Towards a Quantum Computer Project

Peter L. Hagelstein, Norm Margolus and Mike Biafore

Massachusetts Institute of Technology
Research Laboratory of Electronics
Cambridge, Massachusetts 02139

ABSTRACT

P. Shor[1] has recently developed a new factoring algorithm that has the potential to factor large numbers efficiently on a quantum computer. No quantum computers currently exist, and it is not known how to build a practical system that would operate as a quantum computer for other than the most trivial cases. Shor's work has stimulated interest in the area of quantum computers, and we have generated some new ideas that may have the potential to lead to an optical implementation.

The basic idea is to use optical pulses in a nonlinear medium to carry the quantum "bits," and try to design a system that can superpose, switch and set relative phases while maintaining coherence during a calculation. It may be that this will prove to be impossible, but the consideration of the problem has led to some new ideas that appear to be quite interesting.

New ideas that have been generated so far are: (1) the possibility that optical pulses in a nonlinear medium may exhibit quantum interference effects (if true, this would be perhaps the most dramatic example of a macroscopic quantum effect to date); (2) the development of a quantum radiation theory in first quantization that is capable of predicting quantum interference effects in ideal nonlinear media; (3) the conceptualization and initial design of an optical system that would generate a superposition of two states for an optical pulse; (4) the conceptualization of an optical quantum switch; and (5) thoughts about what a quantum computer might look like if the individual parts could be demonstrated to work.

1 Introduction

The possibility of quantum computers was first considered in the early 1980s by Feynman and by Benioff. A quantum computer might be thought of as a quantum system that performs a calculation based on a given input, and then uses the results of that calculation to set the relative quantum phase. A quantum computer can be initialized into a superposition of quantum states, with each quantum state corresponding to a different input; the end result in this case would be a superposition of final states, one corresponding to each of the initial states, with a phase associated with each state.

The outcome of the computation is determined through a measurement performed at the end of the computation. The result of the measurement is probabilistic; the probability for any individual outcome can be determined in principle by summing the individual quantum probability amplitudes and then squaring the result. As envisioned, a quantum computer would reach a different final result depending on which input it received. Placing it in a large superposition of initial states is in some ways akin to using it as a parallel machine, although in the quantum case only one computer pathway is implemented to carry out a very large number of these quantum calculations in parallel. This is in fact what gives a quantum computer the potential to have such interesting behaviour.

Shor's algorithm takes advantage of this property of a quantum computer. The algorithm begins by initializing the computer into a superposition of all possible input states with equal phase. In the case of factoring large numbers, this might be 2^{300} initial states. In the algorithm, the relative phases are adjusted depending on various algebraic products in such a way that strong phase addition occurs for inputs that correspond ultimately to the correct answer at the end of the computation, and extreme cancellation occurs for all inputs that result in other answers. Such an algorithm would not be feasible in principle on a classical computer.

On the microscopic scale, quantum interference effects are well known, and such effects have been observed with photons, electrons, and neutrons. Our proposed device uses macroscopic signals: such an operating quantum computer would be perhaps the first example of a truly macroscopic object exhibiting quantum interference.

Arguing against the possibility of observing macroscopic interference effects is the extreme difficulty of maintaining the phase of a macroscopic system. Any process that perturbs the relative macroscopic quantum phase by even $\pi/2$ on average during the entire computation over all of the bits destroys the quantum interference required for the operation of Shor's algorithm. The general approach that we describe here has the advantage of using nonlinear optics related techniques to manipulate light pulses as quantum objects, but whether it can lead to a solution of the phase problem is unknown. There exist no examples of systems of any sort that work this way; the successful development of a quantum computer would be most notable by far for its solution to this problem.

o ideally, demos macro pulse coherence o practically, within possiblities of materials? o pulse manipulations lead to phase correlations between individual photons.

In light of this discussion, we would ideally like to demonstrate that macroscopic interference of light pulses can be observed. However, constraints imposed by material properties would probably make this impossible to demonstrate experimentally. The demonstration of macroscopic interference is not, however, a prerequisite for construction of a quantum computer; it is only interference between single photons that is ultimately required. As a practical matter, we can simply use the interaction of pulses to establish correlations between single photons, which can then be demonstrated directly.

2 Quantum Computation and Shor's Algorithm

The operation of a quantum computer may be described mathematically at a very elementary level from the quantum postulates. The computer is initialized into a superposition of states with potentially arbitrary amplitudes and phases; in Shor's algorithm the phases are matched and amplitudes are equally weighted

$$\Psi_i = \frac{1}{p-1} \sum_{a=0}^{p-2} \sum_{b=0}^{p-2} |a, b, 0> \tag{2.1}$$

Shor's algorithm is expressed in terms of states with three indices; these states are of the form $|a, b, z>$. It will be convenient for the purposes of discussion here to instead use a vector notation

$$|a, b, z> \longrightarrow \Phi_{\mathbf{k}} \tag{2.2}$$

where we define the vector $\mathbf{k}$ to be

$$\mathbf{k} = \mathbf{i}_x a + \mathbf{i}_y b + \mathbf{i}_z z \tag{2.3}$$

The initial state for Shor's algorithm is then

$$\Psi_i = \sum_{\mathbf{k}} c_{\mathbf{k}} \Phi_{\mathbf{k}} \tag{2.4}$$

where

$$c_{\mathbf{k}} = \frac{1}{\sqrt{N_\perp}} \delta_{0 k_z} \tag{2.5}$$

where $N_\perp$ is the number of states $(p-1)^2$ which have $k_z = 0$.

Computations are then carried out, which result in a change in the state of the computer. This may be represented as

$$\Psi' = e^{\hat{S}} \Psi_i \tag{2.6}$$

In Shor's algorithm, this is a unitary transformation that produces a single $\Phi_{\mathbf{k}'}$ from each initial $\Phi_{\mathbf{k}}$. Consequently, we have

$$\Psi' = \sum_{\mathbf{k}} c_{\mathbf{k}} \Phi_{\mathbf{k}'} \tag{2.7}$$

where $\mathbf{k}'$ is a function of $\mathbf{k}$.

Shor's algorithm might be thought of as a quantum version of sorting algorithm. A subset of all poosible states are input, and through interference effects, only states that correspond to the answer are kept. Consequently, the key to this algorithm is the application of a highly nontrivial unitary operator that does this. We denote this operation as

$$\Psi'' = e^{\hat{\mathcal{A}}} \Psi' \tag{2.8}$$

which implements the unitary operation

$$\Psi'' = \sum_{\mathbf{k}} \sum_{\mathbf{k}''} c_{\mathbf{k}} d_{\mathbf{k}''} e^{2\pi i(\mathbf{k}'_\perp \cdot \mathbf{k}''_\perp)/(p-1)} \Phi_{\mathbf{k}''} \tag{2.9}$$

with

$$d_{\mathbf{k}''} = \frac{1}{\sqrt{N_\perp}} \delta_{k'_z, k''_z} \tag{2.10}$$

A measurement is then performed on this state. The detection scheme implements a measurement of the diagonal operator $\hat{\kappa}$, which has the property

$$\hat{\kappa}\Phi_{\mathbf{k}} = \mathbf{k}\Phi_{\mathbf{k}} \tag{2.11}$$

The expectation value of this operator is

$$< \hat{\kappa} >= \sum_{\mathbf{k}} p_{\mathbf{k}} \mathbf{k} \tag{2.12}$$

where the probability $p_{\mathbf{k}}$ is

$$p_{\mathbf{k}} = |\sum_{\eta} c_\eta d_{\mathbf{k}} e^{2\pi i(\mathbf{k}_\perp \cdot \eta'_\perp)/(p-1)}|^2 \tag{2.13}$$

In light of equation (2.10), here $d_{\mathbf{k}} = \frac{1}{\sqrt{N_\perp}} \delta_{\eta'_z, k_z}$. Shor has arranged the computation so that the probability $p_{\mathbf{k}}$ is strongly maximized for values of $\mathbf{k}$ that correspond to the desired result.

