12

Hamiltonian PDEs

Many physical processes of interest not only evolve continuously in time but also possess a continuous spatial structure and, hence, can be described by partial differential equations (PDEs). Furthermore, many fundamental laws of physics, such as quantum mechanics, electrodynamics, ideal continuum mechanics, can be formulated within an extension of the Hamiltonian framework discussed so far to PDEs. In this chapter we focus on two particular examples of such Hamiltonian PDEs and discuss a number of numerical discretization techniques. The reader should, however, keep in mind that the solution behavior of PDEs is much more complex than that of ODEs and that the choice of an appropriate discretization will depend very much on the anticipated type of solutions. The techniques described in this chapter are very much restricted to *smooth* solutions such as solitons [53, 201] and balanced geophysical flows [169]. This excludes, in particular, the consideration of shocks [201]. A general introduction to numerical methods for PDEs can be found, for example, in [140].

12.1 Examples of Hamiltonian PDEs

12.1.1 The nonlinear wave equation

Let us consider the nonlinear wave equation

$$u_{tt} = \partial_x \sigma'(u_x) - f'(u), \quad u = u(x,t), \tag{12.1}$$

where σ and f are smooth functions. If $\sigma(u_x) = u_x^2/2$, then the semi-linear wave equation

$$u_{tt} = u_{xx} - f'(u)$$

is obtained. Other choices for $\sigma(u_x)$ lead to idealized one-dimensional models for fluids and materials.

Throughout this chapter, solutions $u = u(x,t)$ of (12.1) are assumed to be smooth in the independent variables x and t and we impose periodic boundary conditions $u(x,t) = u(x+L,t)$, $L > 0$.

We introduce the total energy $\mathcal{E}[u]$ by

$$\mathcal{E}[u] = \int_0^L \left[\frac{1}{2}u_t^2 + \sigma(u_x) + f(u)\right] dx,$$

and observe that, using integration by parts,

$$\begin{aligned}\frac{d}{dt}\mathcal{E}[u] &= \int_0^L \left[u_t u_{tt} + \sigma'(u_x)u_{xt} + f'(u)u_t\right] dx \\ &= \int_0^L u_t \left[u_{tt} - \partial_x \sigma'(u_x) + f'(u)\right] dx.\end{aligned}$$

However, the term in brackets is equal to zero along solutions of (12.1) and, hence, the total energy $\mathcal{E}[u]$ is conserved.

Let us denote the space of smooth and L-periodic functions in x by $S = C^\infty[0,L]$. Our assumption then is that $u(.,t) \in S$ for $t \geq 0$ or in short hand $u(t) \in S$. This smoothness assumption explicitly excludes the consideration of shock-type solutions [201].

Furthermore, upon rewriting (12.1) as

$$\begin{aligned}u_t &= v, \\ v_t &= \partial_x \sigma'(u_x) - f'(u),\end{aligned}$$

the wave equation can, formally, be viewed as a Hamiltonian system with phase space $(u,v)^T \in S \times S$, symplectic form,

$$\bar{\omega} = \int_0^L du \wedge dv \, dx, \tag{12.2}$$

and Hamiltonian functional

$$\mathcal{H}[u,v] = \int_0^L \left[\frac{1}{2}v^2 + \sigma(u_x) + f(u)\right] dx. \tag{12.3}$$

The Hamiltonian equations of motion are derived in the following way. First, the gradient ∇_z of classical mechanics is replaced by the "variational" gradient $\delta_z = (\delta_u, \delta_v)^T$, $z = (u,v)^T \in S \times S$. The variational derivative $\delta_u \mathcal{G}$ of a functional $\mathcal{G}[u]$ is defined by

$$\int_0^L (\delta_u \mathcal{G}[u]\, \delta u)\, dx = \lim_{\epsilon \to 0} \frac{\mathcal{G}[u + \epsilon \delta u] - \mathcal{G}[u]}{\epsilon},$$

for any $\delta u \in S$. Let us demonstrate this for $\delta_u \mathcal{H}[u,v]$ which is equivalent to the

variational derivative of $\mathcal{G}[u] = \int_0^L [\sigma(u_x) + f(u)]\,dx$

$$\begin{aligned}\int_0^L (\delta_u \mathcal{G}[u]\,\delta u)\,dx &= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left\{\int_0^L [\sigma([u+\varepsilon\delta u]_x) + f(u+\varepsilon\delta u)]\,dx - \int_0^L [\sigma(u_x)+f(u)]\,dx\right\}\\ &= \int_0^L [\sigma'(u_x)(\delta u)_x + f'(u)\delta u]\,dx\\ &= \int_0^L [-\partial_x\sigma'(u_x)\delta u + f'(u)\delta u]\,dx\\ &= \int_0^L [-\partial_x\sigma'(u_x) + f'(u)]\,\delta u\,dx.\end{aligned}$$

Comparison of the left- and right-hand side yields

$$\delta_u \mathcal{H}[u,v] = \delta_u \mathcal{G}[u] = -\partial_x\sigma'(u_x) + f'(u).$$

One also obtains $\delta_v \mathcal{H}[u,v] = v$.

