Helical Logic

by

Ralph C. Merkle
Xerox PARC
3333 Coyote Hill Road
Palo Alto, CA 94304
merkle@xerox.com
ftp://parcftp.xerox.com/pub/merkle/merklesHomePage.html

and

K. Eric Drexler
Institute for Molecular Manufacturing
123 Fremont Avenue
Los Altos, CA 94022

This article has been submitted for publication to *Nanotechnology*. Further information about *Nanotechnology* is available from URL "ftp://parcftp.xerox.com/pub/nano/IOP.html"

A WWW page on this article (with links to the postscript) is available from URL "ftp://parcftp.xerox.com/pub/nano/helicalIntro.html"

This draft is subject to revision.

Abstract

Helical logic is a theoretical proposal for a future computing technology using the presence or absence of individual electrons (or holes) to encode 1s and 0s. The electrons are constrained to move along helical paths, driven by a rotating electric field in which the entire circuit is immersed. The electric field remains roughly orthogonal to the major axis of the helix and confines each charge carrier to a fraction of a turn of a single helical loop, moving it like water in an Archimedean screw. Each loop could in principle hold an independent carrier, permitting high information density. One computationally universal logic operation involves two helices, one of which splits into two "descendant" helices. At the point of divergence, differences in the electrostatic potential resulting from the presence or absence of a carrier in the adjacent helix controls the direction taken by a carrier in the splitting helix. The reverse of this sequence can be used to merge two initially distinct helical paths into a single outgoing helical path without forcing a dissipative transition. Because these operations are both logically and thermodynamically reversible, energy dissipation can be reduced to extremely low levels. This is the first proposal known to the authors that combines thermodynamic reversibility with the use of single charge carriers. It is important to note that this proposal permits a single electron to switch another single electron, and does not require that many electrons be used to switch one electron. The energy dissipated per logic operation can likely

be reduced to less than 10^{-27} joules at a temperature of 1 Kelvin and a speed of 10 gigahertz, though further analysis is required to confirm this. Irreversible operations, when required, can be easily implemented and should have a dissipation approaching the fundamental limit of ln 2 x *kT*.

Introduction

Trends in computer hardware are leading toward higher density and lower energy dissipation. Ultimately, some approaches should result in packing densities in excess of 10^{17} logic devices in a cubic centimeter[REF17, REF18] (although the current proposal might require a somewhat larger volume). The trend towards higher packing density strongly influences energy dissipation. Conventional devices must dissipate more than ln(2) x *kT* joules in switching; so 10^{17} conventional devices operating at room temperature (ln(2) x kT~3 x 10^{-21} joules for $T = 300$ Kelvins) at a frequency of 10 gigahertz would dissipate >3,000,000 watts; a computer with 1,000 times as many logic elements would still be of reasonable size but would dissipate 3,000,000,000 watts.

Conventional circuits perform more poorly. Even an idealized device which used a one volt power supply and dissipatively discharged a single electron to ground during a switching operation would dissipate one electron volt per switching operation. At T=300 Kelvins, this is 40 x *kT* per switching operation or about 160,000,000 watts for a computer with 10^{20} logic elements operating at 10 gigahertz. If each switching operation involves hundreds of electrons then energy dissipation enters the multigigawatt range.

New thermodynamically reversible circuits (including CMOS, nMOS and CCD-based logic circuits[REF21, REF38, REF58, REF59, REF60]) would fare better, but these circuits still have dissipative losses caused by the resistance of the circuit. While resistance in sufficiently small wires can be very low[REF63], if such wires are connected to each other, to logic elements or to larger structures it is common to find resistances of the order of 13KΩ (half of h/e^2, where h is Planck's constant) (note that no claim is made that the successful operation of such circuits must fundamentally require resistances of this magnitude, we simply note that shrinking current circuits to a small scale would result in such resistances: further research in this area might be successful in dealing with this problem). Assuming that 100 electrons were required to charge and discharge the wires and capacitive loads in each logic element, and assuming a resistance of approximately 13KΩ, we would still find our 10^{20} gate computer dissipating over 50 megawatts even using these particular thermodynamically reversible methods.

If the exponential trends of recent decades continue, energy dissipation per logic operation will reach *kT* (for T=300 Kelvins) early in the next century[REF06]. Either energy dissipation per logic operation will be reduced significantly below 3 x 10^{-21} joules, or we will fail to achieve computers that simultaneously combine high packing densities with gigahertz or higher speeds of operation. There are only two ways that energy dissipation can be reduced below 3 x 10^{-21} joules: by operating at temperatures below room temperature (thus reducing *kT*), or by using thermodynamically reversible logic. Low temperature operation doesn't actually reduce total energy dissipation, it just shifts it from computation to refrigeration. Thermodynamically reversible logic elements, in contrast, can reduce total energy dissipation per logic operation to << *kT*. This paper analyzes a proposed thermodynamically reversible single electron logic system. To achieve high reliability while switching single electrons, we analyze operation at ~1 Kelvin.

Reversible Architectures and Logic Proposals

One concern about this approach is the need to use reversible computer architectures. Such architectures are entirely feasible[REF06, REF07, REF08, REF11, REF15, REF16, REF19, REF21, REF23, REF29, REF38, REF51, REF53, REF55, REF58] and will not be discussed further here. Bennett[REF15] has proposed very general methods for converting arbitrary irreversible computations into time and space efficient reversible computations.

A wide variety of reversible device proposals have been made[REF02, REF09, REF10, REF18, REF23, REF29, REF38, REF58, REF59, REF60]. The present proposal is abstractly similar to billiard ball logic[REF03, REF07], and in particular uses the concept of a "switch gate" or "interaction gate."

In billiard ball logic[REF07], a set of billiard balls are fired into a set of immovable reflectors at a fixed speed. As the billiard balls bounce off each other and off the reflectors, they perform a reversible computation. Provided that the collisions between the billiard balls and between the billiard balls and the reflectors are perfectly elastic, the computation can proceed at a fixed finite speed with no energy loss.

As originally proposed, this "ballistic" model of computation suffers from the shortcoming that the positions of the reflectors and the initial velocity of the billiard balls must be perfectly accurate. In the real world, the computation would rapidly deteriorate into chaos unless some restoring force maintained the alignment of the billiard balls. Landauer[REF53] proposed exactly this, suggesting that the billiard balls should remain in the trough of a moving periodic potential (illustrated in Figure 4).

Billiard ball logic is usually thought of in connection with Fredkin gates[REF07]. Fredkin gates are conservative three-input three-output gates that are logically complete, e.g., any computation can be implemented by an appropriately connected set of Fredkin gates. An appropriately arranged set of reflectors can be used to implement a Fredkin gate from billiard balls. Rather than considering the Fredkin gate we will focus instead on the switch gate[REF07]. This can be used to make a Fredkin gate (and so is also logically universal) but it is physically simpler and easier to implement.

A Fredkin gate is illustrated in Figure 1, a switch gate is illustrated in Figure 2, a method of implementing a Fredkin gate from switch gates is shown in figure 3, while a billiard ball being carried along by a periodic potential is shown in Figure 4.

While the abstract concept of billiard ball logic has been known for some time, the idea of implementing this approach by replacing the billiard balls with charged particles is surprisingly recent[38]. In the present proposal, we replace the billiard ball with a single electron. While other charged particles or charged packets of particles can be used, we will focus on the single electron implementation because of the obvious long term performance advantages. The concept that future electronic devices might use single electrons is becoming more accepted[REF04, REF26, REF54]. This is the first description of a thermodynamically reversible switching device using single electrons. Not only are we switching single electrons, we are also using single electrons to control the switching (unlike many proposals which require many electrons to switch a single electron). It is plausible that thermodynamically reversible single electron switching devices will be the ultimate evolutionary end point of electronic logic devices. The authors emphasize that

