Hamiltonian equations for multiple-collective-variable theories of nonlinear Klein-Gordon equations: A projection-operator approach

R. Boesch, P. Stancioff, and C. R. Willis
Department of Physics, Boston University, Boston, Massachusetts 02215
(Received 18 April 1988)

We consider classical Hamiltonian systems in which there exist collective modes where the motion associated with each collective mode is describable by a collective coordinate. The formalism we develop is applicable to both continuous and discrete systems where the aim is to investigate the dynamics of kink or solitonlike solutions to nonlinear Klein-Gordon equations which arise in field theory and condensed-matter theory. We present a new calculational procedure for obtaining the equations of motion for the collective coordinates and coupled fields based on Dirac's treatment of constrained Hamiltonian systems. The virtue of this new (projection-operator) procedure is the ease with which the equations of motion for the collective variables and coupled fields are derived relative to the amount of work needed to calculate them from the Dirac brackets directly. Introducing collective coordinates as dynamical variables into a system enlarges the phase space accessible to the possible trajectories describing the system's evolution. This introduces extra solutions to the new equations of motion which do not satisfy the original equations of motion. It is therefore necessary to introduce constraints in order to conserve the number of degrees of freedom of the original system. We show that the constraints have the effect of projecting out the motion in the enlarged phase space onto the appropriate submanifold corresponding to the available phase space of the original system. We show that the Dirac bracket accomplishes this projection, and we give an explicit formula for this projection operator. We use the Dirac brackets to construct a family of canonical transformations to the system of new coordinates (which contains the collective variables) and to construct a Hamiltonian in this new system of variables. We show the equations of motion that are derived through the lengthy Dirac bracket prescription are obtainable through the simple projection-operator procedure. We provide examples that illustrate the ease of this projection-operator method for the single- and multiple-collective-variable cases. We also discuss advantages of particular forms of the *Ansatz* used for introducing the collective variables into the original system.

I. INTRODUCTION

The purpose of this paper is to present a complete, exact solution of the problem of the introduction of classical collective variables into nonlinear Klein-Gordon kink equations[1–3] such as the sine-Gordon[3–6] (SG), double sine-Gordon[7–10] (DSG), and ϕ^4 equations[11–13] and obtain a complete Hamiltonian theory for such systems. In the collective-variable approach, one introduces particle-like parameters such as the center of mass of the kink X, or the separation R between the subkinks of a DSG kink, and considers them as Hamiltonian dynamical variables: $X \rightarrow X(t)$, $R \rightarrow R(t)$. The introduction of N additional dynamical variables requires the addition of $2N$ constraints—one for each added dynamical variable and one for its canonically conjugate momentum—in order that the number of degrees of freedom the original problem is conserved. The nature of the possible constraints is quite general, but we will show that a particular form of the constraints naturally suggests itself which has the convenient (additional) feature of requiring much less calculation to obtain the equations of motion than other forms. We show that the theory of Dirac brackets[14] for constrained dynamical systems applies directly to our collective-variable treatment of nonlinear Klein-Gordon kink equations. We use the Dirac bracket formalism to derive the transformation from an unconstrained set of Hamiltonian variables to set a constrained variables. For the particular form of the constraints, the results of the extremely lengthy calculation of the Dirac bracket equations of motion are equivalent to the equations of motion obtained by simply applying projection operators to the original nonlinear Klein-Gordon equations. The explicit form of the projection operators are derived in the proof of the equivalence. We stress that the amount of calculation required by our projection operator approach to obtain the desired equations of motion for the collective variables and accompanying radiation field is an extremely small fraction of the amount of calculation required when using the Dirac brackets.

By introducing a collective variable, such as the center of mass, $X(t)$, one attempts to separate out the localized collective phenomena of the system from the nonlocalized (propagating) phenomena to which the collective variable is coupled in an attempt to render the problem easier to understand physically. Thus, in general, one aims to arrive at an equation of motion that governs the time evolution of the desired collective variable in the presence of the other collective variables and fields. The resulting equations, one for each collective variable, and one for each field, are a set of nonlinear, coupled, differential equations which are usually not analytically solvable because of their complexity. In many problems, however, there is only one propagating field weakly coupled to the collective variables and so it is possible to linearize the field equations and obtain equations for the propagating field whose source depends on the nonlinear motion of the collective modes.

Quantization of such a system, where the collective variables appear explicitly in the theory, is then possible using the Dirac bracket formalism. This is illustrated by Tomboulis[1] where he introduces the center-of-mass variable into a general nonlinear wave equation that admits soliton solutions and uses the Dirac bracket formalism to transform the singular Lagrangian system into a Hamiltonian system in order to quantize the field. We will not deal with quantization in this paper. Flesch *et al.*[3] have used the center-of-mass variable, $X(t)$, in condensed-matter physics for the case of translational invariance that has been broken by the presence of an impurity. Other authors[2–6] have used $X(t)$ as a collective variable in the case where the translational invariance is replaced by the discrete periodicity of a lattice. In addition to $X(t)$ there are other examples of collective variables such as $R(t)$, the separation between the subkinks in a DSG kink;[7–10] the slope $l(t)$ in a ϕ^4 kink;[11–13] particles $X_\alpha(t)$ in the SG kink crystal[15] and the separation $X_r(t)=X_2(t)-X_1(t)$ between two colliding DSG kinks,[16] to name a few. In all of the above cases the important consequence of the collective variable approach is that situations arise where highly nonlinear motion of the collective "particlelike" variable persists with little or no radiation from the propagating (phonon, for example) field. In these situations, the radiation can be computed with linear field equations, as stated earlier, whose sources are determined by the nonlinear motion of the collective variables.

Another class of problems where the collective variable approach has proved useful is the statistical mechanical formulation of kinks interacting with heat baths.[17,18] In these problems the original nonlinear field interacting with a bath is replaced by a "particle" variable which undergoes Brownian motion with a relaxation time determined by the interaction of the kink with the phonon bath. The collective variable approach for kinks in interaction with a heat bath is also important for those magnetic problems where kinks play an important role.

We divide Sec. II into two parts. In Sec. II A we derive the transformation to the collective variable equations for the case of a single collective variable and illustrate the simplicity of the projection operator approach. In Sec. II B we consider the multiple-collective-variable case. This entails focusing on the fundamental nature of the Dirac bracket which we explicitly show to be a projection in symplectic space. Following the procedure in Sec. II A, we derive differential equations that must be satisfied by the original coordinates and momenta as functions of the new coordinates (which are the collective variables and the radiation field) and their conjugate momenta, given the constraints. We solve these equations to arrive at the canonical transformation; then derive the Hamiltonian and consequently, the equations of motion for the collective variables and radiation field. The main point of Sec. II B is then to show that these equations of motion can be very simply obtained by operating on the original equations of motion with projection operators that emerge naturally from the structure of the Dirac bracket and the form of the constraints. In an effort to make things more clear, we point out now that there are *two* projection operators that we will encounter both of which are derived in this paper. The first is the projection operator in symplectic space which we will show depends only on the structure of the Dirac bracket. The second is the projection operator we use to operate on the original nonlinear Klein-Gordon equation under consideration in order to directly obtain the Dirac bracket equations of motion.

We consider the DSG system in Sec. III where we introduce two collective variables into the field, $X(t)$ and $R(t)$ discussed earlier, in order to illustrate the ease with which the equations of motion for the collective variables and radiation field may be derived using the projection operator formalism when there is more than one collective variable present. We incorporate relativity into the problem and discuss some of the problems that arise in doing so. We also expand the potential (the term nonlinear in the radiation field) as a power series in the radiation field and rewrite the resulting equations in a convenient form that explicitly shows their linear and nonlinear contributions from the radiation field. We further discuss the connection between the equivalence of the Dirac bracket equations of motion and those obtained by the projection operator method.

Introducing collective variables to dynamically parametrize a field requires one to formulate an *Ansatz*, for example, $\phi=\sigma+\chi$. Here ϕ is the original field which has been broken up into two functions σ and χ. σ is a function that best represents the shape of the kink and depends on the collective variables. χ may or may not depend on the collective variables. Much of the literature[1–5] postulates σ as the continuum static solution to the original nonlinear field equation. χ then represents the radiation coupled to the kink and any meson dressing for the kink required by the particular conditions in the problem. We show in Sec. IV that one can choose the function σ such that for small oscillation the function χ vanishes. In fact, in a future paper[6] we show that exact agreement is obtained between molecular dynamics and the theory of the present paper for the problem of the small oscillation Peierls-Nabarro frequency of a trapped SG kink. Previous *Ansätze* have led to results for the small oscillation frequency that were off by a factor of 2. Section V contains a discussion and conclusion.

II. COLLECTIVE-VARIABLE FORMALISM

A. Single-collective variable

We are studying systems with kink or solitonlike solutions with one collective mode describable by a single collective coordinate such as X, the center of mass for a SG kink; or l, the length of a ϕ^4 kink in the kink's rest frame (corresponding to the kink's internal mode). The derivation done in this section is for a discrete system but generalizes straightforwardly to a continuum model.

We start with a system of unit masses harmonically coupled to nearest neighbors and subjected to a substrate described by a potential $V_S(Q_i)$. For simplicity the chain is infinite. The Lagrangian and Hamiltonian functions for this system are

$$L=\tfrac{1}{2}\sum_i \dot{Q}_i^2-\tfrac{1}{2}\sum_i (Q_{i+1}-Q_i)^2-\sum_i V_S(Q_i) \qquad (2.1)$$

and

$$H=\tfrac{1}{2}\sum_i P_i^2+\tfrac{1}{2}\sum_i (Q_{i+1}-Q_i)^2+\sum_i V_S(Q_i)\ , \qquad (2.2)$$

where $P_i=\partial L/\partial\dot{Q}_i=\dot{Q}_i$ is the momentum conjugate to $\dot{Q}_i$. Throughout the paper we use dimensionless units.

We are interested in systems which have an approximate static kink solution $Q_i\simeq f_i(X)$ where the parameter X is the position of the center of the kink. In order to obtain exact static solutions or complete dynamic solutions for this system, we introduce new variables, q_i, such that

$$Q_i=f_i(X)+q_i(t)\ . \qquad (2.3)$$

Note here $q_i(t)$ does not depend on the collective variable X. (See discussion after Eq. (2.61) for the case when q_i depends on X.) We note that since the sum f_i+q_i is to be an exact solution of the general dynamical problem, f_i need not be an exact static solution since the presence of the q_i will account for any dressing of the kink or any radiation coupled to the kink's motion. We will discuss more about the choice of f_i in Sec. IV.

We want to treat the position of the center of the kink, X, as a dynamical variable, $X\to X(t)$. This necessitates performing a canonical transformation from what we will call the "old" variables, Q_i, P_i to a "new" set of variables, q_i, p_i, X, P where p_i and P are the momenta conjugate to q_i and X, respectively. However, by introducing the dynamical variables X and P into the system, we have increased the original number of degrees of freedom by two. Therefore, in order to conserve the original number of degrees of freedom we introduce two constraints:

$$C_1=\sum_i f_i'(X)q_i\approx 0\ , \qquad (2.4)$$

$$C_2=\sum_i f_i'(X)p_i\approx 0\ , \qquad (2.5)$$

where a prime denotes differentiation with respect to the argument. The "$\approx$" sign denotes "weak equality." In Dirac's terminology, a quantity which is weakly equal to zero cannot be set to zero until all variations of the quantity with respect to the dynamical variables, to obtain the equations of motion, have been performed. The motivation for the form of the constraints is to minimize the q_l in the vicinity of the kink.

At this point we do not know the form of P_l, the old momentum, in terms of the new variables. It is not arbitrary, as we require the transformation to be canonical. For the constrained system, the requirement of canonicity can be expressed by requiring invariance of the modified Poisson brackets (i.e., the Dirac brackets) under the transformation which will determine the form of P_l in terms of the new variables. We now clarify precisely what we mean by this.

Since the system is constrained when expressed in terms of the new variables, we cannot simply utilize the Poisson brackets but must follow Dirac's prescription and invoke the modified brackets. In the new system of variables the modified bracket of two functions A and B of the variables q_l, p_l, X, and P is

$$\{A,B\}^*=\{A,B\}-\sum_{i,j}\{A,C_i\}C_{ij}^{-1}\{C_j,B\}\ , \qquad (2.6)$$

where the sum on i,j is over all the constraints; the elements C_{ij} of the matrix C are defined to be the Poisson brackets of the constraints $C_{ij}\equiv\{C_i,C_j\}$ and where we define the notation $C_{ij}^{-1}\equiv(C^{-1})_{ij}$. The Poisson brackets (without the asterisk) in Eq. (2.6) are defined by

$$\{A,B\}_{(qpXP)}=\sum_l\left[\frac{\partial A}{\partial q_l}\frac{\partial B}{\partial p_l}-\frac{\partial B}{\partial q_l}\frac{\partial A}{\partial p_l}\right]+\frac{\partial A}{\partial X}\frac{\partial A}{\partial P}-\frac{\partial B}{\partial X}\frac{\partial A}{\partial P}$$

for the discrete system. Note that in the definition of the Poisson bracket, $\{A,B\}$, the summation is over *all* of the dynamical variables, including the collective ones. For the continuous system, we have the analogous quantities with sums replaced by integrals. Invariance under the transformation thus requires that any functions A and B of the positions and momenta satisfy

$$\{A,B\}_{(Q_i,P_i)}=\{A,B\}^*_{(q_i,p_i,X,P)}\ .$$

In particular, if A and B are the Q_l's and P_l's, this requires that

$$\{Q_l,P_n\}^*=\delta_{ln}$$

and (2.7)

$$\{Q_l,Q_n\}^*=\{P_l,P_n\}^*=0\ .$$

We will make no assumptions about the form of the dependence of P_l on the original variables. We assume only that $P_l=P_l(p_i,q_i,X,P)$, and, as we shall show below, the form of P_l is determined by requiring that the canonical brackets, Eqs. (2.7), be satisfied. We now proceed to evaluate the canonical bracket $\{Q_l,P_n\}^*$ using Eq. (2.6). We calculate the elements C_{ij}^{-1} of the matrix C^{-1}, and substitute into Eq. (2.6) to obtain

$$\{Q_l,P_n\}^*=\{Q_l,P_n\}-\frac{1}{M}(\{Q_l,C_2\}\{C_1,P_n\}-\{Q_l,C_1\}\{C_2,P_n\})\ , \qquad (2.8)$$

where $M\equiv\{C_1,C_2\}=\sum_i f_i'^2$ plays the role of the kink mass.

Evaluating the Poisson brackets in Eq. (2.8), and requiring that $\{Q_l,P_n\}^*$ *in the new variables* be equal to the Poisson bracket $\{Q_l,P_n\}$ *in the original variables* (i.e.,

$\{Q_l,P_n\}=\delta_{ln}$), we find that Eq. (2.8) reduces to

$$\sum_s \left[\delta_{ls}-\frac{f_l'f_s'}{M}\right]\frac{\partial P_n}{\partial p_s}+f_l'\frac{\partial P_n}{\partial P}(1-\eta/M)=\delta_{ln}\ , \tag{2.9}$$

where we have defined $\eta=\sum_i f_i''(X)q_i$.

We solve Eq. (2.9) by first multiplying it by f_l' and summing over l to obtain

$$\frac{\partial P_n}{\partial P}=\frac{f_n'}{M(1-\eta/M)}\ . \tag{2.10}$$

Then, substituting Eq. (2.10) into (2.9) we obtain

$$\frac{\partial P_n}{\partial p_l}=\delta_{nl}+f_l'd_n\ , \tag{2.11}$$

where d_n must obey certain conditions given in Appendix A when we require that $\{P_l,P_n\}^*$ vanish. The general formula for the momentum transformation is obtained by integrating Eqs. (2.10) and (2.11). The desired general result is

$$P_n=p_n+\frac{Pf_n'}{M(1-\eta/M)}+h_n(q_i,X)\ . \tag{2.12}$$

The function h_l must also obey certain conditions when we require that $\{P_l,P_n\}^*$ vanish. We have not looked in detail at the functions h_l and d_l but we show near the end of Appendix A that the conditions on h_l and d_l are satisfied if we set $h_l=d_l=0$. That Eq. (2.12) is independent of d_l is not a consequence of choosing $d_l=0$ but is a general result. We also find in Appendix A, that the remaining bracket $\{Q_l,Q_n\}^*=0$ is identically satisfied and, since it is independent of P_l, imposes no conditions on P_l.

We have proved, therefore, that by assuming the form of the constraints, C_i, and the coordinate transformation for the positions of the particles, Q_l, and requiring the transformation to be canonical (thereby assuring that Hamilton's equations in the new system reduce to Hamilton's equation in the old system), the form of the old momentum in terms of the new variables is completely determined up to the function h_l. We show in Appendix D that *if* the momentum is given by some form other than Eq. (2.12), then one of the assumptions in the previous sentence has to be modified.

