SIAM J. APPL. MATH.
Vol. 59, No. 3, pp. 1112–1133

EXACTLY CONSERVATIVE INTEGRATORS*

B. A. SHADWICK†‡, JOHN C. BOWMAN†§, AND P. J. MORRISON†¶

Abstract. Traditional explicit numerical discretizations of conservative systems generically predict artificial secular drifts of any nonlinear invariants. In this work we present a general approach for developing explicit nontraditional algorithms that conserve such invariants exactly. We illustrate the method by applying it to the three-wave truncation of the Euler equations, the Lotka–Volterra predator-prey model, and the Kepler problem. The ideas are discussed in the context of symplectic (phase–space-conserving) integration methods as well as nonsymplectic conservative methods. We comment on the application of our method to general conservative systems.

Key words. conservative, integration, numerical, symplectic

AMS subject classifications. 65L05, 34-04, 34A50

PII. S0036139995289313

1. Introduction.

For many years now symplectic integrators have been the subject of much productive study. (See Channell and Scovel [6] for an overview; see also the recent book by Sanz-Serna and Calvo [25].) There are many Hamiltonian systems for which symplectic methods have proven extremely useful, if not essential; but these methods do not constitute the last word on integration techniques. As Ge and Marsden show [11], exact energy conservation is, in general, not possible with a symplectic method. Since the energy error is typically not secular but rather oscillatory, it is commonly believed that exact energy conservation is not as important a benefit as preserving the phase-space structure.

Less is known about the numerical preservation of more general constants of motion. Based on the work of Cooper [8], Sanz-Serna [25, 23] has shown that a restricted class of quadratic invariants will be conserved by certain symplectic Runge–Kutta schemes. For the Runge–Kutta methods studied by Cooper [8], conservation of quadratic invariants necessarily requires that the method be implicit. One technique for ensuring the preservation of any constant of motion is to use the constant to reduce the number of equations that must be solved. If the constants are in involution, then an entire degree of freedom (one coordinate and one momentum) can be removed from the dynamics for each such constant. This is seldom practical since the relationship between the constants of motion and a given dynamical variable may well be noninvertible (see the discussion in Gear [12]). The net result is that the reduced equations tend to be more complicated than the original system (hence the "force" terms are more expensive to compute); thus, in a system with a large number of degrees of freedom, little advantage is gained. Furthermore, if the constants of motion are not in involution, the system obtained by eliminating these invariants will

*Received by the editors November 16, 1995; accepted for publication (in revised form) March 17, 1997; published electronically March 23, 1999. This work was supported by the U.S. Department of Energy under contracts DE-FG05-80ET-53088 and PDDEFG-03-95ER-40936.
http://www.siam.org/journals/siap/59-3/28931.html
†Department of Physics and Institute for Fusion Studies, University of Texas at Austin, Austin, TX 78712–1081.
‡Present address: Physics Department, 366 Le Conte Hall University of California at Berkeley, Berkeley, CA 94720–7300 (shadwick@physics.berkeley.edu).
§Present address: Department of Mathematical Sciences, University of Alberta, Edmonton, AB, Canada T6G 2G1 (bowman@math.ualberta.ca).
¶(morrison@hagar.ph.utexas.edu).

be noncanonical [22, 21], resulting in even greater complexity.

It may be that the system of interest is most naturally described by variables that give rise to a noncanonical Hamiltonian structure. For noncanonical systems, Ge and Marsden [11] have provided a general construction for integrators that preserve both momentum maps and the structure of the Poisson manifold. Channell and Scovel [7] have shown how to implement these algorithms without the need of coordinating the configuration space group. This notwithstanding, they report that, with the exception of certain special (albeit important) forms of the Hamiltonian, such methods tend to be computationally expensive.

There is a further class of dynamical systems that is of interest, namely those systems that are not Hamiltonian (canonical or otherwise) but still possess constants of motion. Noteworthy examples of such systems are transport equations, such as the Boltzmann equation. A further example is the truncated Fourier-transformed Euler fluid equation. The untruncated equation constitutes an infinite-dimensional Hamiltonian field theory; however, when the number of Fourier modes is reduced to a finite set, the Hamiltonian structure is typically lost, even though energy and enstrophy are still conserved. (The overall effect of such truncations on the dynamics is very much an open question.) Given that these systems are not Hamiltonian, symplectic methods, per se, are of little relevance, whereas the preservation of constants of motion is still of great interest.

A variety of methods for enforcing conservation of general invariants has been proposed. Bayliss [3] and Isaacson [16] have proposed a two-stage algorithm in which the approximate solution, obtained by standard methods in the first stage, is projected in the second stage onto the constraint surface defined by the invariants. Brasey and Hairer [5] have proposed a "half-explicit" method in which the projection (via a Lagrange multiplier) and integration stages are merged together. LaBudde and Greenspan [18, 19] have developed an algorithm for central force problems that conserves both energy and angular momentum. Gear [12, 13] advocates an approach that amounts to an embedding of the original system into a higher-dimensional space, yielding a set of differential-algebraic equations, the solution of which coincides with the solution of the original equations and preserves the invariants.

Our purpose in this paper is to present another approach to the development of exactly conservative algorithms. Beginning with a simple model problem with two quadratic invariants, of interest in both fluid mechanics and plasma physics, we develop explicit integrators that conserve both invariants exactly. We then further illustrate our method by applying it to the Lotka–Volterra predator-prey model and to the Kepler problem.

2. A model problem. Our original interest in the issue of exact preservation of constants of motion arose in the study of two-dimensional inviscid fluid turbulence. As an illustration, consider the "three-wave" problem obtained by restricting the Fourier-transformed Euler equations to three modes [2, 9, 4]:

$$\frac{d\psi_K}{dt} = M_K\,\psi_P\,\psi_Q \equiv S_K(\psi)\,, \tag{1a}$$

$$\frac{d\psi_P}{dt} = M_P\,\psi_Q\,\psi_K \equiv S_P(\psi)\,, \tag{1b}$$

$$\frac{d\psi_Q}{dt} = M_Q\,\psi_K\,\psi_P \equiv S_Q(\psi)\,, \tag{1c}$$

where $\psi = (\psi_K, \psi_P, \psi_Q)$, K, P, and Q are the magnitudes of the Fourier wavenumbers of the three modes, and the mode coupling coefficients M_K, M_P, and M_Q satisfy

$$M_K + M_P + M_Q = 0 \tag{2}$$

and

$$K^2\, M_K + P^2\, M_P + Q^2\, M_Q = 0\,. \tag{3}$$

This system possess two invariants: the total energy

$$E = \frac{1}{2}\left(\psi_K^2 + \psi_P^2 + \psi_Q^2\right) \tag{4}$$

and the total enstrophy

$$Z = \frac{1}{2}\left(K^2\,\psi_K^2 + P^2\,\psi_P^2 + Q^2\,\psi_Q^2\right). \tag{5}$$

The constancy of these quantities follows directly from properties of S_k:

$$\sum_k \psi_k\, S_k = 0\,, \tag{6a}$$

$$\sum_k k^2\, \psi_k\, S_k = 0\,, \tag{6b}$$

where k ranges over the set $\{K, P, Q\}$. (These equations are identical to Euler's equations for a rigid body, in which case the second invariant is the norm of the total angular momentum.)

When (1) is integrated numerically using standard explicit methods, neither E nor Z is exactly conserved. This behavior is made apparent by applying Euler's method with a time step τ:

$$\psi_k(t+\tau) = \psi_k(t) + \tau\, S_k\,, \qquad k \in \{K, P, Q\}\,. \tag{7}$$

The energy at the new time is given by

$$\begin{aligned} E(t+\tau) &= \frac{1}{2}\sum_k \left[\psi_k(t) + \tau\, S_k\right]^2 \\ &= \frac{1}{2}\sum_k \left[\psi_k^2 + 2\,\tau\, S_k\,\psi_k + \tau^2\, S_k^2\right] \\ &= E(t) + \frac{1}{2}\,\tau^2 \sum_k S_k^2\,, \end{aligned} \tag{8}$$

where we have used (6a) in the last step. Thus the total energy is *always* increasing. A similar calculation for the enstrophy gives

$$Z(t+\tau) = Z(t) + \frac{1}{2}\,\tau^2 \sum_k k^2\, S_k^2\,, \tag{9}$$

which is likewise always increasing. For extremely long runs these results imply that a very small time step is required to keep the accumulated error down to a given level—clearly an undesirable situation.

Many authors have noted that the lack of preservation of constants of motion potentially introduces significant nonphysical effects and as such these errors are, in some sense, more important than those numerical errors that do not alter constants of motion. As de Frutos and Sanz-Serna [10] point out, one can think of the local error in a numerical integration as having two "components": one which leads to unphysical changes in the constants of motion and another which does not. When these local errors accumulate over many time steps, the former component might be significantly more harmful than the latter in that errors which lead to changes in the constants of motion can affect the *qualitative* nature of the solution, whereas other errors may only affect the *quantitative* results. (A similar observation regarding the accumulation of error in area-preserving maps has been made by Greene [15].) In essence it is suggested that nonconservative integrators have the potential to make "structural" errors in the solution—an observation which agrees well with one's physical intuition. In the context of our model problem the implication is clear: keeping the time step small enough to maintain a reasonable level of energy and enstrophy conservation is likely to use more computational resources to obtain a given accuracy in the solution than would otherwise be necessary with a conservative integrator.