The unitary transformation prepares the state of the quantum computer so that a diagonal operator may be used for the measurement. It would also be possible to include the unitary operator in the definition of the measurement. For example, consider the operator

$$\hat{\mathbf{Q}} = e^{\hat{\mathcal{A}}^\dagger} \hat{\kappa} e^{\hat{\mathcal{A}}} \tag{2.14}$$

A measurement performed on this operator using Ψ' is equivalent to a measurement on $\hat{\kappa}$ using Ψ''

$$< \Psi'|\hat{\mathbf{Q}}|\Psi' >=< \Psi''|\hat{\kappa}|\Psi'' >= \sum_{\mathbf{k}} p_{\mathbf{k}} \mathbf{k} \tag{2.15}$$

The use of this identity implies that there may be more than one architecture that could be used to implement the basic algorithm. It is not yet obvious that this leads to a reduction in the size or complexity of the quantum computer.

It seems clear that Shor's algorithm is an example of a more general class of algorithm that takes advantage of interesting unitary transformations. The simplest immediate genearlization might be through generalizing $\hat{\mathcal{A}}$ to give

$$e^{\hat{\mathcal{A}}} \Phi_{\mathbf{k}} = \sum_{\mathbf{k}'} d_{\mathbf{k}'} e^{i\theta(\mathbf{k},\mathbf{k}')} \Phi_{\mathbf{k}'} \tag{2.16}$$

In Shor's algorithm, the phase function is quadratic, and selected to add phases for one set of $\mathbf{k}$ while randomizing phases for all other $\mathbf{k}$. This functionality is quite general, and we may expect to see other algorithms of this type in the future.

3 Towards Using Nonlinear Optics

Implementations of quantum computers that have been proposed to date are based on accessing individual electronic or nuclear spins, or working with single photons. The difficulties *a priori* associated with either approach are extreme, and it is not clear to those working in the field that a quantum computer will ever become a reality.

In discussions at LCS, we pursued the question of whether nonlinear optics could be used to advantage on this problem. For example, although it is difficult (if not impossible) to switch with single photons, perhaps schemes could be arranged where a few photons might be present so that the chances of obtaining a nonlinear response might improve. The major problem with using more than one photon for switching is that many photons do not generally produce the type of interference effects that a single quantum object would, and some modification either in the quantum algorithm or in the nonlinear optics scheme is required.

For example, if a single photon could be used to represent a bit, perhaps by propagating along one path or another, then at the end of the computation an ideal measurement will find the photon in only one of the two paths. If two photons were used for one bit, then a measurement of both photons at the end of the computation in one path or the other can be interpreted as a bit. Yet there exists the possibility that one photon will be found in each path, which does not have a clear interpretation in this scheme. The use of a sufficiently large number of photons required for nonlinear effects to occur will in turn result in the generation of an enormous multiplicity of states; whether algorithms can be generated to survive such fractionation of quantum states is unknown (but unlikely).

At some point, the question came up as to whether it was possible to have many photons behave as a single quantum entity. This is roughly equivalent to suggesting that one could have one's proverbial cake and eat it too. Although there exist no examples of this in nonlinear optics, it is not obvious that this cannot be done in principle. What would be required to implement such an approach would be new optical elements that would operate on pulses, never allowing any fractionation of states.

This question led to rather intense discussions during the past several weeks. It seems to be obvious that a group of photons can in fact exhibit nontrivial collective quantum effects, and be manipulated as a quantum object without a necessary fractionation of states. This is discussed in Section 5.

4 The Phase Problem

One motivation for considering using photons for an implementation of a quantum computer is that it is relatively straightforward to produce optical interference effects. Conventional optical interference is fundamentally a quantum effect involving the interference of single photon probability amplitudes. The quantum interference proposed as the basis of quantum computation is very closely related, and would be identical in the limit that the computation involved a single photon. A possible advantage of using photons for quantum computers is that optical systems can be constructed with quite small phase errors, and extremely complicated interference patterns have been generated. In some ways, quantum computation might be thought of as constructing very complicated interference patterns involving correlations between many-particle systems.

We can pursue these ideas further, by considering a hypothetical quantum computer that is capable of manipulating single photons. To make the example as simple as possible, we consider a two-bit quantum computer that manipulates two photons to give one of four possible outcomes as determined by a measurement performed at the end of the computation. We assume two possible paths for each photon, and assign arbitrarily 0 to one path and 1 to the other for each pair of paths; a binary indexing scheme for the two photon states could then be developed to yields states $j = 1$ through $j = 4$ for states Φ_{00} through Φ_{11}.

The initial state required by Shor's algorithm is

$$\Psi_i = \frac{1}{2}\left[\Phi_1 + \Phi_2 + \Phi_3 + \Phi_4\right] \tag{4.1}$$

After propagating through the optical system of the quantum computer, the final state is produced

$$\Psi_f = c_1\Phi_1 + c_2\Phi_2 + c_3\Phi_3 + c_4\Phi_4 \tag{4.2}$$

A measurement is then performed to determine the answer. In this case, the measurement process would correspond to a matrix operator

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \\ A_{41} & A_{42} & A_{43} & A_{44} \end{bmatrix} \tag{4.3}$$

This matrix may in general be complex, but it must be Hermitian since it corresponds to an observable quantity.

The measurement in this case would be an example of a hypothetical two-photon correlation measurement. The different states that we have used might be taken to correspond to products of electric fields: $\Phi_1 \sim E_0^{(1)}E_0^{(2)}$, $\Phi_2 \sim E_1^{(1)}E_0^{(2)}$, $\Phi_3 \sim E_0^{(1)}E_1^{(2)}$, and $\Phi_4 \sim E_1^{(1)}E_1^{(2)}$. In this idealized example, we might imagine using a two photon converter to generate second harmonic fields that correspond to the four states, and then interfere appropriate combinations of these fields to generate interference patterns appropriate to the measurement A.

The initial states we have considered here (equations (4.1) and (4.2)) are in a binary representation (built up of 1-photon states); the states corresponding to $E_0^{(1)}E_0^{(2)}$ etc., imply a change in representation to 2-photon states. The proposed implementation of Shor's unitary transformation discussed later (in Section 6) will employ a change of basis that corresponds to an n-photon representation, implemented using logic.