We now obtain the Hamiltonian by substituting the new variables, Eq. (2.3) for Q_l and Eq. (2.12) for P_n (with $h_n=0$), into the old Hamiltonian Eq. (2.2):

$$H=\frac{P^2}{2\overline{M}}+\tfrac{1}{2}\sum_i p_i^2+V\ , \tag{2.13a}$$

where the dressed mass is given by $\overline{M}\equiv M(1-\eta/M)^2$. See the discussion after Eq. (2.47) for the reason why the cross terms in the new variable momenta do not appear in the Hamiltonian. In Eq. (2.13a) the potential V includes the harmonic coupling term as well as the substrate potential, i.e.,

$$V\equiv\sum_i \tfrac{1}{2}(f_{i+1}+q_{i+1}-f_i-q_i)^2+\sum_i V_S(f_i+q_i)\ . \tag{2.13b}$$

The equations of motion are derived by taking the Dirac bracket of desired quantities with the Hamiltonian, i.e.,

$$\dot{\alpha}=\{\alpha,H\}^*=\{\alpha,H\}-\sum_{i,j}\{\alpha,C_i\}C_{ij}^{-1}\{C_j,H\}\ . \tag{2.14}$$

The Dirac bracket equations of motion are derived for the general N collective variable case in Appendix B. For the single-collective-variable case they are

$$\dot{q}_l=p_l-\frac{\eta\dot{X}}{M}f_l'\ , \tag{2.15a}$$

$$\dot{p}_l=-\sum_n(\delta_{ln}-\mathcal{P}_{ln})\frac{\partial H}{\partial q_n}-\frac{\dot{X}f_l'}{M}\sum_i f_i''p_i\ , \tag{2.15b}$$

$$\dot{X}=\frac{P}{\overline{M}}\ , \tag{2.15c}$$

$$\dot{P}=-\frac{\partial H}{\partial X}+\frac{\eta}{M}\sum_i f_i'\frac{\partial H}{\partial q_i}\ , \tag{2.15d}$$

where we have defined the projection operator $\mathcal{P}_{ln}$ to be

$$\mathcal{P}_{ln}=\frac{f_l'f_n'}{M}\ . \tag{2.16a}$$

Eliminating the momenta from these equations to get second-order equations for the position variables is a fairly tedious procedure. The results are

$$\sum_n(\delta_{ln}-\mathcal{P}_{ln})\left[\ddot{q}_n+f_n''\dot{X}^2+\frac{\partial V}{\partial q_n}\right]=0 \tag{2.16b}$$

and

$$\ddot{X}+\frac{1}{M}\sum_n f_n'\left[\ddot{q}_n+\dot{X}^2f_n''+\frac{\partial V}{\partial q_n}\right]=0\ . \tag{2.16c}$$

It should be emphasized here that one of the main goals of our formalism, developed in detail below, is to show that the second-order equations (2.16b) and (2.16c) can be obtained without having to explicitly find the canonical transformation from the old variables to the new and therefore without having to construct a Hamiltonian in terms of the new variables. The first-order equations of motion (2.15a)–(2.15d) therefore never need to be derived in order to find the equivalent second-order equations of motion (2.16b) and (2.16c). We only need to know the *Ansatz* function f_l in Eq. (2.3), and the projection operator, $\mathcal{P}_{ln}$, defined in Eq. (2.16a) in terms of the *Ansatz* function f_l. We also require the constraints to be of the form of Eqs. (2.4) and (2.5). Then Eqs. (2.16b) and (2.16c) may be obtained directly. To be explicit, note that Eq. (2.16b) is obtained by substituting the *Ansatz*, Eq. (2.3), into the original equations of motion, $\ddot{Q}_n+\partial V(Q_n)/\partial Q_n=0$, operating with $\sum_n(\delta_{ln}-\mathcal{P}_{ln})$ and using $\sum_n(\delta_{ln}-\mathcal{P}_{ln})f_n'=0$. Likewise, Eq. (2.16c) is obtained by substituting the *Ansatz* Eq. (2.3) into the original equations of motion, operating with $\sum_n\mathcal{P}_{ln}$ and using $\sum_n\mathcal{P}_{ln}f_n'=f_l'$. This is truly a drastic reduction in the amount of work needed to be done in order to obtain the Dirac bracket equations of motion. The full derivation of the equivalence of this simple procedure to the Dirac bracket method is given below.

B. Multiple-collective coordinates

In problems such as the double sine-Gordon kink or the multikink solutions of the SG system, it is necessary to define more than one collective variable. In the following, we formulate the problem of multiple-collective variables along the lines of the preceding derivation.

The derivation below proceeds in several steps, some of which parallel the single-collective-variable derivation. First, the multivariable problem is cast into symplectic notation.[19] In this notation the Poisson bracket and equations of motion take on a relatively simple form. We then show that, in the symplectic notation, the modified Poisson bracket, or the Dirac bracket, can be written using a projection operator defined in symplectic space (which is different from the projection operator $\mathcal{P}_{ln}$ we have just encountered at the end of Sec. II A). The projection operator in symplectic space then simplifies considerably when the constraint equations are linear (but the form of the constraints are chosen for reasons other than just to simplify the projection operator in symplectic space).

We then proceed to derive the canonical transformation to the new variables. First the form of the momentum is calculated in much the same procedure as the calculation of Eq. (2.12) of Sec. II A. The Hamiltonian is then expressed in terms of the new variables and the equations of motion are calculated using the Dirac brackets.

We conclude the section by providing the statement made at the end of Sec. II A, that the equations of motion for the new set of variables are, in general, easily obtained by, first, substituting the new set of variables directly into the original equations of motion rather than going through the lengthy and tedious procedure of the canonical transformation. Then, the correct equations of motion for the collective variables as well as the new "phonon" variables q_l are derived from this substitution by appropriate application of the projection operator $\mathcal{P}_{ln}$ as described at the end of Sec. II A (but appropriately generalized to the multiple-collective-variable case).

We start, as before, with an infinite chain of particles on a substrate, but now we investigate solutions containing for example, multiple kinks, kinks with internal modes, or both. We thus have a system with N collective coordinates. We define the collective coordinates X_i and the new particle coordinates q_l by the ansatz

$$Q_l = f_l(X_1(t), X_2(t), \ldots, X_N(t)) + q_l(t) \ . \tag{2.17}$$

We designate P_l to be the momentum conjugate to Q_l and call Q_l and P_l the *old variables.* The equations of motion for the old coordinates, derivable from the original Lagrangian

$$L = \tfrac{1}{2}\dot{Q}_l\dot{Q}_l - V(Q_l) \ , \tag{2.18}$$

are

$$\ddot{Q}_l + V'(Q_l) = 0 \ , \tag{2.19}$$

where *the summation over repeated indices is implied for the rest of the paper.* As in Sec. II A, $V(Q_l)$ includes the harmonic coupling potential as well as the substrate potential. There are no constraints associated with the Lagrangian in Eq. (2.18); consequently $P_l = \partial L/\partial \dot{Q}_l = \dot{Q}_l$. The first time derivative of Q_l is

$$\dot{Q}_l = \frac{\partial f_l}{\partial X_i}\dot{X}_i + \dot{q}_l = f_{l,i}\dot{X}_i + \dot{q}_l \ . \tag{2.20}$$

In our notation an index after a comma stands for partial differentiation with respect to the collective coordinate denoted by the index. Any sum on an index *before* a comma is over all particles; any sum on an index *after* a comma is over collective variables.

As before, introducing the collective variables leads to a system with more variables than degrees of freedom. We therefore introduce $2N$ constraints,

$$C_i \approx 0 \ , \tag{2.21}$$

$i = 1, \ldots, 2N$, where the C_i are functions of the dynamical variables.

For the case of many collective variables, we have found that casting the problem into symplectic notation makes the problem tractable. Noting that there are M particles (with $M \to \infty$) and N collective variables we define the $2(M+N)$ components of a symplectic column matrix $\mathbf{n}$ as

$$\mathbf{n}^T = [q_1, \ldots, q_M, X_1, \ldots, X_N, p_1, \ldots, p_M, Y_1, \ldots, Y_N] \ , \tag{2.22}$$

where the T superscript denotes the transpose matrix. The quantities p_i and Y_i are the momenta conjugate to the phonon variables q_i and the collective variables X_i, respectively. We also define the derivative of a function with respect to $\mathbf{n}$ as

$$\left[\frac{\partial A}{\partial \mathbf{n}}\right]^T = \left[\frac{\partial A}{\partial q_1}, \ldots, \frac{\partial A}{\partial q_M}, \frac{\partial A}{\partial X_1}, \ldots, \frac{\partial A}{\partial X_N}, \frac{\partial A}{\partial p_1}, \ldots, \frac{\partial A}{\partial p_M}, \frac{\partial A}{\partial Y_1}, \ldots, \frac{\partial A}{\partial Y_N}\right] \tag{2.23}$$

and the $2(N+M)\times 2(N+M)$ matrix:

$$\mathbf{J} = \begin{bmatrix} 0 & \mathbf{I} \\ -\mathbf{I} & 0 \end{bmatrix} , \tag{2.24}$$

where each submatrix has dimension $(N+M)\times(N+M)$ and $\mathbf{I}$ is the identity. $\mathbf{J}$ has the property

$$\mathbf{J}\mathbf{J}^T = \mathbf{J}\mathbf{J}^T = \mathbf{I} \ .$$

In this notation the Poisson bracket of any two functions

A and B of the dynamical variables is

$$\{A,B\}=\left[\frac{\partial A}{\partial \mathbf{n}}\right]^T \mathbf{J}\frac{\partial B}{\partial \mathbf{n}} . \tag{2.25}$$

The usual Hamilton's equations therefore have the simple form

$$\dot{\mathbf{n}}=\mathbf{J}\frac{\partial H}{\partial \mathbf{n}} \tag{2.26}$$

if the coordinates in the simplectic column matrix $\mathbf{n}$ are independent.

We now use this result to put the modified, or Dirac bracket, as defined in Eq. (2.6), into the symplectic form,

$$\begin{aligned}\{A,B\}^*&=\left[\frac{\partial A}{\partial \mathbf{n}}\right]^T \mathbf{J}\frac{\partial B}{\partial \mathbf{n}}\\ &\quad-\left[\frac{\partial A}{\partial \mathbf{n}}\right]^T \mathbf{J}\frac{\partial C_i}{\partial \mathbf{n}}C_{ij}^{-1}\left[\frac{\partial C_j}{\partial \mathbf{n}}\right]^T \mathbf{J}\frac{\partial B}{\partial \mathbf{n}}\\ &=\left[\frac{\partial A}{\partial \mathbf{n}}\right]^T\left[\mathbf{I}-\mathbf{J}\frac{\partial C_i}{\partial \mathbf{n}}C_{ij}^{-1}\left[\frac{\partial C_j}{\partial \mathbf{n}}\right]^T\right]\mathbf{J}\frac{\partial B}{\partial \mathbf{n}} .\end{aligned} \tag{2.27}$$

We define the operator $\mathbf{P}$ by rewriting Eq. (2.27) as

$$\{A,B\}^*=\left[\frac{\partial A}{\partial \mathbf{n}}\right]^T(\mathbf{I}-\mathbf{P})\mathbf{J}\frac{\partial B}{\partial \mathbf{n}} , \tag{2.28}$$

where

$$\mathbf{P}\equiv\mathbf{J}\frac{\partial C_i}{\partial \mathbf{n}}C_{ij}^{-1}\left[\frac{\partial C_j}{\partial \mathbf{n}}\right]^T . \tag{2.29}$$

Now using the definition

$$C_{ij}\equiv\{C_i,C_j\}\equiv\left[\frac{\partial C_i}{\partial \mathbf{n}}\right]^T \mathbf{J}\frac{\partial C_j}{\partial \mathbf{n}} ,$$

we find that $\mathbf{P}$ has the following properties:

$$\mathbf{P}^2=\mathbf{P} , \tag{2.30a}$$

$$\left[\frac{\partial C_k}{\partial \mathbf{n}}\right]^T\mathbf{P}=\left[\frac{\partial C_k}{\partial \mathbf{n}}\right]^T , \tag{2.30b}$$

$$\mathbf{PJ}\frac{\partial C_k}{\partial \mathbf{n}}=\mathbf{J}\frac{\partial C_k}{\partial \mathbf{n}} . \tag{2.30c}$$

We see that $\mathbf{P}$ is a projectionlike operator. That is, when operating to the right $\mathbf{P}$ projects in the direction of $\mathbf{J}(\partial C_k/\partial \mathbf{n})$ and when operating to the left it projects in the direction of $(\partial C_k/\partial \mathbf{n})^T$. Therefore, we call $\mathbf{P}$ a *symplectic projection operator.* We see that if $\mathbf{P}$ were not present in Eq. (2.28), the expression on the right-hand side would reduce to the usual Poisson bracket. The presence of $\mathbf{P}$ explicitly projects out the extra degrees of freedom that arise when the collective variables are introduced. We may express this more concisely by using Eq. (2.26) to write

$$\dot{\mathbf{n}}_D=(\mathbf{I}-\mathbf{P})\dot{\mathbf{n}}_P , \tag{2.31}$$

where D and P refer to Dirac and Poisson. That is, calculate the time dependence of the coordinates and momenta according to the Poisson prescription and then operate with $(\mathbf{I}-\mathbf{P})$ to obtain the physically meaningful time dependence of the variables. This general result is a property of the definition of the Dirac bracket and is independent of the form of the constraints.

Expressing Dirac's formula for his modified Poisson brackets in symplectic notation has naturally led to a *symplectic projection operator* interpretation of the new bracket. Ultimately, when we write the Dirac equations of motion explicitly in terms of the coordinates and momenta in the more familiar function space (that is *not* in symplectic space), out of $\mathbf{P}$ will emerge, for a particular set of constraints, an ordinary projection operator—ordinary in the sense that it projects in the same direction whether operating to the left or to the right. This more familiar projection operator is the multiple-collective-coordinate generalization of $\mathcal{P}_{ln}$ encountered at the end of Sec. II A, and we will show that $\mathcal{P}_{ln}$ leads to an enormous simplification in the derivation of the Dirac equations of motion. (We will use the same symbol, namely $\mathcal{P}_{ln}$, regardless of the number of collective variables.)

Our next step is to calculate the projection operator $\mathbf{P}$ with a particular set of constraints. The form of the constraints is motivated by the same argument as for the case of the single-collective variable: to minimize the q_l in the vicinity of the kink. We write the $2N$ constraints as

$$C_i=f_{l,i}q_l\approx 0, \quad C_{i+N}=f_{l,i}p_l\approx 0 , \tag{2.32}$$

where $i=1,\ldots,N$.

Equation (2.32) is a convenient ordering of the constraints for the following reason. In order to calculate the symplectic projection operator as defined in Eq. (2.29) we must invert the following $2N\times 2N$ matrix formed by the Poisson brackets of the constraints:

$$C=\begin{bmatrix} \{C_1,C_1\} & \cdots & \{C_1,C_N\} & \{C_1,C_{N+1}\} & \cdots & \{C_1,C_{2N}\} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \{C_N,C_1\} & \cdots & \{C_N,C_N\} & \{C_N,C_{N+1}\} & \cdots & \{C_N,C_{2N}\} \\ \{C_{N+1},C_1\} & \cdots & \{C_{N+1},C_N\} & \{C_{N+1},C_{N+1}\} & \cdots & \{C_{N+1},C_{2N}\} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \{C_{2N},C_1\} & \cdots & \{C_{2N},C_N\} & \{C_{2N},C_{N+1}\} & \cdots & \{C_{2N},C_{2N}\} \end{bmatrix} . \tag{2.33a}$$

Because of the ordering of the constraints in Eq. (2.32), when one evaluates the Poisson brackets, the matrix C takes the block form,

$$\begin{bmatrix} 0 & \cdots & 0 & M_{11} & \cdots & M_{1N} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & M_{N1} & \cdots & M_{NN} \\ -M_{11} & \cdots & -M_{1N} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ -M_{N1} & \cdots & -M_{NN} & 0 & \cdots & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{M} \\ -\mathbf{M} & \mathbf{0} \end{bmatrix} , \tag{2.33b}$$

where the components of the $N\times N$ matrix $\mathbf{M}$, on the right-hand side of Eq. (2.33b), are

$$M_{\alpha\beta}=f_{l,\alpha}f_{l,\beta} , \tag{2.34}$$

where $\alpha,\beta=1,\ldots,N$. The inverse of this matrix is then

$$C^{-1}=\begin{bmatrix} \mathbf{0} & -\mathbf{M}^{-1} \\ \mathbf{M}^{-1} & \mathbf{0} \end{bmatrix} . \tag{2.35}$$

The existence of C^{-1}, made up of only second-class constraints, was proved by Dirac; consequently $\mathbf{M}^{-1}$ exists. We use roman boldface letters to denote matrices in both the collective variable space (such as $\mathbf{M}$) and in symplectic space (such as $\mathbf{J}$ or $\mathbf{n}$). No confusion will arise since matrices in the collective variable space never appear in the same equation with matrices in the symplectic space.

We now substitute this expression for C^{-1} into Eq. (2.27). The block form of Eq. (2.35) now allows us to separate the sum over the $2N$ constraints in Eq. (2.27) into two sums, each over N constraints. To this end we now label the constraints as

$$C_{1\alpha}=f_{l,\alpha}q_l \quad \text{and} \quad C_{2\alpha}=f_{l,\alpha}p_l , \tag{2.36}$$

where $\alpha=1,\ldots,N$ and Eq. (2.27) becomes

$$\{A,B\}^*=\left[\frac{\partial A}{\partial \mathbf{n}}\right]^T\left[\mathbf{I}+\mathbf{J}\frac{\partial C_{1\alpha}}{\partial \mathbf{n}}M_{\alpha\beta}^{-1}\left[\frac{\partial C_{2\beta}}{\partial \mathbf{n}}\right]^T -\mathbf{J}\frac{\partial C_{2\alpha}}{\partial \mathbf{n}}M_{\alpha\beta}^{-1}\left[\frac{\partial C_{1\beta}}{\partial \mathbf{n}}\right]^T\right]\mathbf{J}\frac{\partial B}{\partial \mathbf{n}} . \tag{2.37}$$

Equation (2.37) is the form of the Dirac bracket we use to generate the equations of motion for the new variables. Before we can derive the equations of motion however, we must derive the canonical transformation to the new variables. This is done by requiring invariance of the Dirac brackets under the transformation.

First, we must introduce some notation that will make the forthcoming formulas less cumbersome. The constraints in Eq. (2.36) can be written in column matrix form:

$$\mathbf{C}_1=\mathbf{f}_l'q_l, \quad \mathbf{C}_2=\mathbf{f}_l'p_l , \tag{2.36$'$}$$

where the αth element in the column matrices, $\mathbf{f}_l'$, $\mathbf{C}_1$, and $\mathbf{C}_2$ are $f_{l,\alpha}$, $C_{1\alpha}$, and $C_{2\alpha}$, respectively. As previously noted, an index after a comma, such as in $\mathbf{C}_{1,\alpha}$, symbolizes taking the derivative of each element with respect to X_α. We will write the $N\times N$ matrix formed from the columns $\mathbf{C}_{1,\alpha}$ as $\mathbf{C}_1'$. (Note that $\mathbf{C}_1'$ is symmetric if q_l is independent of the collective variables, as we have assumed.) $\mathbf{Y}$ represents a column matrix whose elements are the momenta, Y_i. A derivative with respect to a column matrix, such as $\partial A/\partial\mathbf{Y}$ is again a column matrix whose elements are $\partial A/\partial Y_i$. The same applies to $\mathbf{X}$ and $\partial A/\partial\mathbf{X}$. The particle index [such as l in Eq. (2.36$'$)] is left explicit in order to avoid mixing notation.

