INSTITUTE OF PHYSICS PUBLISHING REPORTS ON PROGRESS IN PHYSICS

Rep. Prog. Phys. **68** (2005) 1703–1759 doi:10.1088/0034-4885/68/8/R01

Smoothed particle hydrodynamics

J J Monaghan

School of Mathematical Sciences, Monash University, Vic 3800, Australia

E-mail: joe.monaghan@sci.monash.edu.au

Received 25 February 2005, in final form 26 May 2005
Published 5 July 2005
Online at stacks.iop.org/RoPP/68/1703

Abstract

In this review the theory and application of Smoothed particle hydrodynamics (SPH) since its inception in 1977 are discussed. Emphasis is placed on the strengths and weaknesses, the analogy with particle dynamics and the numerous areas where SPH has been successfully applied.

Contents

1. Introduction

Smoothed particle hydrodynamics (SPH) is a method for obtaining approximate numerical solutions of the equations of fluid dynamics by replacing the fluid with a set of particles. For the mathematician, the particles are just interpolation points from which properties of the fluid can be calculated. For the physicist, the SPH particles are material particles which can be treated like any other particle system. Either way, the method has a number of attractive features. The first of these is that pure advection is treated exactly. For example, if the particles are given a colour, and the velocity is specified, the transport of colour by the particle system is exact. Modern finite difference methods give reasonable results for advection but the algorithms are not Galilean invariant so that, when a large constant velocity is superposed, the results can be badly corrupted. The second advantage is that with more than one material, each described by its own set of particles, interface problems are often trivial for SPH but difficult for finite difference schemes. The third advantage is that particle methods bridge the gap between the continuum and fragmentation in a natural way. As a consequence, the best current method for the study of brittle fracture and subsequent fragmentation in damaged solids is SPH (see, e.g. Benz and Asphaug (1994, 1995)). A fourth advantage is that the resolution can be made to depend on position and time, which makes the method very attractive for most astrophysical and many geophysical problems. Fifth, SPH has the computational advantage, particularly in problems involving fragments, drops or stars, that the computation is only where the matter is, with a consequent reduction in storage and calculation. Finally, because of the close similarity between SPH and molecular dynamics, it is often possible to include complex physics easily.

Although the idea of using particles is natural, it is not obvious which interactions between the particles will faithfully reproduce the equations of fluid dynamics or continuum mechanics. One way of doing this was proposed by Bob Gingold and myself (Gingold and Monaghan (1977) where the term SPH was coined) and independently by Lucy (1977). Gingold and Monaghan derived the equations of motion using a kernel estimation technique, pioneered by statisticians, to estimate probability densities from sample values (Rosenblatt (1956), Parzen (1962) and, for a general discussion, see Boneva *et al* (1971)). When applied to interpolation, this yielded an estimate of a function at any point using the values of the function at the particles. This estimate of the function could be differentiated exactly provided the kernel was differentiable. In this way, the gradient terms required for the equations of fluid dynamics could be written in terms of the properties of the particles. Because of its close relation to the statistical ideas, Gingold and Monaghan (1977) described the method as a Monte Carlo method, as did Lucy (1977) who had, in effect, re-discovered the statistical technique. However, in subsequent papers (e.g. Gingold and Monaghan (1978)), it was discovered that the errors were much smaller than the predicted probability estimates. Gingold and Monaghan realized that the particle number density was not equivalent to a probability density because the fluctuations predicted by probability theory require energy, which is not available from the equations of motion. This is particularly easy to see in the case of static equilibrium as the system moves to a minimum energy state in which large voids do not occur, since they require higher energy. In a dynamical problem more disorder can occur but only to the extent allowed by the dynamical equations.

The original papers (Gingold and Monaghan 1977, Lucy 1977) proposed numerical schemes which did not conserve linear and angular momentum exactly, but gave good results for a class of astrophysical problems that were considered too difficult for the techniques available at the time. The basic SPH algorithm was improved to conserve linear and angular momentum exactly using the particle equivalent of the Lagrangian for a compressible non-dissipative fluid (Gingold and Monaghan 1982). In this way, the similarities between SPH and

molecular dynamics were made clearer. Recent studies by Hoover (1998) and Hoover *et al* (2004) explore the correspondence between SPH and molecular dynamics.

Since SPH models a fluid as a mechanical and thermodynamical particle system, it is natural to derive the SPH equations for non-dissipative flow from a Lagrangian. The equations for the early SPH simulations of binary fission and instabilities were derived from Lagrangians (Gingold and Monaghan 1978, 1979, 1980). These Lagrangians took into account the smoothing length (the same for each particle) which was a function of the coordinates. More recent examples include Lagrangians which incorporate a resolution length for each particle (Springel and Hernquist 2002, Monaghan 2002), a relativistic Lagrangian (Monaghan and Price 2001), a Lagrangian for MHD problems (Price and Monaghan 2004a, 2004b) and a Lagrangian for SPH compressible turbulence (Monaghan 2002). In addition, Bonet and his colleagues (Bonet and Lok 1999, Bonet and Kulasegaram 2000, 2001) have used Lagrangians for the SPH simulation of elastic materials. The advantage of a Lagrangian is that it not only guarantees conservation of momentum and energy, but also ensures that the particle system retains much of the geometric structure of the continuum system in the phase space of the particles. This includes Liouville's theorem and the Poincare invariants. In addition, as noted by Dirac, basing the equations of motion on a Lagrangian allows new physical interactions to be included consistently.

The comments made by Von Neumann in 1944 (see Von Neumann (1944)), in connection with the use of the particle methods to model shocks, are relevant to SPH. To paraphrase his remarks:

> *The particle method is not only an approximation of the continuum fluid equations, but also gives the rigorous equations for a particle system which approximates the molecular system underlying, and more fundamental than the continuum equations.*

When combined with a simple but effective viscosity, and a form of the thermal energy equation that guarantees that the viscous dissipation increases both the thermal energy and the entropy, a variety of shock problems have been studied (Monaghan and Gingold 1983, Monaghan 1997, Price and Monaghan 2004a). The SPH algorithm gives very satisfactory results for shocks though they are not as accurate as those obtained from well-designed Riemann solvers and other modern techniques—although these have their own set of problems, especially when approximate Riemann solvers are used (Quirk 1994). Sharpness is often overated as a measure of the fidelity of simulations. Real shocks are only a few mean free paths thick so that, in a typical shock tube of 2 m length, $\sim 10^7$ finite difference cells, in each direction, would be required in a finite difference code to resolve the shock. However, most codes can afford only 10^3 cells along each coordinate so that their numerical shock widths are 10^4 times greater than the actual shock width. Therefore, the discussion about which code gives the sharpest shocks is irrelevant; they are all outstandingly bad. What are relevant are the pre- and post-shock values of the physical variables. SPH is able to obtain these as accurately as desired in one dimension, but in two and three dimensions SPH shocks, using current algorithms, can be noisy. In astrophysical problems, this should not be a cause for concern because the flows are invariably turbulent and the noise created in an SPH shock is small relative to that owing to turbulence.

In problems involving very small perturbations, the lower accuracy of SPH makes finite difference methods preferable. However, it has advantages which show up in those fluid problems where the perturbations are large. The first of these is that complex physics can often be included with little effort and effective codes produced in days, whereas finite difference codes would take many months or years to write. The second is that the SPH method can be easily extended to include a resolution which varies in space and time. That is, each particle has

its own resolution length (Gingold and Monaghan (1982), see their section 4). It is, therefore, ideal for astrophysical problems where enormous variations in the relevant length scales are common (see, e.g. the simulation of the formation of the Moon (Benz *et al* (1986), or the star formation studies of Bate *et al* (1995) and Bate *et al* (2003) or the binary neutron star collisions of Rosswog and Davies (2002)). Furthermore, the SPH method combines easily with the particle methods used for star systems and is a natural tool for cosmological simulations, in particular (see, e.g. Hernquist and Katz (1989), Couchman *et al* (1995), Springel and Hernquist (2002) and Marri and White (2003)).

Because SPH is essentially a technique for approximating the continuum equations, it can be used for a wide range of fluid dynamical problems. Although the initial applications were to gas dynamic problems, it has also been applied to problems in incompressible flow by treating that flow as slightly compressible with an appropriate equation of state (Monaghan 1994). Using the same idea waves, breaking on arbitrary structures (Monaghan *et al* 2004, Colagrossi and Landrini 2003) as well as the more classical problems of waves on beaches (Monaghan and Kos 1999) could be simulated. Colagrossi (2004) has made a detailed study of the application of SPH to breaking waves, where an accurate boundary element method could be used up to the point where the wave curls over to touch the water surface in the front. The SPH calculations agree with the boundary element method up to the point that it can be used, and thereafter the SPH method gives good agreement with the experiment. Colagrossi (2004) also shows that the SPH simulation of sloshing tanks and the bow waves produced by certain ship hulls are in good agreement with the experiment. Simulations of liquid metal moulding (Cleary and Ha 2002) also show good agreement with the experiment.

Another class of problems suitable for the SPH algorithm arise in elasticity and fracture. Libersky and Petschek (1991) derived and applied the SPH equations for elasticity. Benz and Asphaug (1994, 1995) showed how SPH could be applied to the fracture of brittle solids, where it gives much better results than the finite element or the finite difference methods. These methods have been applied to the breakup of planetesimals and the formation of asteroid families (Michel *et al* 2004). In these simulations, the ease with which the SPH particles can describe the transition from a continuum to a set of fragments gives it a computational edge over other numerical methods. Commerical software packages (e.g. Dyna3D and Autodyn) for simulating impact now incorporate SPH. Elastic SPH also provides a simple and robust technique for simulating complex fracture in geological rock formations and in brittle materials (Gray *et al* 2001, Gray and Monaghan 2004). SPH is also being used in virtual reality surgery (see, e.g. the work of M Mueller, S Schirm and M Teschner at the Computer Graphics Laboratory ETH, Zurich).

