2\pi iabk/2^n}}{1-e^{2\pi iak/2^n}}|k\rangle_x|f(j)\rangle_y \\
&= \frac{1}{2^n} \sum_{k=0}^{2^n-1} \sum_{j=0}^{a-1} e^{2\pi ijk/2^n} \frac{e^{\pi iabk/2^n}}{e^{\pi iak/2^n}} \frac{\sin(\pi abk/2^n)}{\sin(\pi ak/2^n)}|k\rangle_x|f(j)\rangle_y \\
&\approx \frac{1}{2^n} \sum_{k=0}^{2^n-1} \sum_{j=0}^{a-1} e^{i\phi_{jk}} \frac{\sin(\pi k)}{\sin(\pi ak/2^n)}|k\rangle_x|f(j)\rangle_y,
\end{aligned} \tag{781}$$

where ϕ_{jk} is a phase factor. In general, the product ab is not quite equal to 2^n, so the factor written as $\sin(\pi k)$ is not quite zero, but is merely a small number.

If we now measure the state of the system, we find the x lines to be number k, and the j lines to be the number $f(j)$ with probability[173]

$$P \approx \frac{a}{2^{2n}} \frac{\sin^2(\pi k)}{\sin^2(\pi a k/2^n)} \,. \tag{783}$$

This probability is very small unless the denominator is also small, which happens when $k \approx l2^n/a$ for positive integer $l \leq a$,

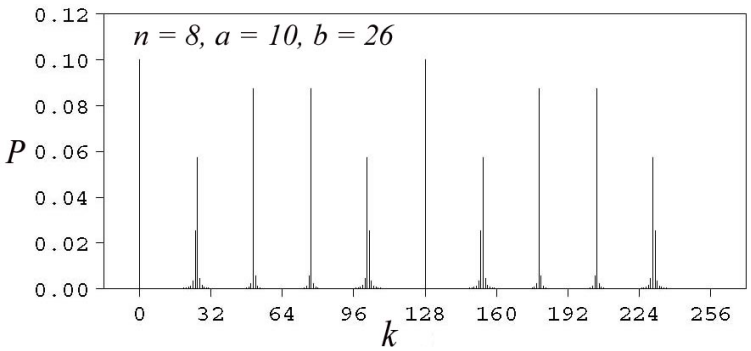
$$P(k \to l2^n/a) \to \frac{a}{2^{2n}} \left(\frac{2^n}{a}\right)^2 = \frac{1}{a} \,. \tag{784}$$

Hence, we infer from a single quantum computation of $\mathbf{\Phi}_x \mathsf{U}_f \mathsf{H}_x^{\otimes n} |0\rangle_x |0\rangle_y$ in which the final state of $|x\rangle_n$ is observed to be $|k\rangle_n$ that the period a of function f is

$$a = \frac{l2^n}{k} \,, \tag{785}$$

where l is a positive integer. With a few repetitions of the computation, we can deduce a unique value for the period a. The solution is readily verified to be valid by checking that $f(x + a) = f(x)$.

For example, in the case of an 8-bit function with period $a = 10$, we have $2^n = 256$ and $b = \text{int}(256/10) + 1 = 26$. The probability distribution (783) of the values of k is shown below.[174]



[173] If we only measure the x lines, the probability that we find number k is the product of eq. (783) times the expectation value of the y state. Since that state is $\sum_{j=0}^{a-1} e^{i\phi_{jk}} |f(j)\rangle_y$, the expectation value is

$$\sum_{j=0}^{a-1} \sum_{l=0}^{a-1} e^{i(\phi_{jk} - \phi)_{lk})} \langle f(l)|f(j)\rangle_y = a, \tag{782}$$

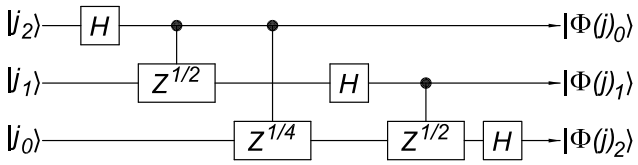
provided that function f is a one-to-one map over its period a (so that $\langle f(l)|f(j)\rangle_y = \delta_{jl}$). In this case, there is no essential difference in the nature of the measurements of the x lines whether or not the y lines are measured as well (A. Hook, 4/7/05).

[174] From http://kirkmcd.princeton.edu/examples/QM/shor_siamjc_26_1484_97.pdf

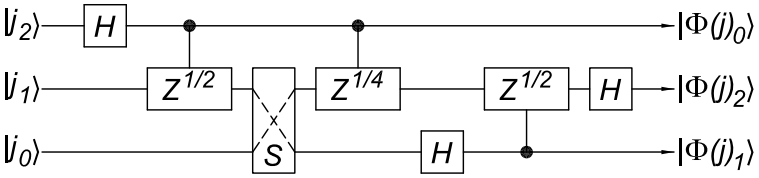
18. Nearest-Neighbor Algorithms

(a) Fourier Transform.

The bit flow diagram for a 3-bit quantum Fourier transform $\mathbf{\Phi}$, based on prob. 17(b), is

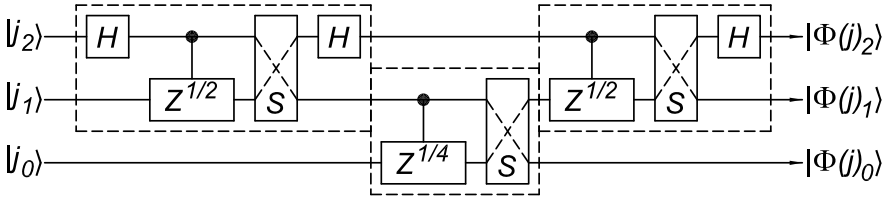


Since this has only one gate that does not involve neighboring bits, we could convert this to an all-nearest-neighbor diagram with only a single swap of input lines 0 and 1,



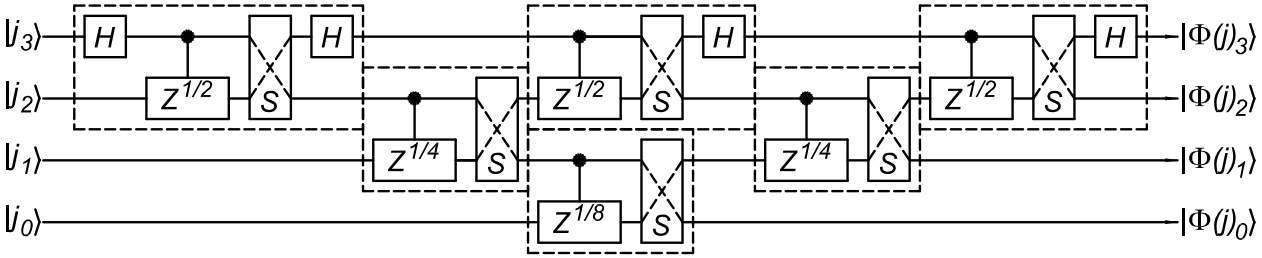
However, this diagram is not very elegant in its placement of the output lines.

To accomplish the goal of reversing the output lines (to bring them back into alignment with the input lines) it is better to begin swapping lines 1 and 2, followed by a swap of lines 0 and 2, and finally a swap of lines 0 and 1. As noted in prob. 17(b), the target and control bits can be interchanged on a Controlled-Z^p gate, so the SWAP gates could be positioned before or after the corresponding Controlled-Z^p gates.



The dashed lines show possible groupings of the logical gates into physical gates.

Extending the nearest-neighbor Fourier-transform circuit to the case of 4 bits, we have



This rather pleasing form of the quantum Fourier transform was first suggested by Griffiths and Niu,[175] who, however, were not very explicit about the need for SWAP gates.

[175] http://kirkmcd.princeton.edu/examples/QM/griffiths_prl_76_3228_96.pdf

Although the nearest-neighbor Fourier transform circuit appears to be left-right symmetric, this does not imply that the Fourier transform is its own inverse. Note that input line j connects with output line $n-j-1$, so the symmetry of the circuit is only partial.

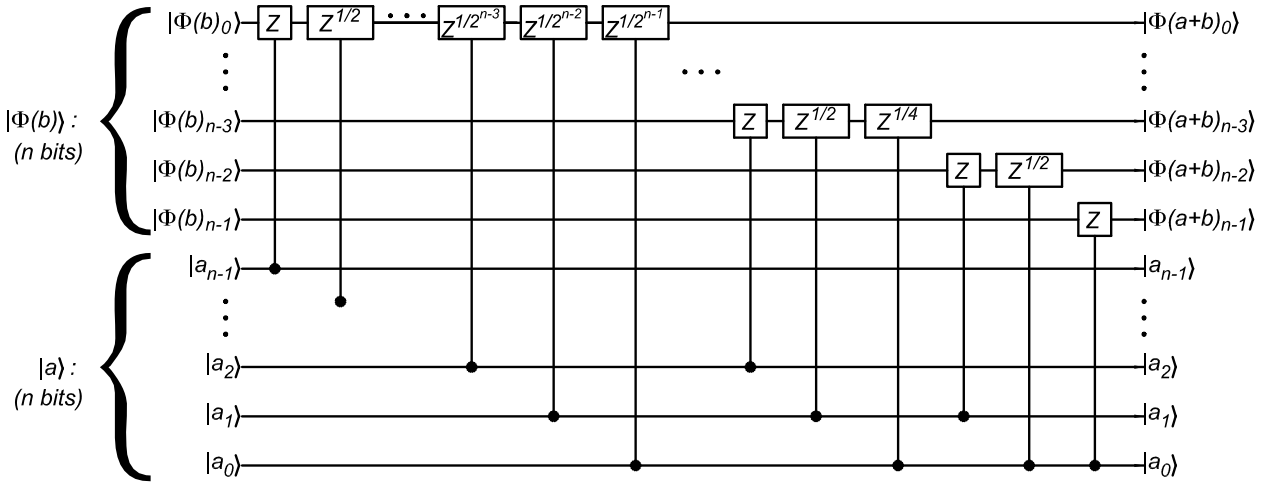
(b) **Fourier Addition.**

Recalling, eq. (775), we can expand the quantum Fourier transform of two $(n-1)$-bit numbers a and b as

$$\begin{aligned}
\mathbf{\Phi}|a+b\rangle_n &= \prod_{l=0}^{n-1} \frac{|0\rangle + \prod_{m=0}^{n-l-1} e^{\pi i (a+b)_m 2^{m+l-(n-1)}}|1\rangle}{\sqrt{2}} \\
&= \prod_{l=0}^{n-1} \frac{|0\rangle + \prod_{k=0}^{n-l-1} e^{\pi i a_k 2^{k+l-(n-1)}} \prod_{m=0}^{n-l-1} e^{\pi i b_m 2^{m+l-(n-1)}}|1\rangle}{\sqrt{2}} \\
&= \prod_{l=0}^{n-1} \frac{|0\rangle + \prod_{k=0}^{n-l-1} \left(\mathsf{Z}^{2^{k+l-(n-1)}}\right)^{a_k} \prod_{m=0}^{n-l-1} e^{\pi i b_m 2^{m+l-(n-1)}}|1\rangle}{\sqrt{2}} \\
&= \prod_{l=0}^{n-1} \prod_{k=0}^{n-l-1} \left(\mathsf{Z}^{2^{k+l-(n-1)}}\right)^{a_k} \frac{|0\rangle + \prod_{m=0}^{n-l-1} e^{\pi i b_m 2^{m+l-(n-1)}}|1\rangle}{\sqrt{2}} \\
&= \prod_{l=0}^{n-1} \prod_{k=0}^{n-l-1} \left(\mathsf{Z}^{2^{k+l-(n-1)}}\right)^{a_k} |\Phi(b)_m\rangle_n \equiv \mathbf{\Phi}_a^+(b), \qquad (786)
\end{aligned}$$

noting that any power of the operator Z has no effect on the state $|0\rangle$. This tells us that the lth bit of $\mathbf{\Phi}|a+b\rangle_n$ is the lth bit of $\mathbf{\Phi}|b\rangle_n$ after it has been acted upon by the operator product $\prod_{k=0}^{n-l-1} \left(\mathsf{Z}^{2^{k+l-(n-1)}}\right)^{a_k}$ based on number $a = \sum_{k=0}^{n-1} a_k 2^k$. Implementing these factor is very similar to the task of implementing the original Fourier transform, with the simplification that the H gates are just Z gates now.

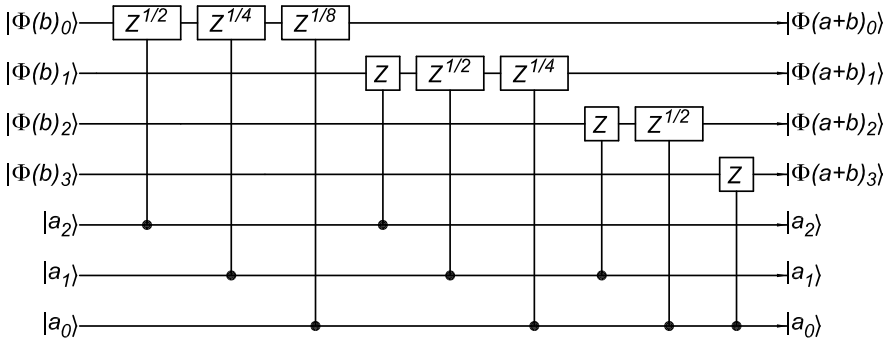
A general bit-flow diagram for the Fourier addition $|\mathbf{\Phi}(a+b)_n\rangle$ is thus



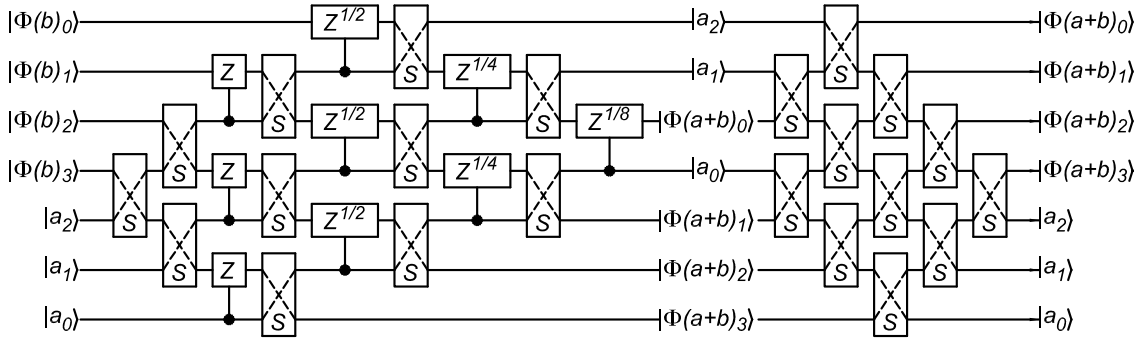
Since number a actually has only $n-1$ bits, the line labeled $|a_{n-1}\rangle$ can be eliminated. Note, however, that the n-bit Fourier transform $|\mathbf{\Phi}(b)\rangle_n$ has nontrivial structure on all of its n lines.

Note also that since the Controlled-Z^p gates commute with one another, the gates in the above circuit could be applied in any order.

