Kinematical bounds on evolution and optimization of mixed quantum states

M. D. Girardeau,[1] M. Ina,[1] S. Schirmer,[2] and T. Gulsrud[3]
[1]*Department of Physics and Institutes of Theoretical Science and Chemical Physics, University of Oregon, Eugene, Oregon 97403*
[2]*Department of Mathematics, University of Oregon, Eugene, Oregon 97403*
[3]*Department of Physics, University of Oregon, Eugene, Oregon 97403*

(Received 20 May 1996)

Upper and lower bounds are established for time-dependent ensemble averages of observables of driven quantum systems in mixed states. They limit controllability of observables independently of the control fields. Narrower bounds are established when the observable is a projector onto a pure quantum state or subspace. They are optimal in the sense of being kinematically achievable. Calculations on nonlinear optimal control of a four-level model indicate that these kinematical bounds are dynamically achievable asymptotically with increasing control pulse strength. [S1050-2947(97)51203-0]

PACS number(s): 03.65.Bz, 05.30.−d, 31.70.Hq

Kinematical constraints on quantum systems. The structure of Hilbert space imposes nontrivial constraints on the temporal variation of observable properties of a driven quantum system in a mixed (statistical) state. Being purely kinematical, these constraints limit the extent to which the system can be controlled by application of time-dependent external fields such as laser pulses, a topic of widespread current interest.

Statistical evolution. We start from the definition of the time-dependent quantum statistical average $\langle A(t)\rangle$ of any observable $\hat{A}$:

$$\langle A(t)\rangle=\sum_{\alpha} w_{\alpha}\langle\Psi_{\alpha}(t)|\hat{A}|\Psi_{\alpha}(t)\rangle=\mathrm{Tr}\,[\hat{A}\hat{\rho}(t)]. \quad (1)$$

Here the $|\Psi_{\alpha}(t)\rangle$ are solutions of the time-dependent Schrödinger equation satisfying initial conditions $|\Psi_{\alpha}(t_i)\rangle=|\alpha\rangle$ at some initial time t_i and the w_{α} are statistical weights for these possible initial states, satisfying $w_{\alpha}\geqslant 0$ and $\Sigma_{\alpha}w_{\alpha}=\mathrm{Tr}\hat{\rho}(t)=1$. The initial states $|\alpha\rangle$ are the complete orthonormal set of eigenstates of the initial statistical operator $\hat{\rho}(t_i)$. Its spectral representation for $t\geqslant t_i$ is

$$\hat{\rho}(t)=\sum_{\alpha} w_{\alpha}|\Psi_{\alpha}(t)\rangle\langle\Psi_{\alpha}(t)| \quad (2)$$

and the w_{α} are its time-independent eigenvalues, some of which may be zero. If more than one eigenvalue is nonzero, $\hat{\rho}$ is said to represent a *mixed state*. The state space is assumed to be a Hilbert space (purely discrete) of either finite or infinite dimension. If the dimension is infinite, then $\hat{\rho}$ has no nonzero smallest eigenvalue but instead an infinite sequence of decreasing eigenvalues with limit zero, but a largest eigenvalue exists in all cases. In the applications we have in mind, the $|\alpha\rangle$ are eigenstates of an unperturbed Hamiltonian representing the system in the absence of the driving laser pulse, but the derivation of kinematical bounds is not restricted to that case.

Bounds: The time-evolved states $\{|\Psi_{\alpha}(t)\rangle\}$ are also complete and orthonormal. Expand the observable $\hat{\mathrm{A}}$ in terms of them:

$$\hat{A}=\sum_{\alpha\beta} A_{\alpha\beta}(t)|\Psi_{\alpha}(t)\rangle\langle\Psi_{\beta}(t)|, \quad (3)$$

where $A_{\alpha\beta}(t)=\langle\Psi_{\alpha}(t)|\hat{A}|\Psi_{\beta}(t)\rangle$. Then with (1) we have

$$\langle A(t)\rangle=\sum_{\alpha} w_{\alpha}A_{\alpha\alpha}(t). \quad (4)$$

It is assumed that the operator $\hat{A}$ is of trace class (finite trace). It is also assumed to be positive semidefinite, having a non-negative expectation value in any state; then the same is true of $A_{\alpha\alpha}(t)$. (The case where the $A_{\alpha\alpha}$ are not non-negative, but nevertheless bounded below, can be reduced to this case by a shift of origin of $\hat{A}$.) Denote the greatest lower bound and least upper bound of the w_{α} by $w_{\min}$ and $w_{\max}$, respectively. Then the sum (4) will not be increased if w_{α} is replaced by $w_{\min}$, and it will not be decreased if w_{α} is replaced by $w_{\max}$. It follows that

$$w_{\min}\mathrm{Tr}\hat{A}\leqslant\langle A(t)\rangle\leqslant w_{\max}\mathrm{Tr}\hat{A}. \quad (5)$$