In many of these problems *a priori* estimates of the accuracy of SPH interpolation suggest that the simulations would give results which would be too inaccurate for most problems. As a consequence, a technique called Moving Least Squares (Dilts 1999) was developed to produce a particle code with perfect linear interpolation. However, the disadvantages are that conservation is lost and the method is considerably slower than the standard SPH. Furthermore, in practice, as noted earlier, the low accuracy predicted from interpolation errors usually does not occur. For example, Colagrossi (2004) shows that, for the complex evolution of a patch of fluid, the SPH results are as good as those from the level set method, and often surpass them. Part of the reason may be that, for non-dissipative problems, the equations follow directly from a Lagrangian, which retains many of the properties of the original continuum Lagrangian.

In problems involving heat conduction, Cleary and Monaghan (1999) showed that the SPH simulations, which conserve thermal energy and guarantee that the entropy increased, were very accurate even though the particles' positions were disordered and the thermal conductivity discontinuous. These results together with those mentioned earlier show that if SPH equations

are set up such that they satisfy the fundamental conservation laws, the results are much better than would be deduced from consideration of the interpolation alone.

The reader may find the early reviews of SPH (Monaghan 1992, Benz 1990) useful. A different aspect of SPH is detailed in the website www.nextlimit.com, which shows a wide variety of SPH simulations of fluids for both scientific problems and for video and film special effects (In the third film of the trilogy 'Lord of the Rings', Nextlimit used SPH to simulate Gollum falling into the lava.)

2. Interpolation

The equations of fluid dynamics have the form

$$\frac{\mathrm{d}A}{\mathrm{d}t} = f(A, \nabla A, \boldsymbol{r}), \tag{2.1}$$

where

$$\frac{\mathrm{d}}{\mathrm{d}t} = \frac{\partial}{\partial t} + \boldsymbol{v} \cdot \nabla \tag{2.2}$$

is the Lagrangian derivative, or the derivative following the motion. It is worth noting that the characteristics of this differential operator are the particle trajectories.

In the equations of fluid dynamics, the rates of change of physical quantities require spatial derivatives of physical quantities. The key step in any computational fluid dynamics algorithm is to approximate these derivatives using information from a finite number of points. In finite difference methods, the points are the vertices of a mesh. In the SPH method, the interpolating points are particles which move with the flow, and the interpolation of any quantity, at any point in space, is based on kernel estimation.

2.1. Integral and summation interpolants and their kernels

SPH interpolation of a quantity A, which is a function of the spatial coordinates, is based on the integral interpolant

$$A_{\mathrm{I}}(\boldsymbol{r}) = \int A(\boldsymbol{r}')W(\boldsymbol{r} - \boldsymbol{r}', h)\,\mathrm{d}\boldsymbol{r}', \tag{2.3}$$

where the function W is the kernel and $\mathrm{d}\boldsymbol{r}'$ is a differential volume element . The interpolant reproduces A exactly if the kernel is a delta function. In practice, the kernels are functions which tend to the delta function as the length scale h tends to zero. They are normalized to 1 so that the constants are interpolated exactly. An example in one dimension x is the Gaussian kernel $W(x, h) = \exp(-x^2/h^2)/(h\sqrt{\pi})$. The Gaussian kernel was used by Gingold and Monaghan (1977) and a kernel with continuous second derivatives of the form $W(r, h) = (105/(16\pi h^3))(1 - r/h)^3(1 + 3r/h)$ in $0 \leqslant r \leqslant h$, and zero otherwise, was used by Lucy (1977) for his three-dimensional calculations. The most commonly used kernels are based on Schoenberg (1946) M_n splines, which are piece-wise continuous functions with compact support having the derivatives up to $(n - 2)$ continuous. They can be defined by the Fourier transform

$$M_n(x, h) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\frac{\sin kh/2}{kh/2}\right)^n \cos(kx)\,\mathrm{d}k \tag{2.4}$$

and algebraic forms are given by Schoenberg (1946) and Monaghan (1985b). The M_2 spline with $q = |x|/h$, is

$$M_2(x) = \begin{cases} 1 - q, & \text{for } 0 \leqslant q \leqslant 1, \\ 0, & \text{for } q \geqslant 1. \end{cases} \tag{2.5}$$

M_2 gives linear interpolation but its first derivative is discontinuous. In its product form it gives what is called equal area interpolation (Hockney and Eastwood 1988). A commonly used kernel is the M_4 kernel (called the cubic spline because it is a piecewise cubic polynomial). It has the form,

$$M_4(x) = \begin{cases} \frac{1}{6}[(2-q)^3 - 4(1-q)^3], & \text{for } 0 \leqslant q \leqslant 1, \\ \frac{1}{6}(2-q)^3, & \text{for } 1 \leqslant q \leqslant 2, \\ 0, & \text{for } q > 2. \end{cases} \tag{2.6}$$

The SPH kernel associated with $M_n(x)$ in one dimension is $W(x,h) = M_n(x)/h$. In $\hat{d}$ dimensions the same functional forms are used but they are multiplied by $1/h^{\hat{d}}$ and by a constant to ensure they are normalized in the new space. For example, the factor $1/6$ in the cubic spline (2.6) is replaced by $15/(14\pi)$ in two dimensions and by $1/(4\pi)$, in three dimensions. Higher order interpolation using splines was studied by Monaghan (1985a). The higher order kernels perform very well for equi-spaced particles, but they require a cancellation of positive and negative contributions which is less likely when the particles are disordered. Furthermore, many of the desirable features of SPH involving positive definite dissipation terms are lost when higher order kernels are used because the gradient of the kernel changes sign. Schoenberg (1946) also discusses a class of smoothing kernels with Fourier transforms which have a Gaussian decay with increasing k. These have not been used in simulations.

Alternative kernels have been studied by Fulk and Quinn (1996) in one dimension. According to their measures, no kernel is significantly better than the cubic spline. Price (2004a) has studied the effect of changing the joining points in one dimension without finding a kernel significantly better than the cubic spline. In higher dimensions it is not clear whether optimum interpolation is obtained with equi-spaced joining points for the piece-wise polynomials of the M_n functions. It would be interesting to study either the functions, or their Fourier transforms, when the joining points are allowed to be arbitrary. It may be that equal volumes should be cut by the slices between the joining points.

To apply this interpolation to a fluid, we divide it into a set of small mass elements. The element a will have a mass m_a, density ρ_a and position $\boldsymbol{r}_a$. The value of A at particle a is denoted by A_a. The interpolation integral can be written as

$$\int \frac{A(\boldsymbol{r}')}{\rho(\boldsymbol{r}')}\rho(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}', \tag{2.7}$$

where an element of mass is $\rho\,\mathrm{d}\boldsymbol{r}'$. The integral can then be approximated by a summation over the mass elements. This gives the summation interpolant

$$A_{\mathrm{s}}(\boldsymbol{r}) = \sum_b m_b \frac{A_b}{\rho_b} W(\boldsymbol{r} - \boldsymbol{r}_{\mathrm{b}}, h), \tag{2.8}$$

where the summation is over all the particles but, in practice, it is only over near neighbours because W falls off rapidly with distance. Typically, h is close to the particle spacing and the kernel W is effectively zero beyond a distance $2h$ (as in the case of the kernel based on the cubic spline M_4). In practice, we choose kernels which have compact support, i.e. they vanish at a finite distance.

As an example of the use of kernel estimation, suppose A is the density ρ. The interpolation formula then gives the following estimate for the density at a point $\boldsymbol{r}$

$$\rho(\boldsymbol{r}) = \sum_b m_b W(\boldsymbol{r} - \boldsymbol{r}_{\mathrm{b}}, h), \tag{2.9}$$

which shows how the mass of a set of particles is smoothed to produce the estimated density. The reader who is familiar with the technique of estimating probability densities from sample

points (Rosenblatt 1956, Parzen 1962) will see that our formula for the density is the same as their formulae for the probability density but with m_b replaced by $1/N$, where N is the number of sample points.

If h is constant, we can integrate the density estimate to give

$$\int \rho(\boldsymbol{r})\,\mathrm{d}\tau = \sum_b m_b = M, \tag{2.10}$$

which shows that mass is conserved exactly. If we allow h to vary, the integral is no longer exactly M, but the total mass is conserved because it is carried by the particles.

2.2. *First derivatives*

The SPH formulation allows derivatives to be estimated easily. If W is a differentiable function then (2.8) can be differentiated exactly to give

$$\frac{\partial A_\mathrm{s}}{\partial x} = \sum_b m_b \frac{A_b}{\rho_b}\frac{\partial W}{\partial x}. \tag{2.11}$$

In SPH the derivative is, therefore, found by an *exact derivative* of an approximate function. However, this form of the derivative does not vanish if A is constant. A simple way to ensure that it does vanish if A is constant is to write

$$\frac{\partial A}{\partial x} = \frac{1}{\Phi}\left(\frac{\partial(\Phi A)}{\partial x} - A\frac{\partial \Phi}{\partial x}\right), \tag{2.12}$$

where Φ is any differentiable function. The SPH form of (2.12) is

$$\left(\frac{\partial A}{\partial x}\right)_a = \frac{1}{\Phi_a}\sum_b m_b \frac{\Phi_b}{\rho_b}(A_b - A_a)\frac{\partial W_{ab}}{\partial x_a}, \tag{2.13}$$

which vanishes if A is constant. In this expression, and elsewhere, W_{ab} denotes $W(\boldsymbol{r}_a - \boldsymbol{r}_b, h)$. Different choices of Φ give all the versions of derivatives in the literature. For example, choosing $\Phi = 1$ gives

$$\frac{\partial A_a}{\partial x_a} = \sum_b \frac{m_b}{\rho_b}(A_b - A_a)\frac{\partial W_{ab}}{\partial x_a} \tag{2.14}$$

and choosing $\Phi = \rho$,

$$\frac{\partial A_a}{\partial x_a} = \frac{1}{\rho_a}\sum_b m_b(A_b - A_a)\frac{\partial W_{ab}}{\partial x_a}. \tag{2.15}$$

These results have immediate application to the convergence equation (often called the continuity equation, but in this review it will be called the *convergence* equation since $-\nabla\cdot\boldsymbol{v}$ is the opposite of *divergence*)

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -\rho\nabla\cdot\boldsymbol{v}. \tag{2.16}$$

Generalizing the previous expressions for derivatives to approximate $\nabla\cdot\boldsymbol{v}$ we find

$$\frac{\mathrm{d}\rho_a}{\mathrm{d}t} = \rho_a\sum_b \frac{m_b}{\rho_b}\boldsymbol{v}_{ab}\cdot\nabla_a W_{ab} \tag{2.17}$$

and

$$\frac{\mathrm{d}\rho_a}{\mathrm{d}t} = \sum_b m_b \boldsymbol{v}_{ab}\cdot\nabla_a W_{ab}, \tag{2.18}$$

where $\boldsymbol{v}_{ab}$ denotes $\boldsymbol{v}_a - \boldsymbol{v}_b$ and ∇_a denotes the gradient taken with respect to the coordinates of particle a. This equation is the time derivative of the summation form of the density (2.9).