Specializing to the case that a and b are 3-bit numbers, the bit-flow diagram for this is:



We convert this to a nearest-neighbor algorithm by inserting the appropriate 2-bit SWAP operations:



About 2/3 of the way through the circuit the computation of $|\mathbf{\Phi}(a+b)_n\rangle$ is complete, but the output lines don't match the input lines. The final group of SWAP operations restores the alignment of inputs and outputs.

(c) **Fourier Subtraction.**

The operation $\mathbf{\Phi}_a^-(b)$ is defined as operation $\mathbf{\Phi}_a^+(b)$ with each of its gates replaced by its inverse and applied in the reverse order. Working from eq. (786), and recalling that the order of gates Z^p is immaterial, we can write

$$\begin{aligned}\mathbf{\Phi}_a^-(b) &= \prod_{l=0}^{n-1}\prod_{k=0}^{n-l-1}\left(\mathsf{Z}^{-2^{k+l-(n-1)}}\right)^{a_k}|\Phi(b)_m\rangle_n \\ &= \prod_{l=0}^{n-1}\frac{|0\rangle+\prod_{m=0}^{n-l-1}e^{\pi i(-a+b)_m 2^{m+l-(n-1)}}|1\rangle}{\sqrt{2}}\,. \end{aligned} \tag{787}$$

We recall that the quantum Fourier transform $\mathbf{\Phi}(b)$ has only been defined for non-negative integers b. Comparing with expansion (775), we see that eq. (787) is indeed the quantum Fourier transform of $b-a$ so long as $b \geq a$.

If however, $b < a$, then we consider the 2's complement representation of $b-a$ where a and b are $(n-1)$-bit numbers, namely

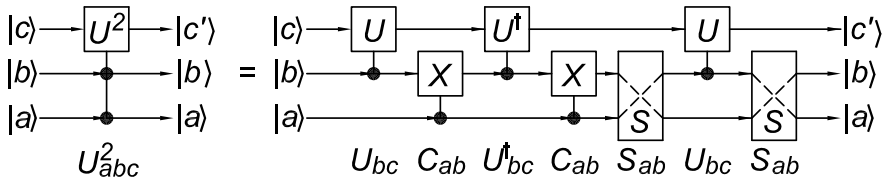
$$b-a \to 2^n-(a-b) \qquad (a-b>0,\ 2'\text{s complement}). \tag{788}$$

We also learned in the derivation of eq. (775) that the n-bit quantum Fourier transform "ignores" higher bits as these lead to phase shifts that are multiples of

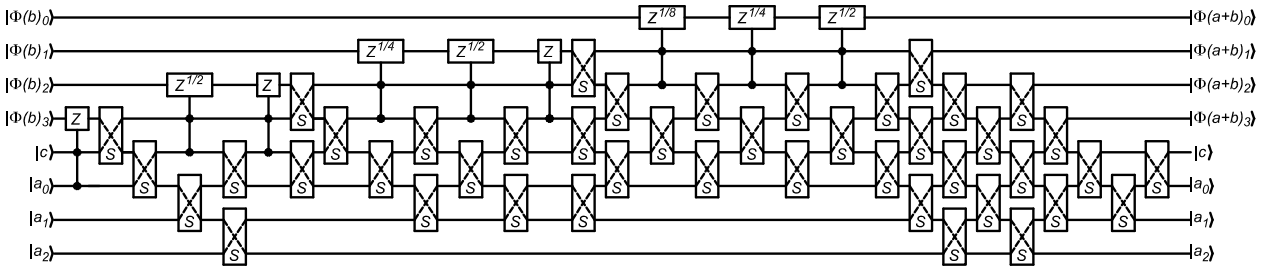
2π. Thus, the quantum Fourier subtraction operator $\mathbf{\Phi}_a^-(b)$ "naturally" performs 2's complement subtraction.

(d) **Controlled Fourier Addition.**

The circuit for a Controlled-Controlled-U^2 operation that was given in prob. 11(c) becomes a nearest-neighbor algorithm with the addition of two SWAP gates:



To implement a Controlled-Fourier addition using Controlled-Controlled-Z^p gates, I found it convenient to group together all such operations on a given bit of $\mathbf{\Phi}(b)$. Also, it seemed simplest if the control bit $|c\rangle$ is between $|a\rangle$ and $|\mathbf{\Phi}(b)\rangle$, and the order of the bits of $|a\rangle$ is reversed (otherwise 6 more layers of SWAP gates are required):



19. **Spin Control**

(a) **Single-Qbit Gates via Pulsed Magnetic Fields**

We wish to evaluate the rotating-frame Hamiltonian,

$$\mathsf{h}_{\rm rot} = e^{-i\frac{\omega_0 t}{2}\boldsymbol{\sigma}_z}\left(\mathsf{h}_{\rm lab} + \frac{\omega_0}{2}\boldsymbol{\sigma}_z\right)e^{i\frac{\omega_0 t}{2}\boldsymbol{\sigma}_z} \tag{236}$$

for a spin-1/2 particle at rest in the time-dependent magnetic field

$$\mathbf{B} = B_0\hat{\mathbf{z}} + B_u[u_x(\cos\omega_0 t\ \hat{\mathbf{x}} - \sin\omega_0 t\ \hat{\mathbf{y}}) + u_y(\sin\omega_0 t\ \hat{\mathbf{x}} + \cos\omega_0 t\ \hat{\mathbf{y}}) + u_z\ \hat{\mathbf{z}}]. \tag{237}$$

The lab-frame (reduced) interaction Hamiltonian $\mathsf{h}_{\rm lab} = \mathcal{H}/\hbar$ for a spin-1/2 particle with magnetic moment $\boldsymbol{\mu} = \gamma\mathbf{s} = \hbar\Gamma\boldsymbol{\sigma}/2$ in the magnetic field (237) is given by

$$\begin{aligned}
\mathsf{h}_{\rm lab} &= -\frac{\Gamma}{2}\boldsymbol{\sigma}\cdot\mathbf{B} = -\frac{\Gamma B_0}{2}\boldsymbol{\sigma}_z \\
&\quad -\frac{\Gamma B_u}{2}[u_x(\cos\omega_0 t\ \boldsymbol{\sigma}_x - \sin\omega_0 t\ \boldsymbol{\sigma}_y) + u_y(\sin\omega_0 t\ \boldsymbol{\sigma}_x + \cos\omega_0 t\ \boldsymbol{\sigma}_y) + u_z\ \boldsymbol{\sigma}_z] \\
&= -\frac{\omega_0}{2}\boldsymbol{\sigma}_z \\
&\quad -\frac{\omega_u}{2}[(u_x\cos\omega_0 t + u_y\sin\omega_0 t)\boldsymbol{\sigma}_x + (-u_x\sin\omega_0 t + u_y\cos\omega_0 t)\boldsymbol{\sigma}_y + u_z\boldsymbol{\sigma}_z]
\end{aligned} \tag{789}$$

where $\omega_0 = \Gamma B_0$ and $\omega_u = \Gamma B_u$. On using this in eq. (236), we need to evaluate the products

$$\begin{aligned}
e^{-i\frac{\omega_0 t}{2}\boldsymbol{\sigma}_z}\boldsymbol{\sigma}_j e^{i\frac{\omega_0 t}{2}\boldsymbol{\sigma}_z} &= \left(\cos\frac{\omega_0 t}{2}\mathbf{I} - i\sin\frac{\omega_0 t}{2}\boldsymbol{\sigma}_z\right)\boldsymbol{\sigma}_j\left(\cos\frac{\omega_0 t}{2}\mathbf{I} + i\sin\frac{\omega_0 t}{2}\boldsymbol{\sigma}_z\right) \\
&= \cos^2\frac{\omega_0 t}{2}\boldsymbol{\sigma}_j - i\cos\frac{\omega_0 t}{2}\sin\frac{\omega_0 t}{2}(\boldsymbol{\sigma}_z\boldsymbol{\sigma}_j - \boldsymbol{\sigma}_j\boldsymbol{\sigma}_z) + \sin^2\frac{\omega_0 t}{2}\boldsymbol{\sigma}_z\boldsymbol{\sigma}_j\boldsymbol{\sigma}_z \\
&= \begin{cases} \cos\omega_0 t\ \boldsymbol{\sigma}_x + \sin\omega_0 t\ \boldsymbol{\sigma}_y & (j=x), \\ -\sin\omega_0 t\ \boldsymbol{\sigma}_x + \cos\omega_0 t\ \boldsymbol{\sigma}_y & (j=y), \\ \boldsymbol{\sigma}_z & (j=z). \end{cases}
\end{aligned} \tag{790}$$

Combining eqs. (236), (789) and (790), we find the Hamiltonian in the rotating frame to be

$$\mathsf{h}_{\rm rot} = -\frac{\omega_u}{2}\hat{\mathbf{u}}\cdot\boldsymbol{\sigma}. \tag{791}$$

If the lab-frame magnetic field is

$$\mathbf{B} = B_0\hat{\mathbf{z}} + B_x(\cos\omega t\ \hat{\mathbf{x}} - \sin\omega t\ \hat{\mathbf{y}}), \tag{240},$$

then the (reduced) lab-frame interaction Hamiltonian is

$$\begin{aligned}
\mathsf{h}_{\rm lab} &= -\frac{\Gamma}{2}\boldsymbol{\sigma}\cdot\mathbf{B} = -\frac{\Gamma B_0}{2}\boldsymbol{\sigma}_z - \frac{\Gamma B_x}{2}(\cos\omega t\ \boldsymbol{\sigma}_x - \sin\omega t\ \boldsymbol{\sigma}_y) \\
&= -\frac{\omega_0}{2}\boldsymbol{\sigma}_z - \frac{\omega_x}{2}(\cos\omega t\ \boldsymbol{\sigma}_x - \sin\omega t\ \boldsymbol{\sigma}_y) \\
&= -\frac{\omega_0}{2}\boldsymbol{\sigma}_z - \frac{\omega_x}{2}\left(e^{i\omega t}\frac{\boldsymbol{\sigma}_x + i\boldsymbol{\sigma}_y}{2} + e^{-i\omega t}\frac{\boldsymbol{\sigma}_x - i\boldsymbol{\sigma}_y}{2}\right) \\
&= -\frac{\omega_0}{2}\boldsymbol{\sigma}_z - \frac{\omega_x}{2}\left(e^{i\omega t}\mathsf{a}^\dagger + e^{-i\omega t}\mathsf{a}\right),
\end{aligned} \tag{792}$$

where $\omega_x = \Gamma B_x$, and

$$\mathsf{a} = \frac{\boldsymbol{\sigma}_x - i\boldsymbol{\sigma}_y}{2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \text{and} \qquad \mathsf{a}^\dagger = \frac{\boldsymbol{\sigma}_x + i\boldsymbol{\sigma}_y}{2} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{793}$$

are the annihilation and creation operators introduced in prob. 9. *We could also use eqs. (123) and (125) to write* $\boldsymbol{\sigma}_z = \mathsf{a}\mathsf{a}^\dagger - \mathsf{a}^\dagger\mathsf{a}$ *to express the Hamiltonian (792) entirely in terms of annihilation and creation operators, but this is not really necessary.*

Then, inserting the trial solution

$$|\psi(t)\rangle = Ae^{i\alpha t}|0\rangle + Be^{-i\beta t}|1\rangle \tag{241}$$

into Schrödinger's equation, $i\partial_t|\psi\rangle = \mathsf{h}_{\rm lab}|\psi\rangle$, usinig eq. (792) we find (for the coefficients of $|0\rangle$ and $|1\rangle$)

$$-\alpha Ae^{i\alpha t} = -\frac{\omega_0}{2}Ae^{i\alpha t} - \frac{\omega_x}{2}e^{i\omega t}Be^{-i\beta t}, \tag{794}$$

$$\beta Be^{-i\beta t} = \frac{\omega_0}{2}Be^{-i\beta t} - \frac{\omega_x}{2}e^{-i\omega t}Ae^{i\alpha t}, \tag{795}$$

or

$$\alpha = \frac{\omega_0}{2} + \frac{\omega_x}{2}\frac{B}{A}e^{-i(\alpha+\beta-\omega)t}, \tag{796}$$

$$\beta = \frac{\omega_0}{2} - \frac{\omega_x}{2}\frac{A}{B}e^{i(\alpha+\beta-\omega)t}. \tag{797}$$

These equations can only be satisfied if

$$\alpha + \beta = \omega, \tag{798}$$

in which case eqs. (796)-(797) become

$$\alpha = \frac{\omega_0}{2} + \frac{\omega_x}{2}\frac{B}{A}, \tag{799}$$

$$\beta = \frac{\omega_0}{2} - \frac{\omega_x}{2}\frac{A}{B}. \tag{800}$$

Substituting eqs. (799)-(800) into eq. (798), we have

$$\omega_0 + \frac{\omega_x}{2}\frac{B}{A} - \frac{\omega_x}{2}\frac{A}{B} - \omega = 0, \tag{801}$$

$$\frac{A}{B} - 2\frac{\omega_0 - \omega}{\omega_x} - \frac{B}{A} = 0, \tag{802}$$

$$\frac{A^2}{B^2} - 2\frac{\omega_0 - \omega}{\omega_x}\frac{A}{B} - 1 = 0 = \frac{B^2}{A^2} + 2\frac{\omega_0 - \omega}{\omega_x}\frac{B}{A} - 1, \tag{803}$$

$$\frac{A}{B} = \frac{\omega_0 - \omega \pm \sqrt{(\omega_0 - \omega)^2 + \omega_x^2}}{\omega_x}, \tag{804}$$

$$\frac{B}{A} = \frac{\omega - \omega_0 \pm \sqrt{(\omega_0 - \omega)^2 + \omega_x^2}}{\omega_x}, \tag{805}$$

$$\alpha = \frac{\omega_0}{2} + \frac{\omega_x}{2}\frac{B}{A} = \frac{\omega \mp \sqrt{(\omega_0 - \omega)^2 + \omega_x^2}}{2}, \tag{806}$$

$$\beta = \frac{\omega_0}{2} - \frac{\omega_x}{2}\frac{A}{B} = \frac{\omega \pm \sqrt{(\omega_0 - \omega)^2 + \omega_x^2}}{2}, \tag{807}$$

and indeed,

$$\alpha + \beta = \omega. \tag{798}$$

We have found two solutions, which must be added to give the general solution to Schrödinger's equation.