Errors in the relative phase affect the measurement. For example, if associated with each individual photon there occurs a random phase error $\Delta\phi$, for an n-photon state j the total phase error would be on the order of

$$\Delta\theta_j \sim \sqrt{n}\Delta\phi \tag{4.4}$$

It would be reasonable to expect that a moderately complicated optical system could be developed that could maintain relative phase coherence for single photons to a relatively small fraction of a wave, consistent with the general requirements under consideration here. This is one of the primary reasons for considering an optical implementation initially. Unfortunately, it would be quite difficult to develop nonlinear optics switching capability that would work with single photons.

While the switching problem may be solved using photon pulses and nonlinear optics, the phase problem is tremendously exacerbated for pulses. A pulse that contains N photons and that acts as a single quantum object has a very large momentum and correspondingly small wavelength. The phase errors induced in the probability amplitude for the center of mass of the pulse will be on the order of $\sqrt{N}\Delta\phi$. The total phase error that would arise for an n-pulse state would then be

$$\Delta\theta_j \sim \sqrt{Nn}\Delta\phi \tag{4.5}$$

For pulses containing on the order of 10^9 photons, this translates into requirements on the quality of the optical system that may not be realistic.

From this discussion, it is clear that another option is required if the general approach is to have any chance of working. It may be that a hybrid scheme might offer a solution to this problem. We know that to switch we require a large number of photons; we also know that interference will be easiest to produce based on combinations of single photon correlations. These two requirements may not be incompatible. For example, if all of the switching is done with pulses in a nonlinear medium, and if a final state is generated that is suitable for measurements involving interference, then if the last step is carried out in a linear medium the resulting interference will be based on simple products of single photon correlations.

5 A Radiation Theory for Pulses in a Nonlinear Medium

We have proposed that pulses propagating in a nonlinear medium can exhibit quantum interference effects, and have the potential to lead to an implementation of the part of a quantum computer where quantum switching occurs (as outlined in Section 4). The description of quantum interference of pulses is highly nontrivial in second quantization, and a first quantized theory did not exist, at least at the beginning of our discussions. A first quantized theory is readily obtained by analogy with the Dirac theory in first quantization. The associated mathematical details are described briefly in the Appendix; here we summarize the basic issues.

Neutrons and protons, and also photons, may be described in a second quantized theory. Both matter and light diffract as individual quanta. In the case of neutrons and protons, attractive forces can exist to establish particle-particle correlations, and collections of neutrons and protons are known to diffract as nuclei. The description of this is convenient in first quantization, but not in second quantization.

In first quantization for particles, the first quantization description describes the position and momentum of a given set of particles. Interacting particles in free space are described by separable Hamiltonians of the form

$$H = \sum_i H_i + \sum_{i,j} V_{i,j} \tag{5.1}$$

Many-particle wavefunctions in this case are separable into center of mass and relative mass coordinates

$$\Psi(\mathbf{r}_1, \cdots, \mathbf{r}_N) = \psi(\mathbf{R})\Phi(\xi_1, \cdots, \xi_N) \tag{5.2}$$

Typically, the Hamiltonian for the center of mass coordinate describes a non-interacting particle that propagates freely. In the case of nuclei, there might be 100 nucleons, with very complicated strong force interactions arising from pion exchange. Yet the center of mass Hamiltonian for most applications can be taken to be

$$H_{\mathbf{R}} = -\frac{\hbar^2 \nabla^2}{2M} \tag{5.3}$$

An equivalent picture for photons does not appear in the literature.

In the case of light, propagation in a nonlinear medium under certain conditions corresponds to propagation with attractive interactions between photons, and photons are known to coalesce into localized pulses (for example, in the case of solitons, and also in self-focusing). The description of the quantum diffraction of such objects is technically difficult in second quantization.

A many-photon Hamiltonian in first quantization has been obtained that is suitable for our purposes. The Hamiltonian is of the form

$$H = \sum_i \alpha_i \cdot c\mathbf{p}_i + \sum_{i,j} v_{i,j} \tag{5.4}$$

This Hamiltonian is separable in the case of a uniform isotropic medium, and wavefunctions may be separated into center of mass and relative mass parts

$$\Psi(\mathbf{r}_1, \cdots, \mathbf{r}_N) = \psi(\mathbf{R})\Phi(\xi_1, \cdots, \xi_N) \tag{5.5}$$

The wavefunction for the center of mass satisfies equations that are very much like Maxwell's equations; they are in fact high spin generalizations of Maxwell's equations, and in the case of an inhomgeneous optical medium can be written as

$$\frac{\partial}{\partial t}\phi = -\frac{1}{N}(i\Sigma) \cdot c\nabla\chi \tag{5.6}$$

$$\frac{\partial}{\partial t}\chi = \frac{1}{N}(i\Sigma) \cdot c\nabla\phi + (i\Sigma) \cdot (c\nabla \ln c)\phi \tag{5.7}$$

where ϕ and χ are very much like electric and magnetic fields and where N is the number of photons in the pulse. The operator Σ is an N-photon spin operator. Whereas $\mathbf{E}$ and $\mathbf{H}$ have three components apiece, since they are vectors, the ϕ and χ fields have $3N$ components each. The spin operators are large $3N \times 3N$ matrices. Although complicated, in the case where all photons are linearly polarized, these equations can be reduced down to a scalar wave equation very much like

$$\frac{\partial^2}{\partial t^2}\mathbf{E} = -c^2\nabla \times \nabla \times \mathbf{E} \tag{5.8}$$

that is easy to work with.

The center of mass propagates at essentially the speed of light in the medium, however the frequency and wavelength of the ϕ and χ fields will be those of a very high energy "photon" (this is sensible as the wavefunction describes an N-photon system which will have N times the energy of a single photon). The equivalent dielectric constant for confinement or steering that is seen by these fields is the sum of the dielectric constants seen by each photon individually.

6 The Architecture of a Quantum Computer

We are considering the physical realizability of a computer which makes essential use of quantum superposition and interference. The major difficulty in this endeavor will be achieving simultaneously all of the following:

- creating quantum signals
- creating superpositions of large numbers of (global) computational states
- causing correlated superpositions of signals to interact computationally
- maintaining (global) superpositions during the computation (i.e., avoiding decoherence)
- avoiding and/or correcting phase and amplitude errors during the computation
- using interference effects to enhance the amplitudes of desired computations

Any proposed architecture must address all of these issues. We will discuss these problems in the order that the computing elements would be used in a quantum computer: first we must generate signals, then create superpositions of signals, then do ordinary reversible logic on these superpositions (without recombining elements of the superposition), then we must generate new superpositions from each computational state, allow all of these results of the computation to interfere, and finally we measure a result.

6.1 Creating signals

Because of our need for long coherence lengths, we are considering an optical implementation, using non-linear optics. The presence or absence of photons at designated places and times will be our binary variables. To achieve nonlinear effects, we will need to use not single photons, but rather, large packets of photons. These packets will be large enough, and our photon losses small enough, that we will not need to regenerate our signals in the course of our computation. In appropriate non-linear media, our photon packets act like single particles.