Although the three-wave problem is both integrable and Hamiltonian, our ultimate interest in this problem concerns the n-wave generalization of this system, which possesses both energy and enstrophy invariants but is not Hamiltonian; hence, we are led to consider methods that do not rely on a particular geometrical structure. One might be tempted to enforce energy and enstrophy conservation by using these invariants to eliminate two modes from the dynamics. In this case the algebraic relations are simple enough to allow this, but there is a compelling physical argument against this approach. The modes are typically associated with different length scales; the choice of which modes to eliminate in favor of the invariants therefore has significant physical implications. Furthermore, putting all of the numerical error into one mode could effectively contribute a nonphysical energy and enstrophy transport between the original modes.

3. Conservative integrators for the model problem. In light of the above discussion, an algorithm that exactly conserves energy and enstrophy is clearly desirable. As we have noted in section 1, a variety of implicit methods are known that preserve quadratic invariants. While implicit methods have noteworthy stability properties, they tend to be less computationally efficient than explicit methods since they typically require multiple evaluations of the "force" terms. We therefore turn our attention to the development of explicit conservative methods for our model problem.

An elegant approach to this problem is found by borrowing from the ideas of backward error analysis [25]. The essential idea is to construct a new system of equations that, under the conventional (nonconservative) integrator, yields a conservative numerical approximation to the original equations. To this end, consider the alternative problem described by three equations of the form

$$\frac{d\psi_k}{dt} = S_k(\psi) + f_k. \tag{10}$$

Our objective is to find an f_k that guarantees exact energy and enstrophy conservation and that vanishes in the limit of small step size. The form of f_k will depend on the integration algorithm. We begin by deriving f_k for Euler's method. We then construct a second-order predictor-corrector scheme.

3.1. Euler's method. As a "proof of principle" test we develop a conservative version of Euler's method. While not particularly useful in practice, Euler's method has the advantage that the algebra associated with constructing the conservative method is quite straightforward.

Application of Euler's method to the modified system yields

$$\psi_k(t+\tau) = \psi_k(t) + \tau(\, S_k + \, f_k). \tag{11}$$

The energy at the new time,

$$\begin{aligned} E(t+\tau) &= \frac{1}{2}\sum_k \left[\psi_k(t) + \tau\,(S_k + f_k)\right]^2 \\ &= E(t) + \frac{1}{2}\sum_k \left[2\,\tau\, f_k\,\psi_k + \tau^2 (S_k + f_k)^2\right], \end{aligned} \tag{12}$$

will be conserved provided that

$$\sum_k \left[2 f_k\,\psi_k + \tau (S_k + f_k)^2\right] = 0\,. \tag{13}$$

There is considerable freedom in satisfying (13). To ensure that our discrete solution approaches the exact solution of the original differential equation in the limit $\tau \longrightarrow 0$, it is necessary that f_k vanish in this limit. That is, in the limit of an infinitesimal time step, we must recover the original integration algorithm (to first order in τ). Moreover, one would prefer that f_k not introduce additional couplings into the differential equations. In light of this observation, let us try to satisfy (13) with the more restrictive condition that each term in the sum must independently vanish:

$$2\, f_k\,\psi_k + \tau\,(S_k + f_k)^2 = 0\,. \tag{14}$$

There is an additional motivation for splitting (13) into three equations, namely, that for f_k satisfying (14), the enstrophy will also be conserved. These equations are easily solved, yielding

$$\tau\, f_k = -(\psi_k + \tau\, S_k) + \sigma_k \sqrt{\psi_k^2 + 2\,\tau\, S_k\,\psi_k}\,, \tag{15}$$

where $\sigma_k \equiv \sigma_k(t,\tau)$ is so far an unknown sign. Evaluation of (15) at $\tau = 0$ implies that $\sigma_k(t,0) = \mathrm{sgn}(\psi_k(t))$. Upon substituting (15) into the Euler integrator, (11), we obtain the following time stepping rule:

$$\psi_k(t+\tau) = \sigma_k \sqrt{\psi_k^2 + 2\,\tau\, S_k\,\psi_k}\,. \tag{16}$$

It is now clear that $\sigma_k(t,\tau)$ must in fact be the sign of $\psi_k(t+\tau)$.

If $\psi_k(t) \neq 0$, then for sufficiently small τ the sign can be expressed explicitly as $\sigma_k = \mathrm{sgn}(\psi_k(t))$. In the $\tau \longrightarrow 0$ limit, f_k then vanishes, or equivalently, (16) reduces to Euler's method:

$$\begin{aligned} \psi_k(t+\tau) &= \mathrm{sgn}(\psi_k(t))\,\sqrt{\psi_k^2 + 2\,\tau\, S_k\,\psi_k} \\ &\approx \psi_k + \tau\, S_k\,. \end{aligned} \tag{17}$$

In this case the new algorithm predicts values of $\psi_k(t+\tau)$ that are quite close to those given by Euler's method—this is exactly what one would expect. The energy

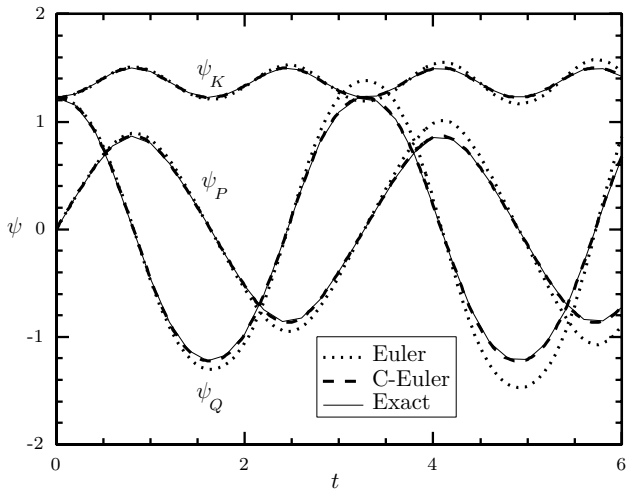

FIG. 1. *Solutions of the three-wave problem for the initial conditions* $\psi_K = \sqrt{1.5}$, $\psi_P = 0.0$, *and* $\psi_Q = \sqrt{1.5}$ *computed using the conventional Euler (Euler) and conservative Euler (C-Euler) methods, with a fixed time step of size* 0.02. *The exact solution (Exact) is also shown. The unphysical energy growth in the conventional Euler algorithm leads to large errors in the amplitudes.*

and enstrophy errors arising from (7) are the result of small (but nontrivial) errors in $\psi_k(t+\tau)$ that can be corrected by making a slight modification to the algorithm.

However, if $\psi_k(t) = 0$, it is seen from (15) that $f_k = -S_k$. Consequently, (16) has a spurious fixed point at $\psi_k(t) = 0$. Moreover, given a fixed time step τ, (17) will break down when $|\psi_k| < 2\tau\,|S_k|$. A related problem with (16) is that the argument of the radical can become negative. In this case, let us rewrite the radical as $\sqrt{\psi_k\,\chi_k}$, where $\chi_k = \psi_k + 2\,\tau\,S_k$ is just the Euler approximation for a step size of $2\,\tau$. The condition $\psi_k\,\chi_k < 0$ implies that Euler's method predicts a sign change of ψ_k between t and $t + 2\,\tau$; hence, we are in the vicinity of $\psi_k = 0$. A modification to (16), discussed in Appendix A, has been devised to circumvent these problems. We give the name "Conservative Euler" (C-Euler) to the resulting algorithm.

In Figure 1 we compare the numerical solutions of the three-wave problem obtained using the conventional Euler method with those obtained using C-Euler and with the exact solution. For these calculations $K = \sqrt{3}$, $P = 3$, $Q = \sqrt{6}$, $M_K = 1$, $M_P = 1$, and $M_Q = -2$. One can discern the effect of energy growth on the amplitudes computed by the Euler method. The errors in the two approximate solutions are shown in Figure 2.

Away from the regions where ψ_k is small, C-Euler is an explicit algorithm. We will see in the next section that the gymnastics described in Appendix A are merely a consequence of the low order of the Euler method and that a fully explicit conservative integrator is possible.