For a transformation to be canonical we require the Poisson brackets of the old variables to remain invariant when the new variables are substituted in for the old. The Poisson bracket is modified in the presence of the constraints and we require

$$\{Q_l,Q_n\}^*=0 , \tag{2.38a}$$

$$\{Q_l,P_n\}^*=\delta_{ln} , \tag{2.38b}$$

$$\{P_l,P_n\}^*=0 , \tag{2.38c}$$

where Q_l is given by Eq. (2.17). We solve for the conjugate momentum P_l as a function of the new variables using Eq. (2.37). The derivation is done in detail in Appendix A. The first condition, Eq. (2.38a), is identically satisfied. The other two conditions lead to requirements on the form of P_l. The result is that Eq. (2.38b) leads to an equation similar in form to the analogous equation for the single-collective-variable case [cf. Eq. (2.9)]

$$(\delta_{ls}-\mathcal{P}_{ls})\frac{\partial P_n}{\partial p_s}+\mathbf{f}_l'^T(\mathbf{I}-\mathbf{b})\frac{\partial P_n}{\partial \mathbf{Y}}=\delta_{ln} , \tag{2.39}$$

where $\mathbf{b}$ is given by

$$\mathbf{b}=\mathbf{M}^{-1}\mathbf{f}_l''q_l . \tag{2.40}$$

We see that $\mathbf{b}$ reduces to η/M for the single-collective-variable case. $\mathcal{P}_{ls}$ is a projection operator defined, in general, by

$$\mathcal{P}_{ls} \equiv f_{l,i} M_{ij}^{-1} f_{s,j} = \mathbf{f}_l'^T \mathbf{M}^{-1} \mathbf{f}_s' \tag{2.41}$$

and using Eq. (2.34) we obtain

$$\mathcal{P}_{rs}\mathcal{P}_{st} = \mathcal{P}_{rt} \ , \tag{2.42a}$$

$$\mathcal{P}_{rs}\mathbf{f}_s' = \mathbf{f}_r' \ , \tag{2.42b}$$

$$\mathcal{P}_{rs} q_s = \mathbf{f}_r'^T \mathbf{M}^{-1} \mathbf{f}_s' q_s = \mathbf{f}_r'^T \mathbf{M}^{-1} \mathbf{C}_1 \approx 0 \ . \tag{2.42c}$$

The expression for $\mathcal{P}_{ln}$ in Eq. (2.41) reduces to that of Eq. (2.16a) for the single-collective-variable case.

Equation (2.39) is a differential equation for the unknown momenta P_l in terms of the new variables. The general solution, derived in Appendix A, is given by

$$\begin{aligned} P_l &= p_l + \mathbf{Y}^T(\mathbf{I}-\mathbf{b})^{-1}\mathbf{M}^{-1}\mathbf{f}_l' + h_l \\ &= p_l + \mathbf{f}_l'^T \mathbf{M}^{-1}(\mathbf{I}-\mathbf{b}^T)^{-1}\mathbf{Y} + h_l \ , \end{aligned} \tag{2.43}$$

where $h_l = h_l(\mathbf{X}, q_i)$ and the second equality in Eq. (2.43) is obtained by taking the transpose of the middle term. In writing the last part of Eq. (2.43) we have used $(\mathbf{M}^T)^{-1} = \mathbf{M}^{-1}$ and have written the transpose of $(\mathbf{I}-\mathbf{b})^{-1}$ as $(\mathbf{I}-\mathbf{b}^T)^{-1}$.

Equation (2.38b) puts no conditions on h_l. For Eq. (2.43) to satisfy Eq. (2.38c), h_l must obey certain conditions as given at the end of Appendix A and one choice of h_l which satisfies these conditions is $h_l = 0$. The canonical transformation is therefore complete, with the momenta given by Eq. (2.43) (with our choice of $h_l = 0$).

We write the Hamiltonian in terms of the new variables by substituting Eq. (2.43) into the original Hamiltonian given by

$$H = \tfrac{1}{2} P_l P_l + V(Q_l) \ . \tag{2.44}$$

We find

$$\begin{aligned} P_l P_l &= \mathbf{Y}^T(\mathbf{I}-\mathbf{b})^{-1}\mathbf{M}^{-1}(\mathbf{I}-\mathbf{b}^T)^{-1}\mathbf{Y} + p_l p_l \\ &= \mathbf{Y}^T \overline{\mathbf{M}}^{-1}\mathbf{Y} + p_l p_l \ , \end{aligned} \tag{2.45}$$

where $\overline{\mathbf{M}}$ is the "dressed" mass and is defined by

$$\overline{\mathbf{M}} \equiv (\mathbf{I}-\mathbf{b}^T)\mathbf{M}(\mathbf{I}-\mathbf{b}) \ . \tag{2.46}$$

Compare this to the single-collective-variable case where

$$\overline{M} = M(1-b)^2 = M\left[1 - \frac{\eta}{M}\right]^2 .$$

Finally, we can write our Hamiltonian in the new variables:

$$H = \tfrac{1}{2}\mathbf{Y}^T \overline{\mathbf{M}}^{-1}\mathbf{Y} + \tfrac{1}{2} p_l p_l + V \ , \tag{2.47}$$

where V is defined by Eq. (2.13b).

The cross terms generated by substituting P_l into the Hamiltonian are proportional to $\mathbf{C}_2$ and are set to zero. We noted earlier that one cannot set the constraints to zero until all variations with respect to the dynamical variables have been performed. But we have yet to perform that variation of the Hamiltonian to get the equations of motion. Herein lies the beauty of the Dirac formalism. The modified bracket guarantees that the constraint will be a constant of the motion, i.e., the Dirac formalism guarantees that for a constraint, C_i, $\{C_i, H\}^* = 0$ as a consequence of the equations of motion. Therefore, adding to the Hamiltonian any terms which are proportional to a constraint, C_i, will not affect the equations of motion. In effect this means that the equations, $C_i = 0$, may now be considered strongly equal to zero (they may be put to zero before varying the Hamiltonian), and further, they need no longer be treated as constraints, but rather as initial conditions. The formalism then guarantees that they will be satisfied for all times.

It is now possible to generate all of the equations of motion by using Eq. (2.37) with this Hamiltonian. We write Eq. (2.37) in the column matrix form of Eq. (2.31):

$$\dot{\mathbf{n}}_D = \left[\mathbf{I} + \mathbf{J}\frac{\partial C_{1\alpha}}{\partial \mathbf{n}} M_{\alpha\beta}^{-1}\left[\frac{\partial C_{2\beta}}{\partial \mathbf{n}}\right]^T - \mathbf{J}\frac{\partial C_{2\alpha}}{\partial \mathbf{n}} M_{\alpha\beta}^{-1}\left[\frac{\partial C_{1\beta}}{\partial \mathbf{n}}\right]^T\right]\mathbf{J}\frac{\partial H}{\partial \mathbf{n}} \ , \tag{2.48}$$

where we have substituted the Hamiltonian H in for B. In Appendix B, we show the details of the calculation of the equations of motion for the new variables. The resulting equations are

$$(\delta_{ln} - \mathcal{P}_{ln})\left[\dot{q}_n - \frac{\partial H}{\partial p_n}\right] = 0 \ , \tag{2.49a}$$

$$(\delta_{ln} - \mathcal{P}_{ln})\left[\dot{p}_n + \frac{\partial H}{\partial q_n}\right] = 0 \ , \tag{2.49b}$$

$$\dot{\mathbf{X}} = \frac{\partial H}{\partial \mathbf{Y}} \ , \tag{2.49c}$$

$$\dot{\mathbf{Y}} + \frac{\partial H}{\partial \mathbf{X}} = \mathbf{C}_1'^T \mathbf{M}^{-1}\mathbf{f}_l'\left[\frac{\partial H}{\partial q_l} + \dot{p}_l\right] + \mathbf{C}_2'^T \mathbf{M}^{-1}\mathbf{f}_l'\left[\frac{\partial H}{\partial p_l} - \dot{q}_l\right] . \tag{2.49d}$$

$\dot{\mathbf{C}}_1 = \dot{\mathbf{C}}_2 = 0$ follows from Eqs. (2.49a)–(2.49d). Thus we need only require $\mathbf{C}_1(t=0) = \mathbf{C}_2(t=0) = 0$.

One can eliminate the momenta, p_n and $\mathbf{Y}$, from Eqs. (2.49a)–(2.49d) to obtain second-order equations for $\ddot{q}_n$ and $\ddot{X}$. This is a fairly tedious procedure. We will now show that the second-order equations can be arrived at by first substituting the *Ansatz*, Eq. (2.17), into the original equations of motion for old variables Q_l, Eq. (2.19). Then, the second-order equations of motion for the new phonon variables, q_l, are obtained by operating on the resulting equations with the projection operator $(\delta_{ln} - \mathcal{P}_{ln})$. The equations for the collective variables are obtained by substituting the *Ansatz* Eq. (2.17) into the original equations of motion (2.19) and operating with $\mathcal{P}_{ln}$. Then, from this $\mathcal{P}_{ln}$ equation, we will be able to generate an equation of motion for each collective variable. (In the next section we provide an example.)

We first consider Eqs. (2.49a) and (2.49b), the equations of motion for the new variables q_n and p_n and start by using a result from Appendix C—that Eqs. (2.49a) and (2.49b) are equivalent to

$$(\delta_{ln}-\mathcal{P}_{ln})\left[\dot{Q}_n-\frac{\partial H}{\partial P_n}\right]=0\ , \qquad (2.50a)$$

$$(\delta_{ln}-\mathcal{P}_{ln})\left[\dot{P}_n+\frac{\partial H}{\partial Q_n}\right]=0\ , \qquad (2.50b)$$

where we mean that the new variables are to be substituted for the old. [This is the only way Eqs. (2.50a) and (2.50b) make sense since the projection operator $\mathcal{P}_{ln}$ is defined only in terms of the "new" coordinates.] We develop, now, an intuitive understanding of the meaning of Eqs. (2.50a) and (2.50b) before proceeding.

We first consider Eq. (2.50a) without the operator $(\delta_{ln}-\mathcal{P}_{ln})$, i.e.

$$\dot{Q}_n-\frac{\partial H}{\partial P_n}=\dot{Q}_n-P_n=0\ . \qquad (2.51a)$$

When we substitute the new variables [Eqs. (2.20) and (2.43)] for the old in Eq. (2.51a) we obtain

$$\mathbf{f}_l'^T\dot{\mathbf{X}}+\dot{q}_l-[\mathbf{Y}^T(\mathbf{I}-\mathbf{b}^T)^{-1}\mathbf{M}^{-1}\mathbf{f}_l'+p_l]=0\ . \qquad (2.51b)$$

The solution set of this equation is larger than the solution set of the original system. This is because Eq. (2.51b) is a function of N more dynamical variables than our original system. The solution set we are looking for is that which satisfies the constraints. Now, by operating with the operator $(\delta_{ln}-\mathcal{P}_{\mathrm{ln}})$ we effectively impose the constraints by picking out only those solutions of Eq. (2.51b) which also satisfy the constraints. The same argument applies to Eq. (2.50b).

Next, we proceed by directly eliminating the momenta from Eqs. (2.50a) and (2.50b). We substitute P_n for $\partial H/\partial P_n$ in Eq. (2.50a) and take the time derivative of the result, which yields

$$(\delta_{ln}-\mathcal{P}_{ln})(\ddot{Q}_n-\dot{P}_n)-\dot{\mathcal{P}}_{ln}(\dot{Q}_n-P_n)=0\ . \qquad (2.52)$$

The last term is independently equal to zero from Eq. (2.51a) and so Eq. (2.52) reduces to

$$(\delta_{ln}-\mathcal{P}_{ln})(\ddot{Q}_n-\dot{P}_n)=0\ . \qquad (2.53)$$

Eliminating $\dot{P}_n$ by using Eq. (2.50b) finally gives

$$(\delta_{ln}-\mathcal{P}_{ln})\left[\ddot{Q}_n+\frac{\partial H}{\partial Q_n}\right]=0\ . \qquad (2.54)$$

Equation (2.54) is the second-order differential equation that is equivalent to the first-order differential equations (2.49a) and (2.49b). Recall that Eq. (2.54) is only meaningful when the new variables are substituted in for the old.

We now turn our attention to Eqs. (2.49c) and (2.49d) and present an argument that will lead us to a second order differential equation equivalent to Eqs. (2.49c) and (2.49d). We note that the set of four equations (2.49a)–(2.49d) must be equivalent to the set of equations:

$$\ddot{Q}_l(q_l,X_i)+\frac{\partial H}{\partial Q_l}(q_l,X_i)=\ddot{Q}_l(q_l,X_i)+\frac{\partial V}{\partial Q_l}(q_l,X_i)=0\ , \qquad (2.55a)$$

$$\mathbf{C}_1=0 \text{ and } \mathbf{C}_2=0\ . \qquad (2.55b)$$

Equation (2.55a) are the original equations of motion in terms of the new coordinates, i.e., with the *Ansatz* substituted in. We have explicitly indicated this dependence on the new variables in Eq. (2.55a). Here, the constraints do *not* follow from the equations of motion but are specified as independent, auxiliary conditions. We will discuss at greater length in the next section the auxiliary nature of the constraints associated with Eq. (2.55a). The point we wish to make now is that from the Dirac equations of motion, Eqs. (2.49a)–(2.49d), we must be able to recover the form of the original equations of motion as a function of the new variables, namely Eq. (2.55a). This is reasonable since the original equations of motion were our starting point. This observation allows us to find a more intuitive form for Eq. (2.49d) which we write in the form [assuming we have eliminated the momenta using Eq. (2.49c)]:

$$\mathcal{P}_{ln}\left[\ddot{Q}_n+\frac{\partial H}{\partial Q_n}\right]+\mathcal{P}_{ln}A_n+(\delta_{ln}-\mathcal{P}_{ln})B_n=0\ , \qquad (2.56)$$

where A_n and B_n are such that Eq. (2.56) is the second-order differential equation that is equivalent to the two first-order differential equations [(2.49c) and (2.49d)]. By adding the $\mathcal{P}_{ln}$ and $(\delta_{ln}-\mathcal{P}_{ln})$ contributions of the Dirac equations of motion [(2.49a)–(2.49d)] we must recover Eq. (2.55a). Since the $\mathcal{P}_{ln}$ and $(\delta_{ln}-\mathcal{P}_{ln})$ contributions of the Dirac equations of motion are contained in Eqs. (2.54) and (2.56) it must follow that

$$(\delta_{ln}-\mathcal{P}_{ln})\left[\ddot{Q}_n+\frac{\partial H}{\partial Q_n}\right]+\mathcal{P}_{ln}\left[\ddot{Q}_n+\frac{\partial H}{\partial Q_n}\right]+\mathcal{P}_{ln}A_n+(\delta_{ln}-\mathcal{P}_{ln})B_n=\ddot{Q}_l+\frac{\partial H}{\partial Q_l}\ .$$

We conclude, therefore, that

$$\mathcal{P}_{ln}A_n+(\delta_{ln}-\mathcal{P}_{ln})B_n=0\ . \qquad (2.57)$$

Therefore, Eqs. (2.49c) and (2.49d) must be equivalent to Eq. (2.56), which in light of Eq. (2.57), reduces to

$$\mathcal{P}_{ln}\left[\ddot{Q}_n+\frac{\partial H}{\partial Q_n}\right]=0\ . \qquad (2.58)$$

Equations (2.54) and (2.58) are crucial to our paper and they state the following. To obtain the Dirac bracket equations of motion for the new variables q_l and X_i in the presence of the given constraints, first substitute the *Ansatz,* Eq. (2.17), into Eq. (2.55a). Then operate with $\mathcal{P}_{ln}$ to project Eq. (2.55a) onto the "$\mathbf{f}_l'$ manifold," thus giving the equation of motion, namely Eq. (2.58), from which will be generated a separate equation of motion for each collective variable X_i. To generate a separate equation for each collective variable X_i from Eq. (2.58), operate *again* on Eq. (2.58) with the "shape mode," $f_{l,i}$ associated with the ith collective variable X_i whose equation of motion is sought. This double operation—first operating with $\mathcal{P}_{ln}$ and then with the appropriate shape mode—is easily condensed into one operation since, by Eq. (2.42b),

$$\mathbf{f}'_l \mathcal{P}_{ln} = \mathbf{f}'_n \ . \tag{2.59}$$

Therefore, we need only operate on Eq. (2.55a) (after substituting in the *Ansatz*) with the shape mode $f_{l,i}$ to obtain the equation of motion for X_i. Next, to obtain the equations of motion for q_l, operate on Eq. (2.55a) (after substitution of the *Ansatz*) with $(\delta_{ln} - \mathcal{P}_{ln})$, which projects Eq. (2.55a) onto a manifold orthogonal to the $\mathbf{f}'_l$ manifold.

The effect of invoking the constraints by operating with $\mathcal{P}_{ln}$ and $(\delta_{ln} - \mathcal{P}_{ln})$ is also apparent from a conservation of energy point of view. In fact, the new variables are shown (at the end of Appendix B) to satisfy the following:

$$\frac{\partial H}{\partial q_l}(\delta_{ln} - \mathcal{P}_{ln})\dot{q}_n + \frac{\partial H}{\partial p_l}(\delta_{ln} - \mathcal{P}_{ln})\dot{p}_n = 0 \ , \tag{2.60a}$$

$$\left[\frac{\partial H}{\partial \mathbf{X}}\right]^T \dot{\mathbf{X}} + \left[\frac{\partial H}{\partial \mathbf{Y}}\right]^T \dot{\mathbf{Y}} + \frac{\partial H}{\partial q_l}\mathcal{P}_{ln}\dot{q}_n + \frac{\partial H}{\partial p_l}\mathcal{P}_{ln}\dot{p}_n = 0 \ . \tag{2.60b}$$

Adding these two equations gives

$$\left[\frac{\partial H}{\partial \mathbf{X}}\right]^T \dot{\mathbf{X}} + \left[\frac{\partial H}{\partial \mathbf{Y}}\right]^T \dot{\mathbf{Y}} + \frac{\partial H}{\partial q_l}\dot{q}_l + \frac{\partial H}{\partial p_l}\dot{p}_l = \frac{dH}{dt} = 0 \ . \tag{2.60c}$$

Equation (2.60c) is the statement of conservation of energy. Its solution set, however, is too large because of the N extra variables. Only those solutions of the variables q_l, X_i, p_l, and Y_i that also satisfy Eq. (2.60a) or (2.60b) as well as (2.60c) are actual solutions to the original system.

In concluding this section, we make a few remarks about the functional dependence of our *Ansatz*. All of our analysis so far has been explicitly for a discrete lattice. All formulas carry over to a continuum description by letting the sum on the discrete index l be replaced by an integral over the independent parameter x with $l \to x$. Also, we have chosen q_l to be independent of the collective variables. There are, however, other *Ansätze* in the literature[1,3] [with $Q_l \to \phi(x,t)$] of the form

$$\phi(x,t) = \sigma(x - X_1, X_2, \ldots, X_N) + \chi(x - X_1, t) \ , \tag{2.61}$$

where σ and χ play the (continuum) role of our f_l and q_l, respectively. Here, $X_1 = X_{\text{c.m.}}$ is the center-of-mass coordinate of the kink. Note that χ now also depends on X_1. (We will always associate X_1 with the center-of-mass coordinate of the kink.) Parenthetically, we have not yet been able to carry out the derivation for the discrete analog of Eq. (2.61), namely

$$Q_l = f_l(l - X_1, X_2, \ldots, X_N) + q_l(l - X_1, t) \ . \tag{2.62}$$

Although certain derivations in this section and in the appendices have to be modified because of the presence of X_1 in χ, all the results of this section are still valid for the *Ansatz* Eq. (2.61) for the continuum. That is, the second order in time differential equations of motion for the N collective variables X_i and χ are generated by substituting the *Ansatz* Eq. (2.61) into the original equation of motion $\ddot{\phi} + \partial V(\phi)/\partial\phi = 0$ and operating with the continuous analog of $\mathcal{P}_{ln}$ and $(\delta_{ln} - \mathcal{P}_{ln})$ [for example, see Eq. (3.23)] in exactly the same manner as for the *Ansatz* in Eq. (2.17). (Note that for the ansatz in Eq. (2.61) we have the correspondence $\dot{q}_l \to \partial\chi/\partial t$ and $\dot{p}_l \to \partial\pi/\partial t$.) We will use the *Ansatz* in Eq. (2.61) for the DSG kink as an example in the next section. We have not investigated continuum *Ansätze* of a more general form than Eq. (2.61), where χ is a function of more than one collective variable.