Note that the trace of $\hat{A}$ is time independent, since it is basis invariant and can be evaluated in a time-independent basis. If the state space has infinite dimension then $w_{\min}=0$ and the lower bound is trivial.

Occupations The simplest case is when $\hat{A}$ is the projector onto a specified state $|\Psi\rangle$, assumed to be normalized. Since the trace of a projector onto any pure quantum state is unity, Eq. (5) implies

$$w_{\min}\leqslant\langle A(t)\rangle=\langle\Psi|\hat{\rho}(t)|\Psi\rangle\leqslant w_{\max}. \quad (6)$$

Much of the work on optimal control of quantum systems is directed toward maximizing the population of a specified state at a specified target time t_f by determination of the optimal time variation of the electric field of a laser pulse during a prior time interval $t_i\leqslant t\leqslant t_f$ [1–7]. The upper bound (6) limits the maximum achievable population in mixed-state (statistical) situations, and the lower bound is also of interest since it implies that during the entire evolution from t_i to t_f the populations of all levels remain in the band between $w_{\min}$ and $w_{\max}$, the minimal and maximal initial populations.

(If any of the w_α are equal to zero, then the lower bound is trivial, $w_{\rm min}=0$. This is also the case if the state space has infinite dimension.) The special role of the initial time arises through the fact that the density operator is diagonal in the basis $\{|\alpha\rangle\}$ only at the initial time, equivalent to the usual assumption [8] of random *a priori* phases. An interesting special case is that of an N-level quantum system with a completely uncertain initial state, in the sense that the statistical weights of all members of the ensemble are equal to $1/N$. Then $w_{\rm min}$ and $w_{\rm max}$ are equal and the subsequent level populations are completely uncontrollable, since they remain equal to $1/N$ independently of any control terms in the Hamiltonian. A somewhat similar result on uncontrollability of a system whose initial state has random initial phase was obtained previously by Dahleh *et al.* [9], although the details of their approach were very different. In the following paragraph we will describe the results of calculations confirming the behavior predicted by (6) and strongly suggesting that the upper bound is asymptotically dynamically achievable, in the sense that it can be approached arbitrarily closely by determination of optimal laser pulses $f(t)$ for a sequence of increasing pulse fluences. We found that the bound $w_{\rm max}$ could be saturated to within a fraction of a percent at fluences low enough that our iterative nonlinear optimization algorithm converged. This can be understood as follows: The expressions (1) and (2) show that if a driving term in the Hamiltonian can be determined during the time interval $t_i \leqslant t \leqslant t_f$ so as to force $|\Psi_{\alpha_{\rm max}}(t)\rangle$ to evolve from its specified initial value $|\Psi_{\alpha_{\rm max}}(t_i)\rangle = |\alpha_{\rm max}\rangle$ at the initial time t_i to the specified target state $|\Psi\rangle$ at the target time t_f, then the target state occupation will assume its kinematical maximum value $w_{\rm max}$ at time t_f. Here $|\Psi_{\alpha_{\rm max}}(t)\rangle$ is the particular state in Eqs. (1) and (2) for which w_α assumes its largest value $w_{\rm max}$. The nonlinear optimization procedure [10,11] is carried out for fixed, finite values of the fluence (driving pulse energy), as was the linear approximation of Wilson and coworkers [6]. Such a formulation is appropriate in view of experimental constraints on laser pulse production, and has mathematical advantages since it leads naturally to an eigenvalue-eigenfunction formulation of the problem [4–6,10]. Our calculations show that the kinematical upper bound $w_{\rm max}$ is approached with increasing pulse fluence, indicating that $|\Psi_{\alpha_{\rm max}}(t)\rangle$ can be made to evolve to a state nearly equal to a specified target state $|\Psi\rangle$ at the target time t_f. This is expected on the basis of a generalized pulse area therorem for multilevel systems [12].