If (2.17) is compared with (2.18) it will be seen that the former involves ρ explicitly in the summation, whereas the latter does not. Both expressions vanish, as they should, when the velocity is constant. However, when the system involves two or more fluids with large density ratios in contact, the expression (2.17) with ρ in the summation is more accurate (Colagrossi 2004). The reason is that near an interface the summation for $\nabla \cdot \boldsymbol{v}$ for one type of fluid SPH particle involves contributions from the other fluid. If we imagine the other fluid being changed for a fluid with exactly the same velocity field, and exactly the same particle positions but different density, we would still want the same estimate of $\nabla \cdot \boldsymbol{v}$. However, with (2.18) the mass elements will be changed and the estimate will be different, but if (2.17) is used the ratio of mass to density will be constant and $\nabla \cdot \boldsymbol{v}$ will not change. In practice, it turns out that either (2.17) or (2.18) can be used for density ratios $\leqslant 2$, but for larger density ratios it is better to use (2.17). The Lagrangian approach, which we consider later, requires that these equations for the rate of change of density with time be included as constraints. As a result, the form of the pressure forces changes with the form chosen for the density convergence equation.

Although the focus in the previous analysis has been on designing interpolation formula to achieve satisfactory accuracy it is natural with particle methods to interpret the formula in terms of interactions between SPH particles. In the present case we expect that as particles get closer their density will increase. In particular, any two particles moving closer together should give a positive contribution to their density. In either form of the convergence equation, we can write

$$\nabla_a W_{ab} = \boldsymbol{r}_{ab} F_{ab}, \tag{2.19}$$

where $F_{ab} \leqslant 0$ is a function of $|\boldsymbol{r}_{ab}|$. The contribution of particle b to the density of particle a in (2.18) is then

$$\rho_a \frac{m_b}{\rho_b} \boldsymbol{v}_{ab} \cdot \boldsymbol{r}_{ab} F_{ab} \tag{2.20}$$

and if the particles a and b are approaching each other (so that $\boldsymbol{v}_{ab} \cdot \boldsymbol{r}_{ab} \leqslant 0$) the contribution to the density change is positive definite as expected. The same is true for (2.17).

2.3. *Second derivatives*

As in the case of first derivatives, second derivatives can be estimated by differentiating an SPH interpolant twice. For example, in a heat conduction problem in one dimension, the second derivative of the temperature T at the position of particle a can be estimated by

$$\left(\frac{\mathrm{d}^2 T}{\mathrm{d}x^2}\right)_a = \sum_b m_b T_b \frac{\mathrm{d}^2 W(x_a - x_b, h)}{\mathrm{d}x_a^2}. \tag{2.21}$$

However, this expression has a number of disadvantages. First, it is very sensitive to particle disorder. Second, the transfer of heat to particle a from particle b may be positive or negative depending on their separation because the second derivative of the kernel can change sign. Physics tells us that a hot particle should transfer heat to a cold particle no matter what the separation. Another disadvantage is that this expression will not result in conservation of thermal energy in an adiabatic enclosure.

A much better approach (Brookshaw 1985, Cleary and Monaghan 1999) is to begin with an integral approximation to the second derivative. For example, starting with

$$I = \int (\kappa(\boldsymbol{r}) + \kappa(\boldsymbol{r}'))(T(\boldsymbol{r}) - T(\boldsymbol{r}'))F(|\boldsymbol{r} - \boldsymbol{r}'|)\,\mathrm{d}\boldsymbol{r}', \tag{2.22}$$

where $\boldsymbol{q}F(|\boldsymbol{q}|) = \nabla W(\boldsymbol{q}, h)$, expanding $\kappa(\boldsymbol{r}')$ and $T(\boldsymbol{r}')$ in a Taylor series about $\boldsymbol{r}$, and keeping up to second order terms, we find

$$I = \nabla \cdot (\kappa \nabla T) + O(h^2). \tag{2.23}$$

The SPH form of I for particle a is

$$I = \sum_b \frac{m_b}{\rho_b}(\kappa_a + \kappa_b)(T_a - T_b)F_{ab}. \tag{2.24}$$

Because $F \leqslant 0$ this expression has the property that if $T_a > T_b$, then heat will flow from particle a to b (that is the contribution to $\mathrm{d}T_a/\mathrm{d}t < 0$) and vice versa. Other second derivatives can be calculated using similar integral expressions.

2.3.1. Second derivatives in two dimensions. To obtain second derivatives integrals of the form

$$J_{xx} = \int \frac{\Delta x \Delta x}{\Delta r^2}(\kappa(\boldsymbol{r}) + \kappa(\boldsymbol{r}'))(T(\boldsymbol{r}) - T(\boldsymbol{r}'))F \,\mathrm{d}\boldsymbol{r}' \tag{2.25}$$

are used (Español and Revenga 2003). Here $\Delta x = x - x'$ and $\Delta r = |\boldsymbol{r} - \boldsymbol{r}'|$. Expanding the κ and T terms in a Taylor series gives, to $O(h^2)$,

$$J_{xx} = \kappa\left(\tfrac{3}{4}T_{xx} + \tfrac{1}{4}T_{yy}\right) + \tfrac{3}{4}\kappa_x T_x + \tfrac{1}{4}\kappa_y T_y \tag{2.26}$$

and

$$J_{yy} = \kappa\left(\tfrac{3}{4}T_{yy} + \tfrac{1}{4}T_{xx}\right) + \tfrac{3}{4}\kappa_y T_y + \tfrac{1}{4}\kappa_x T_x, \tag{2.27}$$

such that

$$J_{xx} + J_{yy} = \nabla \cdot (\kappa \nabla T). \tag{2.28}$$

Furthermore,

$$J_{xy} = \tfrac{1}{4}(2\kappa T_{xy} + T_x\kappa_y + T_y\kappa_x). \tag{2.29}$$

If we construct the J integrals taking $\kappa = 1$, we get estimates for the second derivatives of T in the form (now using tensor notation for the coordinates denoted by x^i and $\Delta x^i = (x^i - x'^i)$)

$$\frac{\partial^2 T}{\partial x^i \partial x^j} = \int \left[4\frac{\Delta x^i \Delta x^j}{\Delta r^2} - \delta^{ij}\right](T(\boldsymbol{r}) - T(\boldsymbol{r}'))F \,\mathrm{d}\boldsymbol{r}' \tag{2.30}$$

or, in SPH form

$$\left(\frac{\partial^2 T}{\partial x^i \partial x^j}\right)_a = \sum_b \frac{m_b}{\rho_b}\left(4\frac{\Delta x^i \Delta x^j}{\Delta r^2} - \delta^{ij}\right)(T_a - T_b)F_{ab}. \tag{2.31}$$

2.3.2. Second derivatives in three dimensions. With the same definition of J_{xx} as before, but now integrating over three dimensions we find

$$J_{xx} = \tfrac{1}{5}\kappa(3T_{xx} + T_{yy} + T_{zz}) + \tfrac{1}{5}(3\kappa_x T_x + \kappa_y T_y + \kappa_z T_z), \tag{2.32}$$

with similar expressions for J_{yy} and J_{zz}. The integral J_{xy} becomes

$$J_{xy} = \tfrac{1}{5}(2\kappa T_{xy} + T_x\kappa_y + T_y\kappa_x). \tag{2.33}$$

These results show that

$$J_{xx} + J_{yy} + J_{zz} = \nabla \cdot (\kappa \nabla T) + O(h^2). \tag{2.34}$$

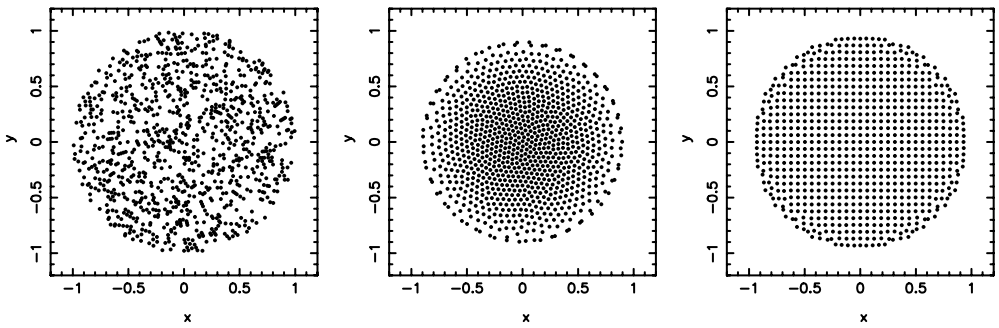

Figure 1. The frame on the left shows SPH particles placed at random according to a constant probability density within a unit circle. Note the large voids. The middle frame shows the positions of equal mass particles settled down in a Toy star potential starting with the particle positions in the left frame. The right-hand frame shows variable mass SPH particles also settled down in a Toy star potential. These latter particles were initially placed on a cubic lattice with cell length Δp and given a mass $\rho(\Delta p)^2$ using the theoretical static density. The cubic spline kernel was used and h was calculated selfconsistently (see section 4).