3.2. Predictor-corrector method. In practice, one would prefer to use a scheme that is of higher order than Euler's method and also has better stability properties. We now turn to a simple second-order predictor-corrector scheme, which we apply to our model problem (1):

$$\widetilde{\psi}_k = \psi_k + \tau\,S_k\,, \tag{18a}$$

$$\psi_k(t+\tau) = \psi_k + \frac{\tau}{2}\left(S_k + \widetilde{S}_k\right)\,, \tag{18b}$$

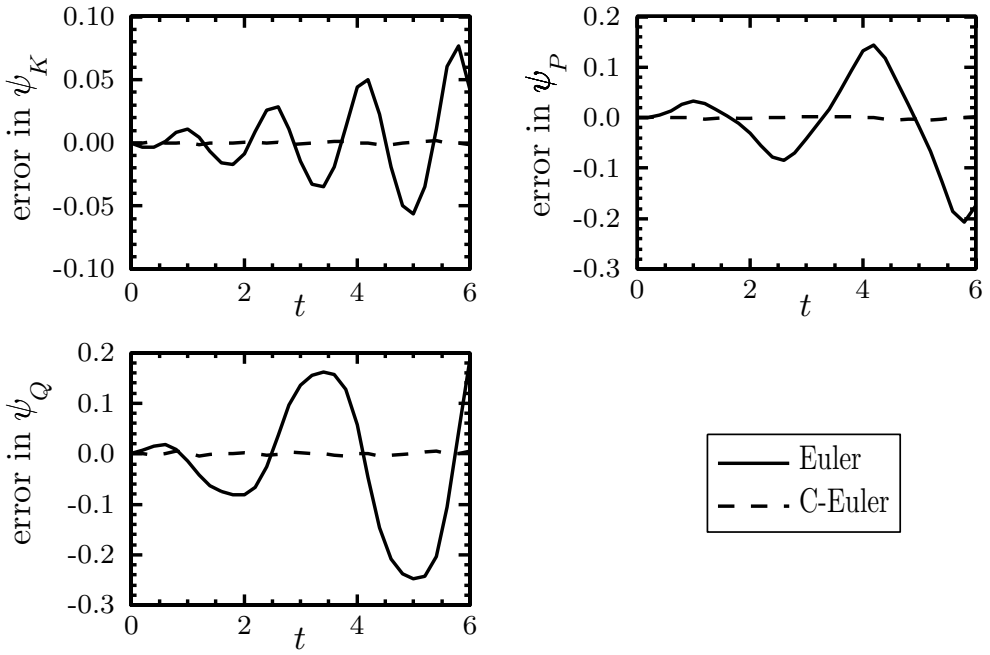

FIG. 2. *Differences between the computed and exact solutions in Fig.* 1.

where $\widetilde{S}_k = S_k(\widetilde{\psi})$. As we will show, using a second-order method overcomes the fixed-point problem that we encountered with Euler's method.

The energy now evolves according to

$$\begin{aligned} E(t+\tau) &= \frac{1}{2}\sum_k \left[\psi_k^2 + \tau\,\psi_k\left(S_k + \widetilde{S}_k\right) + \frac{\tau^2}{4}\left(S_k + \widetilde{S}_k\right)^2\right] \\ &= E(t) + \frac{1}{2}\sum_k \left[\tau(\psi_k\,S_k + \widetilde{\psi}_k\,\widetilde{S}_k) - \tau^2\,S_k\,\widetilde{S}_k + \frac{\tau^2}{4}\,(S_k + \widetilde{S}_k)^2\right] \\ &= E(t) + \frac{\tau^2}{8}\sum_k \left(S_k - \widetilde{S}_k\right)^2, \end{aligned} \tag{19}$$

where we have used the definition of $\widetilde{\psi}_k$ and the properties of S_k in the final step. A similar calculation gives

$$Z(t+\tau) = Z(t) + \frac{\tau^2}{8}\sum_k k^2\left(S_k - \widetilde{S}_k\right)^2. \tag{20}$$

Again we see that the numerical method yields an ever-increasing energy and enstrophy.[1]

[1]One might be tempted to conclude that any conventional method will yield a positive-definite energy growth. While nonconservation is generic, the sign of the energy error is typically indefinite. For example, a second-order Runge–Kutta method (Equation 25.5.7 in Abramowitz and Stegun [1]) gives oscillatory errors in energy and enstrophy, although on average both the energy and enstrophy grow.

To obtain a conservative version of this algorithm, we proceed as above by applying the predictor-corrector method to the modified equation of motion, (10), giving

$$\widetilde{\psi}_k = \psi_k + \tau\,(S_k + f_k)\,, \tag{21a}$$

$$\psi_k(t+\tau) = \psi_k + \frac{\tau}{2}\left(S_k + f_k + \widetilde{S}_k + \widetilde{f}_k\right). \tag{21b}$$

As we commented above, the conservative algorithm makes only small corrections to the values of $\psi_k(t+\tau)$. This immediately brings to mind the underlying philosophy of the predictor-corrector algorithms; in fact, one might suspect that energy and enstrophy conservation can be achieved by modifying *only* the corrector part of the integrator. Since the predictor is merely an intermediate approximation, there is surely no need for it to be conservative. Thus we can replace (21) with the simpler prescription

$$\widetilde{\psi}_k = \psi_k + \tau\, S_k\,, \tag{22a}$$

$$\psi_k(t+\tau) = \psi_k + \frac{\tau}{2}\left(S_k + \widetilde{S}_k + g_k\right). \tag{22b}$$

As before, we determine g_k by demanding conservation of energy and enstrophy. The energy at $t+\tau$ is given by

$$\begin{aligned} E(t+\tau) &= \frac{1}{2}\sum_k \left[\psi_k(t)^2 + \tau\,\psi_k\left(S_k + \widetilde{S}_k + g_k\right) + \frac{\tau^2}{4}\,\left(S_k + \widetilde{S}_k + g_k\right)^2\right] \\ &= E(t) + \frac{\tau}{2}\sum_k \left[g_k\,\psi_k - \tau\, S_k\,\widetilde{S}_k + \frac{\tau}{4}\,(S_k + \widetilde{S}_k + g_k)^2\right], \end{aligned} \tag{23}$$

where the last step follows from the definition of the predictor and the properties of S_k. We see that energy will be conserved provided that

$$\sum_k \left[g_k\,\psi_k - \tau\, S_k\,\widetilde{S}_k + \frac{\tau}{4}\,(S_k + \widetilde{S}_k + g_k)^2\right] = 0\,. \tag{24}$$

Similarly, enstrophy will be conserved if

$$\sum_k k^2\left[g_k\,\psi_k - \tau\, S_k\,\widetilde{S}_k + \frac{\tau}{4}\,(S_k + \widetilde{S}_k + g_k)^2\right] = 0\,. \tag{25}$$

We can satisfy these conditions simultaneously if we can solve

$$g_k\,\psi_k - \tau\, S_k\,\widetilde{S}_k + \frac{\tau}{4}\,(S_k + \widetilde{S}_k + g_k)^2 = 0 \tag{26}$$

for g_k. Some straightforward algebra gives

$$\frac{\tau}{2}\,g_k = -\left[\psi_k + \frac{\tau}{2}\,\left(S_k + \widetilde{S}_k\right)\right] + \sigma_k\sqrt{\psi_k^2 + \tau\left(\psi_k\,S_k + \widetilde{\psi}_k\,\widetilde{S}_k\right)}\,, \tag{27}$$

where we choose $\sigma_k = \pm 1$ such that as $\tau \longrightarrow 0$, g_k vanishes. We consider the limit of small τ in two cases. If ψ_k is nonzero, then for small enough τ, both ψ_k and $\widetilde{\psi}_k$ have the same sign, and we can expand the radical to give

$$\frac{\tau}{2}\,g_k = -\psi_k - \frac{\tau}{2}\left(S_k + \widetilde{S}_k\right) + \sigma_k\,\mathrm{sgn}(\psi_k)\left[\psi_k + \frac{\tau}{2}\left(S_k + \widetilde{S}_k\right)\right] + O(\tau^2)\,, \tag{28}$$

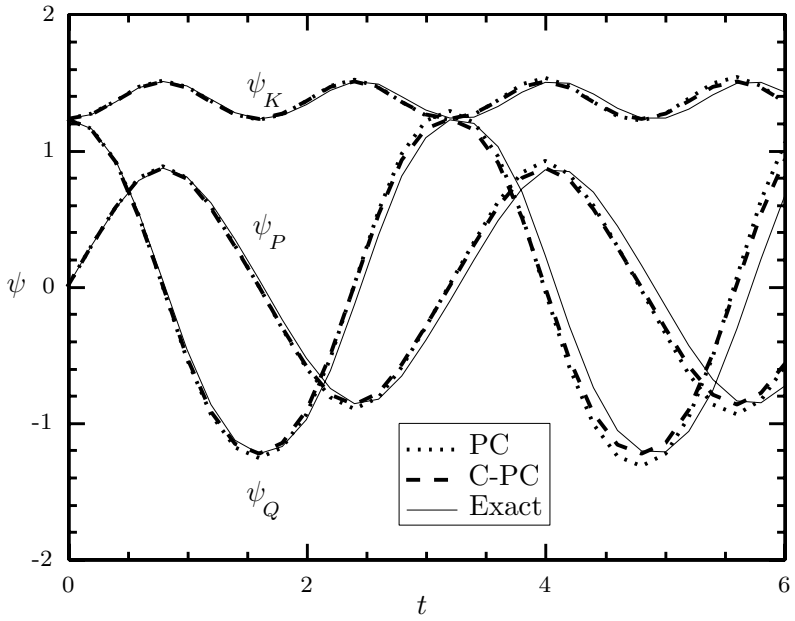

FIG. 3. *Solutions of the three-wave problem for the initial conditions* $\psi_K = \sqrt{1.5}$, $\psi_P = 0.0$, *and* $\psi_Q = \sqrt{1.5}$ *computed using the predictor-corrector (PC) and conservative predictor-corrector (C-PC) methods, with a time step of size* 0.2. *The exact solution (Exact) is also shown.*

leading us to choose $\sigma_k = \mathrm{sgn}(\psi_k)$. Otherwise, if $\psi_k = 0$, then $\widetilde{\psi}_k = \tau\, S_k$ and $\widetilde{S}_k = S_k + O(\tau)$, so that

$$\begin{aligned} \frac{\tau}{2}\, g_k &= -\tau\, S_k + \sigma_k \sqrt{\tau^2\, S_k^2} + O(\tau^2) \\ &= -\tau\, S_k + \tau\, \sigma_k\, \mathrm{sgn}(S_k)\, S_k + O(\tau^2)\,. \end{aligned} \tag{29}$$

In this case we take $\sigma_k = \mathrm{sgn}(S_k) = \mathrm{sgn}(\widetilde{\psi}_k)$. In the previous case, we noted, for small τ, that ψ_k and $\widetilde{\psi}_k$ have the same sign. Therefore, the choice $\sigma_k = \mathrm{sgn}(\widetilde{\psi}_k)$ will always provide the correct limiting behavior.