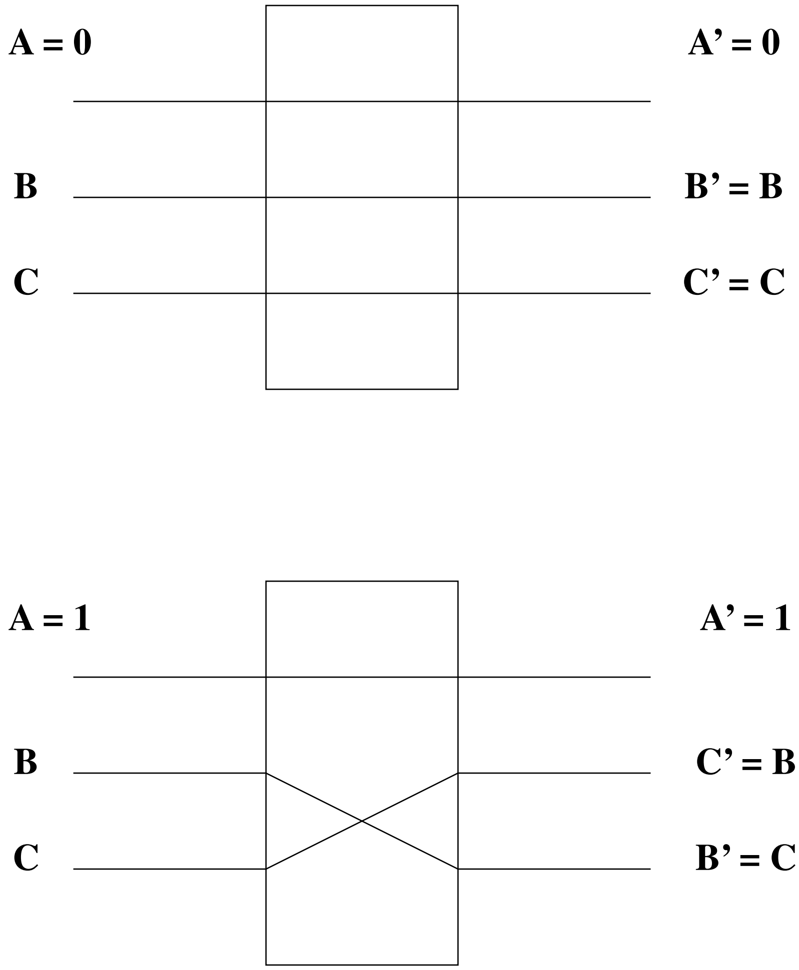


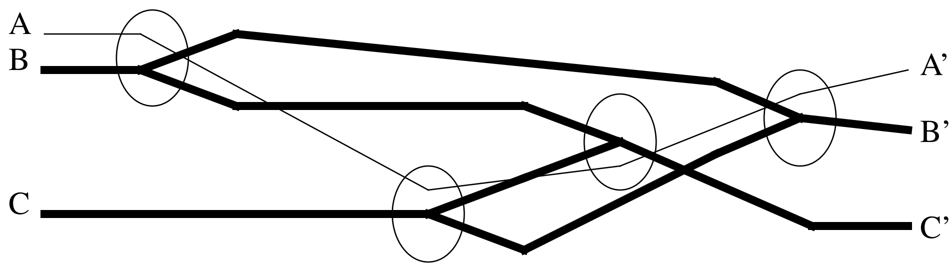
Figure 1.
A Fredkin Gate

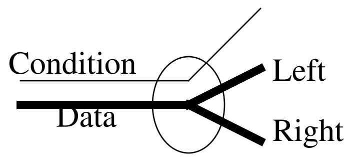
other proposals for single electron thermodynamically reversible logic devices will certainly be advanced, so the present proposal must be viewed as the first in a class rather than "the" one and only design for a single electron thermodynamically reversible logic device.

Clock Distribution using Rotating Electric Fields

The periodic moving potential (illustrated in figure 4) which drives the electrons forwards can be viewed as a periodic clock signal. It is conventional in today's circuits to distribute clock signals using wires. This approach is energetically disadvantageous because the propogation of the clock signal through wires is itself dissipative, as well as requiring a complex (and space consuming) network for distribution. An attractive alternative is to use time varying electric fields. We propose the use of rotating electric fields as these are conceptually simple, easy to produce, and adequate to drive a wide range of logic devices.

The picture of clock distribution being proposed here is new (though it is abstractly similar to





Symbol for Single
Switch Gate

Figure 3.
Implementation of a
Fredkin Gate
from Switch Gates
(after Fredkin & Toffoli[REF7])

the rotating magnetic field used to clock and power magnetic bubble devices[REF24, REF45]). The computing element is immersed in a rotating electric field. The electric field, and hence the clock, is available at every point in the device and yet no wires are required to distribute it. This reduces energy dissipation and eliminates the space that would otherwise be occupied by the clock distribution network.

A rotating electric field is rather different from a periodic moving potential. In order to achieve the same effect (e.g., to move electrons forwards along their path) we propose the use of helical paths, with the electric field rotating at right angles to the axis of the path. The electron will be

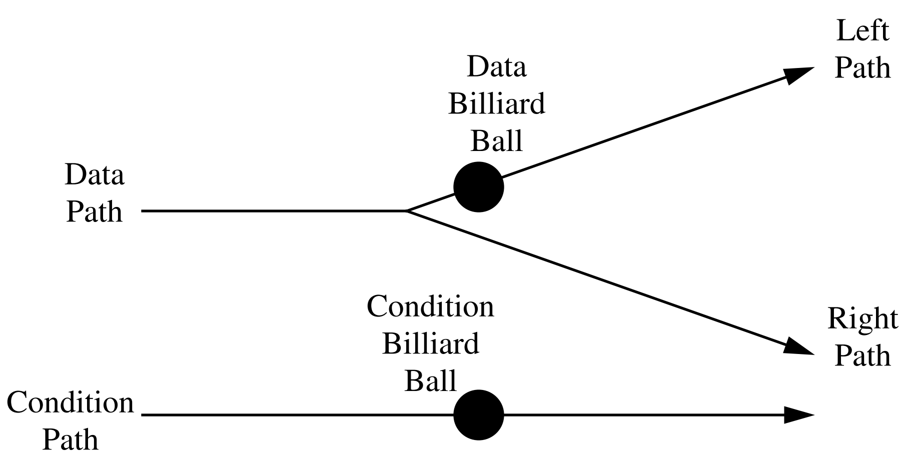


Figure 2.
Schematic diagram of
a switch gate

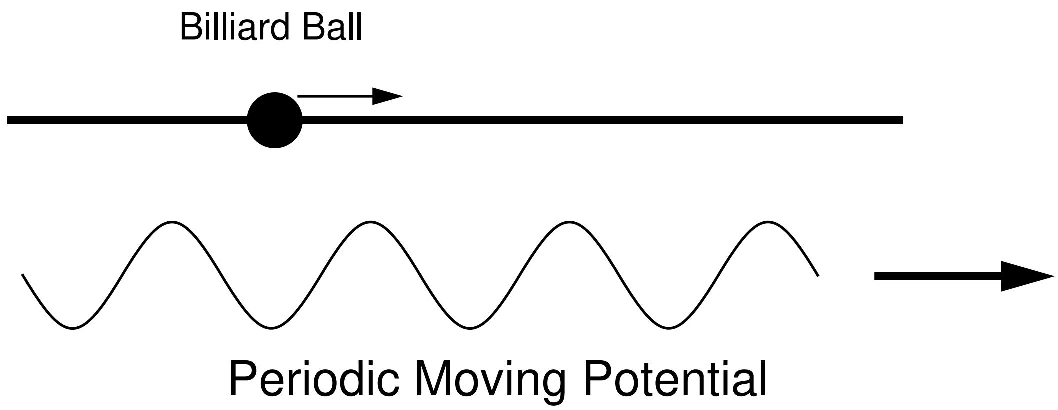


Figure 4
Billiard ball being carried along
by a moving potential

confined to part of a single turn of the helix and as the electric field rotates the electron will be moved along the helical path much as water is moved along an Archimedes screw. This is illustrated in figure 5.

If we were to examine the potential at any point along the helical path we would find that it varied in a sinusoidal fashion much as the potential varies along a straight path under the influence of a moving periodic potential. The use of a periodic moving potential with straight paths (as shown in figure 4) and the use of a simple rotating electric field with helical paths produce very similar results. In either case, the electron is moved along the path.

There are other methods of providing a moving periodic potential that do not involve distribution of clock signals through wires. A simple proposal would be to have a circuit on the surface of a rotating disk, with an opposing disk with fixed charges on it. The relative motion of the two disks would result in a moving potential that would clock the circuit. Alternatively, instead of disks two tubes of differing diameter, one tube placed inside the other, could be used. Circuitry could be placed on the inner surface of the outer tube, while fixed charges could be placed on the outer surface of the inner tube. Again, relative motion of the inner tube with respect to the outer tube could provide a moving potential. This approach also has the advantage that the tubes could be reduced in radius to a very small size (nanometers) and the rotational speed of the inner tube could be made very high (gigahertz) without any fundamental problems. Many such tubes could be stacked adjacent to each other, and charged particles being swept along a potential on one tube could move to an adjacent tube, provided that the movements of the clocking potentials in the two adjacent tubes were appropriately synchronized. A different approach would be to operate the circuit on a piezo-mechanical surface in which surface acoustic waves created a moving electric potential.

We will not consider the many possible alternatives in this paper.

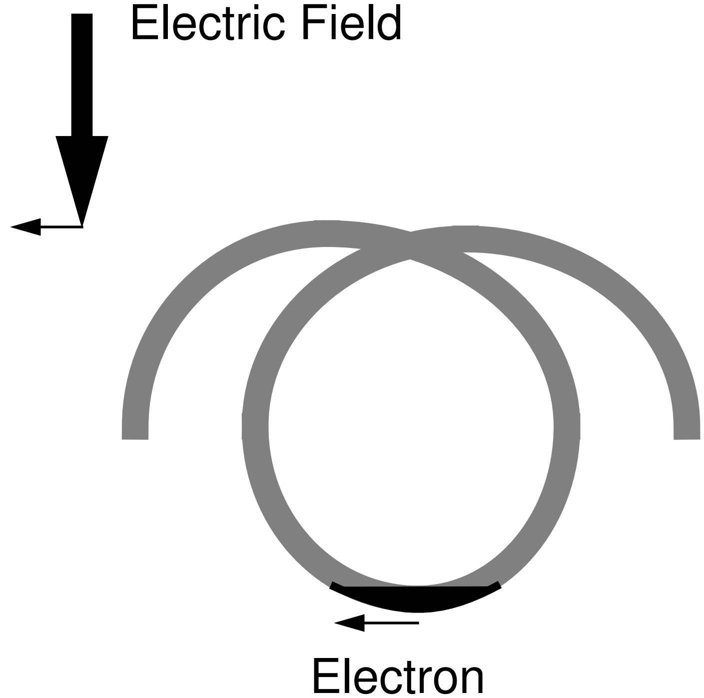


Figure 5.
Electron being carried along a helical
path by a rotating electric field

The Basic Idea

In helical logic, information is encoded in the presence or absence of individual electrons. Electrons are moved from place to place along helical paths. When two helical paths are brought into close proximity the electrons interact through electrostatic repulsion, thus producing an electronic switch gate. The electronic switchgate involves two incoming and three outgoing helical paths. One incoming helical path is termed the “data” path while the other is termed the “condition” path. The two descendants of the data path are called simply the “left” and “right” paths. The electron in the data path will be steered to either the left or right path depending on the presence or absence of an electron in the condition path. This is illustrated schematically in figure 6.