For completeness we give the canonical transformation together with the Hamiltonian induced by it and the constraints [where all integrals are from $-\infty$ to ∞ with respect to $y = (x - X_1)$] for the *Ansatz* in Eq. (2.61). The constraints are

$$C_{1i} = \int \sigma_{,i}(y, X_2, \ldots, X_N)\chi(y,t)dy = 0 \ ,$$
$$C_{2i} = \int \sigma_{,i}(y, X_2, \ldots, X_N)\pi(y,t)dy = 0 \ , \tag{2.63a}$$

where $i = 1, \ldots, N$. $\pi(x - X_1, t)$ and $\Pi(x,t)$ are the momenta conjugate to $\chi(x - X_1, t)$ and $\phi(x,t)$, respectively. $\Pi(x,t)$ and the Hamiltonian are given by

$$\Pi(x,t) = \dot{\phi} = \pi + \left[\mathbf{Y}^T - \int \pi \left[\frac{\partial \chi}{\partial \mathbf{X}}\right]^T dx'\right](\mathbf{I} - \mathbf{b})^{-1}\mathbf{M}^{-1}\frac{\partial\sigma}{\partial \mathbf{X}} + h(x, \mathbf{X}, \chi) \ , \tag{2.63b}$$

$$H = \tfrac{1}{2}\overline{\mathbf{Y}}\,\overline{\mathbf{M}}^{-1}\overline{\mathbf{Y}} + \tfrac{1}{2}\int \pi^2(y,t)dy + V \ , \tag{2.63c}$$

$$\overline{Y}_i = Y_i - \int \pi(y,t)\chi_{,i}(y,t)dy \ , \tag{2.63d}$$

where Y_i is the momentum conjugate to X_i and V is given by

$$V = \tfrac{1}{2}\int \left[\frac{\partial(\sigma + \chi)}{\partial y}\right]^2 dy + \int V_S(\sigma + \chi)dy \ . \tag{2.63e}$$

Also

$$M_{ij} = \int \sigma_{,i}(y)\sigma_{,j}(y)dy \tag{2.63f}$$

$$b_{ij} = M_{ik}^{-1}\int \sigma_{,kj}(y)\chi(y,t)dy \tag{2.63g}$$

and $\overline{\mathbf{M}}$ is defined as in Eq. (2.46). Note $\chi_{,i}(x - X_1, t) = 0$ for $i \geq 2$. At the end of Appendix A we give the conditions that $h(x, \mathbf{X}, \chi)$ must satisfy in order that the transformation defined by Eqs. (2.61) and (2.63b) be canonical. As in the discrete case, $h(x, \mathbf{X}, \chi) = 0$ satisfies these conditions.

We have proved in this section that a powerful method exists for calculating the Dirac equations of motion when the constraints are of the form given in Eq. (2.32) [or Eq. (2.63a)]. The method consists of substituting the *Ansatz*, Eq. (2.17) [or Eq. (2.61)], into the original equations of motion and projecting the equation onto orthogonal manifolds in phase space using $\mathcal{P}_{ln}$ to obtain the equations of motion of the collective variables X_i and using $(\delta_{ln} - \mathcal{P}_{ln})$ to obtain the equations of motion for the q_l. That this is at all possible is a consequence of the general structure of the Dirac bracket which is itself the definition of a projection in symplectic space. We have shown that this procedure can be carried out for the *Ansatz* in Eq. (2.17) and state that similar calculations show that the projection operator procedure works also with the *Ansatz* in Eq. (2.61). We use the *Ansatz* in Eq. (2.61) for illustration in the next section.

III. EXAMPLE: THE DOUBLE SINE-GORDON KINK

In this section we work out the equations of motion for the DSG system where we consider the collective variables $X(t)$ and $R(t)$ to be the center of mass of the kink and the separation between the two subkinks, respectively. We will take $R=\mathcal{R}$ at equilibrium. Note also that $\mathcal{R}$ is the time independent parameter in the DSG substrate potential:[7-10]

$$V_{\mathcal{R}}(\phi)=-4\left[\frac{2\pi}{l_0}\right]\text{sech}^2(\mathcal{R})\left\{\sinh^2(\mathcal{R})(\cos\phi-1)-\left[1+\cos\left[\frac{\phi}{2}\right]\right]\right\}.$$

We take as our *Ansatz*

$$\phi=\sigma(\gamma(x-X),R)+\chi(\gamma(x-X)), \tag{3.1a}$$

where

$$\sigma(\gamma(x-X),R)=\sigma_{\rm SG}\left[\frac{2\pi}{l_0}\gamma(x-X)+R\right]-\sigma_{\rm SG}\left[R-\frac{2\pi}{l_0}\gamma(x-X)\right] \tag{3.1b}$$

and

$$\sigma_{\rm SG}(x)=4\tan^{-1}[\exp(x)]. \tag{3.1c}$$

We have included the center-of-mass variable in the radiation field χ. The constraints are

$$C_X=\int\sigma'(y,R)\chi(y,t)dy=0, \tag{3.2a}$$

$$C_R=\int\sigma_R(y,R)\chi(y,t)dy=0, \tag{3.2b}$$

where the prime in this section denotes the derivative with respect to $y=\gamma(x-X)$ and $\sigma_R=\partial\sigma/\partial R$. The integrals are taken from $-\infty$ to ∞.

In order to obtain equations of motion correct to terms of order $\dot{X}^2$, it is necessary to introduce γ into the *Ansatz* where

$$\gamma=1/(1-\dot{X}^2)^{1/2}.$$

We will show in a forthcoming paper,[20] which investigates the radiation mechanism of an oscillating, discrete SG kink trapped in the Peierls-Nabarro well, that agreement with simulation is much better when γ is included in the *Ansatz*. The presence of γ, however, causes the Lagrangian obtained by substituting the *Ansatz* into

$$L=\tfrac{1}{2}\int\dot{\phi}^2(x,t)dx-\tfrac{1}{2}\int\left[\frac{\partial\phi}{\partial x}(x,t)\right]^2dx-\int V(\phi(x,t))dx, \tag{3.3}$$

to be a (nonlinear) function of the acceleration $\ddot{X}$. As a result, varying the action $S=\int L\,dt$ with respect to X and requiring $\delta S=0$, leads to equations of motion that contain d^4X/dt^4. Hence, the complexity of the resulting equations of motion is greatly increased and further, we have not determined whether or not the projection operator method of the last section is valid for a Lagrangian that is a function of acceleration. Therefore, in the present section we treat γ as a parameter, taking $\dot{\gamma}=0$. This is an excellent approximation for studying problems such as, for example, the process by which untrapped (nonrelativistic) kinks become trapped,[21] the coupling of the radiation field to oscillating trapped kinks,[20] the study of (nonrelativistic) kink-kink or kink-antikink collisions[12,16] or any other process in which $\dot{\gamma}$ may be regarded as small. (Note that the *Ansatz* frees χ from having to dress the kink to account for the kink's Lorentz contraction.)

In the following derivation of the equations of motion for the collective variables and coupled field of the DSG system using the projection-operator method presented in the last section, we will also perform manipulations that yield more convenient or practical expressions for the equations of motion. We will explicitly show what steps are carried out throughout this section.

We begin by substituting the *Ansatz*, Eq. (3.1a), into the original field equation

$$\ddot{\phi}-\frac{\partial^2\phi}{\partial x^2}+V_{\mathcal{R},\phi}(\phi)=0, \tag{3.4}$$

where $V_{\mathcal{R},\phi}=\partial V_{\mathcal{R}}/\partial\phi$. When we bring all terms containing χ on the right-hand side we obtain

$$\gamma\ddot{X}\sigma'+\gamma^2\sigma''-\gamma^2\dot{X}^2\sigma''+2\gamma\dot{X}\dot{R}\sigma'_R-\ddot{R}\sigma_R-\dot{R}^2\sigma_{RR}=\frac{\partial^2\chi}{\partial t^2}-\gamma^2\chi''+V_{\mathcal{R},\sigma}(\sigma+\chi)-\gamma\ddot{X}\chi'+\gamma^2\dot{X}^2\chi''-2\gamma\dot{X}\frac{\partial\chi'}{\partial t}, \tag{3.5}$$

where $\sigma_{RR}=\partial\sigma_R/\partial R$. Now we combine the γ^2 terms:

$$\gamma\ddot{X}\sigma'+2\gamma\dot{X}\dot{R}\sigma'_R-\ddot{R}\sigma_R-\dot{R}^2\sigma_{RR}=\frac{\partial^2\chi}{\partial t^2}-\chi''+[V_{\mathcal{R},\sigma}(\sigma+\chi)-\sigma'']-\gamma\ddot{X}\chi'-2\gamma\dot{X}\frac{\partial\chi'}{\partial t}. \tag{3.6}$$

Before operating on Eq. (3.6) with the projection operator, we express Eq. (3.6) in a more convenient form. We note that σ satisfies

$$\sigma''(x,X,R)=V_{R,\sigma}(\sigma(x,X,R)), \tag{3.7}$$

where the *same* R that appears in σ appears in the potential. We can then rewrite the expression in brackets in Eq. (3.6) in the following manner:

$$V_{\mathcal{R},\sigma}(\sigma+\chi)-\sigma''=V_{\mathcal{R},\sigma}(\sigma+\chi)-V_{R,\sigma}(\sigma) \tag{3.8}$$

and expanding in a power series in χ we obtain

$$V_{\mathcal{R},\sigma}(\sigma)+V_{\mathcal{R},\sigma\sigma}(\sigma)\chi+\mathcal{N}(\sigma,\chi)-V_{R,\sigma}(\sigma)=-S+V_{\mathcal{R},\sigma\sigma}(\sigma)\chi+\mathcal{N}(\sigma,\chi) , \tag{3.9}$$

where

$$\mathcal{N}(\sigma,\chi)\equiv V_{\mathcal{R},\sigma}(\sigma+\chi)-V_{\mathcal{R},\sigma}(\sigma)-V_{\mathcal{R},\sigma\sigma}(\sigma)\chi \tag{3.10a}$$

represents the nonlinear dependence on χ and

$$S\equiv V_{R,\sigma}(\sigma(x,X,R))-V_{\mathcal{R},\sigma}(\sigma(x,X,R)) . \tag{3.10b}$$

When we substitute Eqs. (3.8) through (3.10b) into Eq. (3.6) we obtain

$$\gamma\ddot{X}\sigma'+2\gamma\dot{X}\dot{R}\sigma'_R-\ddot{R}\sigma_R-\dot{R}^2\sigma_{RR}=\frac{\partial^2\chi}{\partial t^2}-\chi''-S+V_{\mathcal{R},\sigma\sigma}(\sigma)\chi+\mathcal{N}(\sigma,\chi)-\gamma\ddot{X}\chi'$$
$$-2\gamma\dot{X}\frac{\partial\chi'}{\partial t}+V_{R,\sigma\sigma}(\sigma)\chi-V_{R,\sigma\sigma}(\sigma)\chi , \tag{3.11}$$

where we have added and subtracted the last two terms on the right-hand side. Then defining the linear operator $\mathcal{L}$:

$$\mathcal{L}\equiv-\frac{\partial^2}{\partial x^2}+V_{R,\sigma\sigma}(\sigma(x,X,R)) ,$$

Eq. (3.11) becomes

$$\frac{\partial^2\chi}{\partial t^2}+\mathcal{L}\chi-S_{,\sigma}\chi+\mathcal{N}-\gamma\ddot{X}\chi'-2\gamma\dot{X}\frac{\partial\chi'}{\partial t}=\gamma\ddot{X}\sigma'+2\gamma\dot{X}\dot{R}\sigma'_R-\ddot{R}\sigma_R-\dot{R}^2\sigma_{RR}+S , \tag{3.12}$$

where

$$S_{,\sigma}\equiv V_{R,\sigma\sigma}(\sigma(x,X,R))-V_{\mathcal{R},\sigma\sigma}(\sigma(x,X,R)) .$$

Equation (3.12) was obtained by substituting the *Ansatz* Eq. (3.1) into the original equation of motion Eq. (3.3) and neglecting $\dot{\gamma}$ terms.

According to the general argument immediately preceding Eq. (2.59), to obtain the equation of motion for the collective variables we must multiply Eq. (3.12) by the shape mode associated with the desired collective variable and integrate. To this end, we define the bracket

$$\langle A \mid B\rangle\equiv\int_{-\infty}^{\infty}A(y,R)B(y,R)dy .$$

When we multiply Eq. (3.12) by $\sigma'(\gamma(x-X),R)$, and integrate with respect to $y=\gamma(x-X)$ we obtain the equation of motion for $\ddot{X}$:

$$\gamma\ddot{X}M_X(R)+\gamma\dot{X}\dot{R}\frac{dM_X(R)}{dR}=\left\langle\sigma'\left|\frac{\partial^2\chi}{\partial t^2}+\mathcal{L}\chi\right.\right\rangle-\left\langle\sigma'\left|S_{,\sigma}\chi-\mathcal{N}+\gamma\ddot{X}\chi'+2\gamma\dot{X}\frac{\partial\chi'}{\partial t}\right.\right\rangle , \tag{3.13}$$

where

$$M_X(R)=\langle\sigma'(y,R)\mid\sigma'(y,R)\rangle$$

and

$$\langle\sigma'\mid\sigma_R\rangle=\langle\sigma'\mid\sigma_{RR}\rangle=\langle\sigma'\mid S\rangle=0$$

because of parity. Equation (3.13), the equation of motion for $\ddot{X}$, is the same equation that the Dirac bracket theory yields for $\ddot{X}$ and we have obtained it with one integration. We now perform some further manipulations on Eq. (3.13) to put it in a more convenient form.

We note that when the linear operator $\mathcal{L}$ operates to the left it gives 0 since σ' is an eigenfunction of $\mathcal{L}$ with eigenvalue 0. Also, $\langle\sigma'\mid S_{,\sigma}\chi\rangle=\langle S'\mid\chi\rangle=-\langle S\mid\chi'\rangle$. Therefore, Eq. (3.13) becomes

$$\gamma\ddot{X}M_X(R)+\gamma\dot{X}\dot{R}\frac{dM_X(R)}{dR}=\left\langle\sigma'\left|\frac{\partial^2\chi}{\partial t^2}\right.\right\rangle+\langle\chi'\,|\,S\rangle-\left\langle\sigma'\left|\gamma\ddot{X}\chi'+2\gamma\dot{X}\frac{\partial\chi'}{\partial t}-\mathcal{N}\right.\right\rangle \tag{3.14}$$

or

$$\gamma\ddot{X}M_X(R)+\gamma\dot{X}\dot{R}\frac{dM_X(R)}{dR}+\gamma\ddot{X}\langle\sigma'\,|\,\chi'\rangle+\gamma\dot{X}\left\langle\sigma'\left|\frac{\partial\chi'}{\partial t}\right.\right\rangle+\gamma\dot{X}\dot{R}\langle\sigma'_R\,|\,\chi'\rangle$$

$$=\left\langle\sigma'\left|\frac{\partial^2\chi}{\partial t^2}\right.\right\rangle+\langle\chi'\,|\,S\rangle-\gamma\dot{X}\left\langle\sigma'\left|\frac{\partial\chi'}{\partial t}\right.\right\rangle+\gamma\dot{X}\dot{R}\langle\sigma'_R\,|\,\chi'\rangle+\langle\sigma'\,|\,\mathcal{N}\rangle\ . \tag{3.15}$$

The term $\gamma\dot{X}\dot{R}\langle\sigma'_R\,|\,\chi'\rangle$ has been added to both sides of Eq. (3.15). The left-hand side of Eq. (3.15) can be written as a total time derivative:

$$\frac{d}{dt}[\gamma M_X(R)\dot{X}+\gamma\langle\sigma'\,|\,\chi'\rangle\dot{X}]\ . \tag{3.16}$$

Integrating the last term in Eq. (3.16) by parts and using

$$b_X=\frac{\langle\sigma''\,|\,\chi\rangle}{M_X(R)}$$

gives finally

$$\frac{d}{dt}[\gamma M_X(R)(1-b_X)\dot{X}]=\langle\chi'\,|\,S\rangle-\gamma\dot{X}\left\langle\sigma'\left|\frac{\partial\chi'}{\partial t}\right.\right\rangle+\gamma\dot{X}\dot{R}\langle\sigma'_R\,|\,\chi'\rangle-\dot{R}\left\langle\sigma'_R\left|\frac{\partial\chi}{\partial t}\right.\right\rangle-\frac{d}{dt}[\dot{R}\langle\sigma'_R\,|\,\chi\rangle]+\langle\sigma'\,|\,\mathcal{N}\rangle \tag{3.17}$$

for the equation of motion for X in a more convenient form where we have eliminated $\langle\sigma'\,|\,\partial^2\chi/\partial t^2\rangle$ by using the fact that second time derivative of the constraint $C_X=\langle\sigma'\,|\,\chi\rangle$ vanishes, namely:

$$\ddot{C}_X=\ddot{R}\langle\sigma'_R\,|\,\chi\rangle+\dot{R}^{\,2}\langle\sigma'_{RR}\,|\,\chi\rangle+2\dot{R}\left\langle\sigma'_R\left|\frac{\partial\chi}{\partial t}\right.\right\rangle+\left\langle\sigma'\left|\frac{\partial^2\chi}{\partial t^2}\right.\right\rangle=\dot{R}\left\langle\sigma'_R\left|\frac{\partial\chi}{\partial t}\right.\right\rangle+\frac{d}{dt}[\dot{R}\langle\sigma'_R\,|\,\chi\rangle]+\left\langle\sigma'\left|\frac{\partial^2\chi}{\partial t^2}\right.\right\rangle=0\ .$$

A similar calculation yields the equation of motion for R. Multiplying Eq. (3.12) by σ_R and integrating gives

$$-\ddot{R}M_R(R)-\tfrac{1}{2}\dot{R}^{\,2}\frac{dM_R(R)}{dR}+\langle\sigma_R\,|\,S\rangle=\left\langle\sigma_R\left|\frac{\partial^2\chi}{\partial t^2}+\mathcal{L}\chi\right.\right\rangle-\frac{\partial}{\partial R}\langle\chi\,|\,S\rangle+\left\langle\sigma_R\left|\mathcal{N}-\gamma\ddot{X}\chi'-2\gamma\dot{X}\frac{\partial\chi'}{\partial t}\right.\right\rangle\ , \tag{3.18}$$

where $M_R(R)=\langle\sigma_R\,|\,\sigma_R\rangle$ and $\langle\sigma_R\,|\,\sigma'_R\rangle=0$ because of parity. We have used $\langle\sigma_R\,|\,S_{,\sigma}\chi\rangle=(\partial/\partial R)\langle\chi\,|\,S\rangle$ to arrive at Eq. (3.18) which is the equation of motion for R. We perform some further manipulations on Eq. (3.18) to put it in a more convenient form. We write Eq. (3.18) as

$$\frac{d}{dt}[M_R(R)(1-b_R)\dot{R}]+\frac{\partial}{\partial R}V_{\text{eff}}(R)=\gamma\ddot{X}\langle\sigma_R\,|\,\chi'\rangle+2\gamma\dot{X}\left\langle\sigma_R\left|\frac{\partial\chi'}{\partial t}\right.\right\rangle+\langle\sigma_R\,|\,\mathcal{L}\chi\rangle+\dot{R}\left\langle\sigma_{RR}\left|\frac{\partial\chi}{\partial t}\right.\right\rangle-\langle\sigma_R\,|\,\mathcal{N}\rangle\ , \tag{3.19}$$

where

$$\frac{\partial}{\partial R}V_{\text{eff}}(R)=-\langle\sigma_R\,|\,S\rangle-\frac{\partial}{\partial R}[\tfrac{1}{2}M_R(R)\dot{R}^{\,2}+\langle\chi\,|\,S\rangle] \tag{3.20}$$

and

$$b_R=\frac{\langle\sigma_{RR}\,|\,\chi\rangle}{M_R(R)}\ .$$

We have eliminated $\langle\sigma_R\,|\,\partial^2\chi/\partial t^2\rangle$ in deriving Eq. (3.19) from Eq. (3.18) by using the fact that second time derivative of the constraint $C_R=\langle\sigma_R\,|\,\chi\rangle$ vanishes, namely

$$\ddot{C}_R=\ddot{R}\langle\sigma_{RR}\,|\,\chi\rangle+\dot{R}^{\,2}\langle\sigma_{RRR}\,|\,\chi\rangle+2\dot{R}\left\langle\sigma_{RR}\left|\frac{\partial\chi}{\partial t}\right.\right\rangle+\left\langle\sigma_R\left|\frac{\partial^2\chi}{\partial t^2}\right.\right\rangle$$

$$=\dot{R}\left\langle\sigma_{RR}\left|\frac{\partial\chi}{\partial t}\right.\right\rangle+\frac{d}{dt}[\dot{R}\langle\sigma_{RR}\,|\,\chi\rangle]+\left\langle\sigma_R\left|\frac{\partial^2\chi}{\partial t^2}\right.\right\rangle=0\ . \tag{3.21}$$

Equation (3.19) is the equation of motion for R in a more convenient form.