Next we rewrite the symplectic form (12.2) as

$$\bar{\omega} = \frac{1}{2}\int_0^L (J_2^{-1}dz)\wedge dz\,dx,$$

with the (local) structure matrix

$$J_2 = \begin{bmatrix} 0 & +1 \\ -1 & 0 \end{bmatrix},$$

and $dz = (du, dv)^T$. Then the wave equation (12.1) becomes equivalent to an abstract Hamiltonian system

$$z_t = J_2\delta_z \mathcal{H}[z]. \tag{12.4}$$

We finally note that the nonappearance of the independent variable x in the functions f and σ implies another conserved functional for the PDE (12.1), namely the total momentum

$$\mathcal{M}[u,v] = \int_0^L v u_x\,dx.$$

Indeed

$$\begin{aligned}\frac{d}{dt}\mathcal{M} &= \int_0^L (v_t u_x + v u_{xt})\,dx\\ &= \int_0^L (u_x[\partial_x\sigma'(u_x) - f'(u)] + v v_x)\,dx\end{aligned}$$

$$\begin{aligned}&= \int_0^L [-\sigma(u_x) - f(u) + v^2/2]_x\,dx\\ &= \left[-\sigma(u_x) - f(u) + v^2/2\right]_{x=0}^{x=L}\\ &= 0.\end{aligned}$$

12.1.2 Soliton solutions

Waves are one of the most important features of fluid dynamics [201]. Particular types of waves are those that travel at a constant speed $c \neq 0$ without changing their shape. These waves are called *traveling waves* or *solitons* [201, 53]. Mathematically a soliton is described by a (smooth) function ϕ such that

$$u(x,t) = \phi(x - ct).$$

Let us introduce the new variable $\xi = x - ct$, then $u_x = \phi_\xi$, $u_t = -c\phi_\xi$, etc. Hence, assuming a solitary solution, the wave equation (12.1) gives rise to a second-order ODE,

$$c^2\phi_{\xi\xi} = \partial_\xi\sigma'(\phi_\xi) - f'(\phi), \tag{12.5}$$

in the independent variable ξ. Equation (12.5) is a Euler–Lagrange equation and to obtain the corresponding Hamiltonian formulation we introduce the new dependent variable (conjugate momentum) $\psi = c^2\phi_\xi - \sigma'(\phi_\xi)$. Let us assume that this relation is invertible, i.e., there is a function $g(\psi)$ such that $\phi_\xi = g'(\psi)$. Hence one can rewrite the second-order ODE (12.5) as a conservative system

$$\psi_\xi = -f'(\phi), \qquad \phi_\xi = g'(\psi), \tag{12.6}$$

with Hamiltonian

$$H = g(\psi) + f(\phi).$$

A solution $\phi(\xi)$ gives rise to a soliton solution if the boundary conditions $\phi_\xi(\pm\infty) = 0$ are satisfied, i.e., $\phi(\xi)$ approaches some constant value as $\xi \to \pm\infty$. In particular, let (ϕ_i, ψ_i) denote the equilibrium solutions of (12.6), then any *homoclinic* or *heteroclinic* solution of (12.6) gives rise to a (not necessarily stable) soliton solution of the nonlinear wave equation (12.1).[1]

[1] A homoclinic solution is a solution connecting an equilibrium point (ϕ_i, ψ_i) with itself, i.e., $\lim_{\xi\to\pm\infty}\phi(\xi) = \phi_i$, and a heteroclinic solution is a solution connecting two different equilibrium points (ϕ_i, ψ_i), (ϕ_j, ψ_j), i.e., $\lim_{\xi\to+\infty}\phi(\xi) = \phi_i$, $\lim_{\xi\to-\infty}\phi(\xi) = \phi_j$.

Example 1 Let us consider the sine-Gordon equation

$$u_{tt} = u_{xx} - \sin(u).$$

Traveling wave solutions $u(x,t) = \phi(x - ct)$ must satisfy the second-order ODE

$$c^2\phi_{\xi\xi} = \phi_{\xi\xi} - \sin(\phi).$$

We introduce the momentum $\psi = c^2\phi_\xi - \phi_\xi$. For $c \neq \pm 1$, this relation can be inverted and we obtain $\phi_\xi = \psi/(c^2-1)$ and $g(\psi) = \frac{1}{2}\psi^2/(c^2-1)$. Hence the Hamiltonian equations are

$$\psi_\xi = -\sin(\phi), \qquad \phi_\xi = \frac{1}{c^2-1}\psi$$

with Hamiltonian

$$H = \frac{\psi^2}{2(c^2-1)} + (1 - \cos\phi).$$

These are the equations of motion for a nonlinear pendulum with mass $m = c^2 - 1$. Hence we first consider the condition $|c| > 1$. The nonlinear pendulum possesses heteroclinic solutions connecting pairs of hyperbolic equilibria $(\phi,\psi) = (k\pi, 0)$ and $(\phi,\psi) = ((k+2)\pi, 0)$ for $k = 1, \pm 3, \pm 5, \ldots$. These heteroclinic connections are easily found as the contour lines of constant energy $H = 0$. Unfortunately, the associated soliton solutions are all unstable [201, 53]. For $|c| < 1$, the equilibrium points $(k\pi, 0)$, $((k+2)\pi, 0)$, $k = 0, \pm 2, \pm 4, \ldots$, become hyperbolic and give rise to stable soliton solutions. An explicit soliton solution with wave speed c, $|c| < 1$, is given by

$$u(x,t) = 4 \arctan \exp\left(\frac{x - ct}{\sqrt{1-c^2}}\right).$$

Because of their special shape, these solutions are called kink solitons. □

12.1.3 The two-dimensional rotating shallow-water equations

Large-scale geophysical flows in the atmosphere and ocean are essentially incompressible and often stratified into nearly two-dimensional layers. Furthermore, the effect of the earth's rotation significantly affects large-scale patterns away from the equator, for example in mid-latitude or near the poles. See ANDREWS [6] and SALMON [169] for an introduction to geophysical fluid dynamics and MORRISON [139] for further details on the Hamiltonian formalism of geophysical fluid dynamics.