We define

$$\Omega = \sqrt{(\omega_0 - \omega)^2 + \omega_x^2}, \tag{808}$$

so that the time-dependent Qbit (241) has the general form

$$|\psi(t)_{\rm lab}\rangle = (A_1 e^{-i\Omega t} + A_2 e^{i\Omega t})e^{i\frac{\omega}{2}t}|0\rangle + (B_1 e^{i\Omega t} + B_2 e^{-i\Omega t})e^{-i\frac{\omega}{2}t}|1\rangle. \tag{809}$$

The coefficients A_j and B_j are related according to eq. (804),

$$A_1 = \frac{\omega_0 - \omega + \Omega}{\omega_x}B_1, \qquad A_2 = \frac{\omega_0 - \omega - \Omega}{\omega_x}B_2. \tag{810}$$

If the initial Qbit is $|\psi(0)\rangle = a|0\rangle + b|1\rangle$, then eq. (809) tells us that

$$A_1 + A_2 = a, \qquad B_1 + B_2 = b. \tag{811}$$

Substituting eq. (810) into the first equation of (811), and then using the second equation of (811) to eliminate for B_2, we have

$$\frac{\omega_0 - \omega + \Omega}{\omega_x}B_1 + \frac{\omega_0 - \omega - \Omega}{\omega_x}(b - B_1) = a, \tag{812}$$

$$\frac{2\Omega}{\omega_x}B_1 = a - \frac{\omega_0 - \omega - \Omega}{\omega_x}b, \tag{813}$$

$$B_1 = \frac{\omega_x}{2\Omega}a - \frac{\omega_0 - \omega - \Omega}{2\Omega}b, \tag{814}$$

$$A_1 = \frac{\omega_0 - \omega + \Omega}{\omega_x}B_1 = \frac{\omega_0 - \omega + \Omega}{2\Omega}a - \frac{(\omega_0 - \omega)^2 - \Omega^2}{2\omega_x\Omega}b, \tag{815}$$

$$B_2 = b - B_1 = -\frac{\omega_x}{2\Omega}a + \frac{\omega_0 - \omega + \Omega}{2\Omega}b, \tag{816}$$

$$A_2 = \frac{\omega_0 - \omega - \Omega}{\omega_x}B_2 = -\frac{\omega_0 - \omega - \Omega}{2\Omega}a + \frac{(\omega_0 - \omega)^2 - \Omega^2}{2\omega_x\Omega}b. \tag{817}$$

Using eqs. (814)-(817) in eq. (809), we have

$$\begin{aligned}|\psi(t)_{\rm lab}\rangle &= \left\{a\cos\Omega t + i\left[-\frac{\omega_0-\omega}{\Omega}a + \frac{(\omega_0-\omega)^2-\Omega^2}{\omega_x\Omega}b\right]\sin\Omega t\right\}e^{i\frac{\omega}{2}t}|0\rangle \\ &\quad + \left\{i\left[\frac{\omega_x}{\Omega}a - \frac{\omega_0-\omega}{\Omega}b\right]\sin\Omega t + b\cos\Omega t\right\}e^{-i\frac{\omega}{2}t}|1\rangle. \qquad (818)\end{aligned}$$

In particular, if $|\psi(0)\rangle = |0\rangle$, then

$$|\psi(t)_{\rm lab}\rangle = \left(\cos\Omega t - i\frac{\omega_0-\omega}{\Omega}\sin\Omega t\right)e^{i\frac{\omega}{2}t}|0\rangle + i\frac{\omega_x}{\Omega}\sin\Omega t e^{-i\frac{\omega}{2}t}|1\rangle. \qquad (819)$$

If in addition the frequency ω of the oscillatory field B_x is equal to the Larmor frequency ω_0 that corresponds to the spin-flip energy, then

$$\Omega = \omega_x \qquad (\omega_0 = \omega), \qquad (820)$$

and the time dependence of the Qbit is

$$|\psi(t)_{\rm lab}\rangle = \cos\omega_x t e^{i\frac{\omega_0}{2}t}|0\rangle + i\sin\omega_x t e^{-i\frac{\omega_0}{2}t}|1\rangle \qquad (\omega_0=\omega,\ |\psi(0)\rangle = |0\rangle), \quad (821)$$

which is commonly called spin resonance or magnetic resonance. Similarly, in the case of spin resonance when the initial state of the Qbit is $|1\rangle$, we have

$$|\psi(t)_{\rm lab}\rangle = -i\sin\omega_x t e^{i\frac{\omega_0}{2}t}|0\rangle + \cos\omega_x t e^{-i\frac{\omega_0}{2}t}|1\rangle \qquad (\omega_0=\omega,\ |\psi(0)\rangle = |1\rangle). \quad (822)$$

If the oscillatory field B_x is applied for a time t such that $\omega_x t = \pi/2$, then there is 100% probability that the Qbit state will have been flipped. That is, a NOT operation (up to a phase) for spin-based Qbits consists of application of a transverse magnetic field of frequency ω_0 for 1/4 of the Rabi period, *i.e.*, for time $t = \pi/2\omega_x$.

(b) **Two-Bit Coupling**

We wish to ascertain the character of the unitary transformation U_{12} that characterizes the time evolution over a period $t = \pi/2\omega_{12}$ of two spin-based Qbits (as viewed in their respective rotating frames) whose (lab-frame) interaction Hamiltonian is

$$\mathsf{h}_{12} = \frac{\omega_{12}}{2}\boldsymbol{\sigma}^{(1)}\cdot\boldsymbol{\sigma}^{(2)}. \qquad (244)$$

We have that

$$|\psi'_{\rm rot}\rangle_{12} = e^{-i\frac{\omega_{12}t}{2}\boldsymbol{\sigma}_x^{(1)}\boldsymbol{\sigma}_x^{(2)}}e^{-i\frac{\omega_{12}t}{2}\boldsymbol{\sigma}_y^{(1)}\boldsymbol{\sigma}_y^{(2)}}e^{-i\frac{\omega_{12}t}{2}\boldsymbol{\sigma}_z^{(1)}\cdot\boldsymbol{\sigma}_z^{(2)}}|\psi_{\rm rot}\rangle_{12} = \mathsf{U}_{12}|\psi_{\rm rot}\rangle_{12}. \quad (245)$$

Since

$$e^{-i\alpha\boldsymbol{\sigma}_j^{(1)}\boldsymbol{\sigma}_j^{(2)}} = \cos\alpha\ \mathbf{I} - i\sin\alpha\ \boldsymbol{\sigma}_j^{(1)}\boldsymbol{\sigma}_j^{(2)}, \qquad (823)$$

where $\alpha = \omega_{12}t/2 = \pi/4$ in the present case, we need to evaluate the operation $(\mathbf{I} - i\boldsymbol{\sigma}_j^{(1)}\boldsymbol{\sigma}_j^{(2)})/\sqrt{2}$ for $j = x, y, z$. Recalling eq. (647) for the tensor product of two 2-Qbit operations, we have

$$\boldsymbol{\sigma}_x^{(1)}\boldsymbol{\sigma}_x^{(2)} = \left(\begin{array}{c|c} 0 & \boldsymbol{\sigma}_x \\ \hline \boldsymbol{\sigma}_x & 0 \end{array}\right) = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \qquad (824)$$

and so

$$e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_x^{(1)}\boldsymbol{\sigma}_x^{(2)}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 0 & -i \\ 0 & 1 & -i & 0 \\ 0 & -i & 1 & 0 \\ -i & 0 & 0 & 1 \end{pmatrix}. \tag{825}$$

Similarly,

$$\boldsymbol{\sigma}_y^{(1)}\boldsymbol{\sigma}_y^{(2)} = \left(\begin{array}{c|c} 0 & -i\boldsymbol{\sigma}_y \\ \hline i\boldsymbol{\sigma}_y & 0 \end{array}\right) = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \tag{826}$$

$$e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(1)}\boldsymbol{\sigma}_y^{(2)}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 0 & i \\ 0 & 1 & -i & 0 \\ 0 & -i & 1 & 0 \\ i & 0 & 0 & 1 \end{pmatrix}, \tag{827}$$

$$\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)} = \left(\begin{array}{c|c} \boldsymbol{\sigma}_z & 0 \\ \hline 0 & -\boldsymbol{\sigma}_z \end{array}\right) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{828}$$

and

$$e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1-i & 0 & 0 & 0 \\ 0 & 1+i & 0 & 0 \\ 0 & 0 & 1+i & 0 \\ 0 & 0 & 0 & 1-i \end{pmatrix}. \tag{829}$$

Then,

$$\begin{aligned} e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_x^{(1)}\boldsymbol{\sigma}_x^{(2)}} e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(1)}\boldsymbol{\sigma}_y^{(2)}} &= \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & -i \\ 0 & 1 & -i & 0 \\ 0 & -i & 1 & 0 \\ -i & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & i \\ 0 & 1 & -i & 0 \\ 0 & -i & 1 & 0 \\ i & 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \end{aligned} \tag{830}$$

This is almost the swap operation (572), except for the phase change between the swaps $|0\rangle|0\rangle \leftrightarrow |1\rangle|1\rangle$ and the swaps $|0\rangle|1\rangle \leftrightarrow |1\rangle|0\rangle$. This discrepancy is fixed by

the operator $e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}}$:

$$\begin{aligned} e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_x^{(1)}\boldsymbol{\sigma}_x^{(2)}}e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(1)}\boldsymbol{\sigma}_y^{(2)}}e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}} &= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1-i & 0 & 0 & 0 \\ 0 & 1+i & 0 & 0 \\ 0 & 0 & 1+i & 0 \\ 0 & 0 & 0 & 1-i \end{pmatrix} \\ &= \frac{1-i}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \frac{1-i}{\sqrt{2}}\mathsf{S}_{12}. \end{aligned} \tag{831}$$

To show that

$$\mathsf{C}_{12} = e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}}e^{i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(2)}}e^{i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(2)}}e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}}e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(2)}}, \tag{247}$$

we accumulate the 4×4 matrix representations of the various operators. Thus,

$$e^{\pm i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(2)}} = \frac{\mathbf{I}\pm i\mathbf{I}^{(1)}\boldsymbol{\sigma}_y^{(2)}}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & \pm 1 & 0 & 0 \\ \mp 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & \pm 1 \\ 0 & 0 & \mp 1 & 1 \end{pmatrix}, \tag{832}$$

$$e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}} = \frac{\mathbf{I}-i\boldsymbol{\sigma}_z^{(1)}\mathbf{I}^{(2)}}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1-i & 0 & 0 & 0 \\ 0 & 1-i & 0 & 0 \\ 0 & 0 & 1+i & 0 \\ 0 & 0 & 0 & 1+i \end{pmatrix}, \tag{833}$$

and

$$e^{i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(2)}} = \frac{\mathbf{I}+i\mathbf{I}^{(1)}\boldsymbol{\sigma}_z^{(2)}}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1+i & 0 & 0 & 0 \\ 0 & 1-i & 0 & 0 \\ 0 & 0 & 1+i & 0 \\ 0 & 0 & 0 & 1-i \end{pmatrix}. \tag{834}$$

Then,

$$\begin{aligned} & e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}}e^{i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(2)}}e^{i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(2)}}e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}}e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(2)}} \\ = \; & \frac{1}{2}e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}}e^{i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(2)}}e^{i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(2)}}\begin{pmatrix} 1-i & 0 & 0 & 0 \\ 0 & 1+i & 0 & 0 \\ 0 & 0 & 1+i & 0 \\ 0 & 0 & 0 & 1-i \end{pmatrix}\begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2\sqrt{2}} e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}} e^{i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(2)}} \begin{pmatrix} 1+i & 0 & 0 & 0 \\ 0 & 1-i & 0 & 0 \\ 0 & 0 & 1+i & 0 \\ 0 & 0 & 0 & 1-i \end{pmatrix} \begin{pmatrix} 1-i & -1+i & 0 & 0 \\ 1+i & 1+i & 0 & 0 \\ 0 & 0 & 1+i & -1-i \\ 0 & 0 & 1-i & 1-i \end{pmatrix} \\
&= \frac{1}{2} e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}} \begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & i & -i \\ 0 & 0 & -i & -i \end{pmatrix} \\
&= \frac{1}{\sqrt{2}} \begin{pmatrix} 1-i & 0 & 0 & 0 \\ 0 & 1-i & 0 & 0 \\ 0 & 0 & 1+i & 0 \\ 0 & 0 & 0 & 1+i \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & -i & 0 \end{pmatrix} \\
&= \frac{1-i}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \frac{1-i}{\sqrt{2}} \mathsf{C}_{12}. \qquad (835)
\end{aligned}
$$

To verify that

$$
\mathsf{C}_{21} = e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(2)}} e^{i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(1)}} e^{i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}} e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}} e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(1)}}, \qquad (248)
$$

we note that

$$
e^{\pm i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(1)}} = \frac{\mathbf{I} \pm i\boldsymbol{\sigma}_y^{(1)}\mathbf{I}^{(2)}}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & \pm 1 & 0 \\ 0 & 1 & 0 & \pm 1 \\ \mp 1 & 0 & 1 & 0 \\ 0 & \mp 1 & 0 & 1 \end{pmatrix}, \qquad (836)
$$

$$
e^{i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}} = \frac{\mathbf{I} + i\boldsymbol{\sigma}_z^{(1)}\mathbf{I}^{(2)}}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1+i & 0 & 0 & 0 \\ 0 & 1+i & 0 & 0 \\ 0 & 0 & 1-i & 0 \\ 0 & 0 & 0 & 1-i \end{pmatrix}, \qquad (837)
$$

and

$$
e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(2)}} = \frac{\mathbf{I} - i\mathbf{I}^{(1)}\boldsymbol{\sigma}_z^{(2)}}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1-i & 0 & 0 & 0 \\ 0 & 1+i & 0 & 0 \\ 0 & 0 & 1-i & 0 \\ 0 & 0 & 0 & 1+i \end{pmatrix}. \qquad (838)
$$

Then,

$$
\begin{aligned}
& e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(2)}} e^{i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(1)}} e^{i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}} e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}} e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(1)}} \\
= & \frac{1}{2} e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(2)}} e^{i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(1)}} e^{i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}} \begin{pmatrix} 1-i & 0 & 0 & 0 \\ 0 & 1+i & 0 & 0 \\ 0 & 0 & 1+i & 0 \\ 0 & 0 & 0 & 1-i \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \\
= & \frac{1}{2\sqrt{2}} e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(2)}} e^{i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(1)}} \begin{pmatrix} 1+i & 0 & 0 & 0 \\ 0 & 1+i & 0 & 0 \\ 0 & 0 & 1-i & 0 \\ 0 & 0 & 0 & 1-i \end{pmatrix} \begin{pmatrix} 1-i & 0 & -1+i & 0 \\ 0 & 1+i & 0 & -1-i \\ 1+i & 0 & 1+i & 0 \\ 0 & 1-i & 0 & 1-i \end{pmatrix} \\
= & \frac{1}{2} e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(2)}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & i & 0 & -i \\ 1 & 0 & 1 & 0 \\ 0 & -i & 0 & -i \end{pmatrix} \\
= & \frac{1}{\sqrt{2}} \begin{pmatrix} 1-i & 0 & 0 & 0 \\ 0 & 1+i & 0 & 0 \\ 0 & 0 & 1-i & 0 \\ 0 & 0 & 0 & 1+i \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & 1 & 0 \\ 0 & -i & 0 & 0 \end{pmatrix} \\
= & \frac{1-i}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = \frac{1-i}{\sqrt{2}} \mathsf{C}_{21}.
\end{aligned} \tag{839}
$$

To verify that the 2-Qbit operation

$$
e^{-i\frac{\pi}{4}\boldsymbol{\sigma}^{(1)}\cdot\boldsymbol{\sigma}^{(2)}} = e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_x^{(1)}\boldsymbol{\sigma}_x^{(2)}} e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(1)}\boldsymbol{\sigma}_y^{(2)}} e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}} \tag{840}
$$

is the SWAP gate via a 'geometrical" argument, we first note some relations between Qbit states $|0\rangle$ and $|1\rangle$ and various units vectors on the Bloch sphere. We recall that a unit vector with polar angle α and azimuthal angle β corresponds to the Qbit

$$
|\psi\rangle = \cos\frac{\alpha}{2}|0\rangle + e^{i\beta}\sin\frac{\alpha}{2}|1\rangle. \tag{42}
$$

Thus we have,

α	β	$\hat{\mathbf{u}}$
$\frac{\pi}{2}$	0	$\|\hat{\mathbf{x}}\rangle = \frac{\|0\rangle+\|1\rangle}{\sqrt{2}}$,
$\frac{\pi}{2}$	π	$\|-\hat{\mathbf{x}}\rangle = \frac{\|0\rangle-\|1\rangle}{\sqrt{2}}$,
$\frac{\pi}{2}$	$\frac{\pi}{2}$	$\|\hat{\mathbf{y}}\rangle = \frac{\|0\rangle+i\|1\rangle}{\sqrt{2}}$,
$\frac{\pi}{2}$	$-\frac{\pi}{2}$	$\|-\hat{\mathbf{y}}\rangle = \frac{\|0\rangle-i\|1\rangle}{\sqrt{2}}$,
0	0	$\|\hat{\mathbf{z}}\rangle = \|0\rangle$,
π	0	$\|-\hat{\mathbf{z}}\rangle = \|1\rangle$.