Before deriving the equation of motion for the radiation field χ, we give an argument that shows it is actually not necessary to do so and that the Dirac equations of motion are indeed equivalent to the three equations of motion (3.12), (3.17), and (3.19).

Consider the symbolic set of equations:

$$E=0\ , \tag{3.22a}$$

$$\mathcal{P}E=0\ , \tag{3.22b}$$

$$(1-\mathcal{P})E=0\ , \tag{3.22c}$$

where Eq. (3.22a) represents the original equations of motion with the *Ansatz* substituted in; Eq. (3.22b), the equation from which the equations of motion for the collective variables may be derived by operating with the N shape modes; and Eq. (3.22c), the equation of motion for the field χ. In general, these three equations correspond, respectively, to Eqs. (2.55a), (2.58), and (2.54). Equations (3.22b) and (3.22c) are the Dirac equations of motion, that is Eqs. (3.22b) and (3.22c) are those equations of motion that the Dirac bracket method yields. [For the DSG case, Eq. (3.22a) corresponds to Eq. (3.12) with all terms brought over to one side.]

Now substituting, say, Eq. (3.22b) into Eq. (3.22c) leads to Eq. (3.22a). Therefore, Eq. (3.22a) may appear to be a consequence of the Dirac equations of motion; but it is not because it does not contain all the information that the Dirac equations of motion contain since it lacks any information about the constraints. It is only the *form* of Eq. (3.22a) that may be obtained by substituting one of the Dirac equations of motion into the other, but Eq. (3.22a) must be accompanied by the constraints (as auxiliary conditions, say) which must be satisfied for all time so that the solution set of Eq. (3.22a) *plus constraints* is completely equivalent to the solution set of the Dirac equations of motion, namely Eqs. (3.22b) and (3.22c). In fact, this is the content of Eqs. (2.55a) and (2.55b) where we required that the *form* of Eq. (2.55a), the original equations of motion with the *Ansatz* substituted in (i.e., our starting point), be recovered from the Dirac equations of motion and where the constraints were applied as auxiliary conditions.

In short, any *pair* of the three equations in (3.22a)–(3.22c) is a necessary and sufficient system of equations whose solution set is to be rigorously equivalent to the solution set of the Dirac equations, since the equation not included in the pair is always derivable from the two that are, and therefore, the effective application of the constraints via operating with $\mathcal{P}$ is insured. [We point out that Eq. (3.22b) is actually equivalent to the N distinct second-order equations of motion for the collective variables that are derivable from it.]

We apply this general argument to the DSG case. The equations of motion for X [Eq. (3.17)] and R [Eq. (3.19)] are, in general, derivable from operating on Eq. (3.12) with $\mathcal{P}$ followed by the corresponding shape modes. Therefore, Eqs. (3.17) and (3.19) are equivalent to Eq. (3.22b) and so the set of equations, (3.12), (3.17), and (3.19) completely determine the solution of the DSG system in terms of the new variables. Therefore, we do not need to calculate an equation of motion for the field χ by operating on Eq. (3.12) with $(1-\mathcal{P})$ according to the general statement immediately following Eq. (2.59).

It is instructive, however, to calculate the form of the projection operator, $\mathcal{P}$, although we did not have to utilize it explicitly. First, we recall the definition of the projection operator and adapt the notation for a continuum description. Recalling Eq. (2.41) we let

$$\mathcal{P}_{ls}\to\mathcal{P}(y,y')\ ,$$

where $y=\gamma(x-X)$. Suppressing the integration variables for brevity we find

$$\mathcal{P}=(\,|\sigma'\rangle\,|\sigma_R\rangle)\begin{bmatrix}\dfrac{1}{M_X(R)} & 0\\ 0 & \dfrac{1}{M_R(R)}\end{bmatrix}\begin{bmatrix}\langle\sigma'|\\ \langle\sigma_R|\end{bmatrix}$$

$$=\frac{|\sigma'\rangle\langle\sigma'|}{M_X(R)}+\frac{|\sigma_R\rangle\langle\sigma_R|}{M_R(R)} \tag{3.23}$$

and identify the projection operators $\mathcal{P}_X$ and $\mathcal{P}_R$ with

$$\mathcal{P}_X=\frac{|\sigma'\rangle\langle\sigma'|}{M_X(R)},\quad \mathcal{P}_R=\frac{|\sigma_R\rangle\langle\sigma_R|}{M_R(R)}\ . \tag{3.24}$$

Thus we see that, because of the diagonal form of the matrix $\mathbf{M}^{-1}$, the projection operator $\mathcal{P}$ splits up into the two projection operators in Eq. (3.24) and so $\mathcal{P}=\mathcal{P}_X+\mathcal{P}_R$.

In this section, we have used the general theory developed in Sec. III to obtain the equations of motion for X [Eq. (3.17)], for R [Eq. (3.19)], and for χ [Eq. (3.12) in the sense described above] by projections of Eq. (3.12) [or equivalently for Eq. (3.5)]. Equation (3.5) was obtained by substituting the *Ansatz,* Eq. (3.1a) into Eq. (3.4). Without the projection operator equivalence to the Dirac bracket method, it would be necessary to carry out an extremely lengthy calculation of all the Dirac brackets (each made up of many Poisson brackets) in order to arrive at the Dirac bracket equations of motion. We note that the system of equations for $\ddot{\chi}$, $\ddot{X}$, and $\ddot{R}$ [namely, Eqs. (3.14), (3.15), and (3.16)] in Ref. 7 are generalized and corrected by the system of equations comprised of Eqs. (3.12), (3.17), and (3.19) in the present paper.

IV. CHOICE OF *ANSATZ* FOR σ

The constrained Hamiltonian Dirac bracket formulation of this paper allows one to choose any *Ansatz* for σ in

$$\phi(x,t)=\sigma(x-X_1,X_2,\ldots,X_N)+\chi(x-X_1,t)\ , \tag{4.1}$$

where only the center-of-mass collective variable appears in χ. We drop the γ for brevity. (Usually, σ will be a stationary solution of the original problem.) However, for each collective variable, the constraints must take the form

$$\left\langle\frac{\partial\sigma}{\partial X_i}\right|\chi\Big\rangle=0\ , \tag{4.2a}$$

$$\left\langle \frac{\partial \sigma}{\partial X_i} \middle| \pi \right\rangle = 0 , \tag{4.2b}$$

i.e., the constraint must be a function of the derivative of σ with respect to the collective variable. If this is not true, the projection operator method is not equivalent to the Dirac bracket formalism—the projection-operator method breaks down. There is no requirement that the shape mode $\partial\sigma/\partial X_i$ be a bound eigenstate of the linearized equation for ϕ. We show, however, that it is advantageous to have the shape mode $\partial\sigma/\partial X_i$ equivalent to a bound state of the linearized equation for ϕ. Then we show that when there is a bound state, ψ_i, of the linearized equation for ϕ, we can find a σ such that

$$\frac{\partial \sigma}{\partial X_i} = \psi_i . \tag{4.3}$$

We illustrate the general argument by considering the particular example of a kink that is invariant under spatial translations, e.g., the continuum SG soliton. Then our *Ansatz*, Eq. (4.1), takes the form:

$$\phi(x,t) = \sigma(x - X(t)) + \chi(x - X(t),t) . \tag{4.4}$$

Consider an infinitesimal displacement $X(t) \to X(t) + \delta X$. Then we can construct an exact solution with $\chi = 0$ for small δX:

$$\phi(x,t) = \sigma(x - X(t)) + \delta X \frac{\partial \sigma(x - X(t))}{\partial X} \tag{4.5}$$

because the shape mode $\partial\sigma/\partial X$ is an exact solution of the linearized SG equation with eigenvalue zero, i.e.,

$$\frac{\partial \sigma}{\partial X} = \psi \sim \text{sech}[x - X(t)] .$$

The eigenfunction ψ is often referred to as the Goldstone mode.

Next consider the case of the DSG kink where we have

$$\phi(x,t,X,R) = \sigma_{\text{DSG}}(x - X(t), R(t)) + \chi(x - X(t),t) . \tag{4.6}$$

The translation of the center of mass $X(t) \to X(t) + \delta X$ leads, in the same manner as above, to a solution for ϕ with $\chi = 0$. If instead we consider a small deviation of the collective variable, $R(t) \to R(t) + \delta R$, then ϕ becomes

$$\begin{aligned}\phi(x,t,X,R) &= \sigma_{\text{DSG}}(x - X(t), \mathcal{R}) \\ &\quad + \delta R \frac{\partial \sigma_{\text{DSG}}}{\partial R}(x - X(t), \mathcal{R}) \\ &\quad + \chi(x - X(t),t) .\end{aligned} \tag{4.7}$$

When the parameter $\mathcal{R}$ is sufficiently large (where, for practical purposes, sufficiently larger means $\mathcal{R} \geq 2$), the function $\partial\sigma_{\text{DSG}}/\partial R$ approaches ψ_2, where ψ_2 is the second of the two exact bound eigenstates[7-9] of the linearized DSG kink equation for ϕ. Consequently, for large $\mathcal{R}$, we have a solution for ϕ for small δR even when $\chi = 0$ in complete analogy with the Goldstone mode case. However, when $\mathcal{R}$ is small ($\mathcal{R} \leq 2$), $\partial\sigma_{\text{DSG}}/\partial R$ begins to differ appreciably from ψ_2 and so $\partial\sigma_{\text{DSG}}/\partial R$ is not an exact eigenstate. If the constrained Hamiltonian Dirac bracket theory of this paper is to be valid in the limit $\delta R \to 0$ for small $\mathcal{R}$, we must have a nonvanishing χ in order that ϕ is a solution.

The nonvanishing $\chi(x,t)$ is a nonpropagating phonon dressing of σ_{DSG}. By inspection of Eq. (4.7) the resulting solution for χ has to be

$$\chi = \left[\psi_2 - \frac{\partial \sigma_{\text{DSG}}}{\partial R} \right] \delta R \tag{4.8}$$

in order that ϕ is the correct solution of the DSG kink equation in the limit $\delta R \to 0$, where the δR term is the nonpropagating part of χ. We expect χ to be nonzero for large nonlinear deviations of the collective variables. However, it is a disadvantage to have to dress even small deviations. (Parenthetically, often σ gives a good representation of many features of a problem[10] for appreciable nonlinear deviations of the collective variables even when $\chi = 0$.)

We now show that we can find a new *Ansatz* for σ in those problems where exact bound states exist such that we have a solution for ϕ with $\chi = 0$ in the limit of small deviations of the collective variables. As an example we consider the DSG kink with a new *Ansatz* $\hat{\sigma}$.

$$\phi(x,t) = \hat{\sigma}_{\text{DSG}}(x - X, R) + \chi(x - X, t) , \tag{4.9}$$

where

$$\hat{\sigma}_{\text{DSG}}(x - X, R) \equiv \sigma_{\text{DSG}}(x - X, \mathcal{R}) + \int_{\mathcal{R}}^{R} \psi_2(R') dR' , \tag{4.10}$$

which has the consequence that the shape mode is the exact eigenstate of the linear problem, i.e.,

$$\frac{\partial \hat{\sigma}_{\text{DSG}}}{\partial R} = \psi_2(R) . \tag{4.11}$$

Consequently, in the limit $\delta R \to 0$ we have a solution

$$\phi(x,t) = \sigma_{\text{DSG}}(x - X, \mathcal{R}) + \delta R \, \psi_2(x - X, \mathcal{R}) \tag{4.12}$$

with $\chi = 0$. Generally,

$$\begin{aligned}\phi(x,t) = \hat{\sigma} + \chi &= \sigma(x - X_1^0, X_2^0, \ldots, X_N^0) \\ &\quad + \sum_{i=1}^{N} \int_{X_i^0}^{X_i} \psi_i(X_i') dX_i' + \chi ,\end{aligned} \tag{4.13a}$$

where χ may or may not be a function of X_1. In the discrete case

$$\begin{aligned}Q_l = \hat{f}_l + q_l &= f(l - X_1^0, X_2^0, \ldots, X_N^0) \\ &\quad + \sum_{i=1}^{N} \int_{X_i^0}^{X_i} \psi_i(X_i') dX_i' + q_l ,\end{aligned} \tag{4.13b}$$

where $q_l = q_l(t)$ and the ψ_i are the eigenfunctions of the linearized discrete equations of motion.

The eigenfunctions, ψ_i, are solutions of the linearized, homogeneous equation for ϕ, and therefore their normalization is not determined. The normalization of the ψ_i is determined by the physics of each case. For example, in the DSG system, for large $\mathcal{R}$ we know that for small oscillations

$$\phi=\sigma_{\rm DSG}((x-X_1),\mathcal{R})+\delta R\frac{\partial\sigma_{\rm DSG}}{\partial R}((x-X_1),\mathcal{R})$$

with $\chi=0$. Consequently, we have the requirement that $\partial\hat{\sigma}_{\rm DSG}/\partial R=\psi_2$ must approach $\partial\sigma_{\rm DSG}(R)/\partial R$ as $\mathcal{R}$ becomes large. Therefore, we first define $\tilde{\psi}_2\equiv N_2^{-1/2}\psi_2$ (where $N_2^{-1/2}$ is a normalization factor) and require $\tilde{\psi}_2$ to satisfy $\langle\tilde{\psi}_2|\tilde{\psi}_2\rangle=1$. The normalization factor is then given by $N_2^{1/2}=\langle\tilde{\psi}_2|\psi_2\rangle$. We must now normalize ψ_2 so that $\psi_2\rightarrow\partial\sigma_{\rm DSG}(R)/\partial R$ as $\mathcal{R}$ becomes large. One way to accomplish this is by requiring

$$\langle\tilde{\psi}_2|\psi_2\rangle=\left\langle\tilde{\psi}_2\left|\frac{\partial\sigma_{\rm DSG}(R)}{\partial R}\right.\right\rangle \tag{4.14a}$$

which leads to

$$\langle\psi_2|\psi_2\rangle=|\langle\tilde{\psi}_2|\psi_2\rangle|^2 =\left|\left\langle\tilde{\psi}_2\left|\frac{\partial\sigma_{\rm DSG}(R)}{\partial R}\right.\right\rangle\right|^2 . \tag{4.14b}$$

As another example, suppose we wish to normalize the eigenfunction for the linearized equation for ϕ for the discrete SG system. Then the normalization is determined by the requirement that the lowest frequency bound state in the discrete problem, ψ_1, approach

$$g_1=1/l_0{\rm sech}[\pi/l_0(l-X)]$$

which is the discretization of the continuum bound state eigenfunction, which agrees well with ψ_1 when l_0 is large. The normalization of ψ_1 for the discrete SG kink can then be invoked by using $\langle\tilde{\psi}_1|\psi_1\rangle=\langle\tilde{\psi}_1|g_1\rangle$.

In Ref. 6 we will show using the appropriate $\hat{\sigma}$ for the discrete SG lattice that we get exact agreement with molecular dynamics simulations for the small oscillations Peierls-Nabarro frequency of a trapped kink. We observe the $\hat{\sigma}$ *Ansatz* is more complicated than the σ *Ansatz* and consequently have to do more work in evaluating quantities that depend on $\hat{\sigma}$. However, in practice, the goal is to choose an *Ansatz* σ that embodies as much physics as possible (including nonlinear variations of the collective variables) so that the more complicated equations for χ, such as Eq. (3.12), can hopefully be treated perturbatively. For example, we have found many situations[10,20] where the radiation which is described by χ and π is very small even though the collective variables in σ undergo very large nonlinear oscillations. In conclusion, all of the derivations in the preceding sections of this paper remain valid when σ is replaced by $\hat{\sigma}$.

V. DISCUSSION AND CONCLUSION

In this paper we have developed a projection operator approach for treating nonlinear field theories in which there exist collective modes. We introduced N coordinates to characterize the N collective modes, thereby introducing $2N$ extra degrees of freedom into the system. $2N$ second-class constraints were applied to conserve the number of degrees of freedom of the original system. The system in terms of the new variables together with the constraints was then able to be treated within the framework of the Dirac bracket formalism in which a family (with respect to the function h_l) of canonical transformations to the new coordinates and momenta was derived.