Using the expression (29) for g_k in our modified predictor-corrector algorithm, (22), we obtain the following conservative integrator:

$$\widetilde{\psi}_k = \psi_k + \tau\, S_k\,, \tag{30a}$$

$$\psi_k(t+\tau) = \widetilde{\sigma}_k \sqrt{\psi_k^2 + \tau \left(\psi_k\; S_k + \widetilde{\psi}_k\, \widetilde{S}_k \right)}\,, \tag{30b}$$

where $\widetilde{\sigma}_k = \mathrm{sgn}(\widetilde{\psi}_k)$. Unlike C-Euler, this algorithm, which we call "conservative predictor-corrector" (C-PC), does not suffer from fixed points (when $\psi_k = 0$, C-PC reduces to (18) as $\tau \longrightarrow 0$). It is still possible that the argument of the radical can become negative; however, this merely indicates that the step size is too large. If S_k has continuous first derivatives, it can be shown that a finite number of time step reductions is sufficient to integrate the system through a negative-argument region.

For our model problem, we now compare the numerical solutions obtained by the conventional predictor-corrector (PC) method with those obtained from C-PC, (30). Our results are summarized in Figures 3–6. In Figure 3 we show $\psi_k(t)$ computed

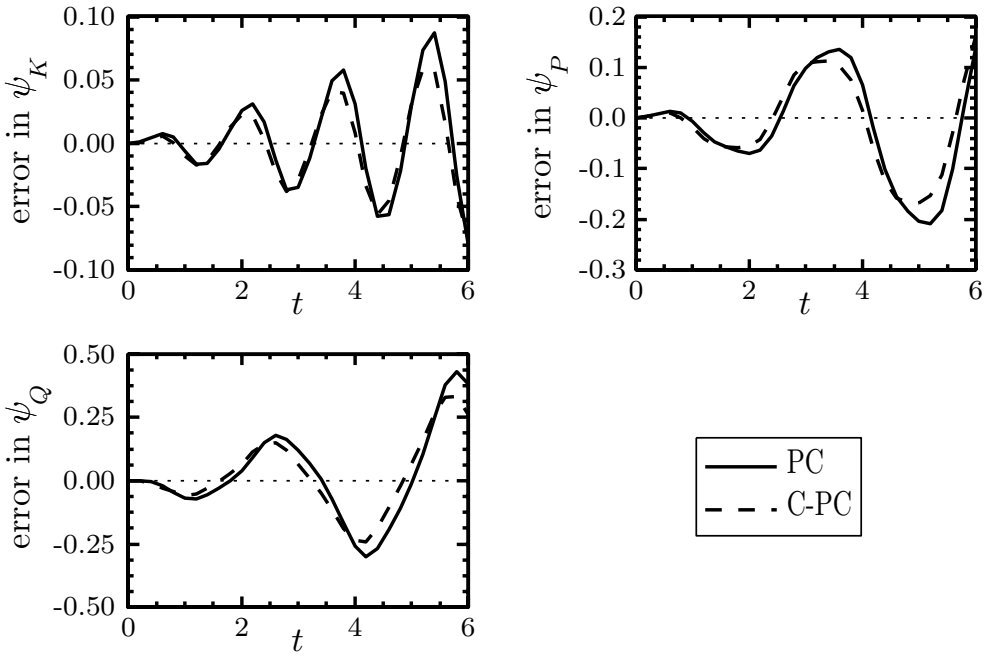

FIG. 4. *Differences between the computed and exact solutions in Fig.* 3.

with both methods as well as the exact solution. The errors in the two approximate solutions are displayed in Figure 4. In Figure 5 we plot $\Delta E = E(t) - E(0)$ and $\Delta Z = Z(t) - Z(0)$ for both methods.

In the limit of infinitesimal step size, C-PC reduces to PC. It is a straightforward exercise to verify analytically that both methods agree with the exact solution to second order in the time step. To illustrate this property we fit the error for a single step of mode K to a power law:

$$\Delta\psi_k = A\,\tau^n\,. \tag{31}$$

The results of this fit, shown for both methods in Figure 6, are consistent with the second-order accuracy of the PC and C-PC methods.

In constructing our conservative algorithms, we have essentially altered the manner in which truncation error enters the solution. Where this error has gone is an important question. It is unreasonable, of course, to expect that the truncation error has vanished. Since our algorithm imposes two independent constraints on the three dynamical variables, all of the truncation error is lumped into the only place left—the phase of the numerical solution with respect to the exact solution. As our ultimate application is to fluid turbulence, the nature of this component of the error in general could be of great importance. There are two possibilities: either this error manifests itself as a global phase shift, with all three waves exhibiting the same phase error, or each wave receives a different phase error, so that relative phase shifts begin to develop. Of the two possibilities, the first is of little consequence in a turbulence simulation, whereas the second could, arguably, be as bad (from a structural point of view) as the energy growth that we sought to eliminate.

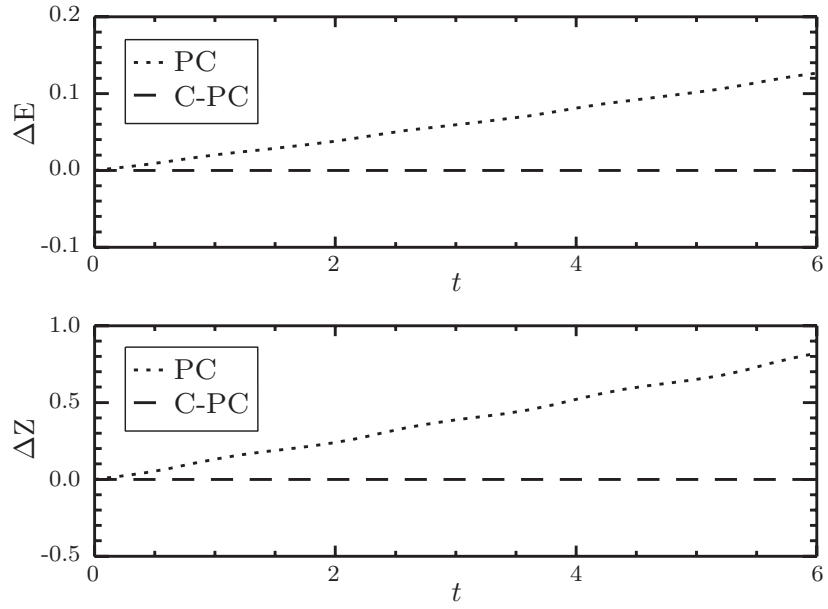

FIG. 5. *Change in energy and enstrophy for the PC and C-PC methods.*

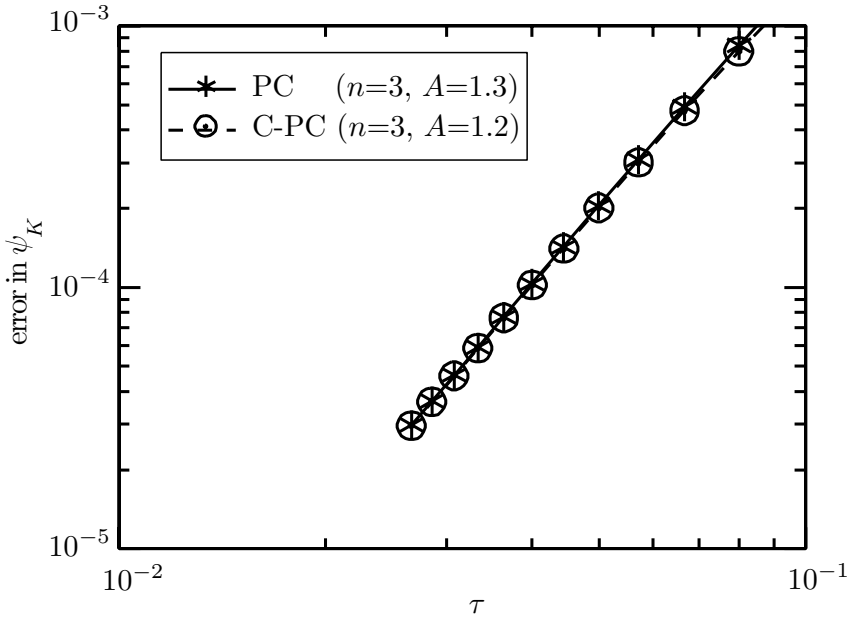

FIG. 6. *Single-step error in mode K for the initial conditions* $\psi_K = \sqrt{1.5}$, $\psi_P = 1.0$, *and* $\psi_Q = \sqrt{1.5}$ *for the PC and C-PC methods. The results of fitting the error to a power law* $A\tau^n$ *are shown, indicating that the C-PC algorithm is of second order, as expected.*