A simple one-layer model system is provided by the two-dimensional rotating shallow-water equations (SWEs) [139, 169]

$$u_t + uu_x + vu_y = +fv - c_0^2 h_x, \tag{12.7}$$

$$v_t + uv_x + vv_y = -fu - c_0^2 h_y, \tag{12.8}$$

$$h_t + uh_x + vh_y = -h(u_x + v_y), \tag{12.9}$$

where $u = (u,v)^T \in \mathbb{R}^2$ is the horizontal velocity field, $c_0 = \sqrt{gH}$, g is the gravitational constant, H is the mean layer depth of the fluid, $h > 0$ is the normalized layer depth with mean value scaled equal to one, and $f > 0$ is twice the angular velocity of the rotating fluid. For simplicity, we will consider the SWEs over a double periodic domain of size $L \times L$ and keep $f = f_0$ constant.

We next introduce the material time derivative of a function $w(x,y,t)$

$$\frac{Dw}{Dt} = w_t + uw_x + vw_y = w_t + u \cdot \nabla_x w,$$

and rewrite the SWEs (12.7)–(12.9) in the form

$$\frac{Du}{Dt} = f_0 J_2 u - c_0^2 \nabla_x \eta, \tag{12.10}$$

$$\frac{Dh}{Dt} = -h\nabla_x \cdot u, \tag{12.11}$$

where $x = (x,y)^T$, $h = 1 + \eta$, and

$$J_2 = \begin{bmatrix} 0 & +1 \\ -1 & 0 \end{bmatrix}.$$

Given a function $w(x,t)$, the material time derivative characterizes the change of w along motion of a fluid particle $X(t) = (X(t), Y(t))^T \in \mathbb{R}^2$, which is passively advected under the velocity field u, i.e.

$$\frac{DX}{Dt} = u. \tag{12.12}$$

As an example consider absolute vorticity

$$\zeta = v_x - u_y + f_0 = \nabla_x \times u + f_0.$$

Using

$$\begin{aligned}
\nabla_x \times \frac{Du}{Dt} &= \frac{\partial}{\partial x}\frac{D}{Dt}v - \frac{\partial}{\partial y}\frac{D}{Dt}u \\
&= (v_t + uv_x + vv_y)_x - (u_t + uu_x + vu_y)_y \\
&= (v_x - u_y)_t + u(v_x - u_y)_x + v(v_x - u_y)_y + (v_x - u_y)(u_x + v_y) \\
&= \frac{D}{Dt}\zeta + (\zeta - f_0)\nabla_x \cdot u,
\end{aligned}$$

it is easy to conclude from (12.10) that absolute vorticity satisfies the continuity equation

$$\frac{D\zeta}{Dt} = -\zeta\nabla_x \cdot u. \tag{12.13}$$

defining the layer-depth h at time t and Eulerian position x. Here δ denotes the Dirac delta function. Using $dX\,dY = |X_a|\,da\,db$ and (12.18), we can pull this integral back to label space, arriving at the relation

$$h(x,t) = \int h_o(a)\,\delta(x - X(a,t))\,da\,db, \tag{12.19}$$

which can be taken as the definition of the layer depth in a Lagrangian description of fluid mechanics.

The SWE (12.10)–(12.11) are now reformulated to

$$\frac{\partial u}{\partial t} = f J_2 u^\perp - c_0^2 \nabla_X h, \tag{12.20}$$

$$\frac{\partial X}{\partial t} = u, \tag{12.21}$$

where h is defined by (12.18) or (12.19). Note that the material time derivative was replaced by the partial derivative with respect to time. This reflects the fact that the material time derivative becomes a partial derivative in a Lagrangian formulation of fluid dynamics, where time t and labels $a \in \mathbb{A}$ are now the independent variables. Next we introduce the canonical momenta

$$p = h_o\,u,$$

and the equations (12.20)–(12.21) become canonical with Hamiltonian

$$\mathcal{H} = \frac{1}{2}\int \frac{\mathbf{p}\cdot\mathbf{p}}{h_o}\,da\,db + \frac{c_0^2}{2}\int h_o h\,da\,db$$

and symplectic two-form

$$\bar{\omega} := \int \left(\frac{h_o f_0}{2} dX \wedge J_2 dX + dp \wedge dX\right) da\,db. \tag{12.22}$$

12.1.4 Noncanonical Hamiltonian wave equations

We have already encountered in Chapter 8 the rigid body as an example of a noncanonical Hamiltonian system of the general form

$$\frac{d}{dt} z = J(z)\nabla_z H(z).$$

For PDEs, this generalizes to

$$u_t = \mathcal{J}(u)\delta_u \mathcal{H}.$$

Here $\mathcal{J}(u)$ is a linear (in general, differential) operator, called the Poisson operator, that has to satisfy certain properties similar to those for the matrix $J(z)$ [139].

A well-known example is provided by the Korteweg–de Vries (KdV) equation [53]

$$u_t + uu_x + u_{xxx} = 0. \tag{12.23}$$

Here the operator $\mathcal{J}$ is equal to

$$\mathcal{J} = -\frac{\partial}{\partial x},$$

and the Hamiltonian functional is given by

$$\mathcal{H} = \int \left[\frac{1}{6}u^3 - \frac{1}{2}u_x^2\right] dx.$$

We can also introduce a noncanonical Poisson bracket

$$\{\mathcal{F},\mathcal{G}\} = -\int (\delta_u \mathcal{F})\partial_x(\delta_u \mathcal{G})\,dx,$$

and the skew-symmetry of $\{\mathcal{F},\mathcal{G}\}$ follows upon integration by parts. Similar to the rigid body, the Poisson operator $\mathcal{J}$ is not invertible and this gives rise to the Casimir function

$$C[u] = \int u\,dx.$$

Indeed, it is easy to verify that

$$\frac{d}{dt} C = \{C,\mathcal{H}\} = 0,$$

along solutions of the KdV equation.

Inviscid fluid dynamics leads to Eulerian equations of motion that are also noncanonical. However, compared with KdV, the situation is made more complicated by the fact that the Poisson operator $\mathcal{J}$ is now no longer constant and independent of the dynamical variables. In fact, one can draw an analogy between two-dimensional incompressible fluid dynamics and an infinite-dimensional version of rigid body dynamics. See the review article by MORRISON [139] for further details.