(841)

We also need the inverse relations,

$$|\hat{\mathbf{z}}\rangle = |0\rangle = \frac{|\hat{\mathbf{x}}\rangle + |-\hat{\mathbf{x}}\rangle}{\sqrt{2}} = \frac{|\hat{\mathbf{y}}\rangle + |-\hat{\mathbf{y}}\rangle}{\sqrt{2}}, \tag{842}$$

$$|-\hat{\mathbf{z}}\rangle = |1\rangle = \frac{|\hat{\mathbf{x}}\rangle - |-\hat{\mathbf{x}}\rangle}{\sqrt{2}} = -i\frac{|\hat{\mathbf{y}}\rangle - |-\hat{\mathbf{y}}\rangle}{\sqrt{2}}. \tag{843}$$

We now examine the effect of operation (840) on the four 2-Qbit states $|0\rangle_1|0\rangle_2 = |\hat{\mathbf{z}}\rangle_1|\hat{\mathbf{z}}\rangle_2$, $|0\rangle_1|1\rangle_2 = |\hat{\mathbf{z}}\rangle_1|-\hat{\mathbf{z}}\rangle_2$, $|1\rangle_1|0\rangle_2 = |-\hat{\mathbf{z}}\rangle_1|\hat{\mathbf{z}}\rangle_2$ and $|1\rangle_1|1\rangle_2 = |-\hat{\mathbf{z}}\rangle_1|-\hat{\mathbf{z}}\rangle_2$.

The first step of operation (840) is $e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}}$ which performs conditional rotations by $\pm 90°$ about the z-axes. Since all of our initial states are aligned along the z axes, the first step merely changes the phases of the initial states, but not their directions. In the geometric view, this step has no effect.

The second step of operation (840) is $e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(1)}\boldsymbol{\sigma}_y^{(2)}}$ which performs a conditional rotation of bit 2 by $\pm 90°$ about the y_2-axis. depending on the state of bit 1 as projected onto the y_1-axis (or equivalently, a conditional rotation of bit 1 by $\pm 90°$ about the y_1-axis. depending on the state of bit 2 as projected onto the y_2-axis). This will result in the initial states, which were aligned along the z axes, being transformed into various combinations of states readily expressed with one bit along its x-axis and the other bit along its y-axis.

The third step of operation (840) is $e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_x^{(1)}\boldsymbol{\sigma}_x^{(2)}}$ which performs conditional rotations $\pm 90°$ about the x-axes. Since the input states to this operation are simply expressed with one of the two bits aligned along its x-axis, it is relatively straightforward to keep track of this step.

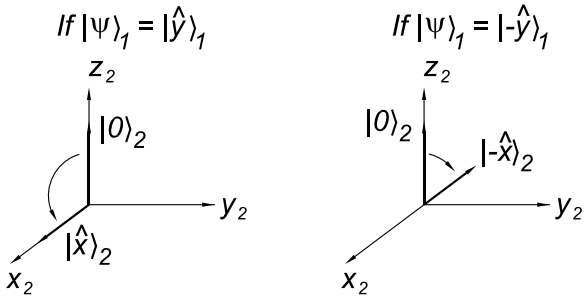
We examine the four initial states one by one:

i. We express the initial state in various equivalent ways,

$$|\psi_0\rangle = |0\rangle_1|0\rangle_2 = |\hat{\mathbf{z}}\rangle_1|\hat{\mathbf{z}}\rangle_2 = \frac{|\hat{\mathbf{y}}\rangle_1 + |-\hat{\mathbf{y}}\rangle_1}{\sqrt{2}}|\hat{\mathbf{z}}\rangle_2 = |\hat{\mathbf{z}}\rangle_1\frac{|\hat{\mathbf{y}}\rangle_2 + |-\hat{\mathbf{y}}\rangle_2}{\sqrt{2}}. \tag{844}$$

To determine the rotation of the second Qbit by the operation $e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(1)}\boldsymbol{\sigma}_y^{(2)}}$, which is conditional on the state of the first Qbit, we need the first Qbit to be described relative to the y_1 axis, as given by the fourth term of eq. (844).

Recalling from eq. (49) that the operator $e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(2)}}$ is a rotation of bit 2 by $+90°$ about the y_2-axis, we learn that when bit 1 is $|\hat{\mathbf{y}}\rangle_1$ bit 2 is rotated from $|\hat{\mathbf{z}}\rangle_2$ to $|\hat{\mathbf{x}}\rangle_2$, and that when bit 1 is $|-\hat{\mathbf{y}}\rangle_1$ bit 2 is rotated from $|\hat{\mathbf{z}}\rangle_2$ to $|-\hat{\mathbf{x}}\rangle_2$.



That is

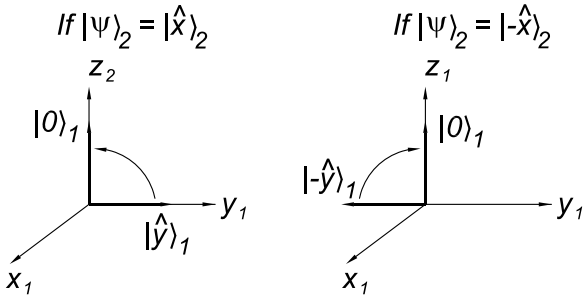
$$|0\rangle_1|0\rangle_2 = \frac{|\hat{\mathbf{y}}\rangle_1 + |-\hat{\mathbf{y}}\rangle_1}{\sqrt{2}}|\hat{\mathbf{z}}\rangle_2 \rightarrow \frac{|\hat{\mathbf{y}}\rangle_1|\hat{\mathbf{x}}\rangle_2 + |-\hat{\mathbf{y}}\rangle_1|-\hat{\mathbf{x}}\rangle_2}{\sqrt{2}}\,. \tag{845}$$

Alternatively, we could consider the 2nd bit to be the control bit, in which case the transformation can be written

$$|0\rangle_1|0\rangle_2 = |\hat{\mathbf{z}}\rangle_1\frac{|\hat{\mathbf{y}}\rangle_2 + |-\hat{\mathbf{y}}\rangle_2}{\sqrt{2}} \rightarrow \frac{|\hat{\mathbf{x}}\rangle_1|\hat{\mathbf{y}}\rangle_2 + |-\hat{\mathbf{x}}\rangle_1|-\hat{\mathbf{y}}\rangle_2}{\sqrt{2}}\,. \tag{846}$$

Using the relations (841) we can verify that the forms (845) and (846) are identical. Indeed, we find the final state to be $(|0\rangle_1|0\rangle_2 + i|1\rangle_1|1\rangle_2)/\sqrt{2}$, which is consistent with the matrix form (827).

To understand the effect of the conditional operation $e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_x^{(1)}\boldsymbol{\sigma}_x^{(2)}}$, one of the two bits should be expressed in terms of its projections onto the x-axis. Both eqs. (845) and (846) are already of the desired form, so we can use either. Specifically, the transformation of eq. (845) is when bit 2 is $|\hat{\mathbf{x}}\rangle_2$ bit 1 is rotated from $|\hat{\mathbf{y}}\rangle_1$ to $|\hat{\mathbf{z}}\rangle_1$, and that when bit 2 is $|-\hat{\mathbf{x}}\rangle_2$ bit 1 is rotated from $|-\hat{\mathbf{y}}\rangle_1$ to $|\hat{\mathbf{z}}\rangle_1$.



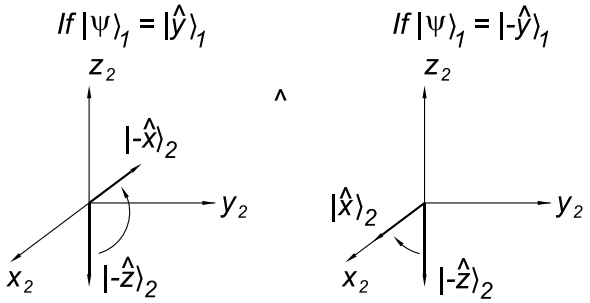
Then,

$$\frac{|\hat{\mathbf{y}}\rangle_1|\hat{\mathbf{x}}\rangle_2 + |-\hat{\mathbf{y}}\rangle_1|-\hat{\mathbf{x}}\rangle_2}{\sqrt{2}} \rightarrow \frac{|\hat{\mathbf{z}}\rangle_1|\hat{\mathbf{x}}\rangle_2 + |\hat{\mathbf{z}}\rangle_1|-\hat{\mathbf{x}}\rangle_2}{\sqrt{2}} = |0\rangle_1|0\rangle_2, \tag{847}$$

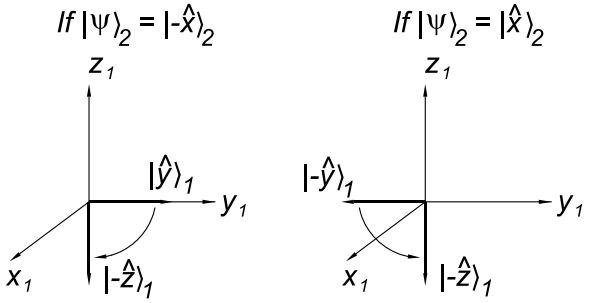
as expected.

ii. Similarly, when

$$|\psi_0\rangle = |0\rangle_1|1\rangle_2 = |\hat{\mathbf{z}}\rangle_1|-\hat{\mathbf{z}}\rangle_2 = \frac{|\hat{\mathbf{y}}\rangle_1 + |-\hat{\mathbf{y}}\rangle_1}{\sqrt{2}}|-\hat{\mathbf{z}}\rangle_2, \tag{848}$$



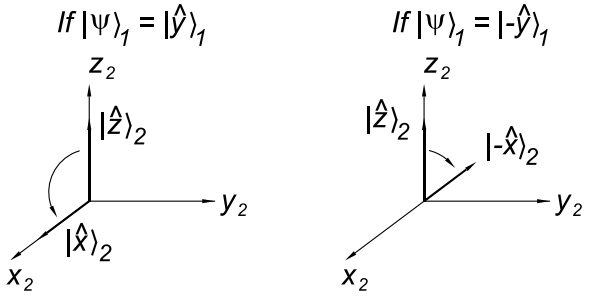
$$e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(1)}\boldsymbol{\sigma}_y^{(2)}}|0\rangle_1|1\rangle_2 = \frac{|\hat{\mathbf{y}}\rangle_1|-\hat{\mathbf{x}}\rangle_2 + |-\hat{\mathbf{y}}\rangle_1|\hat{\mathbf{x}}\rangle_2}{\sqrt{2}}\,, \tag{849}$$



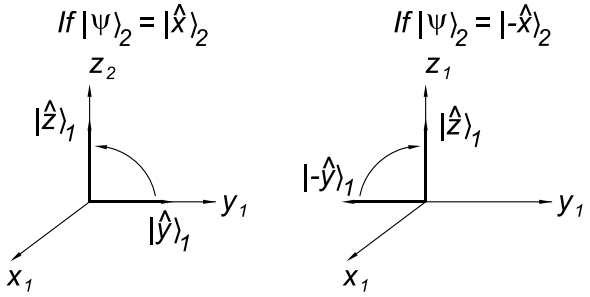
$$e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_x^{(1)}\boldsymbol{\sigma}_x^{(2)}}e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(1)}\boldsymbol{\sigma}_y^{(2)}}|0\rangle_1|1\rangle_2 = \frac{|-\hat{\mathbf{z}}\rangle_1|-\hat{\mathbf{x}}\rangle_2 + |-\hat{\mathbf{z}}\rangle_1|\hat{\mathbf{x}}\rangle_2}{\sqrt{2}} = |-\hat{\mathbf{z}}\rangle_1|\hat{\mathbf{z}}\rangle_2 = |1\rangle_1|0\rangle_2. \tag{850}$$

iii. Likewise,

$$|\psi_0\rangle = |1\rangle_1|0\rangle_2 = |-\hat{\mathbf{z}}\rangle_1|\hat{\mathbf{z}}\rangle_2 = -i\frac{|\hat{\mathbf{y}}\rangle_1 - |-\hat{\mathbf{y}}\rangle_1}{\sqrt{2}}|\hat{\mathbf{z}}\rangle_2, \tag{851}$$



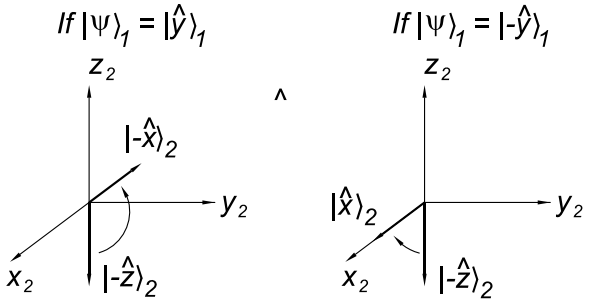
$$e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(1)}\boldsymbol{\sigma}_y^{(2)}}|1\rangle_1|0\rangle_2 = -i\frac{|\hat{\mathbf{y}}\rangle_1|\hat{\mathbf{x}}\rangle_2 - |-\hat{\mathbf{y}}\rangle_1|-\hat{\mathbf{x}}\rangle_2}{\sqrt{2}}\,, \tag{852}$$