The reverse of this operation, in which three incoming helical paths produce two outgoing helical paths, is similar but “run backwards” in time. Because the basic operation of a switchgate is reversible, it can be operated in either the forwards or reverse direction. We will therefore not explicitly analyze this process. Note that in reverse operation it is essential that the presence or absence of a charge carrier on the data path be correctly correlated with the path (left or right) along which the switched charge is entering the switchgate. This constraint can be met (as illustrated in figure 3 showing the design of the Fredkin gate).

It is worth noting that irreversible operations are sometimes convenient. A basic irreversible operation in helical logic merges two incoming helical paths into one outgoing helical path (and dispenses with the data path). Such an irreversible logic operation, if properly designed, should have an energy dissipation which approaches the fundamental limit: ln 2 *kT*. Helical logic has the valuable ability to degrade gracefully when irreversible operations are required.

We will focus primarily on fundamental issues of device performance while neglecting the issues involved in manufacturing any specific device. Thus, we will simply assume that the ability to economically manufacture atomically precise semiconductor material, with dopant atoms placed at atomically precise lattice coordinates, is available. Such a manufacturing technology is not available today but should be available at some point in the future[REF18, REF65]. In the long run, if we are to achieve the maximum performance possible from semiconductor devices, we will have to develop and use some sort of molecular manufacturing technology. This is true almost regardless of the specific details of the device proposal. The more precisely a device can be fabricated, the better the achievable performance. The limit of this trend will be devices in which each atom is in the right place.

There are many choices for the materials that comprise the helical path and the surrounding medium. An obvious choice is GaAs and AlAs[REF01]. Electrons prefer GaAs to AlAs by about 0.3 ev[REF44], so by operating at a sufficiently low temperature complete confinement of the electrons could be achieved. Another choice would be Ge and Si[REF01, REF50]. Other possibilities that should become feasible in the future with the advent of molecular manufacturing[REF18], would be Si and GaP[REF01], Si and ZnS[REF01], or even C (in the form of diamond) and vacuum. The latter would offer significant advantages because of the low dielectric constant of vaccuum, although the negative electron affinity of the hydrogenated diamond (111) surface suggests that either the use of holes as the charge carriers or the modification of the diamond surface to achieve a positive electron affinity (i.e., fluorinated diamond (111)) would be advantageous. If the helical channels are diamond with hydrogenated diamond (111) surfaces, an electron in the channel would tend to escape into the surrounding vacuum. The idea that a hole might perform a similiar feat is less plausible. Diamond channels with fluorinated surfaces in vaccuum using either electrons or holes (or both) as charge carriers might be an attractive choice once our manufacturing

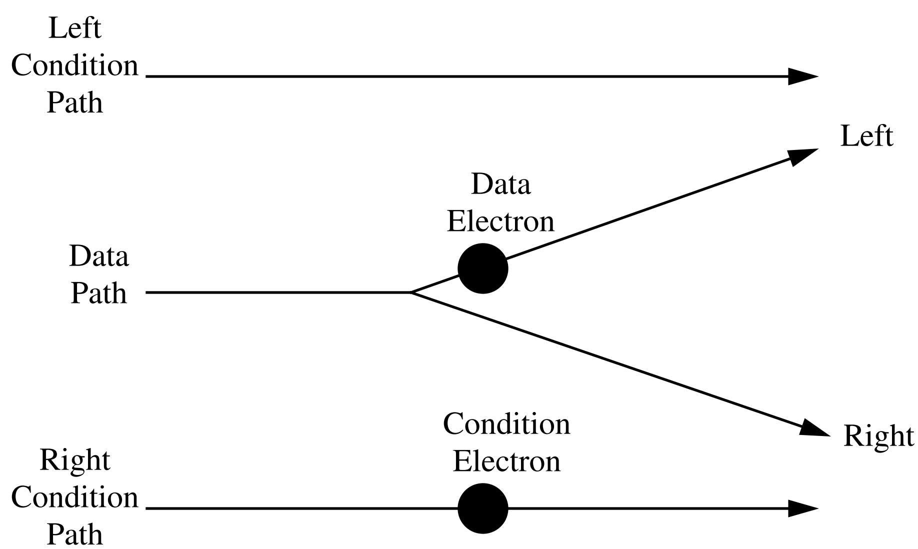


Figure 6.
Schematic diagram of
a two rail switch gate

technology is able to build the required structures.

The negative electron affinity of the hydrogenated (111) diamond surface also suggests the use of an evacuated channel surrounded by walls of diamond. This approach means the channels are holes bored in a block of diamond. The motion of electrons through evacuated channels might prove advantageous by reducing electron/channel interactions.

The use of other charge carriers produces a wider range of possibilities. For example, the charge carrier could in principle be a single proton. The channels would then be pores through which H^+ could easily move. The increased mass of H^+ (as compared with an electron) would simplify confinement by decreasing the effective distance through which the charge carrier could tunnel. Channel size and interchannel distances could both be made much smaller; proton-proton repulsive interactions would therefore be larger and operating temperature could be increased. The smaller size and higher operating temperature of a helical logic system based on hydrogen ions would compensate to some extent for the slower speed, and the result might be advantageous in some applications. Many other small ions could also be used.

The use of charge packets made up of many electrons (or holes) should provide a method of implementing helical logic using today's technology at higher temperatures. The use of charge packets is common in CCDs. It is an open question whether today's technology can economically manufacture the complex structures needed for helical logic (though research devices should be feasible). However, a 2D version of helical logic which used spiral paths should be feasible with today's technology. The implementation of a planar switchgate should likewise be feasible. More research on this point would be worthwhile.

Whatever the materials choice and manufacturing technology, the primary requirement is that a charge carrier be confined to a helical path, and be able to interact with other charge carriers in other helical paths.

Energy Dissipation

Energy dissipation is caused by three main factors: energy loss during charge transport, energy lost to the material in the absence of charge carriers (the dielectric loss factor), and energy loss during switching operations. All three loss mechanisms are a function of clock frequency. We will call these *transport loss*, *dielectric loss* and *switching loss*. Energy loss will also occur when irreversible logic operations occur. We will neglect this source of energy loss in the following analysis, as it will vary widely depending on the logical design of the system and possibly the specific algorithm being used.

Transport Losses

While electrons moving in bulk material are subject to scattering and resistive losses, an electron confined to a portion of a helix at a sufficiently low temperature is almost always in its ground state. The interactions between the electron and the lattice should be greatly diminished, particularly if the energy of the ground state is several times kT below the energy of the first excited state. Conventional resistive losses should be essentially eliminated.

It is important to distinguish between charge transport in helical logic and charge transport along a wire, even a very fine wire. The mobility of an electron confined to a one-dimensional wire can be much higher than in the bulk material[REF63]. However, even in this case the electron is confined in only two dimensions and can move freely along the third. In helical logic, by contrast, the electron is confined in all three dimensions. Viewed from the frame of reference of the moving electron, it is simply sitting at the bottom of a potential well.

Lattice vibrations induced by the moving charge will cause transport losses. Treating the electron as an isolated charge subject to an oscillating force F, we have(Drexler[REF18, page 164]):

$$P_{rad} \approx F^2 \omega^2 \sqrt{\rho} \frac{1}{8\pi M^{3/2}}$$

P_{rad} is the radiated power, F is the force applied to the charge, ω is the frequency in radians per second, and M is a modulus of elasticity.

For an electric field E of 10^8 volts/meter (10^6 volts/centimeter, well below the breakdown strength of diamond) we have a force F on a single electron of $1.6 \times 10^{-19} \times 10^8 = 1.6 \times 10^{-11}$ N. The density ρ of diamond is about 3,500 kg/m^3. The frequency ω in radians per second for a 100 picosecond time to make a single rotation is $2\pi \times 10^{10}$. M is about 10^{12} Pascals. Substituting these values into equation (1) yields a radiated power of roughly 2.4×10^{-18} watts, or 2.4×10^{-28} joules per charge carrier per rotation of the electric field.

This treats the electron as an isolated point charge, and also uses an approximation which should be accurate at frequencies well below the Debye frequency. The frequency selected is below the Debye frequency for diamond, while the error caused by the point charge assumption is conservative (in the sense that this assumption makes the energy dissipation higher than one would otherwise expect).

Lower power dissipation should be feasible. We could maintain approximate charge neutrality by placing a fixed charge in the neighborhood of the electron. A donor atom contributing a positive charge will effectively cancel the negative charge of the electron. This is likely to produce a lower energy dissipation, as acoustic energy radiated from a dipole is lower than that radiated from a monopole (when the wavelength of the radiation is longer than the distance between the charges in the dipole). Alternatively, we could represent a logic “1” by the presence of an electron *and* a hole. The resulting dipole could then be transported along two helices (“double helix logic”) or both the electron and the hole could be confined in a single helix, one of them being half a turn ahead of the other. In the latter case, two types of dipole would be possible: one in which the electron precedes the hole, and another in which the hole precedes the electron. While it might appear paradoxical to have a helix which can simultanesouly confine both an electron and a hole, diamond helices with fluorinated surfaces in vacuum should (as noted earlier) be able to do exactly that.

A frequency of 10^{10} hertz, along with the assumption that a single turn of the helix has a diameter of about 100 nanometers, implies that the electron is moving at a speed of about 3-4,000 meters per second (below the speed of sound in diamond). The dissipation mechanisms that might come into play for significantly higher speeds are not analyzed here.

While the regular motion of charge along a helix should not result in energy dissipation, accelerations caused by variations in the helical structure for the purpose of altering the direction of charge transport will result in accelerations of the charged particle that might well result in radiative losses. Radiative losses are discussed further under switching losses, as accelerations of the charged particles during switching can likewise result in radiative losses.