A similar expression to (2.30) with the factor 4 replaced with 5 in the integral, gives the second derivatives of T. These results generalize those of Español and Revenga (2003) who work out the case where κ is a constant. The application of these derivatives to problems involving viscous effects and thermal conduction will be considered later.

2.4. Errors in the integral interpolant

It is not easy to estimate the errors in the SPH equations from first principles because the particles get disordered during motion. The errors depend on the type of disorder which, in turn, depends on the dynamics. One approach to estimating the errors is to begin with particles on a regular lattice, then give each particle a random shift in position (Colagrossi 2004). However, this kind of short wavelength disorder does not usually occur if the particle spacing is much smaller than the dominant length scales of the motion. For example, if the particles are damped to an equilibrium, they fall into a nearly regular cell structure which depends, in general, on the kernel being used and the masses of the particles (see, e.g. figure 1). If the particles are in motion, for example, in a breaking wave, most of the particles are in a type of nearly ordered array associated with shearing a regular array of particles. The end result is that SPH simulations are much more accurate than the interpolation of quantities from randomly disordered particle arrays would suggest. For that reason, it is better to run carefully designed test cases to assess the accuracy of an SPH simulation. However, it is still interesting to study the kernel interpolation on a regular array of points.

Starting with the integral interpolant in one dimension

$$A_{\rm I}(x) = \int A(x')W(x-x',h)\,{\rm d}x' = A(x) + \int (A(x') - A(x))W(x-x',h)\,{\rm d}x'. \qquad (2.35)$$

The error can be estimated by a Taylor series expansion of $A(x')$. Assuming $W(q,h)$ is an even function of q, the interpolant gives

$$A_{\rm I}(x) = A(x) + \frac{\sigma h^2}{2}\frac{{\rm d}^2A(x)}{{\rm d}x^2}\cdots, \qquad (2.36)$$

where σ is a constant dependent on the kernel. The integral interpolant, therefore, gives at least a second order interpolation. The interpolation is better if σ is zero, in which case higher

order terms in the Taylor series expansion must be included. The third order term vanishes because of symmetry leaving a possible fourth order term. All these results assume that the integrals can be extended to the entire volume within the support of the kernel. If this is not possible, for example, near a boundary, the error is larger.

Monaghan (1985a) gave a technique for constructing higher order kernels from lower order kernels using a variant of Richardson extrapolation. An example is the kernel

$$W(x,h) = \frac{1}{h\sqrt{\pi}}\left(\frac{3}{2} - \frac{x^2}{h^2}\right)\mathrm{e}^{-x^2/h^2}, \tag{2.37}$$

which is based on the Gaussian. For this kernel, the integral interpolant is accurate to $O(h^4)$. This kernel changes sign; a necessary feature of higher order interpolation. Unfortunately this may have unwanted side effects, including the possibility that the density might become negative near a strong shock. It would, however, be possible to use a high order kernel using a switch from high to low order kernels near shocks. Such a technique has been used but not fully explored.

2.5. Errors in the summation interpolant

If the particles are equi-spaced in one dimension, we can easily estimate the errors in the summation interpolant using the Poisson summation formula

$$\sum_{j=-\infty}^{\infty} f(j) = \int_{-\infty}^{\infty} f(j)\,\mathrm{d}j + 2\sum_{r=1}^{\infty}\int_{-\infty}^{\infty}\cos(2\pi r j) f(j)\,\mathrm{d}j, \tag{2.38}$$

where, on the right-hand side, j is treated as a continuous quantity.

Consider the interpolation of the function $g(x) = \alpha + \beta x$ with the particles equi-spaced with spacing Δ along an infinite line so that $\rho = 1$ and $m = \Delta$. The SPH interpolation formula gives, at $x_a = a\Delta$, the following expression for g at the point $x = a\Delta$.

$$\Delta\sum_{j=-\infty}^{\infty}(\alpha + \beta j\Delta)W(a\Delta - j\Delta, h). \tag{2.39}$$

If the origin is shifted to the point $a\Delta$ and the Poisson summation formula is used together with the assumption that the kernel is an even function, (2.39) becomes

$$(\alpha + \beta a\Delta)\left(\int_{-\infty}^{\infty} W(q,h)\,\mathrm{d}q + 2\int_{-\infty}^{\infty}\cos\left(\frac{2\pi q}{\Delta}\right)W(q,h)\,\mathrm{d}q + \cdots\right). \tag{2.40}$$

This formula shows how the error depends on the Fourier transform of the kernel (Schoenberg 1946). If the kernel is a Gaussian, the previous expression becomes

$$(\alpha + \beta a\Delta)(1 + 2\mathrm{e}^{-\pi^2h^2/\Delta^2} + \cdots) \tag{2.41}$$

In this simple case, we conclude that the SPH summation interpolant does not even interpolate a constant exactly, but the error is exponentially small and is negligible if $h > \Delta$. If we have any sufficiently smooth kernel the Fourier transform decreases rapidly and the error can be made negligible. The frequently used cubic spline kernel gives the following expression for the previous interpolation:

$$(\alpha + \beta a\Delta)\left(1 + 2\left(\frac{\sin\pi h/\Delta}{\pi h/\Delta}\right)^4 + \cdots\right). \tag{2.42}$$

In this case the dominant error terms vanish if $h = \Delta$ and are small if $h > \Delta$.

Of greater interest are the errors in the derivative. For the previous case, and using (2.13) we find $\mathrm{d}g/\mathrm{d}x$ is estimated by

$$\beta\Delta \sum_{j=-\Delta\infty}^{\infty} j\Delta \frac{\partial W(a\Delta - j\Delta, h)}{\partial x_a}. \tag{2.43}$$

Using the Poisson summation formula again and shifting the origin of the summation to a, we find the derivative is given by

$$\beta\left(1 - \int_{\infty}^{\infty} q\frac{\partial W}{\partial q}\cos\left(\frac{2\pi q}{\Delta}\right)\mathrm{d}q + \cdots\right). \tag{2.44}$$

The error now involves the Fourier transform of the gradient of the kernel and is larger than in the case of the function interpolation. In the case of the Gaussian, the errors are again exponentially small and are negligible if $h > \Delta$. We conclude from these results (which may be easily extended to 2 or more dimensions) that the SPH interpolation is as accurate as desired provided the particles are equi-spaced in an infinite space. This has led some researchers (Chaniotis *et al* 2002) to use re-meshing strategies for SPH, and their simulations of homogeneous fluids give very good results. At fixed boundaries they use one-sided interpolation which works well. However, boundaries, such as free surface liquid problems, then require special care as do multi-phase and multi-material problems.

2.6. Errors when the particles are disordered

During the course of an SPH calculation the particles become disordered. The exact form of this disorder depends on the dynamics. When Bob Gingold and I first ran the SPH calculations, we thought that the disorder could be described by a probability distribution proportional to the mass density and that the errors could be estimated in the same way as a Monte Carlo estimate. In particular, we expected that the errors arising from fluctuations would be $\sim 1/\sqrt{N}$, where N is the number of particles. However, the errors were much smaller than this estimate would suggest. The reason for the smaller errors, as mentioned earlier, is that the probability estimates allow fluctuations which are inconsistent with the dynamics. The result is that the SPH particles are disordered, but in an *orderly* way.

For example, the left frame of figure 1 shows the positions of 971 particles with equal mass placed at random within a unit circle. The middle frame shows the same particles after they have been allowed to evolve in a simple linear force field, where the equation of motion is

$$\frac{\mathrm{d}\boldsymbol{v}_a}{\mathrm{d}t} = -\nu\boldsymbol{v}_a - \sum_b m_b\left(\frac{P_a}{\rho_a^2} + \frac{P_b}{\rho_b^2}\right)\nabla_a W_{ab} - \boldsymbol{r}_a, \tag{2.45}$$

where the term $-\nu\boldsymbol{v}$ damps the motion, the terms involving the pressure P approximate the pressure gradient (discussed in detail in section 3) and the last term is the body force. In this force field the exact density varies with radius r according to $(1 - r^2)$. The particles are still disordered but the large voids and concentrations appearing in the left frame of figure 1 have disappeared, and the disorder is far from random. The set of particles gives a density field shown in figure 2. Because the density is accurate, we can deduce that the gradients of the pressure field are accurately computed in spite of the disorder in the particles.

The simulation can be set up differently by choosing particle positions for particles which begin on a lattice of square cells with sides of length Δp. The particles have mass $\rho\Delta p^2$, which varies over the domain. Only the particles that have a radius $r < (1 - 0.5\Delta p)$ are kept such that no particle has zero mass. In the previous case the particles had equal mass and their spacing varied. Now the particles have different masses but equal initial spacing. After evolving the particles with the same damping as before, the particle positions settle into the state shown in

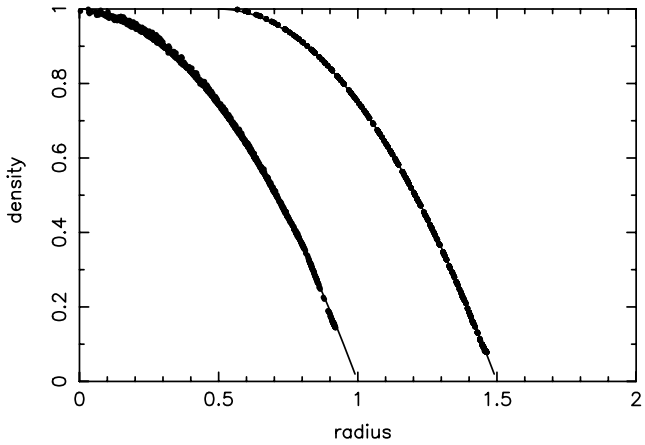

Figure 2. The density radius relation for an SPH simulation of a gas in a linear force field. The exact results are shown by the curved lines and the SPH results by the filled circles. The left-hand curve is for the equal mass particle case and the right-hand curve (shifted for clarity) is for the variable mass particles with nearly equal separation.

the right-hand frame of figure 1. As expected, by analogy with atomic systems, the difference in the force between pairs produces a different particle equilibrium configuration.