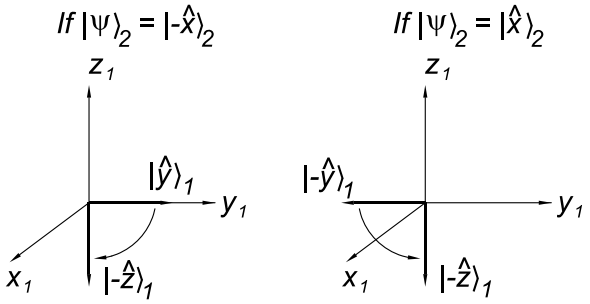
$$e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_x^{(1)}\boldsymbol{\sigma}_x^{(2)}}e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(1)}\boldsymbol{\sigma}_y^{(2)}}|1\rangle_1|0\rangle_2 = -i\frac{|\hat{\mathbf{z}}\rangle_1|\hat{\mathbf{x}}\rangle_2 - |\hat{\mathbf{z}}\rangle_1|-\hat{\mathbf{x}}\rangle_2}{\sqrt{2}} = i|\hat{\mathbf{z}}\rangle_1|-\hat{\mathbf{z}}\rangle_2 = i|1\rangle_1|0\rangle_2. \tag{853}$$

iv. Finally,

$$|\psi_0\rangle = |1\rangle_1|1\rangle_2 = |-\hat{\mathbf{z}}\rangle_1|-\hat{\mathbf{z}}\rangle_2 = -i\frac{|\hat{\mathbf{y}}\rangle_1 - |-\hat{\mathbf{y}}\rangle_1}{\sqrt{2}}|-\hat{\mathbf{z}}\rangle_2, \tag{854}$$



$$e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(1)}\boldsymbol{\sigma}_y^{(2)}}|1\rangle_1|1\rangle_2 = -i\frac{|\hat{\mathbf{y}}\rangle_1|-\hat{\mathbf{x}}\rangle_2 - |-\hat{\mathbf{y}}\rangle_1|\hat{\mathbf{x}}\rangle_2}{\sqrt{2}}, \tag{855}$$



$$e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_x^{(1)}\boldsymbol{\sigma}_x^{(2)}}e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(1)}\boldsymbol{\sigma}_y^{(2)}}|1\rangle_1|1\rangle_2 = -i\frac{|-\hat{\mathbf{z}}\rangle_1|-\hat{\mathbf{x}}\rangle_2 - |-\hat{\mathbf{z}}\rangle_1|\hat{\mathbf{x}}\rangle_2}{\sqrt{2}} = i|-\hat{\mathbf{z}}\rangle_1|-\hat{\mathbf{z}}\rangle_2 = i|1\rangle_1|1\rangle_2. \tag{856}$$

Equations (847), (850), (853) and (856) verify that the operation $e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_x^{(1)}\boldsymbol{\sigma}_x^{(2)}}e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_y^{(1)}\boldsymbol{\sigma}_y^{(2)}}$ performs a SWAP up to a phase, although the final phases appear to be different for the four 2-Qbit basis states. Our algebraic analysis, eq. (831), shows that the additional phase changes introduced by the operations $e^{-i\frac{\pi}{4}\boldsymbol{\sigma}_z^{(1)}\boldsymbol{\sigma}_z^{(2)}}$ result in a common phase for the final states of all four 2-Qbit basis states, as is desirable.

20. **Dephasing**

(a) The density matrix for a Qbit is a 2×2 hermitian matrix whose trace is 1. So we can write

$$\boldsymbol{\rho} = \frac{\mathbf{I} + \mathsf{A}}{2}\,, \tag{857}$$

where matrix A is also hermitian, but with zero trace. Then, we have

$$\mathsf{A} = \begin{pmatrix} z & b \\ c & -z \end{pmatrix} = \mathsf{A}^\dagger = \begin{pmatrix} z^\star & c^\star \\ b^\star & -z^\star \end{pmatrix}. \tag{858}$$

Therefore, z is real, and if $b = x - iy$ then $c = x + iy$, so that

$$\mathsf{A} = x \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + y \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + z \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \mathbf{r} \cdot \boldsymbol{\sigma}, \tag{859}$$

where $\mathbf{r} = (x, y, z)$ is a real 3-vector. Thus

$$\boldsymbol{\rho} = \frac{\mathbf{I} + \mathbf{r} \cdot \boldsymbol{\sigma}}{2}\,, \tag{289}$$

as claimed.

If this density matrix represents a pure state, then

$$\frac{\mathbf{I} + \mathbf{r} \cdot \boldsymbol{\sigma}}{2} = \boldsymbol{\rho} = \boldsymbol{\rho}^2 = \frac{\mathbf{I} + 2\,\mathbf{r} \cdot \boldsymbol{\sigma} + (\mathbf{r} \cdot \boldsymbol{\sigma})^2}{4} = \frac{\mathbf{I}(1 + |\mathbf{r}|^2) + 2\,\mathbf{r} \cdot \boldsymbol{\sigma}}{4}\,, \tag{860}$$

using eq. (36). Hence, $|\mathbf{r}|^2 = 1$ for a pure state.

If the density matrix represents a mixed state,

$$\boldsymbol{\rho} = \sum_i P_i |\psi_i\rangle\langle\psi_i| = \sum_i P_i \boldsymbol{\rho}_i, \tag{276}$$

then each of the component pure-state density matrices can be represented in the form (289) with a corresponding unit vector $\hat{\mathbf{r}}_i$. Thus, the vector $\mathbf{r}$ for the density matrix (276) obeys

$$\mathbf{r} = \sum_i P_i \hat{\mathbf{r}}_i, \qquad \text{where} \qquad \sum P_i = 1. \tag{861}$$

Hence $|\mathbf{r}| \le 1$, and the bound is achieved only if all $\hat{\mathbf{r}}_i$ are identical, in which case the density matrix actually represents a pure state.

For a pure state

$$|\psi\rangle = e^{i\gamma} \left[\cos\frac{\alpha}{2} |0\rangle + e^{i\beta} \sin\frac{\alpha}{2} |1\rangle \right], \tag{42}$$

the density matrix is

$$\begin{aligned} \boldsymbol{\rho} &= \begin{pmatrix} \cos^2\frac{\alpha}{2} & \cos\alpha \sin\alpha e^{-i\beta} \\ \cos\alpha \sin\alpha e^{i\beta} & \sin^2\frac{\alpha}{2} \end{pmatrix} \\ &= \frac{\mathbf{I} + \begin{pmatrix} \cos\alpha & \sin\alpha(cos\beta - i\sin\beta) \\ \sin\alpha(\cos\beta + i\sin\beta) & -\cos\alpha \end{pmatrix}}{2}\,. \end{aligned} \tag{862}$$

From this we read off the components of $\hat{\mathbf{r}}$ as

$$\hat{\mathbf{r}} = (\sin\alpha\cos\beta, \sin\alpha\sin\beta, \cos\alpha), \tag{863}$$

which corresponds to a unit vector in the direction (α, β) in a spherical coordinate system in Bloch space, consistent with our geometric interpretation of a Qbit in prob. 4.

(b) The density operator ρ for the pure states $|\pm\rangle = (|0\rangle_A|\text{vac}\rangle_\text{B} \pm |\text{vac}\rangle_\text{A}|1\rangle_\text{B})/\sqrt{2}$ is

$$\begin{aligned}\rho &= |\pm\rangle\langle\pm| = \frac{|0\rangle_A|\text{vac}\rangle_\text{B} \pm |\text{vac}\rangle_\text{A}|1\rangle_\text{B}}{\sqrt{2}} \frac{\langle 0|_A\langle\text{vac}|_\text{B} \pm \langle\text{vac}|_\text{A}\langle 1|_\text{B}}{\sqrt{2}} \\ &= \frac{|0\rangle_A\langle 0|_A\ |\text{vac}\rangle_\text{B}\langle\text{vac}|_\text{B}}{2} + \frac{|\text{vac}\rangle_\text{A}\langle\text{vac}|_\text{A}\ |1\rangle_\text{B}\langle 1|_\text{B}}{2} \\ &\quad \pm\frac{|0\rangle_A\langle\text{vac}|_\text{A}\ |\text{vac}\rangle_\text{B}\langle 1|_\text{B}}{2} \pm \frac{|\text{vac}\rangle_\text{A}\langle 0|_\text{A}\ |1\rangle_\text{B}\langle\text{vac}|_\text{B}}{2}\,. \end{aligned} \tag{864}$$

Taking the partial trace over subsystem B with the aid of eq. (281), we find

$$\begin{aligned}\rho_A &= \text{tr}_B(\rho) = \frac{|0\rangle_A\langle 0|_A\langle\text{vac}_B|\text{vac}_B\rangle}{2} + \frac{|\text{vac}\rangle_\text{A}\langle\text{vac}|_\text{A}\langle 1_\text{B}|1_\text{B}\rangle}{2} \\ &\quad \pm\frac{|0\rangle_A\langle\text{vac}|_\text{A}\langle\text{vac}_\text{B}|1_\text{B}\rangle}{2} + \frac{|\text{vac}\rangle_\text{A}\langle 0|_\text{A}\langle 1_\text{B}|\text{vac}_\text{B}\rangle}{2} \\ &= \frac{|0\rangle_A\langle 0|_A\langle\text{vac}_B|\text{vac}_B\rangle}{2} + \frac{|\text{vac}\rangle_\text{A}\langle\text{vac}|_\text{A}\langle 1_\text{B}|1_\text{B}\rangle}{2} = \frac{\mathbf{I}}{2}\,. \end{aligned} \tag{865}$$

The pure state (864) of our spatially encoded Qbit is an entangled state of the Qbits of subsystems A and B. However, from the perspective of an observer of subsystem A who is ignorant of subsystem B, the state of A is given by eq. (865) which is a mixed state comprised of $|0\rangle_A$ with 50% probability, and state $|\text{vac}\rangle_\text{A}$ with 50% probability.

I now speculate on additional insights to be obtained from this part.

Einstein often said that the "real" situations of two spatially separated (sub)systems should be independent of one another (meaning that one system cannot affect the other in a way that implies faster-than-light transmission of a signal).

If separate observers A and B (who are at rest with respect to each other) look at their respective systems at the same time (as determined by previously synchronized clocks), the combined results of their observations of the spatially encoded Qbit will be that exactly one of A or B finds a particle present in their system.

Einstein appears to have concluded from this that the initial state of the system AB was not a pure state, but rather a mixed state,

$$\rho_\text{Einstein} = \frac{|0\rangle_A\langle 0|_A\ |\text{vac}\rangle_\text{B}\langle\text{vac}|_\text{B}}{2} + \frac{|\text{vac}\rangle_\text{A}\langle\text{vac}|_\text{A}\ |1\rangle_\text{B}\langle 1|_\text{B}}{2}\,. \tag{866}$$

An interpretation of eq. (866) is that the particle "really" was in subsystem A, or in subsystem B, all along, but we don't know which until we "look".

Note that the formal knowledge of observer A, if ignorant about system B (as seems natural if subsystems A and B are spatially separated), is the same for both the pure state (864) and for Einstein's mixed state (866),

$$\rho_A = \mathrm{tr}_B(\rho) = \mathrm{tr}_B(\rho_{\mathrm{Einstein}}) = \frac{\mathbf{I}}{2}\,. \tag{867}$$

I would like to argue (with Einstein, if he were alive), that this shows how the characterization of knowledge of quantum systems via density operators satisfies the criterion of separability that Einstein insisted upon.[176]

However, it remains that the pure state (864) is a more subtle entity than Einstein's mixed state (866). We claim that pure states such as eq. (864) can exist in Nature, and that the particle is not "really" anywhere until it is observed to be somewhere. And further, that the correlations implied in the form (864) insure that the particle will only be observed in one place, without any signal being sent between the separated subsystems in which it might appear.

(c) The state of interest of the 3-Qbit system used for quantum teleportation is,

$$|\psi_E\rangle = \alpha\frac{|000\rangle + |100\rangle + |011\rangle + |111\rangle}{2} + \beta\frac{|010\rangle - |110\rangle + |001\rangle - |101\rangle}{2}\,. \tag{582}$$

where the initial state of the first bit was $|a\rangle = \alpha|0\rangle + \beta|1\rangle$.

To take the partial trace over bits $|a\rangle$ and $|b\rangle$, the appropriate version of the rule (281) is,

$$\mathrm{tr}_{AB}(|A_1B_1C_1\rangle\langle A_2B_2C_2|) = |C_1\rangle\langle C_2|\ \langle A_1B_1|A_2B_2\rangle. \tag{868}$$

Since the bits $|0\rangle_c$ and $|1\rangle_c$ each appear in two of the basis states of eq. (582) that are multiplied by α and in two that are multiplied by β, the Bob's reduced density matrix simplifies to,

$$\boldsymbol{\rho}_{\mathrm{Bob,E}} = \mathrm{tr}_{ab}(|\psi_E\rangle\langle\psi_E|) = \frac{|\alpha|^2 + |\beta|^2}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) = \frac{\mathbf{I}}{2}\,. \tag{869}$$

So, at the time when the state of the system is $|\psi_E\rangle$, Bob's knowledge of the system, as summarized in eq. (869) includes no information about the amplitudes α and β of the initial state of bit $|a\rangle$.

Only after getting additional information from Alice can he convert his density matrix [at step I of the figure of solution 6(d)] to

$$\boldsymbol{\rho}_{\mathrm{Bob,I}} = \begin{pmatrix} |\alpha|^2 & \alpha\beta^\star \\ \alpha^\star\beta & |\beta|^2 \end{pmatrix} = (\alpha|0\rangle + \beta|1\rangle)(\alpha^\star\langle 0| + \beta^\star\langle 1|). \tag{870}$$

For completeness, I note that the density matrix that represents Alice's knowledge only, when the system is in state $|\psi\rangle_E$, is

$$\boldsymbol{\rho}_{\mathrm{Alice,E}} = \mathrm{tr}_c(\boldsymbol{\rho}_E) = \frac{1}{4}\begin{pmatrix} 1 & 2Re(\alpha\beta^\star) & |\alpha|^2 - |\beta|^2 & 2Re(\alpha\beta^\star) \\ 2Re(\alpha\beta^\star) & 1 & 2iIm(\alpha\beta^\star) & |\alpha|^2 - |\beta|^2 \\ |\alpha|^2 - |\beta|^2 & -2iIm(\alpha\beta^\star) & 1 & -2Re(\alpha\beta^\star) \\ 2Re(\alpha\beta^\star) & |\alpha|^2 - |\beta|^2 & -2Re(\alpha\beta^\star) & 1 \end{pmatrix}, \tag{871}$$

[176] For a variant of this arqument, see http://kirkmcd.princeton.edu/examples/QM/cantrell_pr_43_499_78.pdf

which has nontrivial off-diagonal elements that contain information as to the initial state $|a\rangle = \alpha|0\rangle + \beta|1\rangle$.

(d) **Other types of bit errors**

The error transformation of a single Qbit $|\psi\rangle$ has the form

$$\boldsymbol{\rho}'_\psi(p, \boldsymbol{\sigma}_j) = (1-p)\boldsymbol{\rho}_\psi + p\ \boldsymbol{\sigma}_j\boldsymbol{\rho}_\psi\boldsymbol{\sigma}_j, \tag{301}$$

where p is the probability that the error occurs, and $j = x$, y or z.