Dielectric Loss

The dielectric loss can be quite small. The highest experimental quality factor (Q) for a superconducting cavity resonator with a vacuum interior was 5 x 10^{11} at a frequency of 10.5 GHz and a temperature of 1.3 K. Experimental values for Q for cavity resonators at low temperatures (a few Kelvins) with saphire interiors are above 10^8 at a frequency of a few gigahertz, while Q's of 10^{15} are in principle feasible for defect–free saphire at 2 Kelvins and about 3 gigahertz[REF39]. For defect-free materials well below the Debye temperature, Q can increase as $1/T^5$ or $1/T^4$ as the temperature T is decreased[REF42, REF49, REF56]. It is reasonable to expect that the Q for defect–free diamond (a semiconductor which could be used for electronic devices) would be better than that for saphire. These results are for a solid block of homogeneous material, while an actual computer would have irregularities and inhomogenieties associated with the charge transport and switching devices, as well as the mobile charge carriers needed to propagate information. As a consequence, the Q for a computing device is likely to be different than the Q for a homogeneous material. Some analysis of this is done below, but this area requires futher investigation before definitive statements about the dielectric loss for a real system can be made.

Dielectric losses for bulk silicon and diamond are described in [REF39, REF42, REF49, REF56] by the equation:

$$\tan\delta = \eta\frac{2\pi\omega^2 (kT)^4}{\varepsilon\rho v^5 h (kT_D)^2}$$

Where η is a dimensionless anharmonicity parameter with typical values between 10 and 100, ω is the frequency in radians per second, ε is the dielectric permittivity of the crystal, ρ the density, h is Planck's constant, ν the speed of sound, k is Boltzmann's constant, T the temperature in Kelvins, and T_D the Debye temperature. Note that $Q \approx 1/(\tan\delta)$.

In the case of diamond and approximating the value of η as 100[REF42] we can approximate the loss factor. If we use $\omega = 2\pi \times 10^{10}$, $k = 1.38 \times 10^{-23}$, $T = 1$ K, $\varepsilon = 5.7$, $\rho = 3{,}510$ kg/meter^3, $h = 6 \times 10^{-34}$, $T_D = 2340$, $v \approx 15{,}000$ m/s; then we get $\tan\delta \approx 10^{-20}$. (Experimental $\tan\delta$'s for saphire of less than 10^{-9} at a frequency of about 10 gigahertz at a temperature of 1.5 Kelvins have been observed[REF42]. Saphire has a fundamentally higher $\tan\delta$ because of its crystal symmetry). If we assume (rather generously) that each switch is a cube 100 nanometers on a side, and if we further assume that ten times this volume might be required for interconnections between switches (again rather generous), we have a total volume per switch of 10^7 cubic nanometers. The energy stored by an electricic field of 10^8 volts/meter in this volume results in an energy of about 2.5 x 10^{-15} joules per gate (assuming a dielectric of 5.7, as before). Multiplied by our dielectric loss factor, we get an energy loss of under 10^{-34} joules per cycle per gate.

The form of the equation governing the dielectric loss varies greatly depending on the crystal symmetry. For example, centrosymmetric crystals of the D_{6h} symmetry group have a dielectric loss that falls off as ωT^9 (linear falloff in frequency, but ninth power in the temperature), while dielectric loss for crystals of the C_i symmetry group falls off as $\omega^5 T$[REF56].

The presence of the helical paths creates a non-uniform structure, and hence induces additional losses. A loss mechanism suggested by Soreff[REF67] is caused by forces acting between induced dipoles. Even when charge carriers are completely absent the rotating electric field will induce dipoles in the helical paths because the dielectric constant of the path and of the surrounding medium typically differ. A single half-turn of a helical path of the dimensions considered here can be approximated as a column about 100 nm long by 10 nm deep and 10 nm wide. If the electric field is 10^8 v/m and we assume the disparity of the dielectric constants of the path and of the medium is approximately 5.7 (as would be the case for diamond helical paths), such a volume would have an induced dipole moment of roughly 5 x 10^{-26} C-m. Two such dipoles at a distance of roughly 100 nm (e.g., the opposing halves of a single helical turn) will experience a repulsive force of under 10^{-12} N (using a formula derived from the formula for the interaction energy given in [REF68]). Recall that the helical track is already being subjected to forces caused by the charge carrier, estimated above as 1.6 x 10^{-11} N, and that the energy dissipation is a function of the square of the force. Thus, the inter-track repulsive forces caused by induced dipoles are in this approximation smaller than the forces caused by a charge carrier by at least two orders of magnitude. Because the inter-track spacing is small compared with the acoustic wavelength, dissipation from this mechanism should be further substantially reduced. While the number of charge carriers is likely to be smaller than the number of helical turns, it is unlikely to be several orders of magnitude smaller.

The actual dielectric losses in a complex three dimensional structure consisting of numerous very small logic gates are certain to be different from the value estimated here for a perfect diamond crystal. However, it seems unlikely that such losses must necessarily be much larger and they can likely be made much smaller. The wide variation of dielectric loss among crystal types suggests that careful selection and design of materials can be used to make systems with very low dielectric losses. The uncertainty in the dielectric loss remains substantial, however, and the actual loss could be significantly higher than estimated here.

Switching Losses

Switching times need not be the same as the reciprocal of the clocking frequency. A single switching event could take place as the charges moved a fraction of a single helical turn, or could be spread out over many helical turns.

Switching losses can be made very small: an analysis of a specific case is considered later. The switching losses occur because the two electrons are excited by their interaction and will dissipate energy when they fall back to their ground state. If we can keep the two electrons from entering an excited state, we can avoid energy dissipation.

A rather interesting observation is that switching losses caused by this mechanism can in principle be avoided almost completely even if the two electrons do enter excited states. An electron is in an “excited” state only with respect to some potential energy function. For any given wave function, however, there exists a potential energy function for which that wave function *is* the ground state. Thus, if we knew the wave function of each electron as it left the switch gate, and if we could engineer the path along which the electron moves to have the appropriate potential energy function, then the energy dissipated by the electron could in principle be eliminated.

Put more generally, if we know the state of a system then we need not (in principle) dissipate any energy at all.

Unfortunately, it would seem that the wave function of an electron leaving a switch gate depends on the presence or absence of the other electron. Even if both electrons entered the switch gate in their ground state, and even though the evolution of the wave function is deterministic, if we look only at one of the departing electrons we will find it in one of two states depending on the presence or absence of the other electron.

This can, in general, be solved by the use of the “interaction gate”[REF07] illustrated in figure 7.

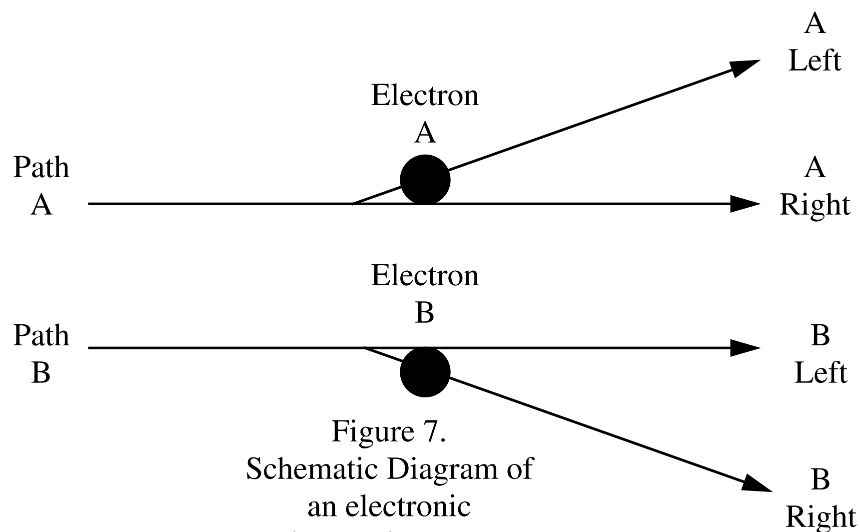


Figure 7.
Schematic Diagram of
an electronic
interaction gate

In the interaction gate two electrons enter along two paths (A and B), and depart by four paths. Again, the interaction gate can be used to implement a Fredkin gate, and so it is logically complete. What makes the interaction gate interesting is the following property: if we know that an electron is present on a particular output path, then we know the inputs to the gate. For example, if an electron leaves the interaction gate along the A left path, then we know that there was an electron on path B (or else the electron would have left the gate on the A right path). Thus, we can engineer the precise geometry and potential of the A left path knowing exactly what state the electron is in when it travels along that path.

Similar observations hold for all other ways in which an electron can leave the interaction gate, and so by appropriate engineering of the outgoing paths we can essentially eliminate energy dissipation from this source.

When a charge is switched down one of two alternate paths it is subjected to an arguably unpredictable acceleration which will in turn generate an arguably unpredictable pattern of radiation. If we assume that the accelerations involved are similar to those that would occur in an oscillating dipole, we can readily estimate the magnitude of this loss by using the formula for the power radiated by an oscillating point charge[REF57, page 32-3]:

$$P_{rad} = \frac{\omega^4 q^2 r^2}{12\pi\varepsilon_0 c^3}$$

Dividing the radiated power by the frequency gives the energy dissipated per cycle, which is:

$$EnergyDissipated/cycle = \frac{\omega^3 q^2 r^2}{6\varepsilon_0 c^3}$$

For f = 10 gigahertz and r = 50 nanometers (and neglecting the dielectric constant of the medium) this results in losses of about 10^{-35} joules per cycle.

Although this is a reasonable approximation to the energy losses caused by the acceleration of a single charge during the switching process, as noted earlier the switching frequency and the frequency of the externally applied electric field need not be exactly the same. Likewise, the distance which the switched and switching charges must traverse during a switching operation is different from the radius of the helix, and the acceleration profiles need not be sinusoidal. Further, modifications to the helical structure to transport charge carriers to various more or less random locations required by the computation will result in accelerations unrelated to switching.

Despite these admittedly rough approximations, the low dissipation computed suggests that this mechanism will not be the dominate source of energy dissipation under the conditions considered in this paper.