Since the disorder depends on the dynamics, it is not possible to make traditional error estimates like those used for finite differences or finite elements. In fact, the previous examples show that, at least for equilibrium configurations, the SPH particles seek the best positions for a given interpolation formula. This is a profoundly different picture from that for finite differences, where the best interpolation formula is sought for a given grid. For this reason, estimates of SPH calculations have had to depend on comparisons with known solutions, comparisons with experiments or by studying how the error varies with particle number (see, e.g. Cleary and Monaghan (1999)). These comparisons show that it is possible to achieve very accurate results with SPH. An example is given in figure 3, where the function $r^2 \exp(-6r^2)$ is interpolated using the cubic spline and the interpolation formula (2.8) with the particle positions shown in the central frame of figure 1.

The calculation of derivatives is less accurate except for the calculation of the density derivatives from the pressure force term in the equation of motion. That derivative is accurate because the particles are forced to move to an equilibrium position where the density gradient is determined accurately to balance the applied force which is $\propto \boldsymbol{r}$. If the derivative with respect to x of $(r^2 - 1)$ is calculated using (2.14), the results are shown in figure 4. The lower curve is for the case of equal mass particles and the upper curve for variable mass (the graphs are shifted by one for clarity). Only the particles within 0.9 of the outer radius were used for these plots. These particles comprise 96% of the mass. The mean square error in the gradient is 0.02.

One reason for the accuracy of SPH despite the disorder in dynamical problems is that it is possible to devise SPH algorithms so that they conserve important quantities like momentum and energy. The importance of this conservation shows up in simple problems involving the integration of ordinary differential equations. Suppose, for example, that we wish to integrate the equations for a binary star system with the stars treated as points and we are offered either a Verlet symplectic integrator (since the system is Hamiltonian) or a standard fourth order Runge–Kutta integrator. The Runge–Kutta scheme is of higher order so that, if we use the same time step in each case, a numerical analyst might argue that the Runge–Kutta will give more *accurate* results than the Verlet integrator. However, the Runge–Kutta scheme

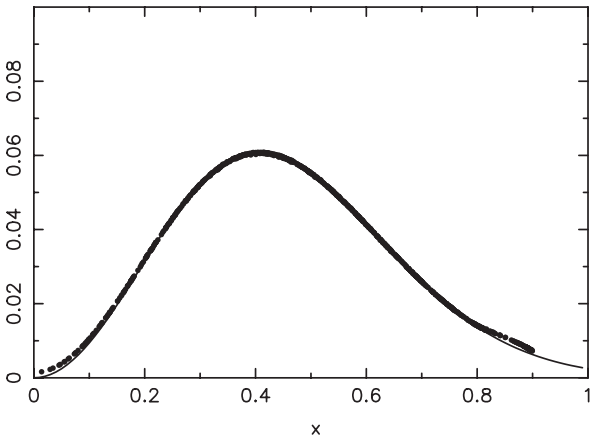

Figure 3. The function $r^2 \exp(-6r^2)$ interpolated using a cubic spline and the particle distribution shown in figure 2. The continuous line is the exact result. The dots show the SPH results.

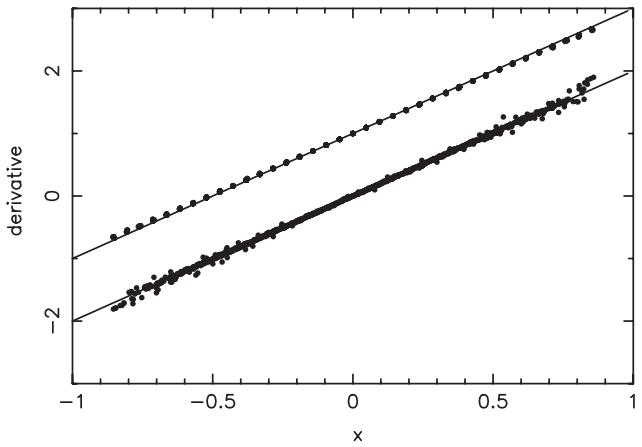

Figure 4. The SPH derivative with respect to x of $(r^2 - 1)$ using (2.14). The upper points are for the particles with initially equal spacing. They have been shifted by 1.0 for clarity. The lower points are for the case of equal mass particles. The points included have $r < 0.9$ and contain 96% of the mass.

produces a less accurate orbit. The effect is more extreme as the eccentricity gets closer to 1. The problem arises because the standard fourth order Runge–Kutta does not conserve angular momentum (it is also not reversible, whereas the system is). On the other hand, the symplectic integrator, which is a lower order integrator, gives much better results because it conserves angular momentum and is reversible. In this example the order of the integrator is less important than the conservation. It turns out that in SPH simulations, and in molecular dynamics, integrators which give very good conservation are to be preferred over higher order integrators which do not have good conservation properties. For these reasons it is preferable to write the gradient terms of SPH algorithms so that conservation is very accurate.

An example of the accuracy of SPH in a complex evolution of a liquid is due to Colagrossi (2004) and shown in figure 5. The liquid is initially in the shape of a square. The initial velocity field for a square with initial side length L is

$$(v_x, v_y) = (V(e^{-(4y/L)^2} - e^{-4}), -V(e^{-(4x/L)^2} - e^{-4})), \tag{2.46}$$

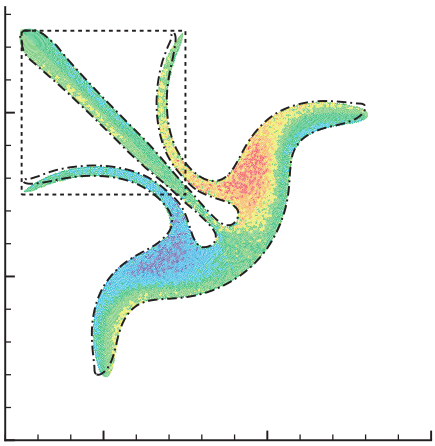

Figure 5. The evolution of an initial square of liquid (- - -) computed using SPH and a combination of level set and finite difference methods (— · —, Colagrossi (2004)).

(This figure is in colour only in the electronic version)

has spatially varying vorticity and produces severe distortion of the original square. The vertices initially have zero velocity and three of the vortices remain at rest. The dash–dot lines show the position of the outer boundary calculated using a combination of level set and finite difference techniques. The agreement between the two methods is remarkably good and shows that SPH is capable of simulating complex flows very satisfactorily.

An alternative approach to accuracy is taken by Vila and Lanson and their colleagues (Ben Moussa *et al* 1999) who extend the idea of Johnson and Beissel (1996) to use normalized kernels, but do so within the framework of an estimate of error bounds. This approach is more mathematical than the applied mathematical approach described in this review. As a result, these authors can get rigorous bounds on errors but with coefficients which cannot be determined accurately. Kahan (1980), in a witty discussion of the problems of estimating errors, comments on the pessimistic nature of error bounds and the options for estimating them accurately, in the following terms:

> 'Both options are often so pessimistic and so costly that most people prefer to take their chances with computations carried out with precisions believed, rightly or wrongly, to exceed by far what is necessary. Their attitude makes sense; they would rather believe the error to be negligible than know how big it isn't'.

However, the analysis of Villa and Lanson has led them to a re-appraisal of the SPH normalization of Johnson and Beissel with promising results.

3. SPH Euler equations

The Euler equations are the equations for the rates of change of velocity, density and position, namely,

$$\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = -\frac{1}{\rho}\nabla P + \boldsymbol{g}, \tag{3.1}$$

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -\rho\nabla\cdot\boldsymbol{v}, \tag{3.2}$$

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = \boldsymbol{v}, \tag{3.3}$$

where $\boldsymbol{v}$ is the velocity, ρ the density, P the pressure and $\boldsymbol{g}$ is the body force per unit mass. In this equation the time derivative is the derivative following the motion. In general, P is a function of ρ and the thermal energy but in the present case where there is no dissipation, the pressure can be taken as a function of ρ and the entropy per unit mass s, which remains unchanged during the motion. In some cases we will assume the entropy is the same for all particles, but, in general, each particle could have a different entropy.

The equation for the rate of change of density and its SPH equivalent have been discussed earlier. The SPH acceleration equation is discussed in the following sections.

3.1. The SPH acceleration equation

The original forms of SPH (Gingold and Monaghan 1977, Lucy 1977) converted the acceleration equation into SPH by writing

$$(\nabla P)_a = \sum_b m_b \frac{P_b}{\rho_b} \nabla_a W_{ab}, \tag{3.4}$$

such that

$$\frac{\mathrm{d}\boldsymbol{v}_a}{\mathrm{d}t} = -\frac{1}{\rho_a} \sum_b m_b \frac{P_b}{\rho_b} \nabla_a W_{ab}. \tag{3.5}$$

However, (3.5) does not conserve linear or angular momentum exactly, since the force on particle a owing to b is not equal and opposite to the force on b owing to a or

$$\frac{m_a m_b P_b}{\rho_a \rho_b} \nabla_a W_{ab} \neq -\frac{m_a m_b P_a}{\rho_a \rho_b} \nabla_b W_{ab}, \tag{3.6}$$

because $P_a \neq P_b$. Note that $\nabla_a W_{ab} = -\nabla_b W_{ab}$.