The density operator for a pure state $|\psi\rangle$ has the form

$$\boldsymbol{\rho}_\psi = \frac{\mathbf{I} + \hat{\mathbf{r}}\cdot\boldsymbol{\sigma}}{2}, \tag{289}$$

where $|\hat{\mathbf{r}}|^2 = 1$. Hence,

$$\begin{aligned}
\boldsymbol{\rho}'_\psi &= (1-p)\frac{\mathbf{I} + \hat{\mathbf{r}}\cdot\boldsymbol{\sigma}}{2} + p\ \boldsymbol{\sigma}_j\frac{\mathbf{I} + \hat{\mathbf{r}}\cdot\boldsymbol{\sigma}}{2}\boldsymbol{\sigma}_j \\
&= \frac{\mathbf{I}}{2} + (1-p)\frac{\hat{\mathbf{r}}\cdot\boldsymbol{\sigma}}{2} + p\ \boldsymbol{\sigma}_j\frac{\hat{\mathbf{r}}\cdot\boldsymbol{\sigma}}{2}\boldsymbol{\sigma}_j \\
&= \frac{\mathbf{I} + \hat{\mathbf{r}}\cdot[(1-p)\boldsymbol{\sigma} + p\ \boldsymbol{\sigma}_j\boldsymbol{\sigma}\boldsymbol{\sigma}_j]}{2}.
\end{aligned} \tag{872}$$

We write $\boldsymbol{\sigma} = \boldsymbol{\sigma}_j + \boldsymbol{\sigma}_\perp$. Then,

$$\boldsymbol{\sigma}_j\boldsymbol{\sigma}\boldsymbol{\sigma}_j = \boldsymbol{\sigma}_j - \boldsymbol{\sigma}_\perp, \tag{873}$$

$$(1-p)\boldsymbol{\sigma} + p\ \boldsymbol{\sigma}_j\boldsymbol{\sigma}\boldsymbol{\sigma}_j = (1-p)(\boldsymbol{\sigma}_j + \boldsymbol{\sigma}_\perp) + p\ (\boldsymbol{\sigma}_j - \boldsymbol{\sigma}_\perp) = \boldsymbol{\sigma}_j + (1-2p)\boldsymbol{\sigma}_\perp, \tag{874}$$

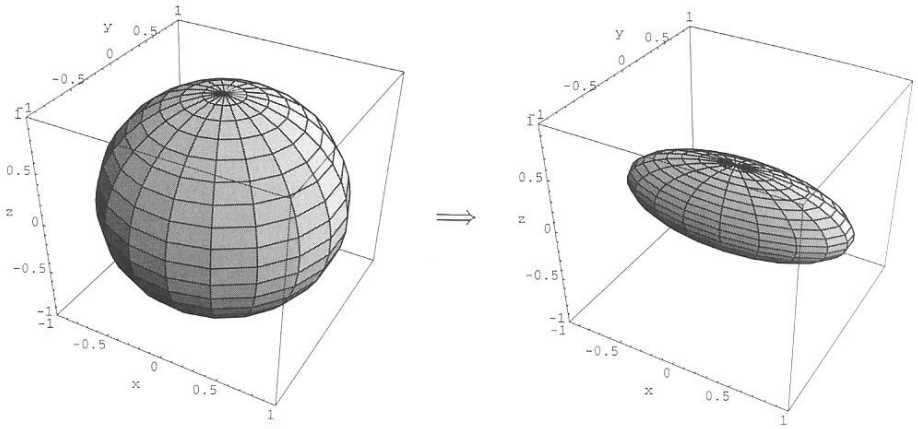
and so

$$\boldsymbol{\rho}'_\psi = \frac{\mathbf{I} + \hat{\mathbf{r}}\cdot[\boldsymbol{\sigma}_j + (1-2p)\boldsymbol{\sigma}_\perp]}{2} = \frac{\mathbf{I} + \mathbf{s}\cdot\boldsymbol{\sigma}}{2}, \tag{875}$$

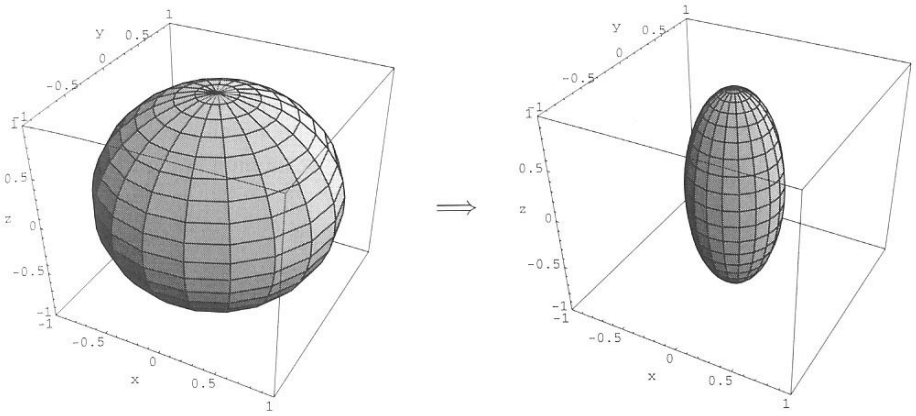
where the vector $\mathbf{s}$ in Bloch space has components $s_j = 1$ and $s_\perp = 1 - 2p$. The surface defined by the vectors $\mathbf{s}$ is an ellipsoid with major axis of length 1 along the j direction, and minor axes of length $1 - 2p$. *For $p > 1/2$ this geometric picture is poorly defined.*

The following figures, from sec. 8.3.3 of Nielsen and Chuang, illustrate the shrinkage of the surface in Bloch space corresponding to a Qbit due to errors of types $\boldsymbol{\sigma}_x$, $\boldsymbol{\sigma}_x$ and $\boldsymbol{\sigma}_z$ when the error probability is $p = 0.3$.

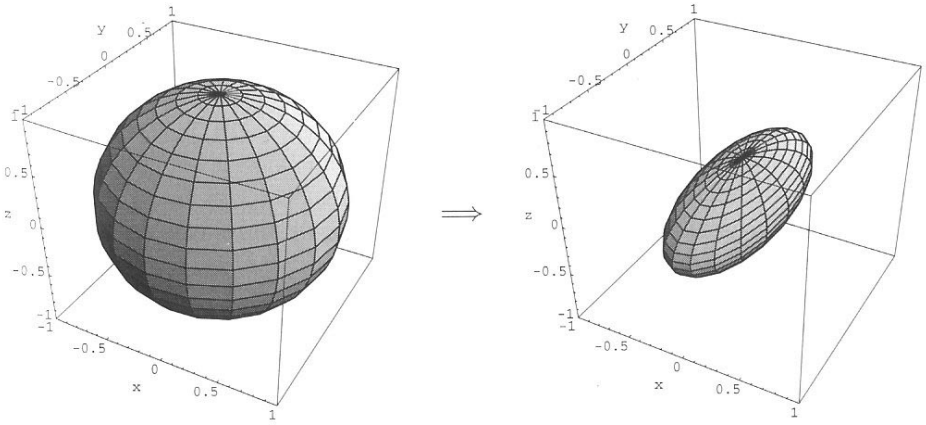
Bit flip ($\boldsymbol{\sigma}_x$):



Bit-phase flip ($\boldsymbol{\sigma}_y$):



Phase flip ($\boldsymbol{\sigma}_z$):



For the depolarizing error transformation,

$$\boldsymbol{\rho}'_\psi = \left(1 - \frac{3p}{4}\right)\boldsymbol{\rho}_\psi + \frac{p}{4}(\boldsymbol{\sigma}_x\boldsymbol{\rho}_\psi\boldsymbol{\sigma}_x + \boldsymbol{\sigma}_y\boldsymbol{\rho}_\psi\boldsymbol{\sigma}_y + \boldsymbol{\sigma}_z\boldsymbol{\rho}_\psi\boldsymbol{\sigma}_z), \tag{302}$$

we recall that the general Qbit density matrix has the form

$$\boldsymbol{\rho}_\psi = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{876}$$

where $a + d = 1$ so that $\mathrm{tr}(\boldsymbol{\rho}_\psi) = 1$. Then,

$$\boldsymbol{\sigma}_x\boldsymbol{\rho}_\psi\boldsymbol{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} d & c \\ b & a \end{pmatrix}, \tag{877}$$

$$\boldsymbol{\sigma}_y\boldsymbol{\rho}_\psi\boldsymbol{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}, \tag{878}$$

$$\boldsymbol{\sigma}_z\boldsymbol{\rho}_\psi\boldsymbol{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} a & -b \\ -c & d \end{pmatrix}, \tag{879}$$

and hence,

$$\boldsymbol{\rho}_\psi + \boldsymbol{\sigma}_x\boldsymbol{\rho}_\psi\boldsymbol{\sigma}_x + \boldsymbol{\sigma}_y\boldsymbol{\rho}_\psi\boldsymbol{\sigma}_y + \boldsymbol{\sigma}_z\boldsymbol{\rho}_\psi\boldsymbol{\sigma}_z = \begin{pmatrix} 2a+2d & 0 \\ 0 & 2a+2d \end{pmatrix} = 2\mathbf{I}. \tag{880}$$

Inserting this in eq. (302), the depolarizing transformation becomes

$$\boldsymbol{\rho}'_\psi = \left(1 - \frac{3p}{4}\right)\boldsymbol{\rho}_\psi + \frac{p}{4}(2\mathbf{I} - \boldsymbol{\rho}_\psi) = p\,\frac{\mathbf{I}}{2} + (1-p)\boldsymbol{\rho}_\psi. \tag{303}$$

21. **Quantum Error Correction**

(a) Using the relation

$$\mathsf{X}_1\mathsf{X}_2\mathsf{X}_3(|000\rangle \pm |111\rangle) = |111\rangle \pm |000\rangle = \pm(|000\rangle \pm |111\rangle), \tag{323}$$

we find that the coded states

$$|\bar{0}\rangle = \frac{1}{2\sqrt{2}}(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle), \tag{307}$$

$$|\bar{1}\rangle = \frac{1}{2\sqrt{2}}(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle). \tag{308}$$

are eigenstates of all three operators

$$\mathsf{X}_1\mathsf{X}_2\mathsf{X}_3\mathsf{X}_4\mathsf{X}_5\mathsf{X}_6, \qquad \mathsf{X}_4\mathsf{X}_5\mathsf{X}_6\mathsf{X}_7\mathsf{X}_8\mathsf{X}_9, \qquad \text{or} \qquad \mathsf{X}_1\mathsf{X}_2\mathsf{X}_3\mathsf{X}_7\mathsf{X}_8\mathsf{X}_9. \tag{324}$$

with eigenvalue $+1$, while a phase-flip error in the first triplet leads to states whose eigenvalues are -1 for operators $\mathsf{X}_1\mathsf{X}_2\mathsf{X}_3\mathsf{X}_4\mathsf{X}_5\mathsf{X}_6$ and $\mathsf{X}_1\mathsf{X}_2\mathsf{X}_3\mathsf{X}_7\mathsf{X}_8\mathsf{X}_9$, but $+1$ for operator $\mathsf{X}_4\mathsf{X}_5\mathsf{X}_6\mathsf{X}_7\mathsf{X}_8\mathsf{X}_9$, *etc.*

Applying, say, the first two operators of the set (324) to the coded state $|\bar{\psi}\rangle = a|\bar{0}\rangle + b|\bar{1}\rangle$ after it may have suffered a single phase-flip error in one of its 9 Qbits, we obtain one of the four results $(\pm,\pm)$. The procedure to diagnose and correct a single phase-flip error is

$(+,+)$ = no error.

$(+,-)$ = phase-flip error in the third triplet; correct using Z_7 (or Z_8 or Z_9 or even $\mathsf{Z}_7\mathsf{Z}_8\mathsf{Z}_9$ if your prefer symmetry at the expense of compactness).

$(-,+)$ = phase-flip error in the first triplet; correct using Z_1.

$(-,-)$ = phase-flip error in the second triplet; correct using Z_4.

If more than one phase-flip error occurs, the above procedure does not necessarily correct the errors. Two phase-flip errors in the same triplet cancel one another, while three phase-flip errors in the same triplet have the same effect as one error, and so would be corrected. But one phase-flip error in one triplet, along with another such error in another triplet, will be misdiagnosed as a single error in the remaining triplet, and the "correction" will leave all three triplets with a phase flip in each.

Thus, if p_z is the probability of a single phase-flip error, the probability that Shor's procedure fails to correct this type of error is, to leading order, $27p_z^2$, since there are 3 ways in which 2 phase-flip errors can occur in the 3 triplets with at most one such error per triplet, and within each triplet that contains an error there are three ways in which that error can occur.

(b) A phase-flip error transforms the Qbit $|\psi\rangle$ into $\mathsf{Y}|\psi\rangle$. Thus, if the first Qbit of Shor's coded Qbits (307)-(308) suffers a phase-flip error, the coded states become

$$\mathsf{Y}_1|\bar{0}\rangle = \frac{i}{2\sqrt{2}}(|100\rangle - |011\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle), \tag{881}$$

$$\mathsf{Y}_1|\bar{1}\rangle = \frac{i}{2\sqrt{2}}(|100\rangle + |011\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle). \tag{882}$$

If we search for bit-flip errors by measuring the operators $\mathsf{Z}_1\mathsf{Z}_2$ and $\mathsf{Z}_2\mathsf{Z}_3$, we will conclude that the first Qbit has been flipped. so, we apply operator X_1 to flip this Qbit.

The coded Qbits are now in the state

$$\mathsf{X}_1\mathsf{Y}_1|\bar{0}\rangle = i\mathsf{Z}_1|\bar{0}\rangle = \frac{i}{2\sqrt{2}}(|000\rangle - |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle), \quad (883)$$

$$\mathsf{X}_1\mathsf{Y}_1|\bar{1}\rangle = i\mathsf{Z}_1|\bar{1}\rangle = \frac{i}{2\sqrt{2}}(|000\rangle + |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle), \quad (884)$$

which are i times coded Qbits that have suffered a phase-flip error on their first Qbit. Measurement of the operators $\mathsf{X}_1\mathsf{X}_2\mathsf{X}_3\mathsf{X}_4\mathsf{X}_5\mathsf{X}_6$, $\mathsf{X}_1\mathsf{X}_2\mathsf{X}_3\mathsf{X}_7\mathsf{X}_8\mathsf{X}_9$ and $\mathsf{X}_4\mathsf{X}_5\mathsf{X}_6\mathsf{X}_7\mathsf{X}_8\mathsf{X}_9$ will show a phase-flip error in the first triplet, which we correct by applying the operator Z_1. The final state of the coded Qbits is

$$\mathsf{Z}_1\mathsf{X}_1\mathsf{Y}_1|\bar{0}\rangle = i\mathsf{Z}_1\mathsf{Z}_1|\bar{0}\rangle = i|\bar{0}\rangle, \quad (885)$$

$$\mathsf{Z}_1\mathsf{X}_1\mathsf{Y}_1|\bar{1}\rangle = i\mathsf{Z}_1\mathsf{Z}_1|\bar{1}\rangle = i|\bar{1}\rangle. \quad (886)$$

Thus the correction of a single bit-flip and a single phase-flip error also corrects for the effects of a single bit-phase-flip error to within an overall phase factor of i, which can be ignored.