12.2 Symplectic discretizations

The basic idea of symplectic discretization methods for Hamiltonian PDEs consists of two steps:

(i) A spatial truncation that reduces the PDE to a system of Hamiltonian ODEs.

(ii) Timestepping of the finite-dimensional Hamiltonian ODE using an appropriate symplectic method.

The crucial new step is the construction of a finite-dimensional ODE model that retains the Hamiltonian character of the given PDE. The most popular approach is based on the introduction of a spatial grid over which the equations of motion can be truncated. Another approach, particularly well suited for Lagrangian fluid dynamics, reduces the PDE to a set of moving particles interacting through an appropriate potential energy function. Both approaches will be described below.

Finally we give a note of warning. Certain noncanonical Hamiltonian PDEs resist a spatial truncation to a finite-dimensional Hamiltonian system. This is true in particular for the Eulerian formulation of inviscid fluid dynamics. The only significant exception is provided by incompressible fluids on a plane with double periodic boundary conditions. See ZEITLIN [208] and MCLACHLAN [128] for a numerical implementation.

12.2.1 Grid-based methods

Consider the nonlinear wave equation (12.1). The first step towards a numerical algorithm is to introduce N grid points $x_i = i\,\Delta x$, $\Delta x = L/N$, $i = 1, \ldots, N$, and to approximate functions $u \in S$ by vectors $u = (u_1, u_2, \ldots, u_N)^T \in \mathbb{R}^N$ with $u(x_i) \approx u_i$. We define $u_{i+N} = u_i$, reflecting the fact that periodic boundary conditions are imposed. The new state space is $z = \{z_i\} \in \mathbb{R}^{2N}$, $z_i = (u_i, v_i)^T \in \mathbb{R}^2$. The symplectic form (12.2) is naturally truncated to

$$\bar{\omega}_N = \sum_{i=1}^{N} du_i \wedge dv_i\,\Delta x = \frac{1}{2}\sum_{i=1}^{N} dz_i \wedge J_2^{-1} dz_i\,\Delta x,$$

and the Hamiltonian functional (12.3) is approximated by the sum

$$H = \sum_{i=1}^{N}\left[\frac{1}{2}v_i^2 + \sigma\left(\frac{u_i - u_{i-1}}{\Delta x}\right) + f(u_i)\right]\Delta x.$$

Hence, we obtain the system of Hamiltonian ODEs

$$\frac{d}{dt}z_i = \frac{1}{\Delta x}J_2\nabla_{z_i}H(z), \qquad i = 1, \ldots, N.$$

Note that

$$\lim_{\Delta x\to 0}\frac{1}{\Delta x}\nabla_u H \to \delta_u\mathcal{H}[u,v] \quad \text{and} \quad \lim_{\Delta x\to 0}\frac{1}{\Delta x}\nabla_v H \to \delta_v\mathcal{H}[u,v],$$

and the spatially discrete equations formally converge to the PDE limit (12.4).

For the specific Hamiltonian, as given above, the finite-dimensional truncation becomes

$$\frac{d}{dt}v_i = \frac{\sigma'(w_{i+1}) - \sigma'(w_i)}{\Delta x} - f'(u_i), \tag{12.24}$$

$$\frac{d}{dt}u_i = v_i, \tag{12.25}$$

$i = 1, \ldots, N$, with

$$w_{i+1} = \frac{u_{i+1} - u_i}{\Delta x}, \qquad w_i = \frac{u_i - u_{i-1}}{\Delta x}.$$

The equations of motion can be integrated in time using any canonical method such as a symplectic Euler method, for example,

$$v_i^{n+1} = v_i^n + \Delta t\left[\frac{\sigma'(w_{i+1}^n) - \sigma'(w_i^n)}{\Delta x} - f'(u_i^n)\right],$$

$$u_i^{n+1} = u_i^n + \Delta t\, v_i^{n+1},$$

where

$$w_i^n = \frac{u_i^n - u_{i-1}^n}{\Delta x}, \qquad w_{i+1}^n = \frac{u_{i+1}^n - u_i^n}{\Delta x}.$$

See MCLACHLAN [129] for further details on this classical approach to the numerical solution of Hamiltonian PDEs.

This might appear to be the end of the story. However, the interpretation of the wave equation (12.1) as an infinite-dimensional Hamiltonian system has masked some of the interesting local features of the PDE. For example, let us have another look at the Hamiltonian functional $\mathcal{H}[u, v]$. We can write this functional as

$$\mathcal{H}[u,v] = \int_0^L E(u,v)\,dx, \qquad E(u,v) = \frac{1}{2}v^2 + \sigma(u_x) + f(u).$$

The function E is called the *energy density*. Let us compute the time derivative of E

$$\begin{aligned} E_t &= vv_t + \sigma'(u_x)u_{xt} + f'(u)u_t \\ &= v\left[\partial_x\sigma'(u_x) - f'(u)\right] + \sigma'(u_x)v_x + f'(u)v \\ &= [v\sigma'(u_x)]_x. \end{aligned}$$

We have obtained what is called an *energy conservation law*

$$E_t + F_x = 0, \tag{12.26}$$

where $F = -v\sigma'(u_x)$ is called the *energy flux*. Under periodic boundary conditions, the conservation law (12.26) immediately implies conservation of total energy since $\int_0^L F_x = [F]_{x=0}^{x=L} = 0$. But the energy conservation law (12.26) is valid independently of any boundary conditions. Hence it is more fundamental than conservation of total energy.