We showed that the lengthy procedure of deriving the $2(N+M)$ coupled equations of motion for the new coordinates and momenta can be circumvented by substituting the *Ansatz* [Eq. (2.17) or (2.61)] into the original equations of motion and operating appropriately with the projection operator $\mathcal{P}_{ls}$ defined by Eq. (2.41). Thus, the second-order equations of motion for the collective variables X_i and the field q_l may be derived without having to explicitly work through a variational procedure in terms of the new variables. The variation of the Lagrangian with respect to the old variables is all that is required which simply gives $\ddot{Q}_l+V'(Q_l)=0$ which are the equations of motion into which the *Ansatz* must be substituted. Such a powerful simplification originates from the general structure of the Dirac bracket which, as was stated earlier, defines a projection in symplectic space.

We must reemphasize the difference between the projection operators **P** and $\mathcal{P}_{ln}$. **P** is the projection operator in symplectic space defined by Eq. (2.29) in terms of which the Dirac bracket may *always* be written. When the equations in symplectic space are decomposed and written in the more usual function space [see, for example, Eqs. (2.49a)–(2.49d)], **P** generaly splits up into a sum of terms, one of which will lead directly to the projection operator $\mathcal{P}_{ln}$ [such as the first term in parentheses on the right-hand side of Eq. (A7)], *only if the constraints are of the form we have assumed them to be.* Therefore, the existence of **P** (which *does not* depend on the form of the constraints) does not imply the existence of $\mathcal{P}_{ln}$ (which *does* depend on the form of the constraints).

We showed in detail the derivation of the coupled equations of motion for a DSG system where the usefulness of the projection operator formalism was made apparent. In that example we included γ in our *Ansatz* and derived the equations of motion in the approximation $\dot{\gamma}=0$. For many problems, this approximation will give satisfactory results. However, in a highly relativistic kink-kink or kink-antikink collision, the $\dot{\gamma}$ terms become important and cannot be neglected. One must then substitute the full *Ansatz* (with γ) into the Lagrangian of Eq. (3.3). The Lagrangian will be a function of $\ddot{X}_1$ and the equation of motion for the coordinate X_1 is given by the Euler-Lagrange equation:

$$\frac{d^2}{dt^2}\left[\frac{\partial L}{\partial\ddot{X}_1}\right]-\frac{d}{dt}\left[\frac{\partial L}{\partial\dot{X}_1}\right]+\frac{\partial L}{\partial X_1}=0 , \tag{5.1}$$

which leads to a fourth-order equation for X, i.e., d^4X/dt^4. In this case, there appears a new constraint in addition to the ones already given for say, the SG case: the inner product of $\partial\sigma/\partial\gamma$ with χ which leads to

$$C_\gamma=\int_{-\infty}^{\infty}y\sigma'(y)\chi(y,t)dy\approx 0 ,$$

i.e., the first moment of the product $\sigma'\chi$, which measures the kink's distortion as it interacts with an impurity or another kink, for example. We are currently investigating further the consequences of the more exact equation of motion for X_1, Eq. (5.1).

We have motivated a particular choice for the *Ansatz* $\sigma=\hat{\sigma}$ that has the property of Eq. (4.3) which we have shown to be advantageous since $\sigma=\hat{\sigma}$ is a solution for the linearized equation of motion for $\phi=\hat{\sigma}+\chi$ in Eq. (4.1) in the limit $\delta X_i \to 0$ and therefore $\chi=0$ in the same limit. Also, since the constraints must be defined with respect to the *Ansatz* according to Eqs. (4.2a) and (4.2b) in order that the projection operator method be equivalent to the Dirac bracket formalism, then setting $\sigma=\hat{\sigma}$ amounts to requiring that the shape mode, even for $\delta X_i \to 0$, is orthogonal to the radiation eigenfunctions.

ACKNOWLEDGMENT

One of the authors (P.S.) would like to thank R. Flesch for pointing out an error in a previous paper, the correction of which prompted much of the discussion in this paper on the canonical transformation.

APPENDIX A

In this appendix we first find the differential equations that must be satisfied by the old momenta in terms of the new variables by requiring the Dirac brackets to be invariant under the canonical transformation of the old variables to the new. Then we solve the resulting differential equations to obtain the form of the canonical transformation. We show that the functional dependence of the old momenta on the new coordinates and momenta is defined to within a function h_l which must obey certain conditions which we calculate explicitly.

We use Eq. (2.37)

$$\{A,B\}^* = \left[\frac{\partial A}{\partial \mathbf{n}}\right]^T \left[\mathbf{I}+\mathbf{J}\frac{\partial C_{1\alpha}}{\partial \mathbf{n}} M^{-1}_{\alpha\beta}\left[\frac{\partial C_{2\beta}}{\partial \mathbf{n}}\right]^T - \mathbf{J}\frac{\partial C_{2\alpha}}{\partial \mathbf{n}} M^{-1}_{\alpha\beta}\left[\frac{\partial C_{1\beta}}{\partial \mathbf{n}}\right]^T\right]\mathbf{J}\frac{\partial B}{\partial \mathbf{n}} \tag{2.37}$$

to derive the differential equations which must be satisfied by the old momenta as functions of the new coordinates and momenta. These differential equations are

$$(\delta_{ls}-\mathcal{P}_{ls})\frac{\partial P_n}{\partial p_s}+\mathbf{f}_l'^T(\mathbf{I}-\mathbf{b})\frac{\partial P_n}{\partial \mathbf{Y}}=\delta_{ln}\ , \tag{A1a}$$

$$\frac{\partial P_l}{\partial q_r}\frac{\partial P_n}{\partial p_r}+\left[\frac{\partial P_l}{\partial \mathbf{X}}\right]^T\frac{\partial P_n}{\partial \mathbf{Y}}+\left[\frac{\partial P_l}{\partial p_s}\mathbf{f}_s'^T+\left[\frac{\partial P_l}{\partial \mathbf{Y}}\right]^T\mathbf{C}_1'^T\right]\mathbf{M}^{-1}\left[\mathbf{f}_r'\frac{\partial P_n}{\partial q_r}-\mathbf{C}_2'\frac{\partial P_n}{\partial \mathbf{Y}}\right]-(l\leftrightarrow n)=0\ , \tag{A1b}$$

where $(l\leftrightarrow n)$ symbolizes the entire preceding expression with the indices l and n interchanged.

Before deriving Eqs. (A1a) and (A1b), it is helpful to explicitly show what the matrix quantities in Eq. (2.37) are:

$$\frac{\partial C_{1\alpha}}{\partial \mathbf{n}}=\begin{bmatrix} f_{1,\alpha}\\ \vdots\\ f_{M,\alpha}\\ C_{1\alpha,1}\\ \vdots\\ C_{1\alpha,N}\\ 0\\ \vdots\\ 0\\ 0\\ \vdots\\ 0\end{bmatrix},\quad \mathbf{J}\frac{\partial C_{1\alpha}}{\partial \mathbf{n}}=\begin{bmatrix} 0\\ \vdots\\ 0\\ 0\\ \vdots\\ 0\\ -f_{1,\alpha}\\ \vdots\\ -f_{M,\alpha}\\ -C_{1\alpha,1}\\ \vdots\\ -C_{1\alpha,N}\end{bmatrix}, \tag{A2}$$

$$\frac{\partial C_{2\alpha}}{\partial \mathbf{n}}=\begin{bmatrix} 0\\ \vdots\\ 0\\ C_{2\alpha,1}\\ \vdots\\ C_{2\alpha,N}\\ f_{1,\alpha}\\ \vdots\\ f_{M,\alpha}\\ 0\\ \vdots\\ 0\end{bmatrix},\quad \mathbf{J}\frac{\partial C_{2\alpha}}{\partial \mathbf{n}}=\begin{bmatrix} f_{1,\alpha}\\ \vdots\\ f_{M,\alpha}\\ 0\\ \vdots\\ 0\\ 0\\ \vdots\\ 0\\ -C_{2\alpha,1}\\ \vdots\\ -C_{2\alpha,N}\end{bmatrix}, \tag{A3}$$

$$\left[\frac{\partial Q_l}{\partial \mathbf{n}}\right]^T=(\delta_{l1},\ldots,\delta_{lM},f_{l,1},\ldots,f_{l,N},0,\ldots,0,0,\ldots,0)\ . \tag{A4}$$

For the transformation to be canonical, we require that Eqs. (2.38a)–(2.38c) be satisfied. We see that

$$\left[\frac{\partial Q_l}{\partial \mathbf{n}}\right]^T\mathbf{J}\frac{\partial Q_n}{\partial \mathbf{n}}=\left[\frac{\partial Q_l}{\partial \mathbf{n}}\right]^T\mathbf{J}\frac{\partial C_{1\alpha}}{\partial \mathbf{n}} =\left[\frac{\partial C_{1\beta}}{\partial \mathbf{n}}\right]^T\mathbf{J}\frac{\partial Q_n}{\partial \mathbf{n}}=0 \tag{A5}$$

by using Eqs. (A2) and (A4). Substituting the expressions in Eq. (A5) into Eq. (2.37) we obtain $\{Q_l,Q_n\}^*=0$. Next, we require $\{Q_l,P_n\}^*=\delta_{ln}$. We substitute Q_l for A and P_n for B in Eq. (2.37). When we then evaluate the second term in parentheses in Eq. (2.37), we obtain zero by virtue of Eq. (A5). Evaluating the third term in parentheses of Eq. (2.37) we find

$$\left[\frac{\partial Q_l}{\partial \mathbf{n}}\right]^T\mathbf{J}\frac{\partial C_{2\alpha}}{\partial \mathbf{n}}=\delta_{ls}f_{s,\alpha}\ , \tag{A6a}$$

$$\left[\frac{\partial C_{1\beta}}{\partial \mathbf{n}}\right]^T\mathbf{J}\frac{\partial P_n}{\partial \mathbf{n}}=f_{s,\beta}\frac{\partial P_n}{\partial p_s}+C_{1\beta,i}\frac{\partial P_n}{\partial Y_i}\ . \tag{A6b}$$

Substituting these results back into Eq. (2.37) we obtain

$$\{Q_l,P_n\}^*=\delta_{ls}\frac{\partial P_n}{\partial p_s}+f_{l,i}\frac{\partial P_n}{\partial Y_i} -\delta_{ls}f_{s,\alpha}M^{-1}_{\alpha\beta}\left[f_{r,\beta}\frac{\partial P_n}{\partial p_r}+C_{1\beta,i}\frac{\partial P_n}{\partial Y_i}\right] . \tag{A7}$$

Using Eq. (2.41) and $b_{\alpha i}=M^{-1}_{\alpha\beta}C_{1\beta,i}$ in Eq. (A7) we obtain

$$\{Q_l,P_n\}^*=(\delta_{ls}-\mathcal{P}_{ls})\frac{\partial P_n}{\partial p_s}+f_{l,\alpha}(\delta_{\alpha i}-b_{\alpha i})\frac{\partial P_n}{\partial Y_i} , \tag{A8}$$

In matrix notation, the right-hand side of Eq. (A8) becomes the left-hand side of Eq. (2.39) [or Eq. (A1a)].

Calculation of the bracket $\{P_l,P_n\}^*$ is made simpler by noticing that the third term in Eq. (2.37) is equal to the second term but with A and B interchanged. To show this, take the third term

$$\left[\left[\frac{\partial A}{\partial \mathbf{n}}\right]^T \mathbf{J}\frac{\partial C_{2\alpha}}{\partial \mathbf{n}}\right]M^{-1}_{\alpha\beta}\left[\left[\frac{\partial C_{1\beta}}{\partial \mathbf{n}}\right]^T \mathbf{J}\frac{\partial B}{\partial \mathbf{n}}\right]$$

and interchange the expressions in parentheses (which are "inner products") while at the same time taking their transpose (which changes nothing) to give

$$\left[\frac{\partial B}{\partial \mathbf{n}}\right]^T \mathbf{J}^T\frac{\partial C_{1\beta}}{\partial \mathbf{n}}M^{-1}_{\alpha\beta}\left[\frac{\partial C_{2\alpha}}{\partial \mathbf{n}}\right]^T \mathbf{J}^T\frac{\partial A}{\partial \mathbf{n}} .$$

Using $\mathbf{J}^T=-\mathbf{J}$ and $M^{-1}_{\alpha\beta}=M^{-1}_{\beta\alpha}$, we obtain the desired result. Since $A=P_l$ and $B=P_n$ this amounts to interchanging l and n. Calculating anew the components of the second term in Eq. (2.37) we find

$$\left[\frac{\partial P_l}{\partial \mathbf{n}}\right]^T \mathbf{J}\frac{\partial C_{1\alpha}}{\partial \mathbf{n}}=-\frac{\partial P_l}{\partial p_s}f_{s,\alpha}-\frac{\partial P_l}{\partial Y_i}C_{1\alpha,i} , \tag{A9a}$$

$$\left[\frac{\partial C_{2\beta}}{\partial \mathbf{n}}\right]^T \mathbf{J}\frac{\partial P_n}{\partial \mathbf{n}}=C_{2\beta,j}\frac{\partial P_n}{\partial Y_j}-f_{r,\beta}\frac{\partial P_n}{\partial q_r} , \tag{A9b}$$

and finally

$$\{P_l,P_n\}^*=\frac{\partial P_l}{\partial q_s}\frac{\partial P_n}{\partial p_s}+\frac{\partial P_l}{\partial X_i}\frac{\partial P_n}{\partial Y_i}+\left[\frac{\partial P_l}{\partial p_s}f_{s,\alpha}+\frac{\partial P_l}{\partial Y_i}C_{1\alpha,i}\right]M^{-1}_{\alpha\beta}\left[f_{r,\beta}\frac{\partial P_n}{\partial q_r}-C_{2\beta,j}\frac{\partial P_n}{\partial Y_j}\right]-(l\leftrightarrow n) . \tag{A10}$$

In matrix notation, the right-hand side of Eq. (A10) becomes the left-hand side of Eq. (A1b).

We now find the general solution of Eqs. (A1a) and (A1b). Operating on Eq. (A1a) with $\mathbf{f}'_l$ from the left yields

$$\mathbf{M}(\mathbf{I}-\mathbf{b})\frac{\partial P_n}{\partial \mathbf{Y}}=\mathbf{f}'_n \tag{A11}$$

since

$$\mathbf{f}'_l(\delta_{ls}-\mathcal{P}_{ls})=0 . \tag{A12}$$

Solving for P_n gives

$$\frac{\partial P_n}{\partial \mathbf{Y}}=(\mathbf{I}-\mathbf{b})^{-1}\mathbf{M}^{-1}\mathbf{f}'_n , \tag{A13a}$$

$$P_n=\mathbf{Y}^T(\mathbf{I}-\mathbf{b})^{-1}\mathbf{M}^{-1}\mathbf{f}'_n+g_n(q_r,p_r,\mathbf{X}) , \tag{A13b}$$

where g_n is a completely arbitrary function of the indicated variables. Next operate again on Eq. (A1a) from the left with $(\delta_{rl}-\mathcal{P}_{rl})$ to obtain

$$(\delta_{rs}-\mathcal{P}_{rs})\frac{\partial P_n}{\partial p_s}=(\delta_{rn}-\mathcal{P}_{rn})$$

or

$$(\delta_{rs}-\mathcal{P}_{rs})\left[\frac{\partial P_n}{\partial p_s}-\delta_{sn}\right]=0 . \tag{A14}$$

This allows us to write

$$\frac{\partial P_n}{\partial p_s}=\delta_{sn}+\mathbf{f}'^T_s\mathbf{d}_n(q_r,p_r,\mathbf{X},\mathbf{Y},t) , \tag{A15a}$$

where $\mathbf{d}_n$ is an arbitrary column matrix. We note that $\mathbf{f}'^T_s\mathbf{d}_n=\mathbf{d}^T_n\mathbf{f}'_s$. [The last term in Eq. (A15a) is zero when operated on with $(\delta_{rs}-\mathcal{P}_{rs})$.] When we integrate Eq. (A15a) with respect to p_s, the second term on the right-hand side will be proportional to $\mathbf{C}_2$ which we strongly set to zero in the presence of the Dirac bracket. Thus $\mathbf{d}_n$ does not enter into the momentum transformation. We may then write

$$P_n=p_n+k_n(q_r,\mathbf{X},\mathbf{Y}) \tag{A15b}$$

for some function k_n. Consistency between Eqs. (A13b) and (A15b) requires

$$P_n=p_n+\mathbf{Y}^T(\mathbf{I}-\mathbf{b})^{-1}\mathbf{M}^{-1}\mathbf{f}'_n+h_n(q_r,\mathbf{X}) . \tag{2.43'}$$

So far there are no conditions on h_l or $\mathbf{d}_l$ but requiring $\{P_l,P_n\}^*=0$, however, puts conditions on h_l and $\mathbf{d}_l$.

Before using Eq. (A1b) to find the conditions on h_l and $\mathbf{d}_l$, we simplify Eq. (A1b). Consider the terms in the first set of parentheses of Eq. (A1b) and use Eq. (A15a) to obtain

$$\left[\frac{\partial P_l}{\partial p_s}\mathbf{f}_s'^T+\left[\frac{\partial P_l}{\partial \mathbf{Y}}\right]^T\mathbf{C}_1'^T\right]\mathbf{M}^{-1}=\mathbf{f}_l'^T\mathbf{M}^{-1}+\mathbf{d}_l^T+\left[\frac{\partial P_l}{\partial \mathbf{Y}}\right]^T\mathbf{C}_1'^T\mathbf{M}^{-1}$$

$$=\mathbf{f}_l'^T\mathbf{M}^{-1}(\mathbf{I}-\mathbf{b}^T)^{-1}(\mathbf{I}-\mathbf{b}^T)+\left[\frac{\partial P_l}{\partial \mathbf{Y}}\right]^T\mathbf{b}^T+\mathbf{d}_l^T$$

$$=\left[\frac{\partial P_l}{\partial \mathbf{Y}}\right]^T(\mathbf{I}-\mathbf{b}^T)+\left[\frac{\partial P_l}{\partial \mathbf{Y}}\right]^T\mathbf{b}^T+\mathbf{d}_l^T=\left[\frac{\partial P_l}{\partial \mathbf{Y}}\right]^T+\mathbf{d}_l^T\ . \tag{A16}$$

Substituting Eq. (A16) back into Eq. (A1b) gives

$$\frac{\partial P_l}{\partial q_r}\frac{\partial P_n}{\partial p_r}+\left[\frac{\partial P_l}{\partial \mathbf{X}}\right]^T\frac{\partial P_n}{\partial \mathbf{Y}}+\left[\left[\frac{\partial P_l}{\partial \mathbf{Y}}\right]^T+\mathbf{d}_l^T\right]\left[\mathbf{f}_r'\frac{\partial P_n}{\partial q_r}-\mathbf{C}_2'\frac{\partial P_n}{\partial \mathbf{Y}}\right]-(l\leftrightarrow n)=0\ . \tag{A17}$$

We note that the term

$$\left[\frac{\partial P_l}{\partial \mathbf{Y}}\right]^T\mathbf{C}_2'\frac{\partial P_n}{\partial \mathbf{Y}}$$

in Eq. (A17) is symmetric in the indices l and n and therefore vanishes by virtue of $(l\leftrightarrow n)$. Using Eq. (A15a) to substitute for $\mathbf{d}_l^T\mathbf{f}'$ in Eq. (A17), we find that Eq. (A17) simplifies to

$$\{P_l,P_n\}^*=\frac{\partial P_l}{\partial q_n}-\frac{\partial P_n}{\partial X_\alpha}\frac{\partial P_l}{\partial Y_\alpha}+\frac{\partial P_n}{\partial q_r}\mathbf{f}_r'^T\frac{\partial P_l}{\partial \mathbf{Y}}$$
$$+\mathbf{d}_n^T\mathbf{C}_2'\frac{\partial P_l}{\partial \mathbf{Y}}-(l\leftrightarrow n)=0\ , \tag{A18}$$

where we have written the second term in Eq. (A18) in component form. The reason is that taking derivatives as indicated by Eq. (A18) would lead to three-dimensional matrices whose manipulation would be cumbersome in the matrix notation we have been using. Also we have interchanged the indices l and n in the second term and have written the inner product in the third term in reverse order. This will make certain forthcoming cancellations easier to identify.