Our computation will begin with the creation of the signals that will be used throughout. In a machine capable of implementing Shor's algorithm for numbers of an interesting size, we would need perhaps 300 photon packets to be produced. Since the speed of the packets in a non-linear medium will depend on the number of photons that comprise each packet, we need to create photon number states, all containing about the same number of photons. To achieve reliable interactions, and to allow phase to be used later for interference purposes, the packets will need to start off at the same time. A phased array of mode-locked lasers is proposed as our source of synchronous pulses. Attention will have to be given to reducing the variance in photon number of the pulses.

6.2 Creating superpositions

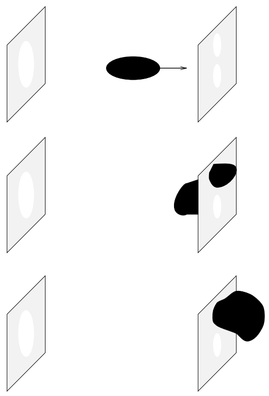

Figure 1: If each many-photon pulse incident on the pair of apertures maintains its integrity, it will have amplitudes to pass through each aperture, but no amplitude to divide itself between the two apertures.

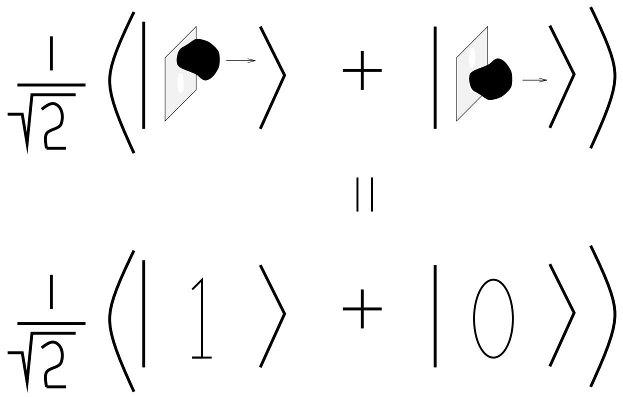

Figure 2: If a superposition of spatially-separated pulses can be produced, they can be used to represent the quantum superposition of binary states 1 and 0.

No one has ever created superpositions of large packets of photons. Our proposal here is to use non-linear optics to put photon solitons into spatial superpositions: if measured, the whole soliton is always found in one fiber or another (see Figures 1 and 2). We would interpret the presence of a soliton in one beam as representing a binary one; its presence in the complementary beam would represent a zero.

There currently exists no optical logic element that will generate a superposition of center of mass states for optical pulses, keeping the pulses intact. A conventional beamsplitter, for example, can set up such a superposition for single photons, but would destroy correlations between photons in an incident photonic "nucleus."

Quantum diffraction can be used, in principle, to do the job. For example, a pulse in a nonlinear medium that is apertured will diffract in the center of mass coordinates. Collecting apertures placed some distance away (also in the same nonlinear medium), can lead to spatially isolated many-photon pulses that are anti-correlated (the basic idea is illustrated in Figures 1 and 2).

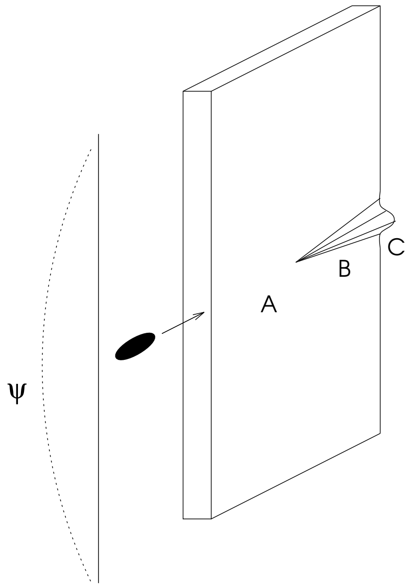

Figure 3: The superposition of spatially-separated pulses may be obtained by projecting the incident pulse onto an effective potential that branches gradually into two separate channels.

How well the idea works can be estimated from scaling arguments. Diffraction in 3-D as illustrated in Figure 1 turns out to be quite difficult, as on the order of 1 MW is required to generate a pulse with 2-D localization. Instead we work with diffraction in 1-D. A pulse propagating a length L in a nonlinear medium diffracts a transverse distance

$$x_d \sim L\Delta\theta \sim \frac{1}{\sqrt{N}}\frac{L\lambda}{d} \tag{6.1}$$

where d is the initial localization of the center of mass. The transverse size of the pulse is roughly

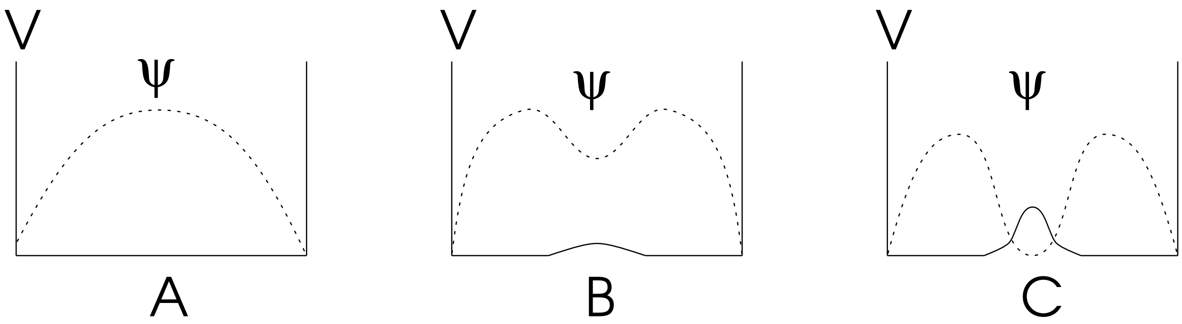

Figure 4: The potential seen by a pulse as it proceeds down the slab in Fig. 3 from A to C.

$$x_t = \frac{3P_c \Delta z W}{N\pi\hbar\omega c} \tag{6.2}$$

where P_c is the critical power of the nonlinear medium (typically 1 MW), where W is the thickness of the medium, and Δz is the length of the pulse. In order to separate pulses, the diffraction width must exceed the transverse size of the pulse. The threshold for this is determined by matching the two expressions, giving the constraint

$$\sqrt{N} = \frac{3P_c \Delta z W d}{\pi\hbar\omega c L\lambda} \tag{6.3}$$

Depending on the choice of parameters, photon numbers on the order of 10^9 seems to correspond to a 10 cm path length. In principle it seems like a plausible approach. If it actually worked, it would likely be the most extreme example of a macroscopic quantum effect to date. This by itself would be an interesting experiment (see Figure 5).

As an implementation of a packet-splitter, this idealized example has the defect that it does not sufficiently reduce the available phase space of the final state. A more practical implementation would switch adiabatically from a single fibre into a pair of fibres.