To write the acceleration equation in a form which conserves linear and angular momentum the original approach was to make use of a Lagrangian (Gingold and Monaghan (1978, 1979) and in more detail Gingold and Monaghan (1982)). However, the same result is obtained by noting that

$$\frac{\nabla P}{\rho} = \nabla \left(\frac{P}{\rho} \right) + \frac{P}{\rho^2} \nabla \rho. \tag{3.7}$$

Using the SPH interpolation rules, (3.7) becomes

$$\frac{\mathrm{d}\boldsymbol{v}_a}{\mathrm{d}t} = -\sum_b m_b \left(\frac{P_b}{\rho_b^2} + \frac{P_a}{\rho_a^2} \right) \nabla_a W_{ab}. \tag{3.8}$$

Writing

$$\nabla_a W_{ab} = \boldsymbol{r}_{ab} F_{ab}, \tag{3.9}$$

where F_{ab} is a scalar function of $|\boldsymbol{r}_a - \boldsymbol{r}_b|$, the force on a owing to b is then

$$m_a m_b \left(\frac{P_b}{\rho_b^2} + \frac{P_a}{\rho_a^2} \right) \boldsymbol{r}_{ab} F_{ab}, \tag{3.10}$$

which is equal and opposite to the force on b owing to a. As a consequence, linear and angular momentum are conserved exactly if h is constant or a symmetric function of a and b. It is possible to maintain this conservation even when h is allowed to vary (see later).

This pair force is actually a disguised many-body force because the pressure and density depend on the distribution of the particles and, in general, the resolution length also depends on the particle number density. The result is that, in general, the dynamics of an SPH system differs from an atomic or molecular system which can be approximated by pure pair forces.

3.2. The energy equations

The assumptions of the Euler equation do not require the time rate of change of thermal energy to be calculated. However, it is convenient to convert the non-dissipative rate of change of thermal energy to its SPH form. From the first law of thermodynamics

$$T\,\mathrm{d}s = \mathrm{d}u + P\,\mathrm{d}V, \tag{3.11}$$

$$= \mathrm{d}u - \frac{P}{\rho^2}\,\mathrm{d}\rho, \tag{3.12}$$

where s is the entropy and all quantities are per unit mass the time rate of change of thermal energy is

$$\frac{\mathrm{d}u}{\mathrm{d}t} = \frac{P}{\rho^2}\frac{\mathrm{d}\rho}{\mathrm{d}t} = -\frac{P}{\rho^2}\nabla\cdot\boldsymbol{v}. \tag{3.13}$$

Using the SPH form for $\nabla\cdot\boldsymbol{v}$ given earlier, the previous equation can be written either as

$$\frac{\mathrm{d}u_a}{\mathrm{d}t} = \frac{P_a}{\rho_a^2}\sum_b m_b\boldsymbol{v}_{ab}\cdot\nabla_a W_{ab} \tag{3.14}$$

or

$$\frac{\mathrm{d}u_a}{\mathrm{d}t} = \frac{P_a}{\rho_a}\sum_b \frac{m_b}{\rho_b}\boldsymbol{v}_{ab}\cdot\nabla_a W_{ab}. \tag{3.15}$$

A good general principle when writing SPH equations is to approximate the same quantity is the same way in all the equations. For example, in the equation for the rate of change of thermal energy, the particular expression for $\nabla\cdot\boldsymbol{v}$ should be the same as that used in the time rate of change of the density.

In addition to an equation for the thermal energy, it is useful to consider the equation for the thermokinetic energy per unit mass defined by

$$\hat{e} = \tfrac{1}{2}v^2 + u. \tag{3.16}$$

The rate of change of $\hat{e}$ with time can be deduced from equations for the acceleration and the rate of change of u. The continuum equation derived in this way is

$$\frac{\mathrm{d}\hat{e}}{\mathrm{d}t} = -\frac{1}{\rho}\nabla\cdot(P\boldsymbol{v}). \tag{3.17}$$

Following the same procedure, but now using the SPH equations we find

$$\frac{\mathrm{d}\hat{e}_a}{\mathrm{d}t} = -\sum_b m_b\left(\frac{P_a\boldsymbol{v}_b}{\rho_a^2} + \frac{P_b\boldsymbol{v}_a}{\rho_b^2}\right)\cdot\nabla_a W_{ab}. \tag{3.18}$$

The continuum limit of this SPH equation is

$$\frac{\mathrm{d}\hat{e}}{\mathrm{d}t} = -\frac{P}{\rho^2}\nabla\cdot(\rho\boldsymbol{v}) - \boldsymbol{v}\cdot\nabla\left(\frac{P}{\rho}\right) = -\frac{1}{\rho}\nabla\cdot(P\boldsymbol{v}). \tag{3.19}$$

Calculations of shock phenomena with finite difference methods often use the thermokinetic energy equation rather than the thermal energy equation because it ensures conservation of the energy. Furthermore, in relativistic problems, it is natural to work with momentum and energy equations which guarantee conservation of momentum and thermokinetic energy. Because of the symmetry of the SPH equation, the rate of change with time of the total thermokinetic energy $\sum_a m_a\hat{e}_a$ is zero.

4. Resolution varying in space and time

In the original calculations of Gingold and Monaghan (1977), each particle had the same h proportional to $(\langle \boldsymbol{r}^2\rangle - \langle \boldsymbol{r}\rangle^2)^{1/2}$ where, for example, $\langle \boldsymbol{r}^2\rangle$ denotes the mass average

$$\langle \boldsymbol{r}^2\rangle = \frac{\sum_b m_b \boldsymbol{r}_b^2}{\sum_b m_b}. \tag{4.1}$$

During a simulation, h is then automatically increased as the particle system expands and decreased as it contracts. In their binary fission calculations, Gingold and Monaghan (1978) used an h proportional to the inverse of the gravitational energy of the system. These two choices were crude attempts to automatically match the resolution length h to the scale of the system. Gingold and Monaghan (1982) suggested that it would be preferable to allow h_a for any particle a to be related to the density according to

$$h_a = \sigma \left(\frac{m_a}{\rho_a}\right)^{1/d}, \tag{4.2}$$

where d is the number of dimensions and σ is a constant $\sim$1.3. This has proved to be a powerful and robust way of specifying the resolution length h. It automatically gives SPH a resolution which varies in time and space and, if used consistently, leads to SPH equations which can be derived from a Lagrangian.

If the density is determined by summation, the density for particle a can be written as

$$\rho_a = \sum_b m_b W_{ab}(h_a). \tag{4.3}$$

The usual approach in the literature is either to calculate h_a at any time using the current value of ρ_a (estimated from the SPH summation), or to calculate h_a from the rate of change of density according to

$$\frac{\mathrm{d}\ln h}{\mathrm{d}t} = -\frac{1}{\mathrm{d}}\frac{\mathrm{d}\ln \rho}{\mathrm{d}t}. \tag{4.4}$$

Various techniques may then be used to adjust the h_a. For example, Steinmetz and Mueller (1993) average the local density and use this to change h. Another often used alternative is to adjust h so that each particle has a constant number of neighbours (Hernquist and Katz 1989).

Ideally, h should be determined from the summation equations so that it is consistent with the density obtained from the summation (Monaghan 2002). Equation (4.3) is a non-linear equation for the single variable ρ_a, which can be solved rapidly by point iteration possibly combined with a Newton–Raphson scheme. For example, in the case of a Toy star potential, starting with random positions in the left frame of figure 1, the mean square error in solving (4.3) is reduced by a factor 10 each point iteration, and one iteration is often sufficient. Further iterations are only required for a sub-set of the particles and the time required for extra iterations is not much (Price 2004b).

In some problems it might be necessary to replace (4.2) by a formula that limits how large or small h can become. For example, an upper bound on h_a when ρ_a becomes very small is desirable to prevent strong interactions between a very low and a very high density region. This can be achieved if (4.2) is replaced by

$$h_a = \sigma \left(\frac{m_a}{A + \rho_a}\right)^{1/d}, \tag{4.5}$$

where A is a suitable constant. A lower bound can be, similarly, included. In all cases (4.3) can be solved consistently.

5. Lagrangian equations

The Lagrangian L for the non-dissipative motion of a fluid in a potential $\Phi(\boldsymbol{r})$ per unit mass is (Eckart 1960)

$$L = \int \rho \left(\frac{1}{2} v^2 - u(\rho, s) - \Phi \right) \mathrm{d}\boldsymbol{r}, \tag{5.1}$$

where $\boldsymbol{v}$ is the velocity, u the thermal energy per unit mass, ρ the density and s is the entropy. We assume that the entropy of each element of fluid remains constant, though each particle can have a different entropy. SPH Lagrangian equations of motion have been obtained by Springel and Hernquist (2002) using a constraint on the mass within a sphere of radius h_a about particle a and by Monaghan (2002) assuming a functional relation between h and ρ. In this review we use the latter approach.

The SPH form of Eckart's Lagrangian is

$$L = \sum_b m_b \left(\frac{1}{2} v_b^2 - u(\rho_b, s_b) - \Phi_b \right). \tag{5.2}$$

From Lagrange's equations for particle a

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \boldsymbol{v}_a} \right) - \frac{\partial L}{\partial \boldsymbol{r}_a} = 0, \tag{5.3}$$

we find

$$\frac{\mathrm{d}\boldsymbol{v}_a}{\mathrm{d}t} = -\sum_b m_b \left(\frac{\partial u}{\partial \rho} \right)_s \frac{\partial \rho_b}{\partial \boldsymbol{r}_a} - \frac{\partial \Phi_a}{\partial \boldsymbol{r}_a}. \tag{5.4}$$

From the first law of thermodynamics

$$\left(\frac{\partial u}{\partial \rho} \right) = \frac{P}{\rho^2}. \tag{5.5}$$

From the SPH summation for the density (2.9) (assuming h is a function of ρ as in (4.2)),

$$\Omega_b \frac{\partial \rho_b}{\partial \boldsymbol{r}_a} = \sum_c m_c \nabla_a W_{ac}(h_a) \delta_{ab} - m_a \nabla_b W_{ab}(h_b), \tag{5.6}$$

where the gradient of W_{ab} is taken keeping h constant, δ_{ab} is the Kronecker delta, and

$$\Omega_b = 1 - H_b \sum_c m_c \frac{\partial W_{bc}(h_b)}{\partial h_b}. \tag{5.7}$$

Here H_b denotes $\partial h_b / \partial \rho_b$.