We now calculate the quantities on the right-hand side of Eq. (A18). Until stated otherwise we take $\mathbf{d}_n=0$. We use

$$\frac{\partial}{\partial R}[\mathbf{A}(R)]^{-1}=-[\mathbf{A}(R)]^{-1}\frac{\partial \mathbf{A}(R)}{\partial R}[\mathbf{A}(R)]^{-1} \tag{A19}$$

to calculate the derivative of the inverse of a matrix $\mathbf{A}$.

We rewrite Eq. (2.43) as

$$P_l=P_l^0+h_l\ , \tag{A20a}$$

$$P_l^0=p_l+\mathbf{Y}^T(\mathbf{I}-\mathbf{b})^{-1}\mathbf{M}^{-1}\mathbf{f}_l'\ , \tag{A20b}$$

and show that P_l^0 satisfies $\{P_l^0,P_n^0\}^*=0$ when $\mathbf{d}_n=0$. The quantities in Eq. (A18) are calculated as follows:

$$\frac{\partial P_l^0}{\partial q_n}=\mathbf{Y}^T(\mathbf{I}-\mathbf{b})^{-1}\frac{\partial \mathbf{b}}{\partial q_n}(\mathbf{I}-\mathbf{b})^{-1}\mathbf{M}^{-1}\mathbf{f}_l'$$
$$=\mathbf{Y}^T(\mathbf{I}-\mathbf{b})^{-1}\mathbf{M}^{-1}\mathbf{f}_n''(\mathbf{I}-\mathbf{b})^{-1}\mathbf{M}^{-1}\mathbf{f}_l'$$
$$=\mathbf{Y}^T(\mathbf{I}-\mathbf{b})^{-1}\mathbf{M}^{-1}\mathbf{f}_n''\frac{\partial P_l}{\partial \mathbf{Y}}\ . \tag{A21}$$

Inserting the identity

$$\mathbf{I}=\mathbf{f}_s'\mathbf{f}_s'^T\mathbf{M}^{-1} \tag{A22}$$

gives

$$\frac{\partial P_l^0}{\partial q_n}=\mathbf{Y}^T(\mathbf{I}-\mathbf{b})^{-1}\mathbf{M}^{-1}\mathbf{f}_s'\mathbf{f}_s'^T\mathbf{M}^{-1}\mathbf{f}_n''\frac{\partial P_l}{\partial \mathbf{Y}}$$
$$=(P_s^0-p_s)\mathbf{f}_s'^T\mathbf{M}^{-1}\mathbf{f}_n''\frac{\partial P_l}{\partial \mathbf{Y}}$$
$$=P_s^0\mathbf{f}_s'^T\mathbf{M}^{-1}\mathbf{f}_n''\frac{\partial P_l}{\partial \mathbf{Y}}\ , \tag{A23}$$

where the second equality follows from Eqs. (A20b) and the third from $\mathbf{C}_2=0$.

We now calculate

$$\frac{\partial P_l^0}{\partial X_\alpha}=\mathbf{Y}^T(\mathbf{I}-\mathbf{b})^{-1}\frac{\partial \mathbf{b}}{\partial X_\alpha}(\mathbf{I}-\mathbf{b})^{-1}\mathbf{M}^{-1}\mathbf{f}_l'-\mathbf{Y}^T(\mathbf{I}-\mathbf{b})^{-1}\mathbf{M}^{-1}\frac{\partial \mathbf{M}}{\partial X_\alpha}\mathbf{M}^{-1}\mathbf{f}_l'+\mathbf{Y}^T(\mathbf{I}-\mathbf{b})^{-1}\mathbf{M}^{-1}\frac{\partial \mathbf{f}_l'}{\partial X_\alpha}$$
$$=\mathbf{Y}^T(\mathbf{I}-\mathbf{b})^{-1}\left[-\mathbf{M}^{-1}\frac{\partial \mathbf{M}}{\partial X_\alpha}\mathbf{M}^{-1}\mathbf{C}_1'+\mathbf{M}^{-1}\frac{\partial \mathbf{C}_1'}{\partial X_\alpha}\right]\frac{\partial P_l}{\partial \mathbf{Y}}-\mathbf{Y}^T(\mathbf{I}-\mathbf{b})^{-1}\mathbf{M}^{-1}\frac{\partial \mathbf{M}}{\partial X_\alpha}\mathbf{M}^{-1}\mathbf{f}_l'+\mathbf{Y}^T(\mathbf{I}-\mathbf{b})^{-1}\mathbf{M}^{-1}\frac{\partial \mathbf{f}_l'}{\partial X_\alpha}$$
$$=-P_s^0\mathbf{f}_s'^T\mathbf{M}^{-1}\frac{\partial \mathbf{M}}{\partial X_\alpha}\mathbf{b}\frac{\partial P_l}{\partial \mathbf{Y}}+P_s^0\mathbf{f}_s'^T\mathbf{M}^{-1}\frac{\partial \mathbf{C}_1'}{\partial X_\alpha}\frac{\partial P_l}{\partial \mathbf{Y}}-P_s^0\mathbf{f}_s'^T\mathbf{M}^{-1}\frac{\partial \mathbf{M}}{\partial X_\alpha}\mathbf{M}^{-1}\mathbf{f}_l'+P_s^0\mathbf{f}_s'^T\mathbf{M}^{-1}\frac{\partial \mathbf{f}_l'}{\partial X_\alpha}\ , \tag{A24}$$

where the last equality follows from insertion of Eq. (A22) in the same manner as for Eq. (A23).

The third term in Eq. (A24) may be written

$$-P_s^0 \mathbf{f}_s'^T \mathbf{M}^{-1} \frac{\partial \mathbf{M}}{\partial X_\alpha} \mathbf{M}^{-1} \mathbf{f}_l' = -P_s^0 \mathbf{f}_s'^T \mathbf{M}^{-1} \frac{\partial \mathbf{M}}{\partial X_\alpha} (\mathbf{I}-\mathbf{b})(\mathbf{I}-\mathbf{b})^{-1} \mathbf{M}^{-1} \mathbf{f}_l' = -P_s^0 \mathbf{f}_s'^T \mathbf{M}^{-1} \frac{\partial \mathbf{M}}{\partial X_\alpha} (\mathbf{I}-\mathbf{b}) \frac{\partial P_l}{\partial \mathbf{Y}} . \tag{A25}$$

The term proportional to **b** in Eq. (A25) cancels the first term in Eq. (A24) and so we have

$$\frac{\partial P_l^0}{\partial X_\alpha} = P_s^0 \mathbf{f}_s'^T \mathbf{M}^{-1} \frac{\partial \mathbf{C}_1'}{\partial X_\alpha} \frac{\partial P_l}{\partial \mathbf{Y}} - P_s^0 \mathbf{f}_s'^T \mathbf{M}^{-1} \frac{\partial \mathbf{M}}{\partial X_\alpha} \frac{\partial P_l}{\partial \mathbf{Y}} + P_s^0 \mathbf{f}_s'^T \mathbf{M}^{-1} \frac{\partial \mathbf{f}_l'}{\partial X_\alpha} . \tag{A26}$$

We now have expressions for $\partial P_l^0/\partial q_n$ [Eq. (A23)] and $\partial P_l^0/\partial X_\alpha$ [Eq. (A26)]. Substituting both of these into Eq. (A18) (with $\mathbf{d}_n=0$) yields

$$\{P_l^0,P_n^0\}^* = P_s^0 \mathbf{f}_s'^T \mathbf{M}^{-1} \mathbf{f}_n'' \frac{\partial P_l}{\partial \mathbf{Y}} - \left[P_s^0 \mathbf{f}_s'^T \mathbf{M}^{-1} \frac{\partial \mathbf{C}_1'}{\partial X_\alpha} \frac{\partial P_n}{\partial \mathbf{Y}} - P_s^0 \mathbf{f}_s'^T \mathbf{M}^{-1} \frac{\partial \mathbf{M}}{\partial X_\alpha} \frac{\partial P_n}{\partial \mathbf{Y}} + P_s^0 \mathbf{f}_s'^T \mathbf{M}^{-1} \frac{\partial \mathbf{f}_n'}{\partial X_\alpha} \right] \frac{\partial P_l}{\partial Y_\alpha}$$
$$+ P_s^0 \mathbf{f}_s'^T \mathbf{M}^{-1} \mathbf{f}_r'' \frac{\partial P_n}{\partial \mathbf{Y}} \mathbf{f}_r'^T \frac{\partial P_l}{\partial \mathbf{Y}} - (l \leftrightarrow n) . \tag{A27}$$

We see that the first term of Eq. (A27) cancels the last term in large parentheses. The first term in large parentheses is symmetric in the indices l and n and so it cancels itself by virtue of $(l \leftrightarrow n)$. That leaves us with

$$\{P_l^0,P_n^0\}^* = P_s^0 \mathbf{f}_s'^T \mathbf{M}^{-1} \frac{\partial \mathbf{M}}{\partial X_\alpha} \frac{\partial P_n}{\partial \mathbf{Y}} \frac{\partial P_l}{\partial Y_\alpha} + P_s^0 \mathbf{f}_s'^T \mathbf{M}^{-1} \mathbf{f}_r'' \frac{\partial P_n}{\partial \mathbf{Y}} \mathbf{f}_r'^T \frac{\partial P_l}{\partial \mathbf{Y}} - (l \leftrightarrow n)$$
$$= P_s^0 \mathbf{f}_s'^T \mathbf{M}^{-1} \left[\frac{\partial \mathbf{f}_r'}{\partial X_\alpha} \mathbf{f}_r'^T + \mathbf{f}_r' \frac{\partial \mathbf{f}_r'^T}{\partial X_\alpha} \right] \frac{\partial P_n}{\partial \mathbf{Y}} \frac{\partial P_l}{\partial Y_\alpha} + P_s^0 \mathbf{f}_s'^T \mathbf{M}^{-1} \mathbf{f}_r'' \frac{\partial P_n}{\partial \mathbf{Y}} \mathbf{f}_r'^T \frac{\partial P_l}{\partial \mathbf{Y}} - (l \leftrightarrow n) ,$$

where we have expanded $\partial \mathbf{M}/\partial X_\alpha$ in the first term. The second term in parentheses is symmetric in l and n as is the *sum* of the remaining two terms and so the right-hand side is zero. Therefore $\{P_l^0,P_n^0\}^*=0$ when $\mathbf{d}_n=0$. If we relax the condition $\mathbf{d}_n=0$ we merely retain the last term in Eq. (A18) and so $\{P_l^0,P_n^0\}^*$ becomes

$$\{P_l^0,P_n^0\}^* = \mathbf{d}_n \mathbf{C}_2' \frac{\partial P_l^0}{\partial \mathbf{Y}} - (l \leftrightarrow n) . \tag{A28}$$

To obtain the general condition which $h_l(q_r,\mathbf{X})$ and $\mathbf{d}_n$ must obey we substitute Eq. (A20) into Eq. (A18). Since we have just proved that P_l^0 satisfies Eq. (A28) we are left with

$$\frac{\partial h_l}{\partial q_n} - \left[\frac{\partial h_n}{\partial \mathbf{X}} \right]^T \frac{\partial P_l^0}{\partial \mathbf{Y}} + \frac{\partial h_n}{\partial q_r} \mathbf{f}_r'^T \frac{\partial P_l^0}{\partial \mathbf{Y}} + \mathbf{d}_n^T \mathbf{C}_2' \frac{\partial P_l^0}{\partial \mathbf{Y}} - (l \leftrightarrow n) = 0 . \tag{A29}$$

Equation (A29) is satisfied if we take $h_l=0$ and $\mathbf{d}_n=0$.

For the transformation defined by Eqs. (2.61) and (2.63a) and (2.63b) where the center of mass appears in the $\mathcal{X}$ field, we find that through an analogous calculation we can derive conditions that must be satisfied by $h(x,\mathbf{X},\mathcal{X})$ and $\mathbf{d}(x,t,\mathbf{X},\mathbf{P},\mathcal{X},\pi)$. We will abbreviate $h(x,\mathbf{X},\mathcal{X})$, $\mathbf{d}(x,t,\mathbf{X},\mathbf{P},\mathcal{X},\pi)$, $\mathcal{X}(x-X_1,t)$, and $\sigma(x-X_1,X_2,\ldots,X_N)$ by $h(x,t)$, $d(x,t)$, $\mathcal{X}(x,t)$, and $\sigma(x)$, respectively. If we define $\Pi^0(x,t)$ by

$$\Pi(x,t) = \Pi^0(x,t) + h(x,t) , \tag{A30a}$$

where

$$\Pi^0(x,t) = \pi + \left[\mathbf{Y}^T - \int \pi \left[\frac{\partial \mathcal{X}}{\partial \mathbf{X}} \right]^T dx' \right] (\mathbf{I}-\mathbf{b})^{-1} \mathbf{M}^{-1} \frac{\partial \sigma}{\partial \mathbf{X}} \tag{A30b}$$

then we are able to show

$$\{\Pi^0(x,t),\Pi^0(y,t)\}^* = \mathbf{d}^T(y,t) \mathbf{C}_2' \frac{\partial \Pi^0(x,t)}{\partial \mathbf{Y}} - (x \leftrightarrow y) . \tag{A31}$$

We find, as above, that requiring $\{\Pi(x,t),\Pi(y,t)\}^*=0$, leads to

$$\frac{\partial h(x)}{\partial \mathcal{X}} - \left[\frac{\partial h(y)}{\partial \mathbf{X}} \right]^T \frac{\partial \Pi^0(x,t)}{\partial \mathbf{Y}} + \left[\int \frac{\partial h(y,t)}{\partial \mathcal{X}(z,t)} \left[\frac{\partial \sigma(z)}{\partial \mathbf{X}} \right]^T dz \right] \frac{\partial \Pi^0(x,t)}{\partial \mathbf{Y}}$$
$$+ \left[\int \frac{\partial h(y,t)}{\partial \mathcal{X}(z,t)} \left[\frac{\partial \mathcal{X}(z,t)}{\partial \mathbf{X}} \right]^T dz \right] \frac{\partial \Pi^0(x,t)}{\partial \mathbf{Y}} + \mathbf{d}^T(y,t) \mathbf{C}_2' \frac{\partial \Pi^0(x,t)}{\partial \mathbf{Y}} - (x \leftrightarrow y) = 0 \tag{A32}$$

which can be satisfied by the choices $h=0$ and $\mathbf{d}=\mathbf{0}$. We notice the presence of the fourth term in Eq. (A32) due to the dependence of $\mathcal{X}$ on X_1. This term is not present in Eq. (A29).

APPENDIX B

We use Eq. (2.48) to derive the Dirac equations of motion Eqs. (2.49a)–(2.49d). We recall Eq. (2.48):

$$\dot{\mathbf{n}}_D=\left[\mathbf{I}+\mathbf{J}\frac{\partial C_{1\alpha}}{\partial \mathbf{n}}M_{\alpha\beta}^{-1}\left[\frac{\partial C_{2\beta}}{\partial \mathbf{n}}\right]^T -\mathbf{J}\frac{\partial C_{2\alpha}}{\partial \mathbf{n}}M_{\alpha\beta}^{-1}\left[\frac{\partial C_{1\beta}}{\partial \mathbf{n}}\right]^T\right]\mathbf{J}\frac{\partial H}{\partial \mathbf{n}} . \tag{2.48}$$

Using Eqs. (A2) and (A3) from Appendix A we find

$$\left[\frac{\partial C_{2\beta}}{\partial \mathbf{n}}\right]^T\mathbf{J}\frac{\partial H}{\partial \mathbf{n}}=C_{2\beta,i}\frac{\partial H}{\partial Y_i}-f_{s,\beta}\frac{\partial H}{\partial q_s} , \tag{B1}$$

$$\left[\frac{\partial C_{1\beta}}{\partial \mathbf{n}}\right]^T\mathbf{J}\frac{\partial H}{\partial \mathbf{n}}=f_{s,\beta}\frac{\partial H}{\partial p_s}+C_{1\beta,i}\frac{\partial H}{\partial Y_i} . \tag{B2}$$

Substituting Eqs. (B1) and (B2) into Eq. (2.48) and using Eqs. (A2) and (A3) and the definition of $\mathbf{n}_D$ whose transpose is given by Eq. (2.22), Eq. (2.48) may be put in the following form:

$$\begin{bmatrix}\dot q_1\\ \vdots\\ \dot q_M\\ \dot X_1\\ \vdots\\ \dot X_N\\ \dot p_1\\ \vdots\\ \dot p_M\\ \dot Y_1\\ \vdots\\ \dot Y_N\end{bmatrix}=\begin{bmatrix}\partial H/\partial p_1\\ \vdots\\ \partial H/\partial p_M\\ \partial H/\partial Y_1\\ \vdots\\ \partial H/\partial Y_N\\ -\partial H/\partial q_1\\ \vdots\\ -\partial H/\partial q_M\\ -\partial H/\partial X_1\\ \vdots\\ -\partial H/\partial X_N\end{bmatrix}+\begin{bmatrix}0\\ \vdots\\ 0\\ 0\\ \vdots\\ 0\\ -f_{1,\alpha}\\ \vdots\\ -f_{M,\alpha}\\ -C_{1\alpha,1}\\ \vdots\\ -C_{1\alpha,N}\end{bmatrix}M_{\alpha\beta}^{-1}\left[C_{2\beta,i}\frac{\partial H}{\partial Y_i}-f_{s,\beta}\frac{\partial H}{\partial q_s}\right]-\begin{bmatrix}f_{1,\alpha}\\ \vdots\\ f_{M,\alpha}\\ 0\\ \vdots\\ 0\\ 0\\ \vdots\\ 0\\ -C_{2\alpha,1}\\ \vdots\\ -C_{2\alpha,N}\end{bmatrix}M_{\alpha\beta}^{-1}\left[f_{s,\beta}\frac{\partial H}{\partial p_s}+C_{1\beta,i}\frac{\partial H}{\partial Y_i}\right] . \tag{B3}$$