The kind of logic scheme we're proposing here, where each signal is represented twice—by both a signal and its complement—is known as a *dual rail* representation. It is used here for two reasons. In the first place, it avoids the need to create or destroy packets of photons. But more importantly, if we keep the rails very close together throughout the computation, we may be able to avoid decoherence caused by photon losses. With signals represented by the presence or absence of a packet in an isolated fiber, a photon absorbed by the fiber would reveal the location of the entire packet, and thus destroy the superposition for that signal. If signals are represented by complementary pairs of adjacent rails (with a very small separation for each pair), then a photon lost from a packet could have come from either rail, and thus perhaps we can arrange to only find out that a signal has passed, but not which rail it was in.[1]

If we have c bits of variable inputs into our computer, this would correspond to c input channels. The first step in our computation involves splitting each channel into two rails, so that we have a superposition

[1] If we use interactions based on photon polarization to do our computation, we could use polarization as our binary variable, and this might be even more stable against decoherence by photon loss. Such materials have other practical problems, however.

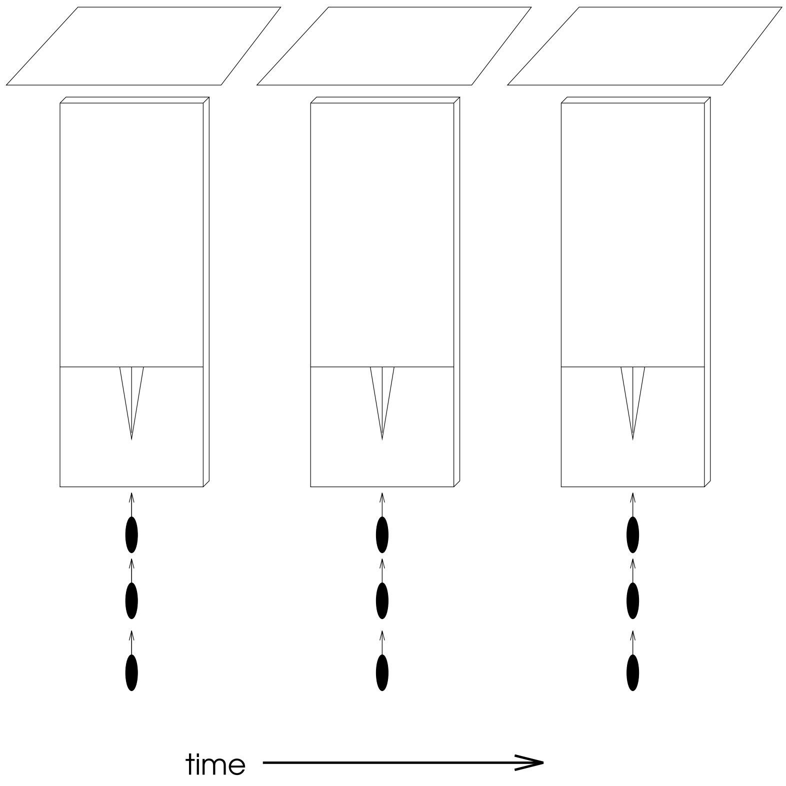

Figure 5: A proposed two-slit experiment to demonstrate that the pulses are exhibiting true quantum interference.

of 2^c input computational states. Note that there would also be some constant inputs to the computer that are needed by the logic of the computation: these would not be split.

One comment on amplitude errors. As is well known, a computer governed by a Schrödinger-equation time evolution is relatively insensitive to small errors in the relative amplitudes of the states making up the initial superposition. This is because the amplitude of each error term in the initial superposition will be exactly the same size as the amplitude of the corresponding error term in the final superposition, since the evolution is linear in this respect. Thus our splitting into pairs of channels doesn't have to be very exact.

6.3 Interaction of signals

A computer is a composition of logic elements. It is necessary both that we have appropriate logic elements available, and that we can interconnect them appropriately. Because the time evolution of our quantum system is unitary (and hence invertible), we must build the computation out of reversible logic. A natural set of reversible logic elements for our use is given by Fredkin's Billiard Ball Model of Computation[2, 3].

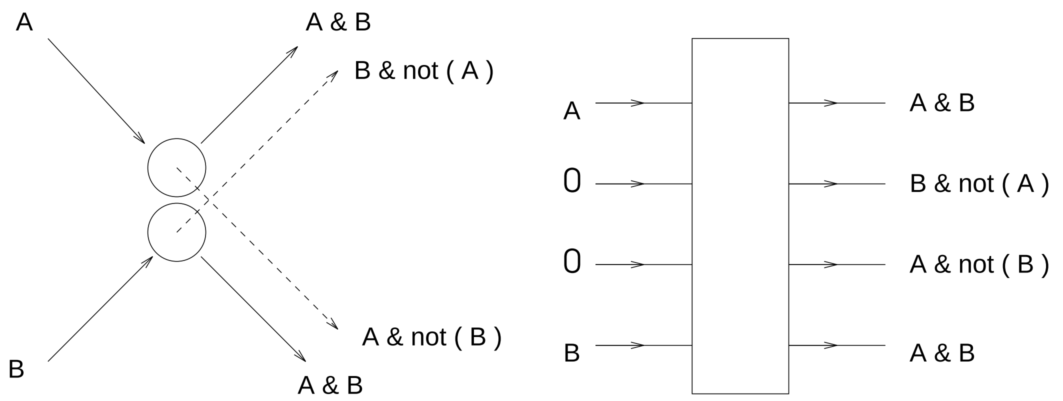

Figure 6: The computation universality of Fredkin's billiard ball model follows from the observation that collisions occuring when two streams of balls cross can be regarded as the operation of a 2-input/4-output logic gate. A ball will be present in each output stream if and only if the indicated logical conditions on the input streams A and B are true.

In the Billiard Ball Model (BBM), logic is performed by the collision of finite diameter billiard balls (Figure 6). The collision resolves the various input cases into distinct output paths—these outputs can then be directed to further collisions, and can eventually produce any desired logical function of a set of binary inputs. Such a collision is about as powerful a logic function as any other universal logic gate normally used for constructing computer logic[2].

In a practical implementation of the BBM, one must worry about keeping signals synchronized so that balls that collide always arrive at collision sites simultaneously; with implementing interactions; and with composing these interactions (i.e., having the outputs of one interaction become inputs to others).

A proposed scheme for implementing an optical billiard ball collision is illustrated in Figure 7. Here an external light source creates an effectively repulsive interaction between our photon packets, since each

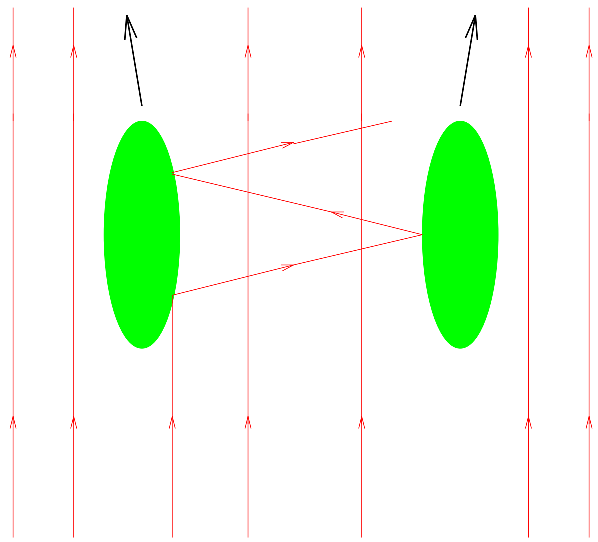

Figure 7: To construct the collision gate of Fig. 6, there must be some effective interation between two optical pulses in the nonlinear medium. A proposed method for inducing a repulsive interaction by an external beam (vertical lines) is depicted schematically here.

scatters some of the light towards the other.