Example. We performed mixed-state nonlinear optimization calculations on a simplified model of the lowest four vibrational levels of the HF molecule, defined by the Hamilton $\hat{H} = \hat{H}_0 + f(t)\hat{V}$ with $\hat{H}_0 = \Sigma_{n=1}^4 E_n |n\rangle\langle n|$ and $\hat{V} = \Sigma_{n=1}^3 d_n(|n\rangle\langle n+1| + |n+1\rangle\langle n|)$, where $E_n/\hbar = \omega_0(n-\frac{1}{2})[1-\frac{1}{2}(n-\frac{1}{2})B]$, $d_n = 0.097 n^{1/2}$ Debye, $\omega_0 = 7.80 \times 10^{14}\ {\rm s}^{-1}$, and $B = 0.0419$. These parameters are appropriate to a Morese-oscillator model studied previously [13] as an example of molecular quantum control. We took the observable $\hat{A}$ to be the projector onto the upper level $n=4$. The statistical weights w_α ($\alpha=1,\ldots,4$) of the unperturbed states were a Boltzmann distribution const$\times e^{-E_\alpha/kT}$ at a temperature such that $kT = E_4 - E_1$, implying $w_1 = 0.3850$, $w_2 = 0.2758$, $w_3 = 0.1976$, and $w_4 = 0.1416$. $f(t)$ is proportional to the electric field of the laser pulse. We determined the optimal $f(t)$ for maximization of the final population of level 4 by a nonlinear eigensystem method similar to one described previously [10]. The optimization was carried out for a fixed pulse length $t_f = 0.841$ ps and a sequence of increasing pulse fluences $\epsilon_p = \int_0^{t_f} F^2(t)dt$, where $F(t) = p_{12} f(t)/\hbar$ and p_{12} is the $1\rightarrow 2$ transition dipole moment. For the largest fluence used, $\epsilon_p = 0.02\ {\rm fs}^{-1}$, the final population $A(t_f)\rangle$ of level 4 for the optimal pulse was 0.3842, only 0.26% less than the upper bound $w_{\rm max} = 0.3850$ of Eq. (6), the initial Boltzmann population of the lowest level 1. Figure 1 shows the populations of all four levels vs time for this case. At $t = t_f$ practically the entire contribution to (1) comes from $|\Psi_1(t)\rangle$ (which belongs to $w_1 = w_{\rm max} = 0.3850$), showing that the optimal pulse causes $|\Psi_1(t)\rangle$ to rotate from $|1\rangle$ at $t=0$ to nearly $|4\rangle$ at $t = t_f$. [Recall that here $\hat{A} = |4\rangle\langle 4|$, and note that the $|\Psi_\alpha(t)\rangle$ form an orthonormal set at all t.] The corresponding optimal pulse $f(t)$ is shown in Fig. 2.

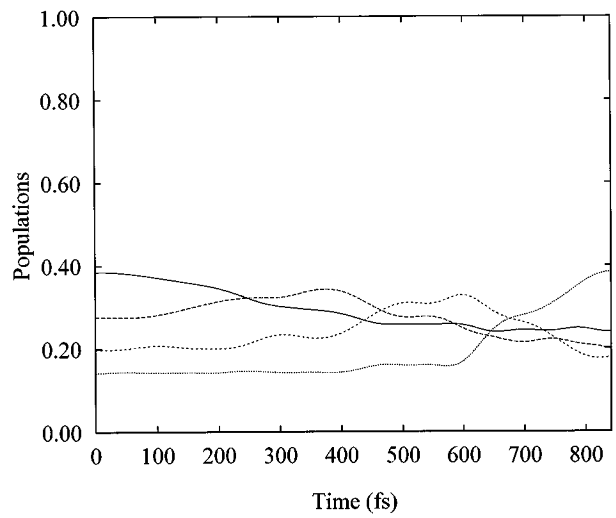

FIG. 1. Populations versus time produced by a pulse optimal for maximizing the population of level 4 at the target time. Pulse fluence was fixed at 0.02 fs^{-1} and initial populations were Boltzmannian. Solid line, ground level 1; long dashes, level 2; short dashed, level 3; dots, upper (target) level 4.