Because the energy dissipated per cycle is a function of the square of the charge, and because it might happen that many switches simultaneously switch charge in the same direction (thus effectively increasing the charge that is being accelerated) the energy dissipation from this mechanism might be increased significantly. This could be reduced by careful design of the switching operations to reduce the likelihood of such an event. Even stronger, if double helical logic is employed then a single switching operation could be so designed that it simultaneously accelerated both an electron and a hole in the same direction and by the same amount. This should effectively reduce energy dissipation from this mechanism by many orders of magnitude.

We assume that acoustic losses caused by the switching operation are similar to or smaller in magnitude than the acoustic losses caused by charge transport.

Quantum Mechanical Analysis of the Operation of a Specific Switch Gate

General

This analysis considers the behavior of a two-rail single-electron switch gate (figure 6). There will always be one switching electron, entering either from the left or the right condition path. The switching electron is assumed to be in a narrow potential well that effectively confines it to a small region, assumed to be a point charge. The switched electron, by contrast, is assumed to be in a relatively wide harmonic potential, and it will be necessary to compute the wave function of the switched electron in this potential during the switching process. Although the switch gate is really a three dimensional structure, we will "unroll" the coordinate systems for this analysis. The z axis will measure "distance along the helix." The wave function of the switched electron is assumed to be confined to an x-y plane at any given moment in time, although the z coordinate of the plane will move forwards as the switching process unfolds. The y coordinate for each path does not vary with time. The condition path is assumed to be a few nanometers "above" (located closer to the principle axis of the helix) the data path, i.e., the y coordinates of the data path and the condition path differ by a few nanometers. This implies that the x coordinates of the condition and data paths can vary independently with no risk that the two paths will collide and merge their contents.

This analysis shows that a "plausible" potential exists which simultaneously provides a low energy dissipation and a low error rate at a reasonable switching speed. The underlieing heuristic guiding its design was to minimize excitation of the switched electron during the switching process. The switched electron is assumed to have a significant probability of being excited, a probability which must be analyzed and minimized. Such excitation leads to both energy dissipation and to errors. By minimizing the excitation, both problems can be reduced.

The authors expect that significantly better potentials are feasible.

Standard SI units are used throughout.

Slow switching

Slow switching processes in single-charge helical logic systems are straightforward: the underlying potential along the paths through the gate (determined by the structure of the device) evolves from one well to two, while the proximity of the switching charge raises the potential energy in the blocked path relative to that of the open path so that the ground state of the switched charge remains bound in the deeper well on the open path. In the limit of slow motion and low temperature, the probability of the charge remaining in the ground state can be made arbitrarily close to one, hence the probability of an error (which requires a higher energy state) can be made arbitrarily small. As will be seen, the energy difference between the ground state and the excited states that can evolve into states occupying the blocked well can be made to exceed $\Delta E = 10^{-21}$ J, hence the probability that the blocked well will be occupied via thermal excitation [on the order of $\exp(-\Delta E/kT)$] is less than 10^{-30} for temperatures of ~1 K. Accordingly, thermally induced errors can be made negligible. The chief source of errors in faster switching processes will be excitation resulting from nonadiabatic evolution of the system.

The switching time proposed is a few orders of magnitude slower than the period of a single oscillation of the switched electron when excited to the first excited state. This speed allows us to use the adiabatic approximation in computing the wave function of the switched electron, in estimating the probability of error, and in estimating the energy dissipation. Faster switching times should be feasible but would require a more detailed (and complex) analysis.

Tailoring the potential function

Minimizing the error rate at a given switching speed requires accurate control of the underlying potential, achieved by some combination of bandgap engineering, modulation of well geometry, and manipulation of the local electrostatic environment (the distribution of charges, dipoles, and local dielectric constants). The following analysis assumes that the potential can be modulated by a tenth of an eV or more in essentially any smooth pattern. Analysis of a related problem (barrier heights of two surfaces sliding against each other) using a simple Monte Carlo technique makes this assumption plausible[18, page 278]. For mathematical convenience, this analysis examines a potential with discontinuous first derivatives; smoother potentials (e.g., the present potential convolved with a suitable Gaussian) can have superior properties.

Variations in potential along the z axis are limited in frequency by the spacing of the atoms, and are further effectively limited by the fact the wave function will have some spread along the z axis. As a consequence, we have limited ourselves to a potential that varies relatively slowly with changes in the z coordinate. Variations in potential along the x axis involve tens of nanometers, and so could reasonably be engineered by appropriate placement of individual atoms. The potential along the y axis is assumed to confine the electron to a fixed distance from the axis of the helix, e.g., the electron is confined to a helical ribbon.

Non-thermal excitation and nearly-harmonic potentials

Each charge is assumed to be in its ground state as it enters the switch. An error is assumed to occur when a charge is excited to an energy level high enough to permit it to occupy a secondary potential well on the blocked path. (This is a conservative criterion: not all charges with sufficient energy will in fact exit along the blocked path.) For error rates to be low excitation probabilities must be small and the evolution of the system must be nearly adiabatic. An expression for the total probability of a transition from the *m*th to the *n*th level under nearly-adiabatic conditions [REF61] is

$$(P_n = |C_{nm}|^2) \cong \frac{4\hbar}{(E_n - E_m)^4}\left|\int v^*{}_n \frac{\partial H}{\partial t} v_m dx\right|^2$$

where P_n is the probability of ending in state n (assuming the initial state is m), C_{nm} is the matrix element, E is the energy of a state, H is the system Hamiltonian, and v_n and v_m are wave functions. This gives an upper bound on the probability when the rate of change in the Hamiltonian ($\partial H/\partial t$) is a constant over some interval and zero before and after. Transition probabilities for systems in which $\partial H/\partial t$ increases and decreases smoothly over a time $\Delta t >> (\bar{\mathrm{h}})/\Delta\mathrm{E})$ are much lower; probabilities can also be larger or smaller after a series of abrupt changes, as phases add or cancel. For the present analysis, the maximum value of P_n encountered in a series of sample times, $P_{n,\max}$, is used to estimate the probability of occupancy of that state at the end of that time interval. A more detailed analysis preserving phase information would be of interest, but should for the present system yield results of the same order as this estimate.

The following analysis describes the Hamiltonian and associated switched-charge wave functions along a one-dimensional coordinate. Wave functions are computed in the adiabatic approximation, neglecting the perturbing potential imposed by acceleration and charge mass (this imposed potential has a magnitude < 5 x 10^{-23} J, significantly smaller than the potentials imposed by either the switching charge or the material of the switch). Changes in this perturbing potential, however, provide a significant component of $\partial H/\partial t$ and hence substantially affect excitation probabilities.

We will call the potential created by the switch gate acting on the switched charge in the absence of the switching charge the "underlieing potential." This potential is created by appropriate design of the switch gate. Because the switch gate is bilaterally symmetric, the underlieing potential must necessarily be bilaterally symmetric as well.

Ideally, the switching charge would steer the switched charge down the chosen path without changing the shape of the potential well. From the frame of reference of the switched electron, almost nothing would have happened and excitation would be minimal. If the underlying potential were harmonic, and the potential imposed by the switching charge had a linear gradient, then the approach of the switching charge would have exactly this effect (a linear gradient applied to a harmonic well results in a new harmonic well of the same width but moved laterally: exactly what we desire).

In practice, the gradient created by a point charge is nonlinear. To some extent, this can be compensated by adjusting the underlieing potential so that the resulting total potential (the underlieing potential plus the potential created by the switching electron) is harmonic. This cannot be done completely, for the underlieing potential is bilaterally symmetric and the wave function of the switched electron will have a significant component on both sides of the center of the switch gate (at least during the earlier phases of switching). However, the region of the underlieing potential in the vicinity of the switched electron and on the same side of the centerline can be made exactly harmonic by appropriate modifications to the underlieing potential. Second order terms can be cancelled by appropriate changes in the underlieing potential, but higher-order terms cannot be cancelled during the early stages of switching, owing to the symmetry constraint on the underlying potential. (Effective cancellation of third-order terms becomes feasible when the switched charge has shifted by more than the characteristic radius of its ground-state wave function.) The residual sources of $\partial H/\partial t$ include third-order terms (during the early stages of switching), together with changes in the acceleration-induced potential and changes in the potential associated with the growth of a barrier as the incoming well in the underlying potential splits to form two outgoing wells.

Structure of the potential

The potential surface analyzed here is not optimized, but is the product of several cycles of modification aimed at correcting obvious defects. The potential surface is defined in terms of two "paths," the path of the switching electron $x_{switching}(z)$ and the path of the switched electron $x_{switched}(z)$ (where z is a dimensionless quantity). While the path of the switching electron is exact, the "path" of the switched electron is only approximate as the wavefunction of the switched electron is diffuse in the x coordinate. The function $x_{switched}(z)$ can therefore be viewed as an arbitrary function used to define the potential.

The potential combines several distinct components. The first is the electrostatic potential imposed by the switching charge:

$$V_{\text{elect}}(x, z) = \frac{e}{4\pi\varepsilon\varepsilon_0}\left[\Delta r^2 + (x - x_{\text{switching}}(z))^2\right]^{-\frac{1}{2}}$$

We assume ε is 5.7 (corresponding to the dielectric constant of diamond) and Δr is 4 nm (a gap in the y direction sufficient to reduce cross-channel tunnelling to low values). The parameter z is a dimensionless measure of distance along the path, entering into the dynamical analysis through a choice of dz/dt.