Packets are kept synchronized by virtue of having the same characteristic speed, so they come into collisions well synchronized in time.[2] Our physical implementation of the collision would have only four output paths, so that small errors in the output angle are not possible, as they are with real billiards. This is equivalent to minimizing the amount of phase space available in the output channel.

Note that we have only discussed the interaction of photon packets, not of superpositions of packets. As long as the superpositions persist, their interactions should just be the superposition of the interactions we have discussed here.

6.4 Interference effects

The key to Shor's algorithm is the use of quantum interference effects to supress the results of some computations present in the superposition, and to enhance the results of others. It is only by virtue of this interference that we distinguish a small subset of interest from the very large set of all possible computations present in our superposition, and allow these results to be manifest in the output of our quantum computer.

Figure 8 shows the simplest example of phase-adjustment based on computational state. In this example, we separate out the four possible input cases for the variables A and B into distinct spatial locations, and then insert appropriate phase adjustments (optical delays) into these paths before recombining the signals.

[2] In principle, packets could be forced to move at a uniform speed by using a sort of *particle-accelerator* mechanism. If we periodically modulate the index of refraction of the medium transmitting the signals in space and in time (in a non-dissipative way), we can create a "groove" that the particles would follow.

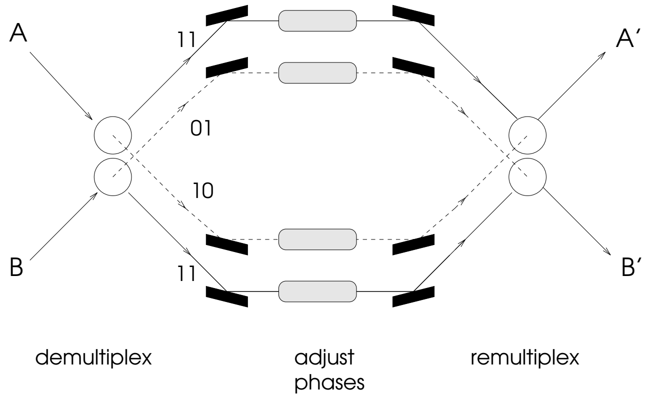

Figure 8: Collision gate is used to spatially demultiplex a 2-bit number into its four cases so that the four phases may in principle be manipulated individually.

Note that once we have encoded information into the phases, we can no longer force the phases to align as we do in the rest of the computation, and so the final analysis step should follow almost immediately.

To implement Shor's algorithm, we not only need to adjust phases, we also need to split computational states into superpositions of states, with appropriate phases. This is the stage in which interference occurs, as the same computational state is generated from several different configurations, each contributing a phase.

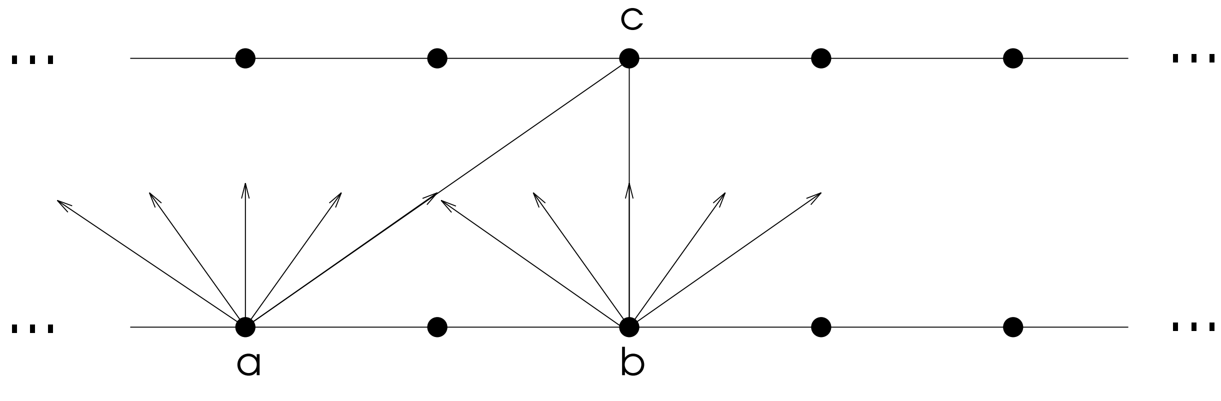

Figure 9: Each point on the line represents a possible configuration of a reversible computation. The bottom line represents the quantum computer in a superposition of such configurations. States are recombined by turning each configuration on the bottom into a superposition of several configurations. On the top we have the resulting state. We show configuration c getting contributions from configurations a and b.

In Figure 9 we illustrate this process. The points at the bottom of the Figure represent possible output configurations produced by the reversible-logic portion of our computer. Each of these distinct output configurations corresponds to a distinct input. Since our initial state was a superposition of input configurations, the reversible-logic produces the corresponding superposition of output configurations. This superposition is represented by the entire bottom line in the Figure. We now split each configuration into a superposition of configurations, with phases that depend on the initial and final configuration in each case. The resulting superposition is represented by the line at the top of the Figure. Each configuration now enters into the new superposition with an amplitude that depends on contributions coming from many configurations on the bottom.

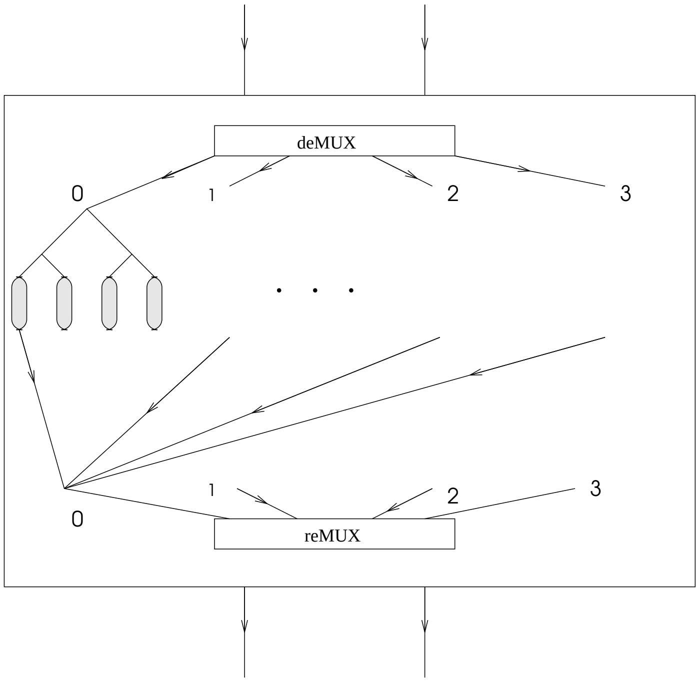

Figure 10: A device for implementing the phase adjustments and state recombination needed for Shor's algorithm. Here we construct an arbitrary unitary transformation on 2 bits in which all matrix elements are of unit magnitude.