Let us repeat the above calculation for the spatially truncated system (12.24)–(12.25). We define the discrete energy density

$$E_i = \frac{1}{2}v_i^2 + \sigma(w_i) + f(u_i),$$

and find:

$$\begin{aligned}\frac{d}{dt}E_i &= v_i\dot{v}_i + \sigma'(w_i)\frac{\dot{u}_i - \dot{u}_{i-1}}{\Delta x} + f'(u_i)\dot{u}_i \\ &= v_i\frac{\sigma'(w_{i+1}) - \sigma'(w_i)}{\Delta x} + \sigma'(w_i)\frac{v_i - v_{i-1}}{\Delta x} \\ &= \frac{v_i\sigma'(w_{i+1}) - \sigma'(w_i)v_{i-1}}{\Delta x} \\ &= -\frac{F_{i+1/2} - F_{i-1/2}}{\Delta x},\end{aligned}$$

where

$$F_{i+1/2} = -v_i\sigma'(w_{i+1}), \quad F_{i-1/2} = -v_{i-1}\sigma'(w_i).$$

Hence we have obtained a semi-discrete energy conservation law

$$\frac{d}{dt}E_i + \frac{F_{i+1/2} - F_{i-1/2}}{\Delta x} = 0,$$

and $F_{i\pm1/2}$ are approximations to the energy flux $F(u(x))$ at $x = i\Delta x \pm \Delta x/2$. Again this local energy conservation law is more fundamental than conservation of total energy H. Applying a symplectic integration method in time, we can now monitor the residual,

$$\mathrm{R}_i^{n+1/2} := \frac{E_i^{n+1} - E_i^n}{\Delta t} + \frac{F_{i+1/2}^{n+1/2} - F_{i-1/2}^{n+1/2}}{\Delta x}, \tag{12.27}$$

of a fully discretized local energy conservation law with

$$E_i^n = \frac{1}{2}(v_i^n)^2 + \sigma(w_i^n) + f(u_i^n),$$

and

$$F_{i+1/2}^{n+1/2} = -\frac{1}{2}\left[v_i^n\sigma'(w_{i+1}^n) + v_i^{n+1}\sigma'(w_{i+1}^{n+1})\right],$$

etc. Similar to the non-conservation of total energy under symplectic time integration, we cannot, in general, expect the residual $\mathrm{R}_i^{n+1/2}$ to be zero. But a formal backward error analysis has been developed by MOORE AND REICH [137, 138] to explain the remarkable global and local energy conservation of symplectic PDE discretizations observed, for example, in [162].

One can also derive methods that exactly conserve energy. See, for example, [96] and [115]. Such methods will, in general, not be symplectic.

Example 2 Let us discuss the sine-Gordon equation

$$u_{tt} = u_{xx} - \sin(u).$$

The energy density is

$$E = \frac{1}{2}((u_t)^2 + (u_x)^2) + (1 - \cos u),$$

and the energy flux is

$$F = -u_t u_x.$$

The spatial discretization (12.24)–(12.25) followed by a symplectic Euler discretization in time yields

$$\frac{v_i^{n+1} - v_i^n}{\Delta t} = \frac{w_{i+1}^n - w_i^n}{\Delta x} - \sin(u_i^n), \quad \frac{u_i^{n+1} - u_i^n}{\Delta t} = v_i^{n+1}, \quad \frac{u_{i+1}^{n+1} - u_i^{n+1}}{\Delta x} = w_{i+1}^{n+1}. \tag{12.28}$$

The semi-discretized energy conservation law is

$$\frac{d}{dt}\left[\frac{1}{2}(v_i^2 + w_i^2) + (1 - \cos(u_i))\right] + \left[\frac{-v_i w_{i+1} + v_{i-1}w_i}{\Delta x}\right] = 0. \tag{12.29}$$

Note that, upon eliminating w_i^n and v_i^n, the method (12.28) is equivalent to the classical centered leap-frog scheme

$$\frac{u_i^{n+1} - 2u_i^n + u_i^{n-1}}{\Delta t^2} = \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{\Delta x^2} - \sin(u_i^n).$$

□

We will come back to the local aspects of Hamiltonian PDEs and their numerical counterparts in Section 12.3. Numerical results will be presented in Section 12.3.4.

Let us briefly discuss a spatial discretization for the KdV equation (12.23). The Hamiltonian is easily discretized to

$$H = \sum_{i=1}^{N}\left[\frac{1}{6}u_i^3 - \frac{1}{2}\left(\frac{u_i - u_{i-1}}{\Delta x}\right)^2\right]\Delta x.$$

The spatial truncation of the Poisson operator $\mathcal{J}$ is potentially more challenging. However, for KdV, any skew symmetric approximation of the differential operator $\mathcal{J} = \partial_x$ is sufficient and will lead to a finite-dimensional Hamiltonian system. This Hamiltonian ODE can be integrated in time by a symplectic integrator. Symplectic methods based on splitting the Hamiltonian into integrable problems are available. See, for example, QUISPEL AND MCLACHLAN [132] and ASCHER AND MCLACHLAN [9].

12.2.2 Particle-based methods

Standard grid-based methods are, in general, not applicable to Lagrangian fluid dynamics. This is due to the ill-conditioning of the map from an initial grid (labels) $\{X_{ij}(0) = a_{ij}\}$, $a_{ij} = (i\Delta x, j\Delta y)$ to the advected particle positions $\{X_{ij}(t)\}$. Since this map determines the layer-depth approximation in a standard mesh-based Lagrangian method via a discrete approximation of (12.18), the quality of the simulation results is usually rather poor and instabilities are observed.

On the other hand, general grid-based methods are very easy to implement for an Eulerian formulation of fluid dynamics. See DURRAN [55] for an overview of such methods. However, none of these methods respects the Hamiltonian nature of the inviscid equations of motion. This is due to the already mentioned difficulty of finding a spatial truncation of the underlying noncanonical formulation of Eulerian fluid dynamics.