All other components are part of the underlying potential and are constrained to be even functions about the midpoint. These include a harmonic term:

$$V_{\text{harm}}(x) = \frac{1}{2}k_s x^2$$

and a recursively-defined correction term:

$$V_{\text{corr}}(x,z) = \begin{cases} \text{if } |x| > 2x_{\text{switched}}(z),\ g(z)\,|x| - V_{\text{elect}}(|x|,z) + g(z)\,|x'| - V_{\text{elect}}(|x'|,z) \\ \text{if } |x| \le 2x_{\text{switched}}(z),\ g(z)\,|x| - V_{\text{elect}}(|x|,z) \end{cases}$$

where

$g(z) = -k_s x_{\text{switched}}(z)$ and $x' = 2x_{\text{switched}}(z) - |x|$

This term cancels the difference between the electrostatic potential and a linear gradient with a slope of g J/m in the region between x = 0 and x = $2x_{\text{switched}}(z)$, yielding a total potential that is exactly harmonic within these bounds. There does not appear to be a unique "right" choice for the value of the total potential outside these bounds. We have (somewhat arbitrarily) chosen the underlieing potential so that $[V_{\text{total}}(x_{\text{switched}}(z)+\text{offset}, z) - V_{\text{total}}(x_{\text{switched}}(z), z)] + [V_{\text{total}}(x_{\text{switched}}(z)-\text{offset}, z) - V_{\text{total}}(x_{\text{switched}}(z), z)] = V_{\text{harm}}(\text{offset}) + V_{\text{harm}}(-\text{offset})$.

Finally, the walls of the potential well are steepened in regions where the ground state wave function has a small amplitude by adding

$$V_{wall}(x,z) = \begin{cases} \text{if } |x| \le x_{switched}(z) + 9 \times 10^{-9}\text{ m}, & 0 \\ \text{if } |x| > x_{switched}(z) + 9 \times 10^{-9}\text{ m}, & 10^{-3} \times [|x|-x_{switched}(z) - 9 \times 10^{-9}]^2 \end{cases}$$

This term substantially raises the energies of the higher excited states without greatly affecting the ground state.

These various potentials have more or less randomly altered the potential at the bottom of the harmonic potential well, e.g., $V_{total}(x_{switched}(z), z)$ might vary as z varies. We desire, however, that the potential at the bottom of the well remain constant. To this end, we define $V_{zero}(z)$ as having that value required to make $V_{total}(x_{switched}(z), z)$ equal 0. Note that $V_{zero}(z)$ is a function of z only and does not change as x changes.

The overall potential is:

$$V_{total}(x,z) = V_{harm}(x) + V_{elect}(x,z) + V_{corr}(x,z) + V_{wall}(x,z) + V_{zero}(z)$$

The total potential has continuous first derivatives so long as the gradient g is numerically equal to $\partial V_{elect}(x,z)/\partial x$ evaluated at $x = 0$. This condition can be maintained during the early evolution of the potential, but must be violated at later times. The difference between these gradients defines the magnitude of a wedge potential (creating a peak at $x = 0$) that forms an implicit component of V_{total}.

Time history of the potential

The evolution of the well potential determines $\partial H/\partial t$. Since the potential is determined by the positions of the switching and switched charges, the evolution of the well potential is specified by choices of $x_{switching}(z)$ and $x_{switched}(z)$. In the model potential, these functions are the natural cubic splines [REF66] passing through the points tabulated in Table 1. The resulting paths and potentials are illustrated in Figures 9 and 10. The speed of switching is determined by dz/dt (assumed to be constant). It is often convenient to use the reciprocal dt/dz, which we call the characteristic switching time or t_{char}. Figures 8, 9 and 10 assume a t_{char} of 5 ps (i.e., dz/dt of 2×10^{11}).

Table 1. Defining points of natural cubic splines [REF66] describing the trajectories of switched and switching charges in the model potential (rounded to three digit precision). The parameter z is dimensionless.

Table 1:

Switching		Switched	
z	x (nm)	z	x (nm)
–2.0	–100.0	–2.0	0.105
–1.0	–50.0	–1.0	0.418

Table 1:

Switching		Switched	
–0.6	–32.0	–0.65	0.872
–0.25	–15.5	–0.6	1.01
–0.1	–11.0	–0.25	4.0
–0.05	–9.9	0.0	10.0
0.2	–15.3	0.4	24.5
0.3	–19.6	2.0	130.0
0.6	–33.8		
1.0	–57.2		

In the initial (pre-wedge) phase of the switching interaction (through z = ~0.05), maintaining a zero wedge potential makes $x_{\text{switching}}$ and x_{switched} functions of one another (a relationship approximated by the given splines). The acceleration profile is chosen to limit the the jerk resulting from the motion of the well minimum. (Note that the "wedge potential" is not an actual potential, but is merely a consequence of combining the potentials already defined. It is useful, however, to give it a name).

In the following (disengagement) phase, the growth of the wedge potential continues the motion of the switched charge away from the midplane, and a substantial and growing region around the well minimum is exactly harmonic. Changes in well shape outside this region cause residual excitation, but these decline as the exactly harmonic region expands into the tails of the ground-state wave function. (In particular, the discontinuity in the potential gradient associated with the peak of the wedge results in a fast rate of change of potential as seen in the well frame.) The disengagement phase ends as the magnitude of the ground-state wave function at $x = 0$ becomes negligible (when $z \approx 0.15$).

In the final (seperation) phase, excitation is small, and (according to the energy-based error bound described above) errors are minimized by choosing a design that maximizes the quantum number n of the lowest excited state centered in a secondary well. Placement of the switching charge so as to create two secondary wells of equal depths is nearly optimal during this phase; in the model potential, this condition determines $x_{\text{switching}}$ for z > 0.2. The lowest state that can become bound in the wrong well, $n = 9$.

Error rates

Error rates depend on dz/dt. Estimating error probabilities P_{err} as the sum from n = 9 to $n = 14$ of $P_{n,\text{max}}$, for values of dz/dt corresponding to t_{char} = 5, 10, 20, and 40 ps yields values of ~ 9.3×10^{-11}, 2.3×10^{-11}, 5.8×10^{-12}, and 1.5×10^{-12}, respectively. Contributions from states with $n \geq 15$ are small. The near proportionallity of P_{err} to $(\mathrm{d}z/\mathrm{d}t)^2$ indicates that jerk contributes little to the excitation, which is instead dominated by consequences of deviations of V_{net} from exact harmonicity. Even with t_{char} = 5 ps, P_{err} remains low enough to enable the construction of digital logic systems of high reliability, with a modest cost in time and energy dissipation resulting from error detection and correction.

Energy dissipation due to switching

Energy dissipation occurs when switched charges undergo excitation and deexcitation (typically with phonon emission). This process is dominated by excitation to $n = 1$, which is > 10^3 times more probable than excitation to any higher state. The associated energy dissipation $E_{\text{diss}} \approx P_{1,\max} \times \hbar\omega \approx P_{1,\max} \times 4.6 \times 10^{-22}$ J. For $t_{\text{char}} = 5$, 10, 20, and 40 ps, $E_{\text{diss}} \approx 1.3 \times 10^{-26}$, 2.3×10^{-28}, 1.7×10^{-28}, and 5.2×10^{-29} J per switching event. Jerk plays a dominant role in energy dissipation—but not error rates—at the faster switching speeds, owing to the greater coupling to low-spatial-frequency features of $\partial H/\partial t$ by $n = 1$, relative to $n \geq 9$. Each of these values of E_{diss} is far smaller than kT at one kelvin (~ 1.4×10^{-23} J). An analysis retaining phase effects would permit the choice of an acceleration profile resulting in substantial cancellation of excitation to the $n = 1$ state, reducing energy dissipation by ~ 1000 relative to the present estimates.

Input and Output

Another issue is communication with the outside world. If the active computing element is embedded in a high intensity rotating electric field then the use of metallic wires poses obvious difficulties (much as placing a fork in a microwave oven might be ill advised). It might still be possible to use wires provided that they are always at right angles to the electric field. It might be difficult to achieve this, however, in which case some pickup from the rotating electric field would still occur. The use of twisted pair would reduce this problem. Despite the obvious drawbacks, the use of metallic wires might prove useful.

Another method would be to use fiber optics. This eliminates concerns about putting wires into an intense rotating electric field. Simply by having an input fiber which started well outside the region of the rotating electric field and terminated within the active region it would be possible to selectively create electron-hole pairs in the semiconductor by injecting light into the other end of the fiber. Absorption of photons can be done in relatively small structures; for example, a one micron thickness of GaAs can absorb roughly a third of the incident light[REF46 page 46]. If the GaAs were part of a helical structure, the generated electrons and holes would then be separated by the electric field and would move to opposite sides of the helix. This method of generating electron-hole pairs will produce packets of irregular size. By making too many electrons and holes and then selectively re-combining unwanted electrons and holes, we could reduce the packet size to the desired value. Re-combination will occur when the diameter of the helix becomes small enough to allow the electrons and holes to recombine. By controlling the diameter of the helix, the number of electrons and holes that remain in the packets can be controlled. An alternative method would be split the helix and allow the packet (through self repulsion) to break apart into smaller pieces. A sufficient number of splits would produce packets that had only a single electron or hole in them. The result would be several helices each one of which contained exactly zero or one charge carrier. This rather error-ridden input could then be converted into a more reliable result using appropriate logic circuitry.

Having once generated a pair of charge carriers, we could simply continue to let the pair be switched throughout further logic operations. Double helical logic would require the use of two helical paths and two switching elements, increasing the volume and component complexity. By using single helical dipole logic (e.g., a single helix carrying both an electron and a hole to represent a “1”) the leading charge carrier would be switched through a switch gate and when the electric field rotated through an additional 180 degrees, the second carrier of the pair would be switched in exactly the same fashion.

For output, we need merely recombine electrons with holes next to an optical fiber. The optical fiber can then carry the resulting photons outside the rotating electric field where conventional detectors and electronics can be used. In single helix dipole logic, a simple method of combining electrons and holes would be to terminate a helical path carrying a pair. The leading charge carrier, upon reaching the end of the helical path, would be unable to continue. The trailing carrier would then catch up. When the two met, the electron and hole would recombine and generate a photon.