To implement this, we take the element of Figure 8, and insert some extra "circuitry" in the middle (Figure 10). After demultiplexing k of our binary variables to produce a photon at one of $q = 2^k$ distinct spatial positions, we use the same *packet-splitter* element that created our initial superposition (see Section 6.2) to split each of these demultiplexed cases into q additional cases. We now simply add in a chosen phase for each of the q^2 cases, and then use the same kinds of elements in reverse to recombine everything back into k bits. Essentially what we've done is take the complete matrix that couples the k inputs to the k outputs, and spatially separate out the cases until each matrix element can be implemented by a phase adjustment element located at a distinct spatial position. This is all that is needed to complete Shor's algorithm.

We'll need several elements of this sort in our computer, but the biggest k that we'll have to deal with in this manner is of order the $\log$ of the number of digits in the numbers we're working with. Once we have finished all the recombining of states, we can measure the bits of our answer. Each measurement of the output bits should yield a consistent result of our reversible computation, and if we control the phases according to Shor's prescription, we will suppress undesired results, and see the results of computations of interest with high probability.

7 Summary

In this document we have considered the possibility of implementing a quantum computer based on the application of nonlinear optics techniques. This discussion is motivated by the recent development by Shor of a new factoring algorithm that makes use of basic quantum mechanical properties of probability amplitudes to achieve its result; the proposed architecture seeks to implement this algorithm.

Maintaining quantum phase coherence is the key requirement that makes the possible development of a quantum computer difficult. The use of photons and high quality optical material minimizes phase errors, but it is impossible to switch with individual photons. Using photon packets allows switching, but it is impossible to maintain phase coherence for a macroscopic wavepacket of photons. The compromise proposed here is to use a hybrid scheme where switching is done with packets and detection is done on single photons.

From our initial considerations, there do not appear to be any obvious difficulties with the scheme that would *a priori* rule out the possibility of the development of such a quantum computer. On the other hand, it is clear that the development of a quantum computer will be a major undertaking that will require a significant theoretical and experimental effort. The potential reward for succeeding is the development of a computational capability that could not be attained through other means even in principle.

So what should be done? On the conceptual and theoretical side, it would make sense to develop optimized designs for quantum devices that generate superpositions and that switch, to develop a formalism that handles the transition between the correlated picture and single photon picture, and to analyze the phase problem in detail. In the area of architecture, there is considerable work to be done in reducing the number and complexity of the quantum circuit elements needed to manipulate the phases. Experimentally, the most important immediate demonstrations that are of interest are: (1) the demonstration that a packet can indeed be put in a superposition of states; (2) the demonstration that single photons derived from a packet that has been in a superposition of states retains sufficient quantum correlation to produce an interference pattern based on the original packet superposition. If these demonstrations are successful, then the door would be open to the experimental demonstration of few bit quantum computers.

Appendix A. First Quantization Theory for Photons

The quantum theory of radiation dates back to 1928 when Dirac used canonical quantization on the electromagentic vector and scalar potentials to quantize Maxwell's equations. This theory was literally born as a second quantized theory. By the late 1940s, quantum radiation theory was successfully coupled to second quantized Dirac theory, producing QED; QED provides the only correct theory describing matter-radiation interactions as far as is currently known.

The starting point for QED is a total Hamiltonian made up of three terms describing the matter field ($\hat{H}_e$), the radiation field ($\hat{H}_r$), and interactions between the two ($\hat{H}_I$).

$$\hat{H} = \hat{H}_e + \hat{H}_r + \hat{H}_I \tag{A.1}$$

It is possible to write the matter and radiation Hamiltonians as summations over plane-wave modes

$$\hat{H}_e = \sum_{\mathbf{k},\sigma} \hbar\omega(\mathbf{k}) \left[\hat{b}^\dagger_{\mathbf{k},\sigma} \hat{b}_{\mathbf{k},\sigma}\right] \tag{A.2}$$

$$\hat{H}_r = \sum_{\mathbf{k},\sigma} \hbar\omega(\mathbf{k}) \left[\hat{a}^\dagger_{\mathbf{k},\sigma} \hat{a}_{\mathbf{k},\sigma}\right] \tag{A.3}$$

where $\hat{b}$ are fermionic operators and where $\hat{a}$ are bosonic operators. The zero-point of the radiation field is suppressed here for pedagogical reasons.

In the case of electrons and positrons, it is a common practice to derive a first quantization theory from the second quantized picture. In fact, essentially all modenr atomic and molecular orbital computations have their basis in such a theory. The first quantized theory is written for one electron as

$$i\hbar\frac{\partial}{\partial t}\psi = \left[\alpha \cdot c\mathbf{p} + \beta mc^2 + V(\mathbf{r})\right]\psi \tag{A.4}$$

The Dirac α matrices can be writte in terms of spin 1/2 Pauli matrices

$$\alpha = \begin{bmatrix} 0 & \sigma \\ \sigma & 0 \end{bmatrix} \tag{A.5}$$

where σ are the Pauli spin matrices

$$\sigma = \mathbf{i}_x\sigma_x + \mathbf{i}_y\sigma_y + \mathbf{i}_z\sigma_z \tag{A.6}$$

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{A.7}$$

In the case of many-electron states, the Hamiltonian is

$$H = \sum_j \alpha_j \cdot \mathbf{p}_j + \sum_j \beta_j mc^2 + + \sum_j V(\mathbf{r}_j) + \sum_{j,j'} \Lambda_{++} V(\mathbf{r}_j, \mathbf{r}_{j'})\Lambda_{++} \tag{A.8}$$

where the projection operators Λ_{++} projects out negative energy states.

No such equivalent theory has been described in the literature for photons. On the one hand, the interactions of photons with matter leads to the creation and destruction of photons, so that it is difficult to find situations where a first quantization description is really justified. On the other hand, significant questions have been raised about what a photon is, and what it means to speak of the position of a photon. In the case of nonlinear optics, propagation in a Kerr medium incurs relatively low losses, so that the use of a Hilbert space description as a first order approximation might be justified. Pulse propagation in low loss fibers seems to be one of the most promising areas for the application of a first quantized radiation theory. We will not address the philosophical questions concerning the position of a photon here, noting that the reduction of the Fock space wavefunction into Hilbert space components can result in the appearance of coordinate space parameters that are equivalent to photon positions.