More general observables. In the case of observables more complicated than the projector onto a pure state, it may be possible to improve the bounds (5). Consider the case where $\hat{A}$ is the projector onto a subspace S of states having some physical property that one might wish to optimize. It can be written as

$$\hat{A} = \sum_\lambda |\lambda,S\rangle\langle\lambda,S|, \tag{7}$$

where $\{|\lambda,S\rangle\}$ is any orthonormal (but incomplete) set spanning S. Then ${\rm Tr}\hat{A} = d_s$ where d_s is the dimension of S, so the upper and lower bounds (6) are multiplied by d_s. To improve these bounds, note that

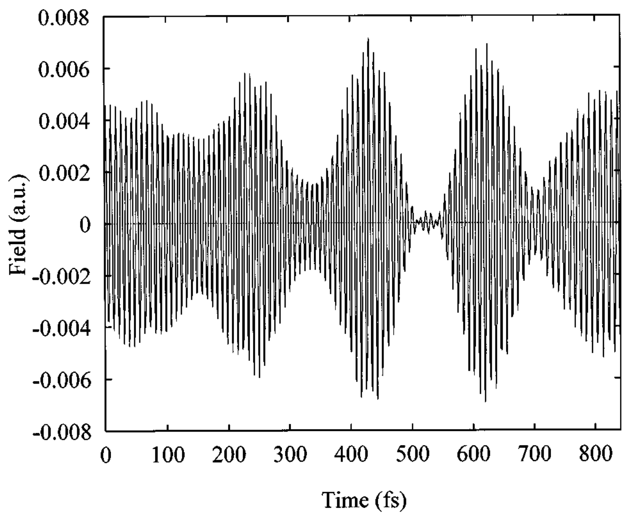

FIG. 2. Optimizing pulse for the case of Fig. 1.

$$\langle A(t)\rangle = \sum_{\alpha} w_{\alpha} x_{\alpha S}, \tag{8}$$

where

$$x_{\alpha S} = \sum_{\lambda} |\langle \Psi_{\alpha}(t)|\lambda, S\rangle|^2. \tag{9}$$

The $x_{\alpha S}$ satisfy $0 \leqslant x_{\alpha S} \leqslant 1$ and the sum rule $\Sigma_{\alpha} x_{\alpha S} = d_S$. Suppose that those $x_{\alpha S}$ multiplying the d_S largest of the w_{α} are unity and all other are zero. Then $\langle A(t)\rangle$ reduces to the sum $w_>(d_S)$ of the d_S largest of the w_{α}. Any decrease in these $x_{\alpha S}$ below unity must be compensated by a corresponding increase of some of the $x_{\alpha S}$ multiplying smaller w_{α}, to maintain satisfaction of the sum rule. Thus $\langle A(t)\rangle \leqslant w_>(d_S)$ for all t. If the state space has finite dimension, then by a similar argument involving the d_S smallest of the w_{α} one concludes that $\langle A(t)\rangle$ cannot be less than the sum of the d_S smallest of the w_{α}, implying

$$w_<(d_S) \leqslant \langle A(t)\rangle \leqslant w_>(d_S), \tag{10}$$

where $w_<(d_S)$ and $w_>(d_S)$ are, respectively, the sums of the d_S smallest and d_S largest of the w_{α}. If the state space has infinite dimension, then there is no nonzero smallest w_{α} and one must define $w_<(d_S)=0$, but the upper bound remains nontrivial so long as $\hat{\rho}$ represents a mixed state. It follows from Eq. (9) that the $x_{\alpha S}$ multiplying a given w_{α} is unity if and only if $|\Psi_{\alpha}(t)\rangle$ lies in S. Since S is of dimension d_S and the $|\Psi_{\alpha}\rangle$ are mutually orthogonal, the upper bound is realized if and only if those associated with the d_S largest of the w_{α} all lie in S. The upper bound (10) is, therefore, optimal in the sense of being kinematically attainable. A similar argument applies to the lower bound in the case in which the state space has finite dimension. However, in what follows it is assumed that the goal is maximization of $\langle A(t_f)\rangle$, in which case the lower bound is not relevant. The question of whether the upper bound is *dynamically* attainable at some target time t_f by suitable driving terms in the Hamiltonian, starting at time $t_i < t_f$ with physically reasonable initial states $|\Psi_{\alpha}(t_i)\rangle = |\alpha\rangle$, remains. In the following paragraph we will describe the results of a calculation confirming the predicted behavior and strongly suggesting that the upper bound is asymptotically dynamically achievable.