All these problems disappear if we give up the grid and work with the Lagrangian instead of the reduced Eulerian formulation of inviscid fluid dynamics. The resulting so-called *mesh-free methods* are based on an approximation of the layer-depth h via the identity (12.19). The most well-known mesh-free method for Lagrangian fluid dynamics simulations is the *Smoothed Particle Hydrodynamics* (SPH) method of LUCY [120] and GINGOLD AND MONAGHAN [71]. Many different variants of the basic SPH method have been proposed over the years. The first application of SPH to the shallow-water equations is due to SALMON [168]. We follow here the general framework of FRANK AND REICH [64].

Any spatial discretization will lead to a finite spatial resolution. For grid-based methods that resolution is directly related to the mesh-size Δx. For a mesh-free method, we have to instead introduce a smoothing or filter length $\alpha > 0$. Any fluid motion below that length scale will not be properly resolved. Hence we may replace the SWEs (12.20)–(12.21) by the 'regularized/smoothed' formulation

$$\frac{\partial}{\partial t} u = f_0 J_2 u - c_0^2 \nabla_X (\mathcal{A} * h), \tag{12.30}$$

$$\frac{\partial}{\partial t} X = u, \tag{12.31}$$

where the convolution $\mathcal{A} * h$ is defined in terms of a smooth kernel function $\psi(x, y) \geq 0$ satisfying

- $\psi(x, y) = \psi(y, x)$ (symmetry),
- $\int \psi(x, y)\, dx\, dy = 1$ (conservation of mass),
- $\psi(x, y) = \psi(\|x - y\|)$ (radial symmetry).

More explicitly, we have

$$\begin{aligned}(\mathcal{A} * h)(x, t) &= \int \left(\psi(x, \bar{x}) \int h_o(a)\, \delta(\bar{x} - X(a, t))\, da\, db \right) d\bar{x}\, d\bar{y} \\ &= \int h_o(a) \left(\int \psi(x, \bar{x})\, \delta(\bar{x} - X(a, t))\, d\bar{x}\, d\bar{y} \right) da\, db \\ &= \int h_o(a)\, \psi(x, X(a, t))\, da\, db.\end{aligned}$$

A kernel often used in the SPH method is the Gaussian

$$\psi(x, y; \alpha) = \frac{1}{\pi^{3/2}\alpha} e^{-\|x-y\|^2/\alpha^2},$$

where $\alpha > 0$ is the smoothing length scale.

To set up the numerical method, we introduce a mesh in label space with equally spaced grid points $\{a_{ij}\}$ and mesh-size $\Delta a = \Delta b$. The grid points are enumerated by integers $k = 1, \ldots, N$, which serve as discrete labels, so the map $s : (i, j) \to k$ is one-to-one. The particle positions at time t are denoted by $X_k(t)$ and initially $X_k(0) = a_{ij}$. Each particle has a "mass" $m_k = h_o(a_{ij})$ and a velocity $u_k(t) \in \mathbb{R}^2$. The layer depth at $X_k(t)$ is then approximated by

$$h_k(t) = \sum_l m_l\, \psi(X_k(t), X_l(t))\, \Delta a \Delta b.$$

Similarly, any integral of the form

$$I = \int h_o(a)\, w(X(a))\, da\, db$$

is approximated by

$$I \approx \sum_k m_k w_k\, \Delta a \Delta b, \qquad w_k = w(X_k).$$

From here on we can follow exactly the same approach as outlined for the symplectic discretization of grid-based methods. The Hamiltonian functional

$$\mathcal{H} = \frac{1}{2} \int \frac{\mathbf{p} \cdot \mathbf{p}}{h}\, da\, db + \frac{c_0^2}{2} \int h_o (\mathcal{A} * h)\, da\, db$$

Taking the wedge product with $J^{-1}dz^n$ and $J^{-1}dz^{n+1}$, respectively, from the left, we obtain

$$J^{-1}dz^{n+1} \wedge dz^n = J^{-1}dz^n \wedge dz^n + \Delta t J^{-1} J H_{zz} \frac{dz^n + dz^{n+1}}{2} \wedge dz^n$$
$$= J^{-1}dz^n \wedge dz^n + \frac{\Delta t}{2} H_{zz} dz^{n+1} \wedge dz^n,$$

and

$$J^{-1}dz^{n+1} \wedge dz^{n+1} = J^{-1}dz^n \wedge dz^{n+1} + \Delta t J^{-1} J H_{zz} \frac{dz^n + dz^{n+1}}{2} \wedge dz^{n+1}$$
$$= J^{-1}dz^{n+1} \wedge dz^n - \frac{\Delta t}{2} H_{zz} dz^{n+1} \wedge dz^n.$$

Here we have made use of Property 3 of the wedge product as stated in Section 3.6 and, in particular, used that

$$dz^n \wedge A dz^n = dz^{n+1} \wedge A dz^{n+1} = 0,$$

for any symmetric matrix $A \in \mathbb{R}^{2d \times 2d}$. Summing up, we arrive at the equality

$$J^{-1}dz^{n+1} \wedge dz^{n+1} = J^{-1}dz^n \wedge dz^n,$$

implying symplecticness of the scheme with respect to the structure matrix J. Observe that this argument required at no point that J be the canonical structure matrix, only that it be constant, skew-symmetric, and invertible, thus we have shown: *the implicit midpoint method preserves any constant symplectic structure*.

4.2 Construction of symplectic methods by Hamiltonian splitting

In the above discussion, we have shown that several integrators are symplectic when applied to integrate Hamiltonian systems. In this section, we show that there is a simple technique that can often be used to produce good symplectic methods.