Using these results (5.4) becomes

$$\frac{\mathrm{d}\boldsymbol{v}_a}{\mathrm{d}t} = -\sum_b m_b \left(\frac{P_a}{\Omega_a \rho_a^2} \nabla_a W_{ab}(h_a) + \frac{P_b}{\Omega_b \rho_b^2} \nabla_a W_{ab}(h_b) \right) + \boldsymbol{g}_a, \tag{5.8}$$

where δ_{ab} is a Kronecker delta, $\boldsymbol{g}_a$ is the force/mass owing to the potential Φ and ∇_a denotes the gradient taken with respect to the coordinates of particle a.

In the case of equi-spaced particles in one dimension, Ω can be estimated using the Poisson summation formula. We find

$$\Omega = 1 + 2h \frac{\partial \tilde{W}}{\partial h}, \tag{5.9}$$

where $\tilde{W}$ is the Fourier transform of W. For the case of a Gaussian kernel

$$\Omega = 1 - \left(\frac{2\pi h}{\Delta} \right)^2 \mathrm{e}^{-(\pi h/\Delta)^2}. \tag{5.10}$$

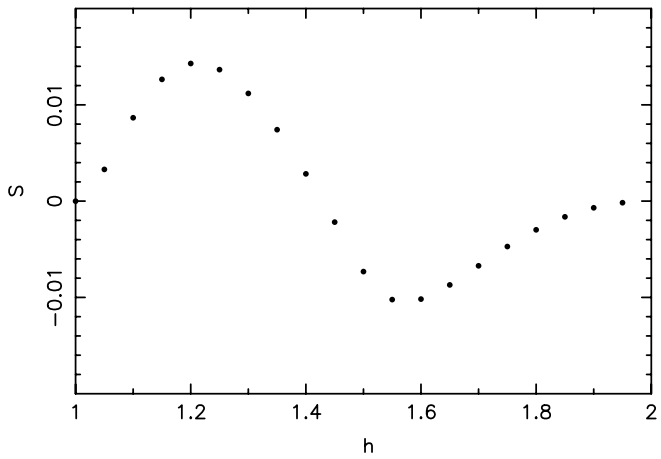

Figure 6. The function S for the cubic spline with equi-spaced particles. The values of h are scaled to the particle spacing.

Since $\pi h/\Delta \sim 4$ this result shows that for the Gaussian kernel and equi-spaced particles Ω is very close to 1. The cubic spline estimate of

$$S = \frac{h}{\rho}\sum_c m_c \frac{\partial W_{bc}(h_b)}{\partial h_b} \tag{5.11}$$

in one dimension is shown in figure 6. The value of S for the cubic spline is larger than for the Gaussian.

However, when ρ varies significantly, Ω can vary significantly and it must be included to give accurate wave propagation. Finally we note (see Monaghan (2002), Price and Monaghan (2004a)) that the rate of change of density with time (2.18), when h is a function of ρ, becomes

$$\frac{\mathrm{d}\rho_a}{\mathrm{d}t} = \frac{1}{\Omega_a}\sum_b m_b \boldsymbol{v}_{ab}\cdot\nabla_a W_{ab}(h_a), \tag{5.12}$$

where the gradient is taken with h_a constant. Similarly, the rate of change of thermal energy per unit mass is

$$\frac{\mathrm{d}u_a}{\mathrm{d}t} = \frac{P_a}{\Omega_a\rho_a^2}\sum_b m_b \boldsymbol{v}_{ab}\cdot\nabla_a W_{ab}(h_a). \tag{5.13}$$

5.1. Conservation laws

The conservation laws can be deduced either from the equations of motion or from the invariance of the Lagrangian to infinitesimal transformations.

5.1.1. Momentum conservation. Linear and angular momentum will be conserved, provided the Lagrangian (5.2) is invariant to translations and rotations. Because the SPH density and therefore the thermal energy term (with constant entropy) is invariant to these transformations, so the Lagrangian, the fluid dynamical terms, therefore, conserve linear and angular momentum. If the force terms owing to the potential are also invariant to the transformations (this is true of self-gravity), the entire system will conserve momentum. The same result follows from (5.8), using

$$\nabla_a W_{ab}(h_a) = \boldsymbol{r}_{ab}F_{ab}(h_a), \tag{5.14}$$

where $F_{ab}(h_a) = F(|\boldsymbol{r}_{ab}|, h_a)$. From the symmetry of the interaction terms the linear and angular momentum are exactly conserved. In addition, because there is no explicit time dependence in L, the energy is conserved. The SPH system, therefore, mimics a system of molecules with forces between their line of centres, but with the difference that the strength of the interaction, through P and ρ, and its geometric dependence through the kernel, depends on the positions of other particles.

5.1.2. Circulation. Kelvin (see, e.g. Lamb (1932)) showed that for an inviscid fluid with $P = P(\rho)$, and conservative body forces, the integral of the velocity around any closed path

$$C_{\mathrm{K}} = \oint \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{r} \tag{5.15}$$

is constant. This conservation law is really infinitely many conservation laws since there are infinitely many closed curves. The constancy of circulation has been found useful in many hydrodynamic and atmospheric problems, and it is also applicable in astrophysical problems involving the dynamics of isothermal or adiabatic gas.

We can recover the circulation conservation directly from our SPH system. Consider a fluid where all the particles have the same mass and imagine a necklace of particles. If the particles have the same entropy (so that the necklace lies in a constant entropy surface) then nothing will change if each particle is shifted to its neighbour's positions always moving in the same sense around the necklace. With a proviso to be considered below, the dynamics should be unchanged. We can interpret this as requiring the change in the Lagrangian to be zero.

In this case, if a particle label on the necklace is ℓ, the change in position and velocity of the ℓth particle will be $\delta\boldsymbol{r}_\ell = (\boldsymbol{r}_{\ell+1} - \boldsymbol{r}_\ell)$ and $\delta\boldsymbol{v}_\ell = (\boldsymbol{v}_{\ell+1} - \boldsymbol{v}_\ell)$, respectively. The change in the Lagrangian to first order is then

$$\delta L = \sum_\ell \left(\frac{\partial L}{\partial \boldsymbol{r}_\ell} \cdot \delta\boldsymbol{r}_\ell + \frac{\partial L}{\partial \boldsymbol{v}_\ell} \cdot \delta\boldsymbol{v}_\ell \right), \tag{5.16}$$

where now the summation only applies to the particles around the necklace. Using the previous expressions for $\delta\boldsymbol{r}_\ell$ and $\delta\boldsymbol{v}_\ell$ together with Lagrange's equations to replace $\partial L/\partial \boldsymbol{r}_\ell$ by $\mathrm{d}(\partial L/\partial \boldsymbol{v}_\ell)/\mathrm{d}t$, and assuming the particle masses are equal, results in

$$\delta L = m \frac{\mathrm{d}}{\mathrm{d}t} \sum_\ell \boldsymbol{v}_\ell \cdot (\boldsymbol{r}_{\ell+1} - \boldsymbol{r}_\ell) = 0, \tag{5.17}$$

which must be zero if there is no change in the dynamics. We conclude that

$$C = \sum_\ell \boldsymbol{v}_\ell \cdot (\boldsymbol{r}_{\ell+1} - \boldsymbol{r}_\ell) \tag{5.18}$$

is constant and this is true regardless of the necklace in the constant entropy surface. This result is a discrete version of Kelvin's theorem. We can get the same result, but with opposite sign, by going around the necklace in the opposite sense. If we combine the two (changing the sign of the second because the integral is in the reverse sense) we get

$$C = \frac{1}{2} \sum_\ell \boldsymbol{v}_\ell \cdot (\boldsymbol{r}_{\ell+1} - \boldsymbol{r}_{\ell-1}), \tag{5.19}$$

which is a more accurate estimate of the circulation.

This result is, in general, only approximate because the changes in position and velocity to get from one place in the necklace to its neighbour are discrete, whereas exact conservation is only true when the transformations are infinitesimal. However, as Frank and Reich (2003) show for the case $P = K\rho^2$, where K is a constant, the SPH equations lead to very accurate

conservation of circulation, provided the necklace is defined by a large set of *tracer particles*. These tracer particles have negligible mass and interact only with the real SPH particles. The pressure force can, therefore, be written as the derivative of a potential and it follows (with ℓ denoting a tracer label and noting that the sum over $\boldsymbol{v}_\ell \cdot (\boldsymbol{v}_{\ell+1} - \boldsymbol{v}_{\ell-1})$ vanishes) that

$$\frac{\mathrm{d}C}{\mathrm{d}t} = -\sum_\ell (\boldsymbol{r}_{\ell+1} - \boldsymbol{r}_{\ell-1}) \cdot \nabla_\ell \Psi_\ell, \tag{5.20}$$

where

$$\Psi_\ell = K \sum_b m_b W(\boldsymbol{r}_\ell - \boldsymbol{r}_b, h). \tag{5.21}$$

If the number of tracer particles is made sufficiently large, the summation over ℓ becomes arbitrarily close to a line integral of a potential function around a closed loop and this vanishes. An interesting conclusion from this result is that the tracer particles have enough information from the real SPH particles to define their velocity and position so that the circulation is constant to high accuracy. The same argument can be extended to more complicated barotropic equations of state and applied to molecules or to clusters of stars.