(c) Shor's error-correction procedure successfully repairs a single bit-flip error in any of the 3 triplets of the encoded states (307)-(308). The leading failure mode for bit-flip errors is that two bit flips occur on different Qbits within the same triplet. So, if p_x is the probability of a single bit-flip error, the probability of two errors in one triplet is $3p_x^2$, to leading order, and the probability of two errors in any one of the three triplets is $9p_x^2$.

We saw in part (a) that the probability of failure to correct phase-flip errors is $27p_z^2$, to leading order, where p_z is the probability of a single phase-flip error.

In the case of bit-phase-flip errors, we see from part (b) that two such errors in a single triplet of the coded states (307)-(308) will not be corrected, since the bit-flip part of such an error would not be properly corrected. So, the probability of this type of failure is $9p_y^2$, to leading order, where p_y is the probability of a bit-phase-flip error. In addition, if one bit-phase flip occurs in one triplet, and another occurs in another triplet, these errors will not be corrected because the phase-flip correction procedure fails here. The probability of this type of error is $27p_y^2$, to leading order, since the combinatorics are the same as for the occurrence of two phase-flip errors. Therefore, the total probability that bit-phase-flip errors go uncorrected is $36p_y^2$, to leading order.

We must also consider the possibility of errors of two different types. Since the procedures for correction of bit-flip and phase-flip errors are independent, we can always correct for the presence of exactly one bit-flip error plus one phase-flip error. Hence, there is no leading-order failure proportional to the product $p_x p_z$.

In the case of one bit-flip error plus one bit-phase-flip error, these errors can be corrected unless they occur on different Qbits within the same triplet. Since there are six ways two different types of errors can occur on different Qbits within a

triplet, and there are three triplets in Shor's code, the probability that we fail to correct a bit-flip + bit-phase-flip error is $6 \cdot 3p_x p_y = 18p_x p_y$.

In the case of one phase-flip error plus one bit-phase-flip error, we can always correct the bit-flip part of the bit-phase-flip error. Because the phase-flip error is distinct from the bit-phase-flip error, there are twice as many ways that we can fail to correct a pair of such errors as there are ways of failing to correct a pair of phase-flip errors. Hence, the probability that we fail to correct a phase-flip + bit-phase-flip error is $54p_y p_z$, recalling part (a).

In sum, the leading-order probability of failure to correct errors of the type bit flip, phase flip, or bit-phase flip is $9p_x^2 + 36p_y^2 + 27p_z^2 + 18p_x p_y + 54p_y p_z$.

(d) A general error on the first bit of Shor's 9-bit coded state $|\bar{\psi}\rangle = \alpha|\bar{0}\rangle + \beta|\bar{1}\rangle$ is described according to eq. (299) by (hermitian) operators E of the form

$$\mathsf{E}|\bar{\psi}\rangle = (a_0\mathbf{I}_1 + a_x\mathsf{X}_1 + a_y\mathsf{Y}_1 + a_z\mathsf{Z}_1)|\bar{\psi}\rangle, \tag{887}$$

which is a superposition of four terms.

After the first measurement during the error-diagnosis procedure (which involves operators for which all four terms of eq. (887) are eigenstates), the coded state is no longer a superposition but is in one of the four states $|\bar{\psi}\rangle$, $\mathsf{X}_1|\bar{\psi}\rangle$, $\mathsf{Y}_1|\bar{\psi}\rangle$ or $\mathsf{Z}_1|\bar{\psi}\rangle$, and the resulting state is normalized. Thus, the process of measurement simplifies the general error state (887) into one of the simple cases of a single type of error for which Shor's correction scheme was designed. That scheme will then find whichever type of error was "selected" by the first measurement process, and correct it.

(e) The error-correction operator U is defined by

$$\mathsf{U}|\bar{\psi}\rangle|0\rangle \equiv \sum_{j\neq 0}(\mathsf{U}_j\mathsf{M}_j|\bar{\psi}\rangle)|j\rangle, \tag{325}$$

$$\mathsf{U}|\bar{\psi}\rangle|j\neq 0\rangle \equiv |\bar{\psi}\rangle|j\rangle. \tag{326}$$

The inner product of two different states $\mathsf{U}|\bar{\psi}\rangle|\phi\rangle$ and $\mathsf{U}|\bar{\psi}'\rangle|\phi'\rangle$, where the ancillary states are written

$$|\phi\rangle = \sum_j \phi_j|j\rangle = \phi_0|0\rangle + \sum_{j\neq 0}\phi_j|j\rangle, \tag{888}$$

is given by

$$\begin{aligned}\langle\phi'|\langle\bar{\psi}'|\mathsf{U}^\dagger\mathsf{U}|\bar{\psi}\rangle|\phi\rangle &= \phi_0'^\star\phi_0\langle 0|\langle\bar{\psi}'|\mathsf{U}^\dagger\mathsf{U}|\bar{\psi}\rangle|0\rangle + \sum_{j\neq 0}\sum_{k\neq 0}\phi_j'^\star\phi_k\langle j|\langle\bar{\psi}'|\mathsf{U}^\dagger\mathsf{U}|\bar{\psi}\rangle|k\rangle \\ &\quad + \sum_{j\neq 0}\phi_0'^\star\phi_j\langle 0|\langle\bar{\psi}'|\mathsf{U}^\dagger\mathsf{U}|\bar{\psi}\rangle|j\rangle + \sum_{j\neq 0}\phi_j'^\star\phi_0\langle j|\langle\bar{\psi}'|\mathsf{U}^\dagger\mathsf{U}|\bar{\psi}\rangle|0\rangle. \end{aligned}\tag{889}$$

The first term in eq. (889) can be rearranged as

$$\begin{aligned}\phi_0'^\star\phi_0\langle 0|\langle\bar{\psi}'|\mathsf{U}^\dagger\mathsf{U}|\bar{\psi}\rangle|0\rangle &= \phi_0'^\star\phi_0\sum_{j\neq 0}\sum_{k\neq 0}\langle j|(\langle\bar{\psi}'|\mathsf{M}_j^\dagger\mathsf{U}_j^\dagger)(\mathsf{U}_k\mathsf{M}_k|\bar{\psi}\rangle)|k\rangle \\ &= \phi_0'^\star\phi_0\sum_{j\neq 0}\langle\bar{\psi}'|\mathsf{M}_j^\dagger\mathsf{U}_j^\dagger\mathsf{U}_j\mathsf{M}_j|\bar{\psi}\rangle = \phi_0'^\star\phi_0\sum_{j\neq 0}\langle\bar{\psi}'|\mathsf{M}_j^\dagger\mathsf{M}_j|\bar{\psi}\rangle \\ &= \phi_0'^\star\phi_0\langle\bar{\psi}'|\bar{\psi}\rangle, \end{aligned}\tag{890}$$

using eq. (325), noting that $\langle j|k\rangle = \delta_{jk}$, and recalling eq. (74). The second term in eq. (889) is

$$\sum_{j\neq 0}\sum_{k\neq 0}\phi_j'^{\star}\phi_k\langle j|\langle\bar{\psi}'|\mathsf{U}^\dagger\mathsf{U}|\bar{\psi}\rangle|k\rangle = \sum_{j\neq 0}\sum_{k\neq 0}\phi_j'^{\star}\phi_k\langle j|\langle\bar{\psi}'|\bar{\psi}\rangle|k\rangle = \sum_{j\neq 0}\phi_j'^{\star}\phi_j\langle\bar{\psi}'|\bar{\psi}\rangle, \tag{891}$$

using eq. (326).

The sum of the first two terms of eq. (889) is therefore

$$\phi_0'^{\star}\phi_0\langle\bar{\psi}'|\bar{\psi}\rangle + \sum_{j\neq 0}\phi_j'^{\star}\phi_j\langle\bar{\psi}'|\bar{\psi}\rangle = \sum_j \phi_j'^{\star}\phi_j\langle\bar{\psi}'|\bar{\psi}\rangle = \langle\phi'|\phi\rangle\langle\bar{\psi}'|\bar{\psi}\rangle. \tag{892}$$

If operator U is indeed unitary, such that $\mathsf{U}^\dagger\mathsf{U} = \mathbf{I}$, then the third and fourth terms of eq. (889) must vanish, since $\langle 0|j\rangle = 0$. In this case, eq. (892) tells us that

$$\langle\phi'|\langle\bar{\psi}'|\mathsf{U}^\dagger\mathsf{U}|\bar{\psi}\rangle|\phi\rangle = \langle\phi'|\phi\rangle\langle\bar{\psi}'|\bar{\psi}\rangle, \tag{893}$$

as desired for a unitary operator.

However, using eqs. (325)-(326), the third term of eq. (889) can be written as

$$\begin{aligned}\sum_{j\neq 0}\phi_0'^{\star}\phi_j\langle 0|\langle\bar{\psi}'|\mathsf{U}^\dagger\mathsf{U}|\bar{\psi}\rangle|j\rangle &= \sum_{j\neq 0}\sum_{k\neq 0}\phi_0'^{\star}\phi_j\langle k|(\langle\bar{\psi}'|\mathsf{M}_k^\dagger\mathsf{U}_k^\dagger)|\bar{\psi}\rangle|j\rangle \\ &= \sum_{j\neq 0}\phi_0'^{\star}\phi_j\langle\bar{\psi}'|\mathsf{M}_j^\dagger\mathsf{U}_j^\dagger|\bar{\psi}\rangle,\end{aligned} \tag{894}$$

which does not appear to vanish, in general.

It would suffice if only the sum of the third and fourth terms of eq. (889) were zero. The fourth term of eq. (889) can be written as

$$\begin{aligned}\sum_{j\neq 0}\phi_j'^{\star}\phi_0\langle j|\langle\bar{\psi}'|\mathsf{U}^\dagger\mathsf{U}|\bar{\psi}\rangle|0\rangle &= \sum_{j\neq 0}\sum_{k\neq 0}\phi_j'^{\star}\phi_0\langle j|(\langle\bar{\psi}'|\mathsf{U}_k)\mathsf{M}_j|\bar{\psi}\rangle|k\rangle \\ &= \sum_{j\neq 0}\phi_j'^{\star}\phi_0\langle\bar{\psi}'|\mathsf{U}_j\mathsf{M}_j|\bar{\psi}\rangle.\end{aligned} \tag{895}$$

It also does not appear that the sum of eqs. (894) and (895) is zero...

22. Fault-Tolerant Quantum Computation

(a) Fault-Tolerant Gates $\bar{\mathsf{X}}_{\rm Steane}$ and $\bar{\mathsf{Z}}_{\rm Steane}$

To show that the transverse logical gates $\bar{\mathsf{X}}_{\rm Steane}$ and $\bar{\mathsf{Z}}_{\rm Steane}$ are fault tolerant, we only need show that a single Qbit error that occurs just before use of these gates can be corrected via application of our usual error correction procedure (prob. 22) immediately after these gates (since a single error that occurs during or after the gate will surely be corrected).

For the gate $\bar{\mathsf{X}}_{\rm Steane}$, we note that the Pauli operators obey the identities

$$\mathsf{XY} = -\mathsf{YX}, \qquad \text{and} \qquad \mathsf{XZ} = -\mathsf{ZX}, \tag{896}$$

Hence, the occurences of an error on a physical Qbit that is associated with operators X (bit flip), Y (bit-phase flip) or Z (phase flip), and which occurs before the $\bar{\mathsf{X}}$ gate, is equivalent to the case that the same type of error occurred after the gate. The post-gate error correction procedure will then correct the error (up to an overall phase factor that can be ignored), so long as only a single error occurred. Hence, the transverse operator $\bar{\mathsf{X}}_{\rm Steane}$ is fault tolerant.

Similarly, the transverse operator $\bar{\mathsf{Z}}_{\rm Steane}$ is fault tolerant in view of the identities

$$\mathsf{ZX} = -\mathsf{XZ}, \qquad \text{and} \qquad \mathsf{ZY} = -\mathsf{YZ}. \tag{897}$$

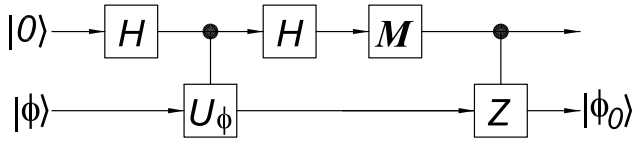
(b) Tranverse Construction for $\mathsf{C\text{-}NOT}_{\rm Steane}$

When the control word is $|\bar{0}\rangle_{\rm Steane}$, its pattern of $|1\rangle$'s leads to flipping of bits in the target word which are equivalent to applying the operator $\bar{\mathbf{0}}_{\rm Steane}$ to the target word. Since both $|\bar{0}\rangle_{\rm Steane}$ and $|\bar{1}\rangle_{\rm Steane}$ are eigenstates of this operator with eigenvalue $+1$, the target word is unchanged.

However, if the control word is $|\bar{1}\rangle_{\rm Steane}$, the flipping of bits in the target word proceeds according to an application of operator $\bar{\mathbf{1}}_{\rm Steane} = \bar{\mathbf{0}}_{\rm Steane}\bar{\mathsf{X}}_{\rm Steane}$. The operator $\bar{\mathsf{X}}_{\rm Steane}$ flips the target state, and then the operator $\bar{\mathbf{0}}_{\rm Steane}$ leaves the target bit in its flipped state, as desired for the operation $\mathsf{C\text{-}NOT}_{\rm Steane}$.

(c) Measurement of an Operator Whose Eigenvalues are ± 1

We are given a unitary single-Qbit operator U_ϕ whose eigenvectors $|\phi_j\rangle$, $j = 0, 1$ obey $\mathsf{U}_\phi|\phi_j\rangle = (-1)^j|\phi_j\rangle$.



Then, the effect of Shor's circuit, as shown above, on a general Qbit $|\phi\rangle = a|\phi_0\rangle + b|\phi_1\rangle$ up to the point of the measurement is

$$\begin{aligned} |\phi\rangle|0\rangle &\rightarrow (a|\phi_0\rangle + b|\phi_1\rangle)\frac{|0\rangle + |1\rangle}{\sqrt{2}} \\ &\rightarrow \frac{(a|\phi_0\rangle + b|\phi_1\rangle)|0\rangle + (a|\phi_0\rangle - b|\phi_1\rangle)|1\rangle}{\sqrt{2}} \\ &\rightarrow a|\phi_0\rangle|0\rangle + b|\phi_1\rangle|1\rangle. \end{aligned} \tag{898}$$

When the control Qbit is measured, if the measured value is 0 the target Qbit will be forced into state $|\phi_0\rangle$, while if the measured value is 1 the target Qbit will be in state $|\phi_1\rangle$.

With the additional relation that $|\phi_0\rangle = \mathsf{Z}|\phi_1\rangle$, as holds for the states (342) and (345)

$$|\phi_j\rangle = \frac{|0\rangle + (-1)^j e^{i\pi/4}|1\rangle}{\sqrt{2}}, \qquad (342\ \&\ 345)$$

the final Controlled-Z operator in the circuit insures that the target Qbit emerges in state $|\phi_0\rangle$.