Suppose that we can split the Hamiltonian H into the sum of $k \geq 2$ Hamiltonians H_i, $i = 1, \ldots, k$, i.e.

$$H(z) = \sum_{i=1}^{k} H_i(z),$$

with each Hamiltonian vector field

$$\frac{d}{dt} z = J \nabla_z H_i(z)$$

explicitly solvable. A symplectic integrator is then derived as an appropriate composition of the corresponding flow maps. Since each flow map is obviously symplectic and any composition of symplectic maps yields a symplectic map, the resulting numerical method is symplectic. See problem 2 in the Exercises.

As a simple example, consider a nonlinear oscillator $H(q,p) = \frac{1}{2}p^2 + \varphi(q)$. The energy can be decomposed into kinetic and potential terms

$$H = H_1 + H_2, \quad H_1 = \frac{1}{2}p^2, \quad H_2 = \varphi(q).$$

Now each term is exactly integrable. The equations of motion for H_1 are

$$\dot{q} = p,$$
$$\dot{p} = 0,$$

which has the flow map

$$\Phi_{t,H_1}\left(\begin{bmatrix} q \\ p \end{bmatrix}\right) = \begin{bmatrix} q + tp \\ p \end{bmatrix}.$$

Similarly, H_2 has flow map

$$\Phi_{t,H_2}\left(\begin{bmatrix} q \\ p \end{bmatrix}\right) = \begin{bmatrix} q \\ p - t\varphi'(q) \end{bmatrix}.$$

Each of these maps is symplectic (each is the flow map of a Hamiltonian system), hence the map defined by

$$\Psi_{\Delta t} := \Phi_{\Delta t,H_1} \circ \Phi_{\Delta t,H_2}$$

is also symplectic.

We still need to show that such a composition method approximates the flow map with at least first order.

FIRST-ORDER SPLITTING

If $H = H_1 + H_2 + \ldots + H_k$ is any splitting into twice differentiable terms, then the composition method

$$\Psi_{\Delta t} = \Phi_{\Delta t,H_1} \circ \Phi_{\Delta t,H_2} \circ \cdots \circ \Phi_{\Delta t,H_k}$$

is (at least) a first-order symplectic integrator.

Let us walk through a proof for the case $k = 2$. In Chapter 5 a more detailed result will be given. We wish to compare the flow map $\Phi_{\Delta t,H}$ with $\Psi_{\Delta t}$. It is enough to compare the images of an arbitrary point z^0 under the two maps. Using a Taylor series expansion in powers of Δt and the definition of the flow map, it is easy to show that

$$\begin{aligned}\Phi_{\Delta t,H}(z^0) &= z^0 + \Delta t \dot{z}(0) + \mathcal{O}(\Delta t^2) \\ &= z^0 + \Delta t J \nabla H(z^0) + \mathcal{O}(\Delta t^2).\end{aligned}$$

We can argue in a similar way that

$$\Phi_{\Delta t,H_i}(z^0) = z^0 + \Delta t J \nabla H_i(z^0) + \mathcal{O}(\Delta t^2),$$

$i = 1, 2$. And then, performing similar expansions in Δt, we arrive at

$$\begin{aligned}\Psi_{\Delta t}(z^0) &= \Phi_{\Delta t,H_1}\left(\Phi_{\Delta t,H_2}(z^0)\right) \\ &= \Phi_{\Delta t,H_2}(z^0) + \Delta t J \nabla H_1\left(\Phi_{\Delta t,H_2}(z^0)\right) + \mathcal{O}(\Delta t^2) \\ &= z^0 + \Delta t J \nabla H_2(z^0) + \Delta t J \nabla H_1(z^0) + \mathcal{O}(\Delta t^2) \\ &= z^0 + \Delta t J \nabla H(z^0) + \mathcal{O}(\Delta t^2) \\ &= \Phi_{\Delta t,H}(z^0) + \mathcal{O}(\Delta t^2).\end{aligned}$$

This estimate of the local error, together with the evident smoothness of the flow map, proves that the composition method is at least first order.

The only apparent drawback of this approach is that it requires the splitting of the given Hamiltonian into explicitly solvable subproblems. This may not always be possible or desirable. In many cases, the system may admit a partitioning, but without the individual terms being exactly integrable. In these cases, one may be able to construct effective schemes by substituting another symplectic integrator for the exact flow map at some stage. The splitting technique may in this way simplify the development of an effective method by breaking down a complicated problem into a series of lesser challenges.

A curious special case arises when H_1 and H_2 are first integrals of each other, i.e., $\{H_1, H_2\}$. Then the two flow maps commute (see problem 4 in the Exercises) and the composition method is exact

$$\Phi_{\Delta t,H_1+H_2} = \Phi_{\Delta t,H_1} \circ \Phi_{\Delta t,H_2} = \Phi_{\Delta t,H_1} \circ \Phi_{\Delta t,H_2}.$$

4.2.1 Separable Hamiltonian systems

The splitting described earlier for the special case of the oscillator is applicable to any separable Hamiltonian of the form

$$H(q, p) = T(p) + V(q).$$

As we have seen in earlier chapters, such systems are ubiquitous in chemical and physical modeling, being the standard form for N-body simulations with a flat (i.e. Euclidean) kinetic energy metric.