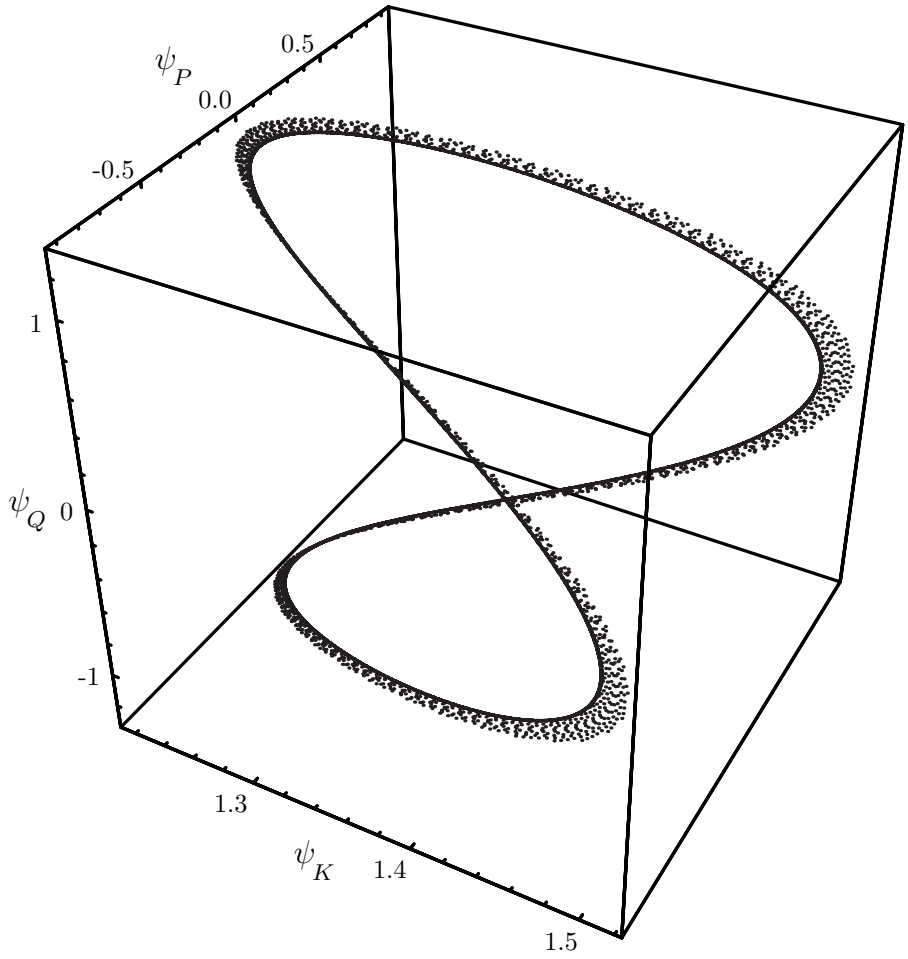

FIG. 7. *Integration of the three-wave problem using a second-order PC method (dotted line) and the C-PC method (solid line). Both methods took approximately* 4000 *time steps of size* 0.05. *Initially* $\psi_K = \sqrt{1.5}$, $\psi_P = 0$, *and* $\psi_Q = \sqrt{1.5}$. *The effect of the* 4% *energy gain by the conventional method is clearly visible.*

Since our model problem is integrable, we can easily distinguish between these cases. In a plot where each of the axes is one of the dynamical variables, the exact solution is a simple closed curve. For our case, such a plot is shown in Figure 7. The solid line is the orbit computed with the C-PC integrator, while the dots represent the solution obtained from the PC method. Since the conservative solution yields a closed curve, we may conclude that the additional phase error introduced is global and thus the relative phases of the waves are not affected by our method. This supports the general observation made by de Frutos and Sanz-Serna regarding the nature of local truncation error in systems with invariants [10].

3.3. Generalizations. The C-PC algorithm has two important straightforward generalizations: to n waves and to complex ψ_k. The n-wave generalization is immediate—nowhere in our derivations of the conservative algorithms have we made use of the number of modes. Both the C-Euler and C-PC methods can be applied to a system

with an arbitrary number of modes, where the energy and enstrophy expressions are the appropriate generalizations of (4) and (5), respectively.

The generalization to complex amplitudes proceeds as follows. Consider a system with n complex-valued amplitudes ψ_k. We split these amplitudes into real and imaginary parts ψ_k^r and ψ_k^i, respectively, which evolve according to

$$\frac{d\psi_k^r}{dt} = S_k^r(\psi)\,, \tag{32a}$$

$$\frac{d\psi_k^i}{dt} = S_k^i(\psi)\,, \tag{32b}$$

where S_k^r and S_k^i are the real and imaginary parts of the source function S_k. For this system the energy and enstrophy are given by

$$E = \frac{1}{2}\sum_k |\psi_k|^2 \tag{33}$$

and

$$Z = \frac{1}{2}\sum_k k^2\ |\psi_k|^2\ , \tag{34}$$

where k ranges over the wavenumbers of the n modes. The properties of the source terms that guarantee conservation of energy and enstrophy are

$$\sum_k \psi_k^r\, S_k^r + \psi_k^i\, S_k^i = 0\,, \tag{35a}$$

$$\sum_k k^2 \left(\psi_k^r\, S_k^r + \psi_k^i\, S_k^i\right) = 0\,; \tag{35b}$$

hence, we see that a system of n complex modes is completely equivalent to a system of $2n$ real modes. Therefore, the complex version of a conservative algorithm follows upon applying the real algorithm separately to each component of the complex amplitudes.

3.4. Discussion. It is worth saying a few words about computational efficiency. There are two sources of computational overhead associated with the conservative algorithms compared to the conventional methods. Here we concentrate on C-PC since C-Euler is not appropriate for practical use. In terms of operations, C-PC requires two additional multiplications and a square-root evaluation over the standard PC method. Importantly, C-PC uses no additional storage. The cost of the extra operations will, in most cases, be negligible compared to the cost of one evaluation of S_k. The square root is a cause for some concern as it may involve a function call. In any event, on modern hardware the square-root operation is only a small number of (typically five to seven) times slower than multiplication. Furthermore, it is reasonable to expect that a conservative integrator will obtain a given global accuracy with a larger time step than the corresponding conventional integrator, thereby ameliorating the overhead problem. The second source of overhead is the occasional need to reduce the time step when the argument of the square root becomes negative. In practice we find that this happens approximately 10% of the time; in light of the above discussion, we feel that this is not significant.

Since the C-PC method reduces to the usual PC algorithm in the limit of an infinitesimal time step, one expects that C-PC will inherit the infinitesimal stability properties of PC. Indeed, in the sense of Skeel [29], one can establish that PC and C-PC are *equally stable*. That is, upon adding a perturbation ξ_k to $\psi_k(t)$ in (18) and (30), one finds that for both discretizations the predicted and corrected values at time $t+\tau$ are perturbed by expressions of the form

$$\xi_k + \tau \sum_j \xi_j \frac{\partial S_k}{\partial \psi_j} + \mathcal{O}(\tau^2). \tag{36}$$

This implies that for sufficiently small τ, both methods have the same stability properties. Although this does not establish the behavior for large time steps, we find in practice that the C-PC algorithm is in fact as numerically robust (e.g., to blow-up of the solution) as the original PC scheme.

In numerical studies of the Euler fluid equations, an artificial viscosity is often added to the dynamical equations to compensate partially for the spurious growth of the energy and enstrophy introduced by the numerical scheme. The viscosity is usually taken to vary as a power of the wavenumber. However, only one of the two invariants can be exactly conserved by such a procedure, and this would require the prescribed viscosity coefficient to be time dependent. Moreover, such a remedy can be shown to contaminate the modal evolution. In contrast, the conservative algorithms developed in this work faithfully reproduce the modal dynamics.

In addition these methods can be applied to dissipative systems where the change in energy has a specific physical origin. The same numerical errors that previously led to nonconservation of energy will now contribute to the net energy change, thus having the effect of altering the underlying physics. For example, in a viscous fluid simulation the amount of energy leaving a mode is determined by the balance between viscosity and nonlinear transfer. It is an open question and a subject of further investigation by the authors as to whether errors of this sort have the same structural effect on the solution as in the dissipationless case.