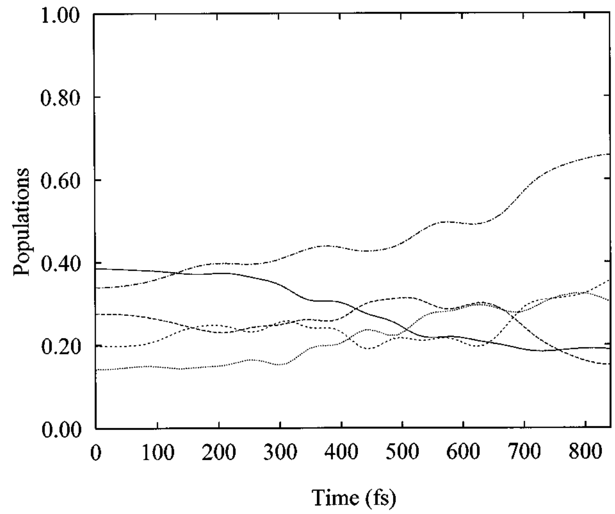

FIG. 3. Populations versus time produced by a pulse optimal for maximizing the sum of populations of levels 3 and 4 at the final time. Pulse fluence was fixed at 0.025 fs^{-1}. Model and initial conditions were the same as for Figs. 1 and 2. Solid line, ground level 1; long dashes, level 2; short dashes, level 3; dots, upper level 4; dash-dots, target population equal to sum of level 3 and 4 populations.

Example. Consider the same molecular vibration model and initial Boltzmann distribution as in the calculations for Figs. 1 and 2, but take the observable to be maximized at time t_f to be the sum of populations of the upper two levels, corresponding to $\hat{A} = |3\rangle\langle 3| + |4\rangle\langle 4|$. The level populations vs time and optimal pulse $f(t)$ are shown in Figs. 3 and 4 for fluence 0.025 fs^{-1}. The final populations of levels 3 and 4 were 0.3553 and 0.3034, summing to $\langle A(t_f)\rangle = 0.6587$, only 0.32% less than the upper bound $w_>(2) = w_1 + w_2 = 0.6608$ of Eq. (10). At $t = t_f$ practically the entire contribution to Eq. (1) comes from $|\Psi_1(t_f)\rangle$ (which belongs to $w_1 = 0.3850$) and $|\Psi_2(t)\rangle$ (which belongs to $w_2 = 0.2758$), and

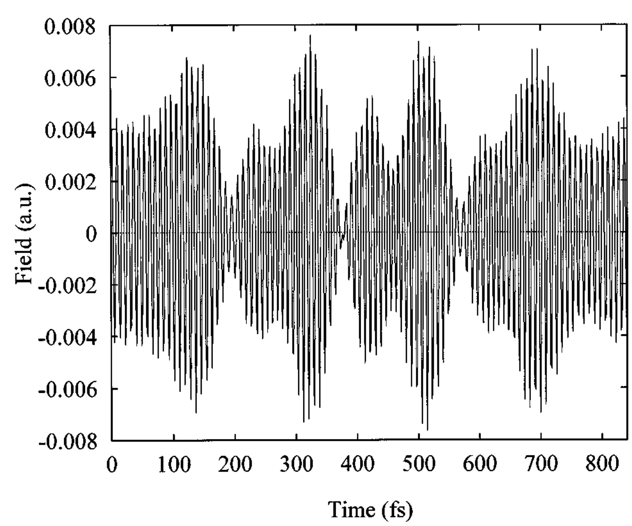

FIG. 4. Optimizing pulse for the case of Fig. 3.