As before, the form of the energy function suggests a natural splitting into kinetic energy

$$H_1(p) = T(p),$$

and potential energy

$$H_2(q) := V(q).$$

The differential equations corresponding to H_2 can be written

$$\begin{aligned}\frac{d}{dt}q &= 0, \\ \frac{d}{dt}p &= -\nabla_q V(q).\end{aligned}$$

These equations are completely integrable, since q is constant along solutions and p therefore varies linearly with time. The flow map is

$$\Phi_{\tau,V}(q, p) = \begin{bmatrix} q \\ p - \tau \nabla_q V(q) \end{bmatrix}.$$

Similarly, we can derive the flow map for the kinetic term (H_1)

$$\Phi_{\tau,T}(q, p) = \begin{bmatrix} q + \tau \nabla_p T(p) \\ p \end{bmatrix}.$$

Now consider the composition of these two maps for $\tau = \Delta t$,

$$\Psi_{\Delta t} := \Phi_{\Delta t,T} \circ \Phi_{\Delta t,V}.$$

Applying this map to a point of phase space (q^n, p^n), we first compute a point $(\bar{q}, \bar{p})$

$$\begin{aligned}\bar{q} &= q^n, \\ \bar{p} &= p^n - \Delta t \nabla_q V(q^n).\end{aligned}$$

Next, apply $\Phi_{\Delta t,T}$ to this point, i.e.

$$\begin{aligned}q^{n+1} &= \bar{q} + \Delta t \nabla_p T(\bar{p}), \\ p^{n+1} &= \bar{p}.\end{aligned}$$

These equations can be simplified by the elimination of the intermediate values, to yield

$$\begin{aligned}q^{n+1} &= q^n + \Delta t \nabla_p T(p^{n+1}), \\ p^{n+1} &= p^n - \Delta t \nabla_q V(q^n).\end{aligned}$$

This is evidently the Euler-B method introduced in the previous section specialized to the case of a mechanical Hamiltonian. While the Euler-B method is normally implicit, it becomes explicit when applied to this special class of Hamiltonians due to the separation of variable dependencies.

By reversing the order in which the two maps are applied, we obtain another composition method, $\Phi_{\Delta t,V} \circ \Phi_{\Delta t,T}$, which reduces after a similar calculation to the Euler-A method.

An interesting point that should be mentioned is that not all symplectic maps are given by a splitting. For the general Hamiltonian, it is clear that the Euler-A and Euler-B methods are not obtained from any splitting. It is only for the special case of a separable Hamiltonian that these methods can be viewed in this special way.

4.2.2 A second-order splitting method

Higher-order splitting methods are also easily constructed. We will take up this theme in more detail in Chapter 6. For now, as an illustration, consider again the mechanical Hamiltonian and the splitting

$$H = H_1 + H_2 + H_3,$$

with

$$H_1 = \frac{1}{2}V(q), \qquad H_2 = T(p), \qquad H_3 = \frac{1}{2}V(q).$$

The associated composition method is equivalent to

$$\Psi_{\Delta t} = \Phi_{\Delta t/2,V} \circ \Phi_{\Delta t,T} \circ \Phi_{\Delta t/2,V}.$$

After simplification, it becomes clear that this integrator is nothing other than the second-order Störmer–Verlet method (2.16)–(2.18) of Chapter 2 written in terms of the canonical coordinates

$$p^{n+1/2} = p^n - \frac{1}{2}\Delta t \nabla_q V(q^n), \tag{4.14}$$

$$q^{n+1} = q^n + \Delta t \nabla_p T(p^{n+1/2}), \tag{4.15}$$

$$p^{n+1} = p^{n+1/2} - \frac{1}{2}\Delta t \nabla_q V(q^{n+1}). \tag{4.16}$$

Since it is a splitting method we may infer: *the Störmer–Verlet method (4.14)–(4.16) is canonically symplectic.*

Just as the implicit midpoint method turned out to be the composition of implicit and explicit Euler steps, notice that the Störmer–Verlet method can be

expressed as a composition of half-steps using the Euler-A and Euler-B methods, i.e.

$$\begin{aligned}\Phi_{\Delta t,\frac{1}{2}V} \circ \Phi_{\Delta t,T} \circ \Phi_{\Delta t,\frac{1}{2}V} &= \Phi_{\Delta t,\frac{1}{2}V} \circ \Phi_{\Delta t,\frac{1}{2}T} \circ \Phi_{\Delta t,\frac{1}{2}T} \circ \Phi_{\Delta t,\frac{1}{2}V} \\ &= \Phi_{\frac{1}{2}\Delta t,V} \circ \Phi_{\frac{1}{2}\Delta t,T} \circ \Phi_{\frac{1}{2}\Delta t,T} \circ \Phi_{\frac{1}{2}\Delta t,V}.\end{aligned}$$

4.3 Time-reversal symmetry and reversible discretizations

An important geometric property of Newton's equations of motion is related to the invariance of a Hamiltonian $H = H(q,p)$ under the *reflection symmetry*

$$p \mapsto -p,$$

i.e. the Hamiltonian is an even function in the momentum p. We consider the consequences of this property for the solution behavior of the corresponding Hamiltonian system

$$\frac{d}{dt}q = \nabla_p H(q,p),$$
$$\frac{d}{dt}p = -\nabla_q H(q,p).$$

The key fact is this: if $(q(t),p(t))$ is a solution then also $(\hat{q}(t),\hat{p}(t)) := (q(-t), -p(-t))$ is.

To prove this observe that

$$\frac{d}{dt}\hat{q}(t) = -\dot{q}(-t) = -\nabla_p H(q(-t),p(-t)) = \nabla_p H(\hat{q}(t),\hat{p}(t)),$$

and

$$\frac{d}{dt}\hat{p}(t) = \dot{p}(-t) = -\nabla_q H(q(-t),p(-t)) = -\nabla_q H(\hat{q}(t),\hat{p}(t)).$$

Here we used the general fact that if $h(x,y)$ is an even function of y, then h_x is also even in y, and h_y is odd in y, meaning that $h_x(x,-y) = h_x(x,y)$, and $h_y(x,-y) = -h_y(x,y)$.

The invariance of the Hamiltonian with respect to $p \mapsto -p$ evidently implies that for every solution of the Hamiltonian system, there is another solution which traverses the same positional curve but in the opposite direction, with a negated momentum.

Another way of saying this is that if we evolve the solution τ units in time forward from a given point (q_0,p_0) to (q_τ,p_τ)