There is a simple interpretation of the C-Euler and C-PC algorithms that sheds light both on their form and on the existence of the two σ_k branches. As numerous authors have observed, most traditional numerical methods conserve the linear invariants of a system. Consequently, one might be led to consider the possibility of transforming ψ_k to new variables, in terms of which the invariants are linear. For the three-wave problem, this can be accomplished by making the transformation $\phi_k = \psi_k^2$. Upon applying the Euler method in the ϕ_k space and transforming back by taking the square root, one immediately obtains (16). This indicates that our restriction of the general constraint (13) to the condition (14) merely ensures that the modal energies evolve in a manner consistent with the Euler discretization of the energy equations. Below we will give derivations of integrators for other systems based on this idea. The C-PC algorithm can be viewed in the same light, except that the predictor is taken to have the simpler, nonconservative form. This also explains the $\psi_k = 0$ fixed point in C-Euler: the modal energies have a second-order zero at $\psi_k = 0$; it is thus no wonder that a first-order method fails at that point.

4. Lotka–Volterra predator-prey model. As a further demonstration, consider the Lotka–Volterra predator-prey equations:

$$\frac{dx}{dt} = -\mu x(1-y)\,, \tag{37a}$$

$$\frac{dy}{dt} = y(1-x)\,. \tag{37b}$$

These equations are surprisingly hard to integrate numerically since they are very susceptible to round-off error. With the exception of Kahan's [17] nontraditional method (which Sanz-Serna [24] has shown to be symplectic), there are virtually no other methods that can integrate this system without eventually failing due to round-off error.

This is a noncanonical Hamiltonian system [24] with Hamiltonian

$$H = x - \log x + \mu y - \mu \log y \tag{38}$$

and Poisson bracket

$$\{f,\, g\} = x\, y \left(\frac{\partial f}{\partial x}\, \frac{\partial g}{\partial y} - \frac{\partial g}{\partial x}\, \frac{\partial f}{\partial y} \right). \tag{39}$$

Just as with the three-wave problem, conventional integrators such as Euler and PC fail to conserve the total energy H. It happens that the dynamics of this system are particularly sensitive to the value of the energy, which explains the difficulty that these methods encounter when integrating (37).

It is possible to derive a conservative algorithm for this system based on the method outlined in section 3.2. However, the transcendental nature of the functions in the energy greatly complicates the procedure and prevents an analytical solution of the relevant equations. In light of these problems, we take an alternative approach to deriving a conservative integrator. We proceed using the observation that standard methods such as PC exactly preserve linear invariants of a system of differential equations. To exploit this behavior, we introduce new variables ξ_1 and ξ_2 defined by

$$\xi_1 = x - \log x\,, \tag{40a}$$

$$\xi_2 = \mu\,(y - \log y)\ . \tag{40b}$$

This transformation was chosen so that H is a linear function of ξ_1 and ξ_2. Using the original equations of motion, we obtain

$$\frac{d\xi_1}{dt} = \mu(x-1)(y-1)\,, \tag{41a}$$

$$\frac{d\xi_2}{dt} = -\mu(x-1)(y-1)\,. \tag{41b}$$

Applying the usual second-order PC to these equations yields

$$\widetilde{\xi}_1 = \xi_1 + \tau\,\mu(x-1)(y-1)\,, \tag{42a}$$

$$\widetilde{\xi}_2 = \xi_1 - \tau\,\mu(x-1)(y-1)\,, \tag{42b}$$

$$\xi_1(t+\tau) = \xi_1 + \frac{\tau}{2}\,\mu\left[(x-1)(y-1) + (\widetilde{x}-1)(\widetilde{y}-1)\right]\,, \tag{42c}$$

$$\xi_2(t+\tau) = \xi_2 - \frac{\tau}{2}\,\mu\left[(x-1)(y-1) + (\widetilde{x}-1)(\widetilde{y}-1)\right]\,. \tag{42d}$$

Strictly speaking, here $\widetilde{x}$ and $\widetilde{y}$ are to be computed from $\widetilde{\xi}_1$ and $\widetilde{\xi}_2$ by inverting (40a) and (40b), respectively. Following the philosophy of section 3.2, we instead compute $\widetilde{x}$

and $\widetilde{y}$ from the original equations of motion to obtain the following conservative integrator:

$$\widetilde{x} = x - \tau\,\mu\,x(1-y)\,, \tag{43a}$$

$$\widetilde{y} = y + \tau\,y(1-x)\,, \tag{43b}$$

$$\xi_1(t+\tau) = \xi_1 + \frac{\tau}{2}\,\mu\,[(x-1)(y-1) + (\widetilde{x}-1)(\widetilde{y}-1)]\,, \tag{43c}$$

$$\xi_2(t+\tau) = \xi_2 - \frac{\tau}{2}\,\mu\,[(x-1)(y-1) + (\widetilde{x}-1)(\widetilde{y}-1)]\,. \tag{43d}$$

Here $x(t+\tau)$ and $y(t+\tau)$ are determined from $\xi_1(t+\tau)$ and $\xi_2(t+\tau)$ by inverting (40). Since this inversion requires solving a transcendental equation, in practice it will have to be carried out iteratively. Although the expressions for $x(t+\tau)$ and $y(t+\tau)$ cannot be written in closed form, (43) is still an explicit scheme.

Notice that ξ_1 and ξ_2 are not one-to-one functions of x and y; ξ_1 has a minimum value of 1 at $x = 1$, and ξ_2 has a minimum of μ at $y = 1$. These minimum values play roles similar to those of the points $\psi_k = 0$ in the three-wave problem. Fortunately, the remedy is similar also: if either ξ_1 or ξ_2 is pushed below its respective minimum, this indicates that the time step is too large. Temporarily reducing the time step alleviates this problem.

To illustrate the effectiveness of our conservative algorithm, we integrate (37) taking $\mu = 1.5$ with an initial condition of $x(0) = 1.0$ and $y(0) = 0.4$. In Figure 8 we show a comparison between the standard PC method and the C-PC algorithm. The C-PC orbit exactly conserves energy and forms a closed curve. The PC orbit spirals outward—a consequence of its energy gain.

We provide this example to illustrate the generality of our method. However, since there is not an explicit expression for the inverse of the transformation (40) and a symplectic algorithm is known [17, 24], this method seems to be of little practical value due to the computational overhead of iteratively determining x and y from ξ_1 and ξ_2.

5. Kepler problem. As a final example we consider the problem of a single particle moving in a gravitational potential [14, 30]. Let $\mathbf{r}$ be the position vector of the particle of mass m and let $\phi(r)$, where $r = |\mathbf{r}|$, be the gravitational potential. The equations of motion for this system are

$$\frac{d\mathbf{r}}{dt} = \mathbf{v}\,, \tag{44a}$$

$$\frac{d\mathbf{v}}{dt} = -\frac{1}{m}\,\nabla\phi \tag{44b}$$

This is a conservative system with the Hamiltonian

$$H = \frac{1}{2}\,m\,\mathbf{v}^2 + \phi(r)\,. \tag{45}$$

As with all central force problems the total angular momentum, $\mathbf{L} = m\,\mathbf{r} \times \mathbf{v}$, is conserved, confining the motion to the plane perpendicular to $\mathbf{L}$. We exploit this feature by aligning our coordinate system with the $\widehat{\mathbf{z}}$ direction parallel to $\mathbf{L}$ and introducing polar coordinates (r, θ) in the plane perpendicular to $\mathbf{L}$.

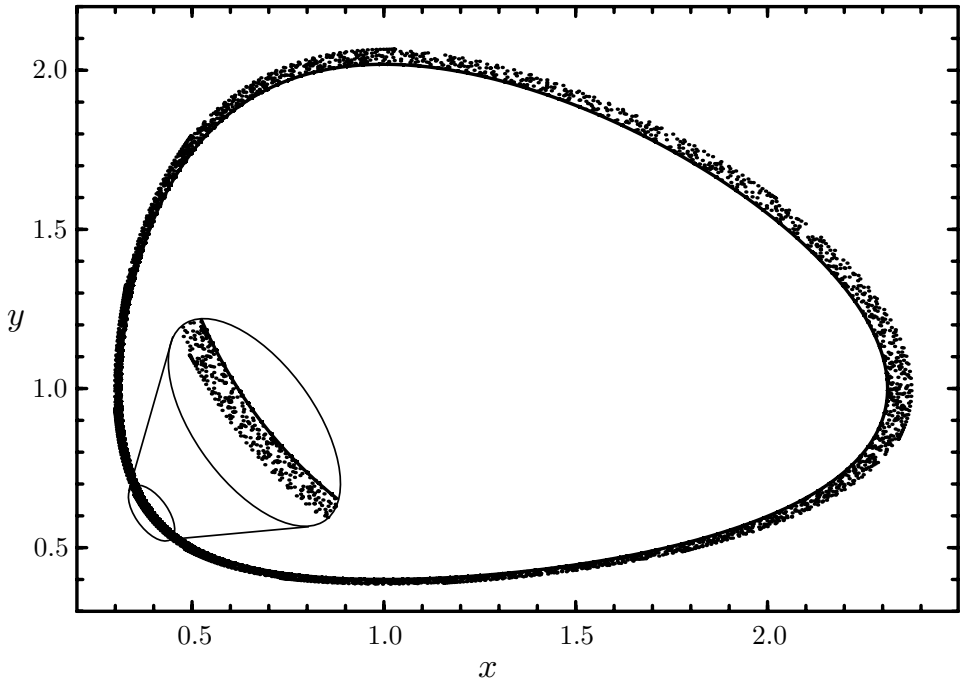

FIG. 8. *Integration of the Lotka–Volterra problem using a standard second-order PC method and a C-PC algorithm, each with* 8×10^5 *time steps of size* 0.02. *A point is plotted every* 200 *time steps. The solid line represents the energy surface containing the initial condition. The points obtained from C-PC all lie on this curve. The dramatic effect of the* 1.2% *energy gain by the standard algorithm is clearly visible.*