these two states lie almost completely in the target subspace spanned by the unperturbed states $|3\rangle$ and $|4\rangle$. This does not mean that either $|\Psi_1(t_f)\rangle$ or $|\Psi_2(t_f)\rangle$ are close to either of these unperturbed states, but rather that they are close to linear combinations of these two states. The freedom of choice of coefficients in these linear combinations is presumably the reason that a single optimal control field $f(t)$ can simultaneously rotate two orthogonal vectors $|\Psi_1(t)\rangle$ and $|\Psi_2(t)\rangle$ from $|1\rangle$ and $|2\rangle$ at $t=0$ to the subspace spanned by $|3\rangle$ and $|4\rangle$ at $t=t_f$. This can be regarded as a nontrivial generalization of the generalized pulse area theorem [12]. In contrast, in the case of Figs. 1 and 2, which target the pure state $|4\rangle$, the populations of levels 3 and 4 sum to 0.5631 at $t=t_f$, whereas $w_>(2)$ is 0.6608. Thus the field of Fig. 2, optimal for targeting level 4, is nonoptimal for targeting the 3,4 subspace. Similarly, the field of Fig. 4, optimal for targeting the 3,4 subspace, is nonoptimal for targeting levels 3 or 4, both of which have a kinematical upper bound $w_{\max}=w_1=0.3850$ on their populations.

These results confirm the relevance and asymptotic dynamical achievability of the upper bounds on evolution and optimization of mixed states given by Eq. (6) and more generally Eq. (12), at least for four-level systems. Since the kinematical bounds do not depend on the Hamiltonian, they remain valid in the presence of stochastic driving terms and should be relevant to models incorporating fluctuating environmental effects via stochastic differential equations [14–16], in addition to the probabilistic initial conditions implicit in the use of statistical ensembles in Eqs. (1)–(10). The question of whether these kinematical bounds are dynamically achievable by suitable controls even in the presence of stochastic driving terms is at present open and should be investigated.

Note added. Recently we have become aware of a very general group-theoretical approach of Hioe and Eberly [17] leading to a generalized coherence vector implying constraints similar to Eq. (10). In view of the similar conclusions, our approach is probably closely related, and detailed exploration of the connection is likely to be fruitful.

One of us (M.D.G.) thanks R. M. Mazo, J. Leahy, N. Deshpande, and R. L. Arnowitt for helpful discussions and suggestions. The computations were performed in the University of Oregon's Computational Science Laboratory on SGI Power Challenge and Onyx systems, whose purchase was funded through NSF Grant No. CDA-9413532 and a corporate donation from Silicon Graphics, Inc.

[1] W. S. Warren, H. Rabitz, and M. Dahleh, Science **259**, 1581 (1993).
[2] P. Brumer and M. Shapiro, Annu. Rev. Phys. Chem. **43**, 257 (1992).
[3] A. P. Peirce, M. A. Dahleh, and H. Rabitz, Phys. Rev. A **37**, 4950 (1988).
[4] S. H. Tersigni, P. Gaspard, and S. A. Rice, J. Chem. Phys. **93**, 1670 (1990).
[5] A. Bartana, R. Kosloff, and D. J. Tannor, J. Chem. Phys. **99**, 196 (1993).
[6] Y. J. Yan, R. E. Gillilan, R. M. Whitnell, K. R. Wilson, and S. Mukamel, J. Phys. Chem. **97**, 2320 (1993).
[7] H. Shen, J. P. Dussault, and A. D. Bandrauk, Chem. Phys. Lett. **29**, 498 (1994).
[8] R. C. Tolman, *The Principles of Statistical Mechanics* (Oxford University Press, Oxford, England, 1938), Sec. 84, pp. 349ff.
[9] M. Dahleh, A. P. Peirce, and H. Rabitz, Phys. Rev. A **42**, 1065 (1990).
[10] M. D. Girardeau and K. G. Kim, Phys. Rev. A **52**, R891 (1995); **53**, 630(E) (1996).
[11] M. Ina, Ph.D. dissertation, University of Oregon, 1996 (unpublished).
[12] M. Holthaus and B. Just, Phys. Rev. A. **49**, 1950 (1994).
[13] S. Chelkowski, A. Bandrauk, and P. B. Corkum, Phys. Rev. Lett. **65**, 2355 (1990).
[14] Z. Schuss, *Theory and Applications of Stochastic Differential Equations* (J. Wiley, New York, 1980).
[15] V. S. Pugachev and I. N. Sinitsyn, *Stochastic Differential Systems* (J. Wiley, New York, 1987).
[16] Regarding subtleties and controversies of interpretations and usage of stochastic differential equations see N. G. van Kampen, J. Stat. Phys. **24**, 175 (1981).
[17] F. T. Hioe and J. H. Eberly, Phys. Rev. Lett. **47**, 838 (1981).