Maxwell's equations in an inhomogeneous isotropic medium can be written as

$$i\hbar\frac{\partial}{\partial t}\psi = \left[\alpha\cdot c(\mathbf{r})\mathbf{p} + V(\mathbf{r})\right]\psi \tag{A.9}$$

where

$$\psi = \begin{bmatrix} D_x/\sqrt{\epsilon_0} \\ D_y/\sqrt{\epsilon_0} \\ D_z/\sqrt{\epsilon_0} \\ iB_x/\sqrt{\mu_0} \\ iB_y/\sqrt{\mu_0} \\ iB_z/\sqrt{\mu_0} \end{bmatrix} \tag{A.10}$$

The Maxwell α matrices can be written in terms of spin 1 matrices

$$\alpha = \begin{bmatrix} 0 & \sigma \\ \sigma & 0 \end{bmatrix} \tag{A.11}$$

where

$$\sigma_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix} \quad \sigma_y = \begin{bmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{bmatrix} \quad \sigma_z = \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{A.12}$$

The potential $V(\mathbf{r})$ in this case is

$$V(\mathbf{r}) = \begin{bmatrix} 0 & \sigma\cdot(\mathbf{p}\ln\frac{1}{\sqrt{\mu}}) \\ \sigma\cdot(\mathbf{p}\ln\frac{1}{\sqrt{\epsilon}}) & 0 \end{bmatrix} \tag{A.13}$$

In this picture, Maxwell's equations might be thought of as a Schrodinger equation for a six-component wavefunction. The 1 photon Hamiltonian looks like a Dirac equation for zero mass, with a spin 1 Maxwell version of the α matrices replacing the spin 1/2 Dirac matrices. As a Schrodinger equation, the number of possible solutions is now too many. Two solutions are physical, corresponding to positive energy fields with transverse polarization. Two solutions correspond to negative energy transverse fields; these are not physical; and two solutions correspond to longitudinal polarization, which are not allowed by Gauss' law.

The Dirac Hamiltonian has the same kind of problem; of four possible solutions, two correspond to negative energy states, and must be avoided. The projection operators Λ_{++} in fact serves to restrict the Hilbert space to only positive energy solutions.

The generalization to a many-photon Hamiltonian is straightforward. In an ideal lossless nonlinear medium, the first order nonlinearity corresponds to the appearance of photon-photon interaction terms in the Hamiltonian. We may write

$$H = \sum_j \alpha_j \cdot c(\mathbf{r}_j)\mathbf{p}_j + \sum_j V(\mathbf{r}_j) + \sum_{j,j'} \Lambda_{++} V(\mathbf{r}_j, \mathbf{r}_{j'}) \Lambda_{++} \tag{A.14}$$

In a Kerr medium, the response of the medium is local in the applied fields, and so the interaction potential here is short range

$$V(\mathbf{r}_j, \mathbf{r}_{j'}) \sim \delta(\mathbf{r}_j - \mathbf{r}_{j'}) \tag{A.15}$$

The resulting physics is in many ways similar to the problem of neutrons and protons in nuclei. The forces between neutrons and protons can be attractive and are of short range. Individually, neutrons and protons diffract; collectively as nuclei, they also diffract. The description of this effect is based on a separation between center of mass and relative mass coordinates. The same coordinate transformation may be applied to the photon case

$$\mathbf{R} = \frac{1}{N} \sum_j \mathbf{r}_j \tag{A.16}$$

$$\xi_j = \mathbf{r}_j - \mathbf{R}_j \tag{A.17}$$

The linearized many-photon Hamiltonian becomes

$$H = \frac{1}{N} \alpha \cdot \mathbf{P} + \sum_j (\alpha_j - \alpha_N) \cdot \pi_j + V_{\mathbf{R}}(\mathbf{R}) + \sum_j V_\xi(\xi_j) + \sum_{j,j'} \Lambda_{++} V_{j,j'} \Lambda_{++} \tag{A.18}$$

where

$$\alpha = \sum_j \alpha_j \tag{A.19}$$

This Hamiltonian is separable, leading to separable solutions of the symbolic form

$$\Psi = \psi(\mathbf{R}) \Phi(\{\xi\}) \tag{A.19}$$

The center of mass motion separates relatively cleanly from the internal structure of a photonic "nucleus".

The Schrodinger equation for the center of mass motion becomes

$$i\hbar \frac{\partial}{\partial t} \psi = \left[\alpha \cdot c\mathbf{P} + V_{\mathbf{R}} \right] \psi \tag{A.20}$$

This is of the general for of Maxwell's equations, but is a high-spin generalization. It can be written to look more like Faraday's and Ampere's law using

$$\psi = \left[\begin{array}{c} \phi \\ i\chi \end{array} \right] \tag{A.21}$$

to give

$$\frac{\partial}{\partial t}\phi = -\frac{1}{N}(i\Sigma) \cdot c\nabla\chi \tag{A.22}$$

$$\frac{\partial}{\partial t}\chi = \frac{1}{N}(i\Sigma) \cdot c\nabla\phi + (i\Sigma) \cdot (c\nabla \ln c)\phi \tag{A.23}$$

for a nonlinear optical material (taking μ to be a constant). The analog large component wavefunction ϕ here corresponds to electric fields (actually to $\mathbf{D}/\epsilon$), and is a column vector of length $3N$. The analog small component χ corresponds to magnetic fields ($\mathbf{B}/\mu$), and is also of length $3N$. We may eliminate χ to obtain a second order high spin wave equation

$$\frac{\partial^2}{\partial t^2}\phi = -\left[\frac{1}{N}(i\Sigma) \cdot c\nabla\right] \left[\frac{1}{N}(i\Sigma) \cdot c\nabla\phi + (i\Sigma) \cdot (c\nabla \ln c)\right] \phi \tag{A.24}$$

This is the analog of the vector wave equation

$$\frac{\partial^2}{\partial t^2}\mathbf{E} = -c^2\nabla \times \nabla \times \mathbf{E} \tag{A.25}$$

The computation of solutions to the center of mass wave equation is nontrivial, since the high spin operators cause coupling between the different polarizations. If it occurs that one of the polarizations is fixed, either by accident or by design, then the coupling problem can be greatly reduced. For example, if we assume that the electric fields of all of the photons are polarized in the x-direction, then only a single polarization is present; in this case we recover a scalar wave equation

$$\frac{\partial^2}{\partial t^2}\phi = c\frac{\partial}{\partial y}(c\frac{\partial}{\partial y}\phi) + c\frac{\partial}{\partial z}(c\frac{\partial}{\partial z}\phi) + N\left[c\frac{\partial}{\partial y}(c\frac{\partial \ln c}{\partial y}) + c\frac{\partial}{\partial z}(c\frac{\partial \ln c}{\partial z})\right]\phi \tag{A.26}$$

An intuitive interpretation of the meaning of this result is that the center of mass wavefunction of a photon pulse evolves more or less according to Maxwell-like equations, with two important differences. Since the center of mass wavefunction describes the location of N photons, the corresponding energy is on the order of $N\hbar\omega$ and the wavelength is on the order of λ/N, where ω and λ refer to the frequency and wavelength of single photons. Additionally, the effects of the dielectric constant are additive in photon number; the effects of gradients in the optical index are greatly amplified.

References

[1] Shor, P., "Algorithms for Quantum Computation: Discrete Log and Factoring", *Proceedings of the 26th Annual ACM Symposium on the Theory of Computation*, **35**, 1–14 (1994).

[2] Fredkin, E. and Toffoli, T., "Conservative Logic", *International Journal of Theoretical Physics*, **21**, 219–253 (1986).

[3] Toffoli, T. and Margolus, N., *Cellular Automata Machines*, MIT Press, Cambridge (1987).