Both input and output, depending on how the’re implemented, might suffer from a relatively high error rate. This can be corrected by using various coding and multiple transmission schemes that provide sufficient redundancy to correct the errors.

Simple, robust I/O using optical methods should be feasible even in the presence of the strong rotating electric field required by helical logic.

The optical methods of input and output described here are highly dissipative. Many entering photons will fail to generate an appropriate electron-hole pair; when a pair is generated in the electric field the resulting current flow will be dissipative; the recombination of electrons with holes to adjust the size of the charge packet is dissipative; and when electrons are combined with holes to generate optical output many of the resulting photons will be lost. While simple, this method of input and output is far from reversible and will impose energy limitations on the I/O bandwidth. It is often feasible in principle for I/O to be done in a reversible fashion. If two reversible processors wish to communicate with each other in a logically reversible manner, then there is no fundamental requirement that the process be dissipative. Further research is needed to provide a good low energy method of input and output.

Three main categories of energy dissipation were considered: transport losses, dielectric losses, and switching losses. Estimates of fundamental losses per cycle when the external field has a frequency of 10 gigahertz were made. Transport losses are primarily acoustic, estimated at below 10^{-27} joules per cycle.

Dielectric losses vary widely depending on the exact crystal symmetry, with even the equation describing the loss being different for different symmetry groups. A rough estimate for the fundamental dielectric loss for diamond is below 10^{-35} joules per cycle per switch with an electric field strength of 10^{8} volts/meter. While pure diamond will almost certainly have a different (and lower) loss than the highly structured complex logic circuitry that would be needed in an actual system, this calculation still suggests that energy dissipation from this source can be made very small (and can likely be reduced to below 10^{-27} joules per cycle, the dissipation estimated from other loss mechanisms). Further research on the dielectric loss from is required to better understand and minimize the losses that can be expected from this mechanism.

Switching losses below 10^{-27} joules per logic operation at switching speeds faster than 10^{-10} s should also be feasible. These losses involve both the excitation of electrons and their dissipative return to the ground state, as well as radiative losses resulting from the accelerations to which the switched electron is subjected during a switching operation.

These estimates support the conclusion that thermodynamically reversible single electron logic operations should eventually be able to achieve very low energy dissipations, very likely below 10^{-27} joules per logic operation at a temperature of 1 K and a speed of 10 gigahertz. The largest energy loss mechanism identified was acoustic radiation, which could be reduced by the use of dipoles. Even though the estimates of dielectric loss and radiative loss were small there is substantial uncertainty in their values: further analysis seems appropriate.

Faster Operation

The reader might have noticed that 10^{-27} J dissipated times 10^{-10} seconds switching time is 10^{-37} J-s, significantly below Planck's constant $h = 6.626 \times 10^{-34}$ J-s. This is not a violation of any fundamental principles, however[REF64]. If, instead of the energy dissipated we consider the magnitude of the change in the switching Hamiltonian ΔH over the course of a switching operation then we can produce an energy-time product which does have significance in the context of the current proposal. The approximate value of ΔH is 10^{-20} J; multiplied by the switching time this gives 10^{-30} J-s which is larger than Planck's constant by over three orders of magnitude. This gap suggests that faster switching should be feasible before any fundamental limits are encountered.

In the present proposal simply increasing the clock frequency would result in faster switching times, though at the cost of increased energy dissipation and error rate since no attempt was made to keep track of phase information. The switching time was simply made long compared with the time of a single oscillation of the first excited state of the switched electron, thus allowing use of the adiabatic approximation. If the switching time were $\sim\Delta H/h$ and if we computed the wave function of the switched electron as it left the switch, it should be possible to "catch" the switched electron in an appropriately designed potential. Switching speeds of $\sim 10^{-14}$ s with minimal energy dissipation and error rates should be feasible using this approach. With lateral movement of the switched electron of $\sim 10^{-8}$ m we have electron speeds of $\sim 10^{6}$ m/s (~.003 c), accelerations of $\sim 10^{20}$ m/s^2, forces of $\sim 10^{-11}$ N, and an electron-electron separation (in vacuum) of $\sim 5 \times 10^{-9}$ m. Relativistic effects and materials limitations should not pose fundamental limitations for operations at these speeds.

A switching speed faster than $\sim 10^{-14}$ s in the context of the current proposal would require creating a larger switching potential by, for example, using more than one switching electron (e.g., a charge packet). Use of a suitable charge packet would also permit reliable room temperature operation. More detailed analysis of rapid switching would be of interest.

Summary

The picture of computing that emerges from this analysis is relatively simple: a solid block of material that is intricately designed at the molecular level is placed into a cavity resonator which has a frequency of several gigahertz. The resonator and the block of material are kept at low temperature, perhaps 1 K. Fiber optic connections allow the block to communicate with the outside world. Photons generated outside the cavity resonator by conventional electronics are carried into the block via fiber optics to create electron hole pairs, which then enter into computations within the block. By combining electrons with holes, photons can be generated within the block which are then carried out of the cavity resonator by fiber optics and detected by conventional electronics.

The cavity resonator has a high Q, and the computations within the block of material produce relatively little additional energy loss. A significant source of energy loss in this particular proposal is in the I/O, which is quite dissipative. If the number of I/O operations is relatively small compared with the number of total logic operations performed or if some method of I/O can be found which dissipates much less energy (which seems likely) then the total energy dissipation per logic operation can be made remarkably small.

Conclusions

Projections of current trends in energy dissipation per gate operation[REF06] suggest that the kT "barrier" will become significant within a few decades. This barrier can be overcome by using reversible logic. Further, the trend towards decreasing device size and increasing device precision will likely result in atomically precise logic elements which employ single electrons, single holes, or both to carry the 0's and 1's of a computation. Helical logic is the first proposed member of the class of single electron thermodynamically reversible atomically precise logic devices. Members of this class might mark the ultimate evolution of electronic devices. Helical logic also eliminates clock distribution. The energetic costs of charging and discharging clock lines might otherwise become a significant source of energy dissipation in future computing devices. A preliminary analysis suggests that the energy dissipation of helical logic can be reduced below 10^{-27} joules per logic operation at a switching speed of 10 gigahertz and an operating temperature of 1 K. Further reductions in energy dissipation might be feasible if double helical logic is used. The estimates of dielectric loss and radiative loss were both well below 10^{-27} J, but had significant uncertainties that require further examination. Further research to develop the highly precise manufacturing technologies needed to economically fabricate this (and other) proposed computing technologies is needed[REF65]. Further investigation of single electron, thermodynamically reversible, atomically precise logic devices seems warranted.

Acknowledgements

The authors would like to thank J. Storrs Hall, Rolf Landauer, Alan Lewis, Norman Margolus, Jim Mikkelsen, John Northrup, Ed Richley, Jeff Soreff, Robert Thornton, Tommaso Toffoli, and Chris van de Walle.

References

REF01.) *High–Speed Semiconductor Devices*, by S. M. Sze, Wiley Interscience 1990.

REF02.) *Rod Logic and Thermal Noise in the Mechanical Nanocomputer*, by K. Eric Drexler, in *Proceedings of the 3rd International Symposium on Molecular Electronic Devices*, F.L. Carter, R.E. Siatkowski, H. Wohltjen, eds., North–Holland 1988.

REF03.) *The Fundamental Physical Limits of Computation*, by Charles H. Bennett and Rolf Landauer, Scientific American, July 1985, pages 48–56.

REF04.) *Single Electronics*, by Konstantin K. Likharev and Tord Claeson, Scientific American Vol 266 No. 6, June 1992, pages 80–85.

REF05.) *The Fundamentals of Physics*, Extended Third Edition, by David Halliday and Robert

Resnick, Wiley 1988.

REF06.) *Dissipation and noise immunity in computation and communication*, by Rolf Landauer, Nature, Vol. 335, October 27 1988, page 779.

REF07.) *Conservative Logic*, by Edward Fredkin and Tommaso Toffoli, International Journal of Theoretical Physics, Vol. 21, Nos. 3/4, 1982, pages 219–253.

REF08.) *The Thermodynamics of Computation − A Review*, by Charles H. Bennett, International Journal of Theoretical Physics, Vol. 21, No. 12, 1982, pages 905–940.

REF09.) *Classical and Quantum Limitations on Energy Consumption in Computation*, by K. K. Likharev, International Journal of Theoretical Physics, Vol. 21, Nos 3/4, 1982, pages 311 to 326.

REF10.) *Reversible Conveyer Computation in Array of Parametric Quantrons*, by K. K. Likharev, S. V. Rylov, and M. K. Semenov, IEEE Transactions on Magnetics, Vol. MAG–21, No. 2, March 1985, pages 947–950.

REF11.) *Notes on the History of Reversible Computation*, by Charles H. Bennett, IBM J. Research and Development, Vol. 32, No. 1, January 1988, pages 16– 23.

REF12.) *Engines of Creation*, by K. Eric Drexler, Doubleday, 1986.

REF13.) *There's Plenty of Room at the Bottom*, a talk by Richard Feynman (awarded the Nobel Prize in Physics in 1965) at an annual meeting of the American Physical Society given on December 29, 1959. Reprinted in "Miniaturization", edited by H. D. Gilbert (Reinhold, New York, 1961) pages 282–296.

REF14.) *A Small Revolution Gets Under Way*, by Robert Pool, Science, Jan. 5 1990.

REF15.) *Time/Space Trade–offs for Reversible Computation*, by Charles H. Bennett, SIAM J. Computing, Vol. 18, No. 4, pages 766–776, August 1989.

REF16.) *Two Types Of Mechanical Reversible Logic*, by Ralph C. Merkle, submitted to Nanotechnology.

REF17.) *Proceedings of the 3rd International Symposium on Molecular Electronic Devices*, F.L. Carter, R.E. Siatkowski, H. Wohltjen, eds., North–Holland 1988.