(d) **Grover's Procedure to Correct Systematic Errors**

To evaluate the improvement of the iteration $\mathsf{U}\mathsf{R}_s\mathsf{U}^\dagger\mathsf{R}_t\mathsf{U}|s\rangle$ over the initial calculation $\mathsf{U}|s\rangle$, it is useful to display two auxiliary relations first.

We decompose $\mathsf{U}|s\rangle$ into components along, and orthogonal to, $|t\rangle$, following eq. (348),

$$\mathsf{U}|s\rangle = \gamma|t\rangle + \delta|v\rangle = \langle t|\mathsf{U}|\mathsf{s}\rangle|t\rangle + \delta|v\rangle, \qquad (899)$$

where $\langle t|v\rangle = 0$. Therefore, $|t\rangle$, following eq. (348),

$$|\delta|^2 = 1 - |\langle t|\mathsf{U}|\mathsf{s}\rangle|^2. \qquad (900)$$

Also,,

$$\mathsf{U}^\dagger\mathsf{U}|s\rangle = |s\rangle = \langle t|\mathsf{U}|\mathsf{s}\rangle\mathsf{U}^\dagger|t\rangle + \delta\mathsf{U}^\dagger|v\rangle, \qquad (901)$$

and hence,

$$\delta\mathsf{U}^\dagger|v\rangle = |s\rangle - \langle t|\mathsf{U}|\mathsf{s}\rangle\mathsf{U}^\dagger|t\rangle. \qquad (902)$$

Similarly,

$$\mathsf{U}^\dagger|t\rangle = \alpha|s\rangle + \beta|u\rangle = \langle s|\mathsf{U}^\dagger|t\rangle|s\rangle + \beta|u\rangle = \langle t|\mathsf{U}|\mathsf{s}\rangle^\star|s\rangle + \beta|u\rangle, \qquad (903)$$

so that

$$\mathsf{U}\mathsf{U}^\dagger|t\rangle = |t\rangle = \langle t|\mathsf{U}|\mathsf{s}\rangle^\star\mathsf{U}|s\rangle +, \qquad (904)$$

and hence,

$$\beta\mathsf{U}|u\rangle = |t\rangle - \langle t|\mathsf{U}|\mathsf{s}\rangle^\star\mathsf{U}|s\rangle. \qquad (905)$$

With these we find

$$\begin{aligned}
\mathsf{U}\mathsf{R}_s\mathsf{U}^\dagger\mathsf{R}_t\mathsf{U}|s\rangle &= \mathsf{U}\mathsf{R}_s\mathsf{U}^\dagger\mathsf{R}_t(\langle t|\mathsf{U}|\mathsf{s}\rangle|t\rangle + \delta|v\rangle) \\
&= \mathsf{U}\mathsf{R}_s\mathsf{U}^\dagger(e^{i\pi/3}\langle t|\mathsf{U}|\mathsf{s}\rangle|t\rangle + \delta|v\rangle) \\
&= \mathsf{U}\mathsf{R}_s(e^{i\pi/3}\langle t|\mathsf{U}|\mathsf{s}\rangle\mathsf{U}^\dagger|t\rangle + \delta\mathsf{U}^\dagger|v\rangle) \\
&= \mathsf{U}\mathsf{R}_s[(e^{i\pi/3}-1)\langle t|\mathsf{U}|\mathsf{s}\rangle\mathsf{U}^\dagger|t\rangle + |s\rangle] \\
&= \mathsf{U}\mathsf{R}_s[(e^{i\pi/3}-1)\langle t|\mathsf{U}|\mathsf{s}\rangle(\langle t|\mathsf{U}|\mathsf{s}\rangle^\star|s\rangle + \beta|u\rangle) + |s\rangle] \\
&= \mathsf{U}\{e^{i\pi/3}[(e^{i\pi/3}-1)\,|\langle t|\mathsf{U}|\mathsf{s}\rangle|^2 + 1]|s\rangle + (e^{i\pi/3}-1)\langle t|\mathsf{U}|\mathsf{s}\rangle\beta|u\rangle\} \\
&= e^{i\pi/3}[(e^{i\pi/3}-1)\,|\langle t|\mathsf{U}|\mathsf{s}\rangle|^2 + 1]\mathsf{U}|s\rangle + (e^{i\pi/3}-1)\langle t|\mathsf{U}|\mathsf{s}\rangle\beta\mathsf{U}|u\rangle \\
&= \{[e^{i\pi/3}(e^{i\pi/3}-1) - (e^{i\pi/3}-1)]\,|\langle t|\mathsf{U}|\mathsf{s}\rangle|^2 + e^{i\pi/3}\}\mathsf{U}|s\rangle \\
&\quad +(e^{i\pi/3}-1)\langle t|\mathsf{U}|\mathsf{s}\rangle|t\rangle \\
&= [(e^{i\pi/3}-1)^2\,|\langle t|\mathsf{U}|\mathsf{s}\rangle|^2 + e^{i\pi/3}]\mathsf{U}|s\rangle + (e^{i\pi/3}-1)\langle t|\mathsf{U}|\mathsf{s}\rangle|t\rangle \\
&\equiv a\mathsf{U}|s\rangle + b|t\rangle = a(\langle t|\mathsf{U}|\mathsf{s}\rangle|t\rangle + \delta|v\rangle) + b|t\rangle, \qquad (906)
\end{aligned}$$

where

$$\begin{aligned} a &= (e^{i\pi/3}-1)^2\,|\langle t|\mathsf{U}|\mathsf{s}\rangle|^2 + e^{i\pi/3} = e^{4i\pi/3}\,|\langle t|\mathsf{U}|\mathsf{s}\rangle|^2 + e^{i\pi/3} \\ &= e^{i\pi/3}(1-|\langle t|\mathsf{U}|\mathsf{s}\rangle|^2), \end{aligned} \tag{907}$$

noting that $e^{i\pi/3}-1 = e^{2i\pi/3}$, and that $e^{4i\pi/3} = -e^{i\pi/3}$.

We are interested in the probability that the state (906) is not $|t\rangle$. From the last form of this equation, the amplitude for this is $a\delta$, and so the probability that the calculation failed to yield $|t\rangle$ is

$$P = |a|^2\,|\delta|^2 = (1-|\langle t|\mathsf{U}|\mathsf{s}\rangle|^2)^2(1-|\langle t|\mathsf{U}|\mathsf{s}\rangle|^2) = \epsilon^3, \tag{908}$$

supposing that $|\langle t|\mathsf{U}|\mathsf{s}\rangle|^2 = 1-\epsilon$.

23. **Quantum Cryptography**

(a) Given that

$$\mathsf{U}|\psi\rangle|a\rangle = |\psi\rangle|b\rangle, \qquad \text{and} \qquad \mathsf{U}|\phi\rangle|a\rangle = |\phi\rangle|c\rangle, \tag{354}$$

we find on taking the inner product,

$$\begin{aligned} \langle a|\langle\psi|\mathsf{U}^\dagger\mathsf{U}|\phi\rangle|a\rangle &= \langle a|a\rangle\langle\psi|\phi\rangle = \langle\psi|\phi\rangle \\ = \langle b|\langle\psi|\phi\rangle|c\rangle &= \langle b|c\rangle\langle\psi|\phi\rangle, \end{aligned} \tag{909}$$

since operator U is presumed to be unitary. Hence, if $\langle\psi|\phi\rangle \neq 0$ then $\langle b|c\rangle = 1$ and states $|\psi\rangle$ and $|\phi\rangle$ cannot be distinguished by this procedure.

(b) Recalling prob. 5(d), we see that in preparation for her "nondemolition" measurement of one of Alice's Qbit $|\psi\rangle$, Eve uses the Controlled-NOT operator and an ancillary Qbit whose initial state is $|0\rangle$ to make as good a copy of $|\psi\rangle$ as possible, namely $\mathsf{C}_{xy}|\psi\rangle|0\rangle$. This procedure refers to a particular basis in which the Controlled-NOT operator is defined. We take this basis to be [0,1] without loss of generality.

If Alice prepared the Qbit $|\psi\rangle$ in the [0,1] basis, then Eve's measurement of the ancillary Qbit (after the Controlled-NOT operation) correctly identifies whether $|0\rangle$ is $|0\rangle$ or $|1\rangle$ without altering state $|0\rangle$. From Alice's public announcement of her choice of bases, Eve knows (after her nondemolition measurement is complete) which of Qbits she has identified correctly. That is, she now knows roughly 50% of the "private" key.

However, Eve's procedure has a nontrivial effect on the 50% of Alice's Qbits that were prepared in the $[+,-]$ basis. Since $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$, Eve's Controlled-NOT operation has the result

$$\mathsf{C}_{xy}|\pm\rangle|0\rangle = \frac{|0\rangle|0\rangle \pm |1\rangle|1\rangle}{\sqrt{2}}\,. \tag{910}$$

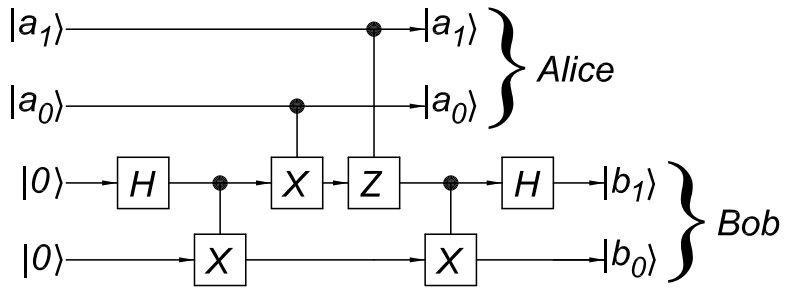
After Eve's measurement of the ancillary (second) Qbit in the [0,1] basis, the original Qbit becomes either $|0\rangle$ or $|1\rangle$ with 50% probability each.

The original Qbit $|\psi\rangle$ will become part of the private key in those cases that Bob measures it in the $[+,-]$ basis. Since Qbit $|\psi\rangle$ is now a $|0\rangle$ or a $|1\rangle$, Bob's measurement of it in the $[+,-]$ basis will produce a $|+\rangle$ or a $|-\rangle$ at random, with 50% probability. This is therefore the probability that Bob's measurement of this Qbit actually agree's with Alice's. Since half of the Qbits used to generate the private key were prepared in the $[+,-]$ basis, we deduce that Alice and Bob will disagree (without knowing it, unless they take further action) as to the value of roughly 25% of the bits in their private key.

(c) **Quantum Dense Coding**

The first two gates of the quantum dense coding circuit (shown on the next page) convert $|0\rangle|0\rangle$ to the entangled state

$$\mathsf{C}_{10}\mathsf{H}_1|0\rangle|0\rangle = \frac{|0\rangle|0\rangle + |1\rangle|1\rangle}{\sqrt{2}}\,. \tag{911}$$



If $|a_1\rangle|a_0\rangle = |0\rangle|0\rangle$, then the X (= NOT) and Z gates are not applied to $|b_1\rangle$, so we have

$$|b_1\rangle|b_0\rangle = \mathsf{H}_1\mathsf{C}_{10}\mathsf{C}_{10}\mathsf{H}_1|0\rangle|0\rangle = |0\rangle|0\rangle, \tag{912}$$

since the Controlled-NOT and Hadamard gates are their own inverses.
If $|a_1\rangle|a_0\rangle = |0\rangle|1\rangle$, then only the X gate is applied to $|b_1\rangle$, so we have

$$\begin{aligned} |b_1\rangle|b_0\rangle &= \mathsf{H}_1\mathsf{C}_{10}\mathsf{X}_1\mathsf{C}_{10}\mathsf{H}_1|0\rangle|0\rangle = \mathsf{H}_1\mathsf{C}_{10}\mathsf{X}_1\frac{|0\rangle|0\rangle+|1\rangle|1\rangle}{\sqrt{2}} = \mathsf{H}_1\mathsf{C}_{10}\frac{|1\rangle|0\rangle+|0\rangle|1\rangle}{\sqrt{2}} \\ &= \mathsf{H}_1\frac{|1\rangle|1\rangle+|0\rangle|1\rangle}{\sqrt{2}} = \frac{(|0\rangle-|1\rangle)|1\rangle+(|0\rangle+|1\rangle)|1\rangle}{2} = |0\rangle|1\rangle. \end{aligned} \tag{913}$$

If $|a_1\rangle|a_0\rangle = |1\rangle|0\rangle$, then only the Z gate is applied to $|b_1\rangle$, so we have

$$\begin{aligned} |b_1\rangle|b_0\rangle &= \mathsf{H}_1\mathsf{C}_{10}\mathsf{Z}_1\mathsf{C}_{10}\mathsf{H}_1|0\rangle|0\rangle = \mathsf{H}_1\mathsf{C}_{10}\mathsf{Z}_1\frac{|0\rangle|0\rangle+|1\rangle|1\rangle}{\sqrt{2}} = \mathsf{H}_1\mathsf{C}_{10}\frac{|0\rangle|0\rangle-|1\rangle|1\rangle}{\sqrt{2}} \\ &= \mathsf{H}_1\frac{|0\rangle|0\rangle-|1\rangle|0\rangle}{\sqrt{2}} = \frac{(|0\rangle+|1\rangle)|0\rangle-(|0\rangle-|1\rangle)|0\rangle}{2} = |1\rangle|0\rangle. \end{aligned} \tag{914}$$

If $|a_1\rangle|a_0\rangle = |1\rangle|1\rangle$, then both the X and Z gates are applied to $|b_1\rangle$. Recalling that

$$\mathsf{ZX} = i\mathsf{Y} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{915}$$

we have

$$\begin{aligned} |b_1\rangle|b_0\rangle &= \mathsf{H}_1\mathsf{C}_{10}\mathsf{Z}_1\mathsf{Z}_1\mathsf{C}_{10}\mathsf{H}_1|0\rangle|0\rangle = \mathsf{H}_1\mathsf{C}_{10}i\mathsf{Y}_1\frac{|0\rangle|0\rangle+|1\rangle|1\rangle}{\sqrt{2}} = \mathsf{H}_1\mathsf{C}_{10}\frac{-|1\rangle|0\rangle+|0\rangle|1\rangle}{\sqrt{2}} \\ &= \mathsf{H}_1\frac{-|1\rangle|1\rangle+|0\rangle|1\rangle}{\sqrt{2}} = \frac{-(|0\rangle-|1\rangle)|1\rangle+(|0\rangle+|1\rangle)|1\rangle}{2} = |1\rangle|1\rangle. \end{aligned} \tag{916}$$

If $|a_1\rangle|a_0\rangle = (|0\rangle|0\rangle+|1\rangle|1\rangle)/\sqrt{2}$, then we can combine eqs. (912) and (916) to give the final state

$$|b_1\rangle|b_0\rangle|a_1\rangle|a_0\rangle = \frac{|0\rangle|0\rangle|0\rangle|0\rangle+|1\rangle|1\rangle|1\rangle|1\rangle}{\sqrt{2}}, \tag{917}$$

which is an entangled state, but which is not the direct product of two entangled states, $(|0\rangle|0\rangle+|1\rangle|1\rangle)(|0\rangle|0\rangle+|1\rangle|1\rangle)/2$, as would hold if the dense coding operation were an exact copy operation.