In these coordinates the equations of motion become

$$\frac{dr}{dt} = v_r, \tag{46a}$$

$$\frac{dv_r}{dt} = \frac{\ell^2}{m^2 r^3} - \frac{1}{m}\,\phi'(r), \tag{46b}$$

$$\frac{d\theta}{dt} = \frac{\ell}{mr^2}, \tag{46c}$$

where ℓ is the magnitude of the angular momentum and the Hamiltonian can be written as

$$H = \frac{1}{2}\,m v_r^2 + \frac{\ell^2}{2mr^2} + \phi(r)\,. \tag{47}$$

Unlike all other central force problems, the Kepler problem has an additional constant of motion known as the Runge–Lenz vector,

$$\mathbf{A} = \mathbf{v} \times \mathbf{L} + \phi\,\mathbf{r}\,. \tag{48}$$

Conservation of the Runge–Lenz vector can be associated with the fact that the orientation of the bound orbits of this system is fixed. It turns out that these orbits are elliptical and oriented with the major axis in the direction of $\mathbf{A}$. We say "associated" here since in central force problems with any other force law, the orientation

of the bound orbits precesses. Furthermore, the Runge–Lenz vector is in some sense redundant: it is *not* needed to integrate the equation of motion, as there are already enough constants of motion to render the problem integrable.

We adopt the initial conditions $r(0) = r_0$ and $\theta(0) = v_r(0) = 0$, so that the vector $\mathbf{A}$ is in the x-direction. Writing the potential as $\phi(r) = -K/r$, where K is a constant, we see that the magnitude of $\mathbf{A}$ is given by

$$A = \frac{\ell^2}{mr_0} - K\,. \tag{49}$$

5.1. A conservative integrator for the Kepler problem. The Kepler problem is an interesting example to consider in a study of conservative integrators, not only as a preliminary to studying multibody problems (which are of astronomical significance), but also because of the Runge–Lenz vector. Although this vector is functionally dependent on the Hamiltonian and on the angular momentum, exact conservation of these invariants neither guarantees conservation of the Runge–Lenz vector nor prevents the computed orbits from exhibiting a spurious precession. Hence numerical conservation of the Runge–Lenz vector is as much a structural issue as is conservation of energy.

We now illustrate a C-PC algorithm for integrating (46) that exactly conserves H and A. The predictor is conventional:

$$\widetilde{r} = r + \tau\, v_r\,, \tag{50a}$$

$$\widetilde{v}_r = v_r + \tau\, \frac{1}{mr^2}\left(\frac{\ell^2}{mr} - K\right), \tag{50b}$$

$$\widetilde{\theta} = \theta + \tau\, \frac{\ell}{mr^2}\,. \tag{50c}$$

To obtain the corrector equations, we transform (r, v_r) to the new variables

$$\xi_1 = -\frac{K}{r}\,, \tag{51a}$$

$$\xi_2 = \frac{1}{2}\, m\, v_r^2 + \frac{1}{2}\, \frac{\ell^2}{mr^2}\,, \tag{51b}$$

so that $H = \xi_1 + \xi_2$. Expressed in these new variables, the Hamiltonian is linear and will be conserved by conventional integrators. The corrector is given by

$$\xi_1(t+\tau) = \xi_1 + \Delta\,, \tag{52a}$$

$$\xi_2(t+\tau) = \xi_2 - \Delta\,, \tag{52b}$$

where

$$\Delta = \frac{\tau}{2}\left(\frac{K v_r}{r^2} + \frac{K\widetilde{v}_r}{\widetilde{r}^2}\right). \tag{53}$$

In terms of the original variables, (52) may be rewritten as

$$r(t+\tau) = \frac{-K}{-K/r + \Delta}\,, \tag{54a}$$

$$v_r(t+\tau) = \operatorname{sgn}(\widetilde{v}_r)\sqrt{v_r^2 + \frac{\ell^2}{m^2}\left(\frac{1}{r^2} - \frac{1}{\widetilde{r}^2}\right) - 2\,\frac{\Delta}{m}}\,. \tag{54b}$$

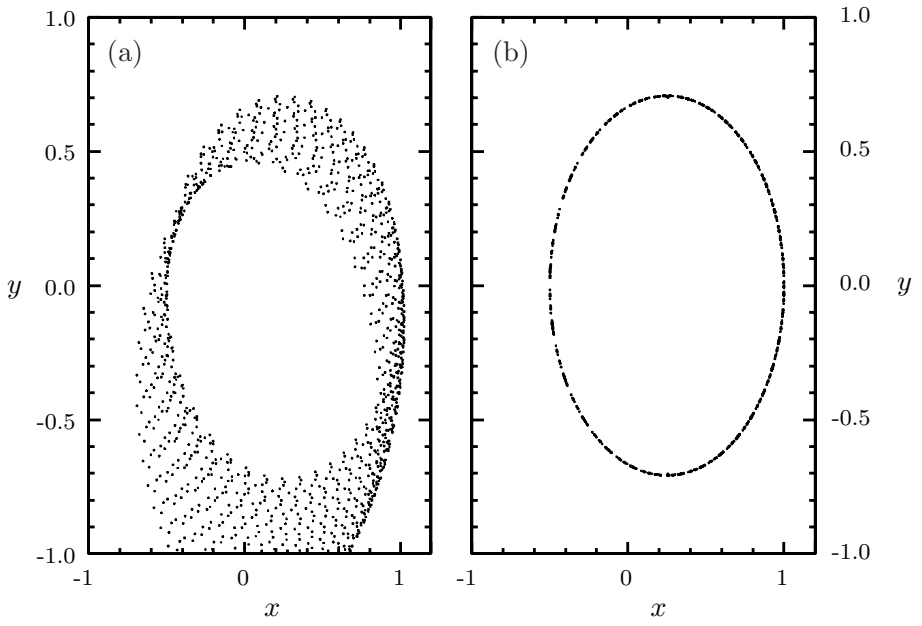

FIG. 9. *Solutions of the Kepler problem:* (a) *computed using the PC algorithm with a total of* 1313 *fixed time steps of size* 0.08*;* (b) *computed using the C-PC algorithm with a total of* 1000 *fixed time steps of size* 0.105.

We still need an equation for θ. Thus far, we have enforced the invariance of H, but not A. Since only one integration variable remains to be determined, the conservation of A enforces the following constraint on θ:

$$A\left(v_r\cos\theta - \frac{\ell}{mr}\sin\theta\right) = -K\,v_r\,, \tag{55}$$

as is seen upon taking the **v**-projection of (48). This equation can be inverted for θ using trigonometric identities and the quadratic formula. However, to avoid the complexities associated with multiply branched solutions, the most convenient method for solving (55) appears to be Newton–Raphson iteration, using $\theta(t)$ for the initial estimate. Convergence is rapid; typically, only three or four iterations are required.

In Figure 9 we present our integration results for the C-PC and PC algorithms, adopting the initial parameters $r = 1$, $v_r = \theta = 0$, $\ell = 1$, $K = 3/2$, and $m = 1$. To allow an even comparison, a slightly larger time step size was chosen for the C-PC run such that the amount of computer time needed to reach the final time was the same in both cases. The new algorithm dramatically outperforms the traditional integrator. The artificial precession of the trajectory exhibited by the PC result does not occur in the C-PC solution, due to the explicit conservation of the Runge–Lenz vector.

5.2. Discussion. We have demonstrated an explicit conservative integrator for the Kepler problem that captures important structural features. LaBudde and Greenspan [18, 19] have constructed integrators for this problem that conserve both energy and angular momentum, but it is unclear whether their method exhibits orbital precession.

The extension of these ideas to multibody problems is a complicated task. Although all of the components of the angular momentum are constants of motion,

they are not in involution. In the simple Kepler problem we are able to avoid any difficulties associated with this behavior because the motion is confined to the plane perpendicular to the angular momentum vector. In the multibody problem this is no longer the case, which significantly complicates matters.

6. Conclusions. We have demonstrated a technique, motivated by the idea of backward error analysis, for deriving *explicit*, exactly conservative integration algorithms. The method consists of modifying the dynamical equations in such a way that when a particular conventional integration algorithm is applied to the modified equations, one obtains a solution consistent with the original equations that exactly conserves a system's invariants. When applied to an explicit conventional algorithm, the method will typically yield an explicit conservative scheme. We have seen that the technique can be interpreted in terms of a transformation to a new set of variables in which the invariants in question are linear. This promises to be a general method for deriving conservative integrators.