REF18.) *Nanosystems: Molecular Machinery, Manufacturing, and Computation*, by K. Eric Drexler, John Wiley, 1992.

REF19.) *The Design of a Conservative Logic Computer and a Graphical Editor Simulator*, by Andrew Lewis Ressler, Masster's Thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, January 1981.

REF20.) *Donor Impurities, N–type Doping, and Impurity Diffusion in Diamond*, by J. Bernholc, S. A. Kajihara, and A. Antonelli, 20th International Conference on Semiconductors, Volume 1, page 332–335.

REF21.) *An Electroid Switching Model for Reversible Computer Architectures*, by J. Storrs Hall, Laboratory for Computer Science Research, Rutgers University, Proceedings of the

Workshop on Physics and Computation, October 2-4 1992, page 237-247, IEEE Press.

REF22.) *Quantum photoyield of diamond (111) —— A Stable negative–affinity emitter*, by F. J. Himpsel, J.A. Knapp, J.A. Van Vechten, and D. E. Eastman, Phys. Rev. B, Vol. 20 No. 2, pages 624–627.

REF23.) *Design Principles for Achieving High–Performance Submicron Digital Technologies*, by Edward Fredkin and Tommaso Toffoli, November 1978 DARPA proposal.

REF24.) *Magnetic Bubbles*, by Andrew H. Bobeck and H.E.D. Scovil, Scientific American, June 1971, Vol. 224, No. 6, pages 78–91.

REF25.) *Ideal MOS Curves for Silicon*, by A. Goetzberger, Bell Systems Technical Journal, Vol 45, September 1966, pages 1097–1122

REF26.) *Single atoms as transistors*, by Sean Washburn, Nature Vol 357, May 21 1992, pages 199–200.

REF27.) *A CCD/CMOS process for integrated image acquistion and early vision signal processing*, by Craig L. Keast and Charles G. Sodini, SPIE Vol. 1242 Charge–Coupled Devices and Solid State Optical Sensors (1990), pages 152–161.

REF28.) *Eliminating Threshold Losses in MOS Circuits by Bootstrapping Using Varactor Coupling*, by Reuben E Joynson, Joseph L. Mundy, and Constantine Neugebauer, IEEE Journal of Solid–State Circuits, Vol SC–7, No. 3, June 1972, pages 217–224.

REF29.) *Hot–Clock nMOS*, by Charles L. Seitz, Alexander H. Frey, Sven Mattisson, Steve D. Rabin, Don A. Speck, Jan L.A. van de Sneepscheut, 1985 Chapel Hill Conference on VLSI, Computer Science Press, pages 1–17.

REF30.) *Semiconductor Surface Varactor* by R. Lindner, The Bell System Technical Journal, Vol. 41 No. 3, May 1962, pages 803–831.

REF31.) *Energy–Time Trade–Offs in VLSI Computations*, by Akhilesh Tyagi, pages 301–311; in Lecture Notes in Computer Science, 1989, Volume 405: Foundations of Software Technology and Theoretical Computer Science, Ninth Conference, Bangalore India, edited by C. E. Veni Madhaven, December 1989.

REF32.) *A Proof about Molecular Bearings*, by Ralph C. Merkle, Nanotechnology 4 (1993) pages 86-90.

REF34.) *Digital Charge-Coupled Logic (DCCL)*, by T. A. Zimmerman, R. A. Allen, and Robert W. Jacobs, IEEE Journal of Solid State-Circuits, Vol. SC-12, No. 5, October 1977, pages 473-485.

REF35.) *Combinatorial Digital Logic Using Charge–Coupled Devices*, by J. Greg Nash, Journal of Solid–State Circuits, Vol SC–17, No. 5, October 1982, pages 957–963.

REF36.) *Basic C.C.D. Logic Gates*, by J. H. Montgomery and H. S. Gamble, The Radio and Electronic Engineer, Vol. 50, No. 5., pages 258-268, May 1980.

REF37.) *A New Concept in Computing*, by R. L Wigington, Proceedings of the IRE, April 1959, Vol 47, No 4, page 516-523.

REF38.) *Reversible Electronic Logic Using Switches*, by Ralph C. Merkle, Nanotechnology 4 (1993) pages 21-40.

REF39.) *Systems with Small Dissipation*, by V. B. Braginsky, V. P. Mitrofanov, and V. I. Panov, University of Chicago Press, 1985.

REF40.) *Transmission through a bend in an electron waveguide*, by Craig S. Lent, Appl. Phys. Lett. 56 (25), 18 June 1990, pages 2554-2556.

REF41.) *Drift and Diffusion in Reversible Computation*, by R. Landauer and M. Buttiker, Physica Scripta Vol. T9, pages 155-164, 1985.

REF42.) *Experimental observation of fundamental microwave absorption in high-quality dielectric crystals*, by V.B. Braginsky, V.S. Ilchenko, and Kh.S. Bagdassarov, Physics Letters A, Volume 160, No. 6, March 2, 1987, pages 300-305.

REF43.) *Electronic Structure of the Superatom: A Quasiatomic System Based on a Semiconductor Heterostructure*, by Takeshi Inoshita, Shuhei Ohnishi, and Atsuchi Oshiyama, Physical Review Letters, Vol. 57 No. 20, November 17, 1986, pages 2560-2563.

REF44.) *Band lineups and deformation potentials in the model-solid theory*, by Chris G. Van de Walle, Physical Review B, volume 39, Number 3, January 15 1989-II, pages 1871-1883.

REF45.) *On Magnetic Bubble Logic Circuits*, by Kozo Kinoshita, Tsutomu Sasao, and Jun Matsuda, IEEE Transactions on Computers, Vol. C-25, No. 3, March 1976 pages 247-253.

REF46.) *Heterostructure Lasers*, by H.C. Casey and M.B. Panish, Academic Press, 1978.

REF47.) *Theory of the Quantized Adiabatic Particle Transport*, by Qian Niu, Modern Physics Letters B, Vol. 5, Nos. 14&15 (1991) 923-931.

REF48.) *Possible Applications of the Single Charge Tunneling*, by D. V. Averin and K. K. Likharev, chapter 9 from *Single Charge Tunneling*, H. Grabert and M.H. Devoret, eds, Plenum Press New York, 1992.

REF49.) *Dielectric Losses in Crystals*, by V.L. Gurevich, Sov. Phys. Solid State 21(11), November 1979, pages 1993-1998

REF50.) *Silicon-Based Semiconductor Heterostructures: Column IV Bandgap Engineering*, by John C. Bean, Proceedings of the IEEE, Vol. 80 No. 4, April 1992.

REF51.) *Irreversibility and Heat Generation in the Computing Process*, by Rolf Landauer, IBM Journal of Research and Development, Vol. 6, July 1961, pages 183-191.

REF52.) *Equipment Simulation: Part I*, by J. Ignacio Ulacia F., Christoph Werner, Solid State Technology, November 1990, pages 107 et sequitur.

REF53.) *Fundamental Physical Limitations of the Computational Process*, by Rolf Landauer, in *Sixth International Conference on Noise in Physical Systems*, P.H.E. Mejier, R.D. Mountain, F.J.Soulen Jr., eds [Nat. Bur. Stand. (U.S.), Publ. 614, 1981] page 12.

REF54.) *Single Charge Tunneling*, by H. Grabert and M.H. Devoret (eds), Plenum Press, 1992.

REF55.) *Logical Reversibility of Computation*, by Charles H. Bennett, IBM Journal of Research and Development, 17, 525-532, November 1973.

REF56.) *Intrinsic dielectric loss in crystals*, by V. L. Gurevich and A. K. Tagantsev, Advances in Physics, 1991, Vol. 40, No. 6, pages 719-767.

REF57.) *The Feynman Lectures on Physics, Vol. I*, by Richard Feynman, Robert Leighton, and Matthew Sands, Addison Wesley 1963.

REF58.) *Practical Implementation of Charge Recovering Asymptotically Zero Power CMOS*, by Saed G. Younis and Tom Knight, AI Laboratory, MIT, preprint.

REF59.) *Adiabatic Switching, Low Energy Computing, and the Physics of Storing and Erasing Information*, by J. G. Koller and W. C. Athas, Proceedings of the Workshop on Physics and Computation, October 2-4 1992, page 267-270, IEEE Press.

REF60.) *Towards Practical Reversible Logic*, by Ralph C. Merkle, Proceedings of the Workshop on Physics and Computation, October 2-4 1992, page 227-228, IEEE Press.

REF61.) *Quantum Theory*, by David Bohm, Prentice Hall 1951.

REF62.) *A Framework for Practical Low-Power Digital CMOS Systems Using Adiabatic Switching Principles*, by W.C. Athas, L. J. Svensson, J.G. Koller, N. Tzartzanis and E. Chou, Proceedings of the 1994 International Workshop on Low Power Design, April 1994, page 189.

REF63.) *Scattering Suppression and High-Mobility Effect of Size-Quantized Electrons in Ultrafine Semiconductor Wire Structures*, by Hiroyuki Sakaki, Japanese Journal of Applied Physics, Vol. 19 No. 12, December 1980, pages L735-L738.

REF64.) *Quantum Mechanical Computers*, by Richard P. Feynman, Optics News, February 1985, pages 11-20.

REF65.) *Self Replicating Systems and Low Cost Manufacturing*, by Ralph C. Merkle, in: *The Ultimate Limits of Fabrication and Measurement*, M.E. Welland, J.K. Gimzewski, eds.; Kluwer, Dordrecht, 1994.

REF66.) *Numerical Recipes in FORTRAN: the art of scientific computing,* second ed. by Press WH, Teukolsky SA, Vetterling WT, and Flannery BP. 1992, Cambridge University Press

REF67.) Jeffrey Soreff, private communication.

REF68.) *Intermolecular and surface forces,* second ed. by Jacob Israelachvili 1992, Academic Press