The equations of motion can now be read off of the above matrix equation and written concisely as

$$\dot q_l=\frac{\partial H}{\partial p_l}-\mathbf{f}_l'^T\mathbf{M}^{-1}\left[\mathbf{f}_s'\frac{\partial H}{\partial p_s}+\mathbf{C}_1'\frac{\partial H}{\partial \mathbf{Y}}\right] , \tag{B4a}$$

$$\dot p_l=-\frac{\partial H}{\partial q_l}-\mathbf{f}_l'^T\mathbf{M}^{-1}\left[\mathbf{C}_2'\frac{\partial H}{\partial \mathbf{Y}}-\mathbf{f}_s'\frac{\partial H}{\partial q_s}\right] , \tag{B4b}$$

$$\dot{\mathbf{X}}=\frac{\partial H}{\partial \mathbf{Y}} , \tag{B4c}$$

$$\dot{\mathbf{Y}}=-\frac{\partial H}{\partial \mathbf{X}}-\mathbf{C}_1'^T\mathbf{M}^{-1}\left[\mathbf{C}_2'\frac{\partial H}{\partial \mathbf{Y}}-\mathbf{f}_s'\frac{\partial H}{\partial q_s}\right] +\mathbf{C}_2'^T\mathbf{M}^{-1}\left[\mathbf{f}_s'\frac{\partial H}{\partial p_s}+\mathbf{C}_1'\frac{\partial H}{\partial \mathbf{Y}}\right] . \tag{B4d}$$

We now manipulate Eqs. (B4a)–(B4d) into the form of Eqs. (2.49a)–(2.49d). Operating on Eqs. (B4a) and (B4b) with $\mathbf{f}_l'$ gives, respectively,

$$\mathbf{f}_l'\dot q_l=-\mathbf{C}_1'\frac{\partial H}{\partial \mathbf{Y}}=-\mathbf{C}_1'\dot{\mathbf{X}} , \tag{B5a}$$

$$\mathbf{f}_l'\dot p_l=-\mathbf{C}_2'\frac{\partial H}{\partial \mathbf{Y}}=-\mathbf{C}_2'\dot{\mathbf{X}} , \tag{B5b}$$

Bringing everything onto the left-hand side of these two equations gives

$$\mathbf{f}_l'\dot q_l+\mathbf{C}_1'\dot{\mathbf{X}}=\dot{\mathbf{C}}_1=0 , \tag{B6a}$$

$$\mathbf{f}_l'\dot p_l+\mathbf{C}_2'\dot{\mathbf{X}}=\dot{\mathbf{C}}_2=0 , \tag{B6b}$$

Equations (B6a) and (B6b) verify that the constraints are indeed constants of the motion.

Using Eqs. (B5a) and (B5b) to eliminate the explicit constraint terms in Eqs. (B4a) and (B4b) leads to

$$\dot q_l=\frac{\partial H}{\partial p_l}-\mathbf{f}_l'^T\mathbf{M}^{-1}\mathbf{f}_s'\frac{\partial H}{\partial p_s}+\mathbf{f}_l'^T\mathbf{M}^{-1}\mathbf{f}_s'\dot q_s , \tag{B7a}$$

$$\dot p_l=-\frac{\partial H}{\partial q_l}+\mathbf{f}_l'^T\mathbf{M}^{-1}\mathbf{f}_s'\dot p_s+\mathbf{f}_l'^T\mathbf{M}^{-1}\mathbf{f}_s'\frac{\partial H}{\partial q_s} , \tag{B7b}$$

and these equations may be written in the following form:

$$(\delta_{ls}-\mathcal{P}_{ls})\left[\dot q_s-\frac{\partial H}{\partial p_s}\right]=0 , \tag{B8a}$$

$$(\delta_{ls}-\mathcal{P}_{ls})\left[\dot p_s+\frac{\partial H}{\partial q_s}\right]=0 , \tag{B8b}$$

where we see $\mathcal{P}_{ls}$ is the projection operator defined in Eq. (2.41). Equations (B8a) and (B8b) are Eqs. (2.49a) and (2.49b).

Equation (B4c) is a comparatively simple equation. This is because the constraints are independent of $\mathbf{Y}$ and so the terms containing the constraints in Eq. (B3) (for the X_i rows) have zero as their coefficients.

When we use Eqs. (B5a) and (B5b) to eliminate the terms $\mathbf{C}_1'\partial H/\partial\mathbf{Y}$ and $\mathbf{C}_2'\partial H/\partial\mathbf{Y}$, respectively, in Eq. (B4d) we obtain

$$\dot{\mathbf{Y}}+\frac{\partial H}{\partial \mathbf{X}}=\mathbf{C}_1'^T\mathbf{M}^{-1}\mathbf{f}_l'\left[\frac{\partial H}{\partial q_l}+\dot{p}_l\right]+\mathbf{C}_2'^T\mathbf{M}^{-1}\mathbf{f}_1'\left[\frac{\partial H}{\partial p_l}-\dot{q}_l\right] \tag{B9}$$

which is identical to Eq. (2.49d).

So far we have derived the Dirac equations of motion for the canonical variables. We end this appendix with a derivation of Eqs. (2.60a)–(2.60c) which project out the N unmeaningful degrees of freedom from the equation for the conservation of energy. First, multiply Eq. (B8a) by $\dot{p}_l$ and Eq. (B8b) by $\dot{q}_l$ and subtract. This lead directly to

$$\frac{\partial H}{\partial q_l}(\delta_{ln}-\mathcal{P}_{ln})\dot{q}_n+\frac{\partial H}{\partial p_l}(\delta_{ln}-\mathcal{P}_{ln})\dot{p}_n=0 \;. \tag{2.60a}$$

Writing this as

$$\frac{\partial H}{\partial q_l}\dot{q}_l+\frac{\partial H}{\partial p_l}\dot{p}_l=\frac{\partial H}{\partial q_l}\mathcal{P}_{ln}\dot{q}_n+\frac{\partial H}{\partial p_l}\mathcal{P}_{ln}\dot{p}_n \tag{B10}$$

and substituting into

$$\left[\frac{\partial H}{\partial \dot{X}}\right]^T\dot{\mathbf{X}}+\left[\frac{\partial H}{\partial \mathbf{Y}}\right]^T\dot{\mathbf{Y}}+\frac{\partial H}{\partial q_l}\dot{q}_l+\frac{\partial H}{\partial p_l}\dot{p}_l=\frac{dH}{dt}=0$$

which must *always* be true (even on the manifold in phase space that subsumes the trajectories of the N unmeaningful degrees of freedom), we obtain

$$\left[\frac{\partial H}{\partial \mathbf{X}}\right]^T\dot{\mathbf{X}}+\left[\frac{\partial H}{\partial \mathbf{Y}}\right]^T\dot{\mathbf{Y}}+\frac{\partial H}{\partial q_l}\mathcal{P}_{ln}\dot{q}_n+\frac{\partial H}{\partial p_l}\mathcal{P}_{ln}\dot{p}_n=0 \;. \tag{2.60b}$$

APPENDIX C

We show in this appendix that the Dirac equations of motion

$$(\delta_{ln}-\mathcal{P}_{ln})\left[\dot{q}_n-\frac{\partial H}{\partial p_n}\right]=0 \;,$$

$$(\delta_{ln}-\mathcal{P}_{ln})\left[\dot{p}_n+\frac{\partial H}{\partial q_n}\right]=0 \;,$$

may be written as

$$(\delta_{ln}-\mathcal{P}_{ln})\left[\dot{Q}_n-\frac{\partial H}{\partial P_n}\right]=0 \;, \tag{C1a}$$

$$(\delta_{ln}-\mathcal{P}_{ln})\left[\dot{P}_n+\frac{\partial H}{\partial Q_n}\right]=0 \;, \tag{C1b}$$

First, we write the fundamental *Ansatz* Eq. (2.17) as

$$q_n=Q_n-f_n \tag{C2}$$

and take its time derivative:

$$\dot{q}_n=\dot{Q}_n-\mathbf{f}_n'^T\dot{\mathbf{X}} \;. \tag{C3}$$

Second, we write

$$\frac{\partial H}{\partial p_n}=\frac{\partial H}{\partial P_s}\frac{\partial P_s}{\partial p_n}=\frac{\partial H}{\partial P_s}\delta_{sn}=\frac{\partial H}{\partial P_n} \tag{C4}$$

and substitute Eqs. (C3) and (C4) into Eq. (2.49a) which gives Eq. (C1a).

Dealing with Eq. (2.49b) is slightly more complicated. We recall the momentum transformation with $h_n=0$:

$$P_n=p_n+\mathbf{f}_n'^T\mathbf{M}^{-1}(\mathbf{I}-\mathbf{b}^T)^{-1}\mathbf{Y} \;. \tag{C5}$$

Solving this for p_n and taking the time derivative yields

$$\dot{p}_n=\dot{P}_n-\mathbf{X}^T\mathbf{f}_n''^T\mathbf{M}^{-1}(\mathbf{I}-\mathbf{b}^T)^{-1}\mathbf{Y}-\mathbf{f}_l'^T\frac{d}{dt}[\mathbf{M}^{-1}(\mathbf{I}-\mathbf{b}^T)^{-1}\mathbf{Y}] \;. \tag{C6}$$

From Eq. (2.49c) and the Hamiltonian in Eq. (2.47) we calculate

$$\dot{\mathbf{X}}^T=\mathbf{Y}^T\overline{\mathbf{M}}^{-1} \tag{C7}$$

and substituting Eq. (C7) into Eq. (C6) gives

$$\dot{p}_n=\dot{P}_n-\mathbf{Y}^T\overline{\mathbf{M}}^{-1}\mathbf{f}_n''^T\mathbf{M}^{-1}(\mathbf{I}-\mathbf{b}^T)^{-1}\mathbf{Y}-\mathbf{f}_n'^T\frac{d}{dt}[\mathbf{M}^{-1}(\mathbf{I}-\mathbf{b}^T)^{-1}\mathbf{Y}] \;. \tag{C8}$$

We also need the following expression for $\partial H/\partial q_n$:

$$\begin{aligned}\frac{\partial H}{\partial q_n}&=\tfrac{1}{2}\mathbf{Y}^T\frac{\partial\overline{\mathbf{M}}^{-1}}{\partial q_n}\mathbf{Y}+\frac{\partial V}{\partial q_n}=-\tfrac{1}{2}\mathbf{Y}^T\overline{\mathbf{M}}^{-1}\frac{\partial\overline{\mathbf{M}}}{\partial q_n}\overline{\mathbf{M}}^{-1}\mathbf{Y}+\frac{\partial V}{\partial Q_s}\frac{\partial Q_s}{\partial q_n}=-\tfrac{1}{2}\mathbf{Y}^T\overline{\mathbf{M}}^{-1}\frac{\partial}{\partial q_n}[(\mathbf{I}-\mathbf{b}^T)\mathbf{M}(\mathbf{I}-\mathbf{b})]\overline{\mathbf{M}}^{-1}\mathbf{Y}+\frac{\partial V}{\partial Q_n}\\&=\tfrac{1}{2}\mathbf{Y}^T\overline{\mathbf{M}}^{-1}\left[\frac{\partial\mathbf{b}^T}{\partial q_n}\mathbf{M}(\mathbf{I}-\mathbf{b})+(\mathbf{I}-\mathbf{b}^T)\mathbf{M}\frac{\partial\mathbf{b}}{\partial q_n}\right]\overline{\mathbf{M}}^{-1}\mathbf{Y}+\frac{\partial H}{\partial Q_n}\\&=\tfrac{1}{2}\mathbf{Y}^T\overline{\mathbf{M}}^{-1}[\mathbf{f}_n''\mathbf{M}^{-1}\mathbf{M}(\mathbf{I}-\mathbf{b})+(\mathbf{I}-\mathbf{b}^T)\mathbf{M}\mathbf{M}^{-1}\mathbf{f}_n'']\overline{\mathbf{M}}^{-1}\mathbf{Y}+\frac{\partial H}{\partial Q_n}\\&=\mathbf{Y}^T\overline{\mathbf{M}}^{-1}\mathbf{f}_n''(\mathbf{I}-\mathbf{b})\overline{\mathbf{M}}^{-1}\mathbf{Y}+\frac{\partial H}{\partial Q_n}=\mathbf{Y}^T\overline{\mathbf{M}}^{-1}\mathbf{f}_n''\mathbf{M}^{-1}(\mathbf{I}-\mathbf{b}^T)^{-1}\mathbf{Y}+\frac{\partial H}{\partial Q_n} \;.\end{aligned} \tag{C9}$$

Substituting Eqs. (C8) and (C9) into Eq. (2.49b) gives Eq. (C1b).

APPENDIX D

In this appendix we investigate modifications that need to be made of the statements in Sec. II A *if* the momentum is given by some form other than Eq. (2.12). We do not want to change the coordinate transformation, the constraint C_1, or the canonicity requirement. We may however wish to change C_2 in order to allow us to put the momentum into some desired form. There is no apparent *a priori* reason why we should not do this. For instance, we are free to choose the momentum to be

$$P_l = p_l - \frac{Pf_l'}{M}$$

(where $M = \sum_i f_i'^2$) as was incorrectly derived and used in Eq. (2.7) of Ref. 4. The requirement of canonicity determines the form of the second constraint. For this particular choice of P_l it turns out that requiring $\{Q_l, P_n\}^* = \delta_{ln}$ leads to the result that C_2 must satisfy the following *Poisson* bracket equation:

$$f_l'\{C_1, C_2\} = -M(1-\eta/M)\{q_l + f_l, C_2\} \ .$$

Explicitly writing out the derivatives and simplifying, we obtain

$$f_l' \sum_i f_i' \frac{\partial C_2}{\partial p_i} + M(1-\eta/M)\frac{\partial C_2}{\partial p_l} = -f_l' M \frac{\partial C_2}{\partial P} \ .$$

Then multiplying by f_l', summing over l, and solving for $\sum_i f_i' \partial C_2/\partial p_i$, we obtain

$$\sum_i f_i' \frac{\partial C_2}{\partial p_i} = -\frac{M}{(2-\eta/M)}\frac{\partial C_2}{\partial P} \ .$$

This equation is satisfied if we let the second constraint be

$$C_2 = \sum_i f_i'(X)p_i - P\left[2 - \frac{\eta}{M}\right] \ . \tag{D1}$$

The bracket of the constraints $\{C_1, C_2\}$ is now what we define to be the *dressed* mass, $\overline{M}$, instead of the *bare* mass, $M: \{C_1, C_2\} = M(1-\eta/M)^2 \equiv \overline{M}$.

In our paper, Ref. 4, we asserted [see Eq. (2.10) of Ref. 4] that in the new variables the canonical brackets are $\{q_l, p_n\}^* = \delta_{ln} - f_l' f_n'/M$, $\{X, P\}^* = 1$ and that *all* other brackets vanish, i.e., brackets between X and q_l, p_l and p_n, P and p_l, etc. This is not the case. We actually must explicitly calculate them all according to the Dirac prescription. We have worked them out and find the following results for the brackets.

$$\{X, P\}^* = 1/(1-\eta/M)^2$$

$$\{q_l, p_n\}^* = \delta_{ln} - \frac{f_l' f_n'}{M(1-\eta/M)^2} = \delta_{ln} - \frac{f_l' f_n'}{\overline{M}} \ ,$$

where $\overline{M}$ appears instead of M. Some of the other brackets do not vanish.

$$\{q_l, P\}^* = -\frac{f_l'\eta}{\overline{M}} \ ,$$

$$\{p_l, P\}^* = \frac{f_l'}{\overline{M}}\frac{\partial C_2}{\partial X} + \frac{f_l'' P}{\overline{M}}\frac{\eta}{M} \ ,$$

$$\{p_l, p_n\}^* = \frac{P}{M\overline{M}}(f_l'' f_n' - f_l' f_n'') \ ,$$

and

$$\{p_l, X\}^* = -\frac{f_l'}{\overline{M}}(2-\eta/M) \ .$$

The remaining brackets do vanish.

We could carry out the derivation of a rigorous Hamiltonian theory using Eq. (D1). However, we prefer to use the simpler constraint Eq. (2.5) instead of Eq. (D1); consequently using Eq. (2.12) for the momentum transformation instead of $P_l = p_l - Pf_l'/M$.

[1]E. Tomboulis, Phys. Rev. D **12**, 1178 (1975).
[2]A. Combs and S. Yip, Phys. Rev. B **28**, 6873 (1983).
[3]R. J. Flesch, S. E. Trullinger, and B. Horovitz, in *Nonlinearity in Condensed Matter,* edited by A. Bishop, D. Campbell, P. Kumar, and S. Trullinger (Springer-Verlag, Berlin, 1987).
[4]C. R. Willis, M. El-Batanouny, and P. Stancioff, Phys. Rev. B **33**, 1904 (1986).
[5]P. Stancioff, C. R. Willis, M. El-Batanouny, and S. Burdick, Phys. Rev. B **33**, 1912 (1986).
[6]R. Boesch and C. R. Willis (unpublished).
[7]D. Campbell, M. Peyrard, and P. Sodano, Physica **19D**, 165 (1986).
[8]C. R. Willis, M. El-Batanouny, S. Burdick, R. Boesch, and P. Sodano, Phys. Rev. B **35**, 3496 (1987).
[9]P. Sodano, M. El-Batanouny, and C. R. Willis, Phys. Rev. B **34**, 4936 (1986).
[10]S. Burdick, C. R. Willis, and M. El-Batanouny, Phys. Rev. B **36**, 6920 (1987).
[11]M. J. Rice, Phys. Rev. B **28**, 3587 (1983).
[12]D. Campbell, J. F. Schonfeld, and C. A. Wingate, Physica **9D**, 1-32 (1983).
[13]P. Stancioff and C. R. Willis (unpublished).
[14]P. A. M. Dirac, *Lectures on Quantum Mechanics* (Academic, New York, 1964).
[15]P. Stancioff and C. R. Willis (unpublished).
[16]R. Ravelo, M. El-Batanouny, C. R. Willis, and P. Sodano, Phys. Rev. B **38** (1988).
[17]Y. Wada and J. R. Schrieffer, Phys. Rev. B **18**, 3897 (1978).
[18]F. Marchesoni and C. R. Willis, Phys. Rev. A **36**, 4559 (1987).
[19]H. Goldstein, *Classical Mechanics,* 2nd ed. (Addison-Wesley, Reading, Mass., 1980).
[20]R. Boesch, C. R. Willis, and M. El-Batanouny (unpublished).
[21]R. Ravelo and C. R. Willis (unpublished).