In section 3, we saw that for a system with quadratic invariants, conservative integrators can be developed that are simple and computationally efficient. The case of quadratic invariants is of particular interest. The invariants in Lie–Poisson systems are typically quadratic Casimirs. Furthermore, for Hamiltonian systems with Lie group symmetry, a Lie–Poisson system is the natural result of reduction; thus, our methods are applicable to integrating the dynamics on the Poisson manifold of such systems. For the integration of canonical Hamiltonian systems where the configuration space is a Lie group, Simo, Lewis, and coworkers [26, 28, 27, 20] have developed a series of methods that are symplectic and conserve momentum. One could imagine a hybrid of these algorithms: a conservative integrator of the type discussed above for integrating the dynamics on the Poisson manifold coupled to the algorithms of Simo et al. for reconstruction of the full phase-space flow.

In addition to the desirable physical aspects of exact energy conservation there is some evidence [26] that such conservation leads to nonlinear numerical stability. Furthermore, any conservative integration method developed for general systems could certainly be applied to Hamiltonian systems, providing an interesting comparison with symplectic methods. For example, this might shed some light on the relative merits of preserving phase-space structure and conserving nonlinear invariants. In fact, one could envision using the local change in phase-space volume as a diagnostic of the performance of a conservative integrator. These ideas will be the subject of a future paper.

Appendix A. Conservative Euler algorithm. One of the requirements in our derivation of C-Euler was that the new algorithm should reduce to the Euler method in the limit $\tau \longrightarrow 0$. This amounts to demanding that (16) have a power series expansion in τ such that the first two terms are given by Euler's formula, requiring that $|2\tau S_k/\psi_k| < 1$. (For $|2\tau S_k/\psi_k| > 1$, (16) has a series expansion, but all terms involve fractional powers of τ.) As noted in section 3.1, (16) will fail to meet this requirement in a neighborhood of $\psi_k = 0$. We now show that it is possible to devise a conservative algorithm that circumvents this problem.

We exploit the fact that (16) conserves energy for any value of the time step τ. This means that the scheme

$$\psi_k(t+\tau) = \sigma_k \sqrt{\psi_k^2(t) + \mu\,\tau\,S_k(t)\,\psi_k(t)} \tag{A1}$$

obtained by substituting $\mu\tau$ for 2τ in (16) is also conservative. Let $\epsilon_k = \tau S_k/\psi_k$. If we choose $\mu = 2 + \epsilon_j$, (A1) is equivalent to (7) for mode j. To make the modified scheme match the Euler algorithm as closely as possible, we choose mode j such that

$$|\epsilon_j| = \max_{k\in\{K,P,Q\}} |\epsilon_k|. \tag{A2}$$

We use (A1) to evolve the system if $|\epsilon_j| < 1/2$. However, if $|\epsilon_j| \geq 1/2$, we advance mode j with (7) and evolve the other modes by stepping *backwards* in time from $t+\tau$ to t. The substitutions $\tau \to -\tau$ and $t \to t+\tau$ in (A1) yield an implicit conservative algorithm for evolving backwards:

$$\psi_k(t) = \mathrm{sgn}(\psi_k(t))\sqrt{\psi_k^2(t+\tau) - \mu\,\tau\,S_k(t+\tau)\,\psi_k(t+\tau)}\,. \tag{A3}$$

Again, the equation for mode j will reduce to the one given by (7) if we choose

$$\mu = \frac{[2\psi_j + \tau\,S_j(t)]S_j(t)}{S_j(t+\tau)\,\psi_j(t+\tau)}. \tag{A4}$$

These modifications introduce a second-order correction to (16) that does not affect the reduction expressed in (17). In either case, if the radical associated with some mode (other than j) has a negative argument, it is clear from the form of (A1) and (A3) that a finite reduction of the time step can always be found such that the second term becomes dominated by the first.

Acknowledgments. The authors would like to thank G. Tarkenton and R. Fitzpatrick for helpful conversations while developing these methods.

REFERENCES

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, Dover, New York, 1965.
[2] J. A. Armstrong, N. Bloembergen, J. Ducuing, and P. S. Pershan, *Interactions between light waves in a nonlinear dielectric*, Phys. Rev., 127 (1962), pp. 1918–1939.
[3] A. Bayliss and E. Isaacson, *How to make your algorithm conservative*, Notices Amer. Math. Soc., 22 (1975), pp. A594–A595.
[4] J. C. Bowman, J. A. Krommes, and M. Ottaviani, *The realizable Markovian closure.* I: *General theory, with application to three-wave dynamics*, Phys. Fluids B, 5 (1993), pp. 3558–3589.
[5] V. Brasey and E. Hairer, *Symmetrized half-explicit methods for constrained mechanical systems*, Appl. Numer. Math., 13 (1993), pp. 23–31.
[6] P. J. Channell and J. C. Scovel, *Symplectic integration of Hamiltonian systems*, Nonlinearity, 3 (1990), pp. 231–259.
[7] P. J. Channell and J. C. Scovel, *Integrators for Lie–Poisson dynamical systems*, Phys. D, 50 (1991), pp. 80–88.
[8] G. J. Cooper, *Stability of Runge–Kutta methods for trajectory problems*, IMA J. Numer. Anal., 16 (1987), pp. 1–13.
[9] R. C. Davidson and A. N. Kaufman, *On the kinetic equation for resonant three-wave coupling*, J. Plasma Phys., 3 (1969), pp. 97–105.
[10] J. de Frutos and J. M. Sanz-Serna, *Erring and being conservative*, in Numerical Analysis 1993, Pitman Res. Notes Math. Ser., D. F. Griffiths and G. A. Watson, eds., Longman Scientific and Technical, Harlow, UK, 1994, pp. 75–88.
[11] Z. Ge and J. E. Marsden, *Lie–Poisson Hamilton–Jacobi theory and Lie–Poisson integrators*, Phys. Lett. A, 133 (1988), pp. 134–139.
[12] C. W. Gear, *Maintaining solution invariants in the numerical solution of ODEs*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 734–743.
[13] C. W. Gear, *Invariants and numerical methods for ODEs*, Phys. D, 60 (1992), pp. 303–310.

[14] H. Goldstein, *Classical Mechanics*, Addison–Wesley, Reading, MA, 1982.
[15] J. M. Greene, *Two-dimensional measure-preserving mappings*, J. Math. Phys., 9 (1968), pp. 760–768.
[16] E. Isaacson, *Integration schemes for long term calculation*, in Advances in Computer Methods for Partial Differential Equations II, R. Vichnevetsky, ed., IMACS (AICA), New Brunswick, NJ, 1977, pp. 251–255.
[17] W. Kahan, *Unconventional numerical methods for trajectory calculations*, unpublished lecture notes, University of California, Berkeley, CA, 1993.
[18] R. A. LaBudde and D. Greenspan, *Discrete mechanics—A general treatment*, J. Comput. Phys., 15 (1974), pp. 134–167.
[19] R. A. LaBudde and D. Greenspan, *Energy and momentum conserving methods of arbitrary order for the numerical integration of equations of motion*, Numer. Math., 25 (1976), pp. 323–346.
[20] D. Lewis and J. C. Simo, *Conservative algorithms for the dynamics of Hamiltonian systems on Lie groups*, Nonlinear Sci.: Theory Appl., 4 (1994), pp. 253–299.
[21] J. E. Marsden, *Lectures on Mechanics*, Cambridge University Press, Cambridge, UK, 1992.
[22] P. J. Morrison, *Hamiltonian description of the ideal fluid*, in Rev. Mod. Phys., 70 (1998), pp. 467–521.
[23] J. M. Sanz-Serna, *Runge–Kutta schemes for Hamiltonian systems*, BIT, 28 (1988), pp. 877–883.
[24] J. M. Sanz-Serna, *An unconventional symplectic integrator of W. Kahan*, Appl. Numer. Math., 16 (1994), p. 245.
[25] J. M. Sanz-Serna and M. P. Calvo, *Numerical Hamiltonian Problems*, Appl. Math. Math. Comput. 7, Chapman & Hall, London, 1994.
[26] J. C. Simo and N. Tarnow, *A new energy and momentum conserving algorithm for the nonlinear dynamics of shells*, Internat. J. Numer. Methods Engrg., 37 (1994), pp. 2527–2549.
[27] J. C. Simo, N. Tarnow, and K. K. Wong, *Exact energy-momentum conserving algorithms and symplectic schemes for nonlinear dynamics*, Comput. Methods Appl. Mech. Engrg., 100 (1992), pp. 63–116.
[28] J. C. Simo and K. K. Wong, *Unconditionally stable algorithms for rigid body dynamics that exactly preserve energy and momentum*, Internat. J. Numer. Methods Engrg., 31 (1991), pp. 19–52.
[29] R. Skeel, *Analysis of fixed-stepsize methods*, SIAM J. Numer. Anal., 13 (1976), pp. 664–685.
[30] E. T. Whittaker, *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, 4th ed., Cambridge University Press, Cambridge, UK, 1959.