The circulation of a fluid also appears in the work of Feynman on vortices in liquid helium and the necklace transformation was used by him to determine the quantization of circulation. For our present purposes we follow Feynman's review article (Feynman 1957). In that review he suggests a simple form of the wave function for a set of N identical helium atoms. If the entire system moves as a rigid body then the wave function Ψ is given by

$$\Psi = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\sum_j \boldsymbol{r}_j} \Phi, \tag{5.22}$$

where $\boldsymbol{r}_j$ is the position vector of particle j and $N\hbar\boldsymbol{k}$ is the momentum of the system. The function Φ is the ground state wave function. Feynman then argues that if the velocity is varying slowly then the wave function in a region must be close to the wave function of the atoms moving at a uniform velocity. As a result, the wave function for the entire fluid is expected to be similar to

$$\Psi = \mathrm{e}^{\mathrm{i}\sum_j m\boldsymbol{v}_j\cdot\boldsymbol{r}_j} \Phi, \tag{5.23}$$

where m is the mass of a helium atom. Feynman argues that the wave function must be invariant to the necklace transformation. When the particles are shifted around the necklace he finds that the change in the phase is given by

$$\frac{1}{\hbar}\sum_j m\boldsymbol{v}\cdot\Delta\boldsymbol{r}_j. \tag{5.24}$$

The wave function will be invariant if this phase is a multiple of 2π. Accordingly, we can write

$$\sum_j \boldsymbol{v}_j \cdot \Delta\boldsymbol{r}_j = \frac{2\pi\hbar n}{m}, \tag{5.25}$$

where n is an integer. Thus, circulation is quantized.

Finally we note that the circulation invariant contains a topological quantity, the loop around which the circulation is calculated and a dynamical quantity, the velocity. Many numerical codes in astrophysics can guarantee satisfactory accuracy for the velocity, but few can guarantee the same accuracy for the circulation because the numerical codes cannot follow the tangling of the loop.

5.2. The Lagrangian with constraints

In the simplest form of the SPH equations, ρ is defined by a summation over kernels. However, as suggested in section 2, there may be advantages in working with other forms of the density convergence equation; for example,

$$\frac{\mathrm{d}\rho_a}{\mathrm{d}t} = \rho_a \sum_b \frac{m_b}{\rho_b} \boldsymbol{v}_{ab} \cdot \nabla_a W_{ab}. \tag{5.26}$$

The action principle requires that

$$S = \int L \, \mathrm{d}t \tag{5.27}$$

is stationary for arbitrary and infinitesimal variations $\delta \boldsymbol{r}$ in the coordinates and corresponding variations $\delta \boldsymbol{v}$ in the velocities. These variations are related by

$$\frac{\mathrm{d}\delta \boldsymbol{r}}{\mathrm{d}t} = \delta \boldsymbol{v}. \tag{5.28}$$

Suppose that the only non-zero variation is $\delta \boldsymbol{r}_a$. The first order change in S is

$$\delta S = \int \left(m_a \boldsymbol{v}_a \cdot \delta \boldsymbol{v}_a - \sum_b m_b \frac{\partial u(\rho_b, s)}{\partial \rho_b} \frac{\delta \rho_b}{\delta \boldsymbol{r}_a} \cdot \delta \boldsymbol{r}_a \right) \mathrm{d}t, \tag{5.29}$$

where $\delta\rho_b/\delta \boldsymbol{r}_a$ denotes the Lagrangian change in ρ_b when the position of particle a changes by $\delta \boldsymbol{r}_a$ at time t. From (5.12) the change in ρ_b (assuming the variation in h can be neglected) is

$$\delta\rho_b = \rho_b \sum_c \frac{m_c}{\rho_c} (\delta \boldsymbol{r}_b - \delta \boldsymbol{r}_c) \cdot \nabla_b W_{bc}(h_b) \tag{5.30}$$

and, therefore,

$$\frac{\delta\rho_b}{\delta \boldsymbol{r}_a} = \rho_b \sum_c \frac{m_c}{\rho_c} (\delta_{ab} - \delta_{ac}) \nabla_b W_{bc}(h_b). \tag{5.31}$$

If this expression is substituted into the integral for δS, and the velocity term is integrated by parts (recalling that $\mathrm{d}\delta \boldsymbol{r}/\mathrm{d}t = \delta \boldsymbol{v}$), the variational principle gives

$$\frac{\mathrm{d}\boldsymbol{v}_a}{\mathrm{d}t} = -\sum_b \frac{m_b}{\rho_a \rho_b} \left(P_a \nabla_a W_{ab}(h_a) + P_b \nabla_a W_{ab}(h_b) \right). \tag{5.32}$$

This is the acceleration equation that is consistent with the convergence equation (5.21). This procedure can be generalized (for details see Price (2004a)) by writing the convergence equation as

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = \left(\frac{\rho}{\Phi} \right) \Phi \nabla \cdot \boldsymbol{v}, \tag{5.33}$$

where Φ is an arbitrary function. We can write (5.28) as

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = \frac{\rho}{\Phi} (\nabla \cdot (\Phi \boldsymbol{v}) - \boldsymbol{v} \cdot \nabla \Phi). \tag{5.34}$$

If the SPH form of (5.29) is used as a constraint, the action principle gives

$$\frac{\mathrm{d}\boldsymbol{v}_a}{\mathrm{d}t} = -\sum_b \frac{m_b}{\rho_a \rho_b} \left(\frac{P_a \Phi_b}{\Phi_a} \nabla_a W_{ab}(h_a) + \frac{P_b \Phi_a}{\Phi_b} \nabla_a W_{ab}(h_b) \right). \tag{5.35}$$

If $\Phi = \rho$, then the first form of the acceleration equation is recovered. If $\Phi = 1$, the second form is recovered. If we choose $\Phi = \sqrt{P}$, then the acceleration equation becomes

$$\frac{\mathrm{d}\boldsymbol{v}_a}{\mathrm{d}t} = -\sum_b m_b \frac{\sqrt{P_a P_b}}{\rho_a \rho_b} \left(\nabla_a W_{ab}(h_a) + \nabla_a W_{ab}(h_b) \right). \tag{5.36}$$

The advantages of choosing this last form consistently with $\Phi = \sqrt{P}$ in the convergence equation have not been analysed. These various forms of the acceleration equation have the same conservation properties.

5.3. Time integration in the absence of dissipation

Because the SPH algorithm reduces the original continuum partial differential equations to sets of ordinary differential equations, any stable time stepping algorithm for ordinary differential equations can be used. However, when there is no dissipation, the properties of the Lagrangian description can be preserved using a symplectic integrator (see, e.g. Leimkuhler *et al* (1997)). A simple example is the Verlet second order integrator which, for the one-dimensional system

$$\frac{\mathrm{d}q}{\mathrm{d}t} = v, \tag{5.37}$$

$$\frac{\mathrm{d}v}{\mathrm{d}t} = f(q), \tag{5.38}$$

takes the form (for constant time step δt)

$$q^{1/2} = q^0 + \tfrac{1}{2}\delta t v^0, \tag{5.39}$$

$$v^1 = v^0 + \delta t f(q^{1/2}), \tag{5.40}$$

$$q^1 = q^{1/2} + \tfrac{1}{2}\delta t v^1, \tag{5.41}$$

where a^0, $a^{1/2}$ and a^1 denote the values of a at the start of a step, halfway and at the end of the step, respectively.

In the case where there are n coordinates $q_1, q_2, q_3, \ldots, q_n$ with velocities $v_1, v_2, v_3, \ldots, v_n$ we get

$$q_i^{1/2} = q_i^0 + \tfrac{1}{2}\delta t v_i^0, \tag{5.42}$$

$$v_i^1 = v^0 + \delta t f(q_1^{1/2}, q_1^{1/2}, q_1^{1/2}, \ldots), \tag{5.43}$$

$$q_i^1 = q_i^{1/2} + \tfrac{1}{2}\delta t v_i^1. \tag{5.44}$$

In an SPH calculation, δt will depend on the speed of sound which, for a non-dissipative system depends on the density and, therefore, on the coordinates. In this case (5.39)–(5.41) are replaced with the following steps where the first half step has a different time step from the second half step.

$$q_i^{1/2} = q_i^0 + \tfrac{1}{2}\delta t^0 v_i^0, \tag{5.45}$$

$$v_i^{1/2} = v_i^0 + \tfrac{1}{2}\delta t^0 f(q_1^{1/2}, q_2^{1/2}, q_3^{1/2}, \ldots), \tag{5.46}$$

$$v_i^1 = v_i^{1/2} + \delta t^1 f(q^{1/2}), \tag{5.47}$$

$$q_i^1 = q_i^{1/2} + \tfrac{1}{2}\delta t^1 v_i^1, \tag{5.48}$$

where, for example,

$$\delta t^{1/2} = \tfrac{1}{2}(\delta t^0 + \delta t^1), \tag{5.49}$$

or, the frequently used

$$\frac{2}{\delta t^{1/2}} = \frac{1}{\delta t^0} + \frac{1}{\delta t^1}, \tag{5.50}$$

so that, with $\delta t^{1/2}$ calculated from the mid-point coordinate values, δt^1 can be calculated for the second half of the time stepping. This algorithm is reversible in time. The time stepping

for the n coordinate system (5.42)–(5.44) can be replaced in the same way. An alternative form with the same accuracy is

$$v^{1/2} = v^0 + \tfrac{1}{2}\delta t f^0, \tag{5.51}$$

$$q^1 = q^0 + \delta t v^{1/2}, \tag{5.52}$$

$$v^1 = v^{1/2} + \tfrac{1}{2}\delta t f^1, \tag{5.53}$$

which can be compared with (5.39)–(5.41). The latter is often referred to as the drift–kick–drift form, whereas the steps (5.51)–(5.53) are referred to as the kick–drift–kick form. The kick is the change in the velocity by the force. The drift is the change in the coordinate moving with the initial velocity. In some cases it may be useful to have the forces evaluated at the end of the step as in the kick–drift–kick form.

It is possible to show that the symplectic integrator equations (5.42)–(5.44) are equivalent to using the Lagrangian

$$L = \sum_i \frac{1}{2} m_i v_i^2 - \Phi - \frac{\delta t^2}{12}\left(\sum_j m_j f_j^2 + \frac{1}{2}\sum_j \sum_k m_j \dot{q}_j \dot{q}_k \frac{\partial f_j}{\partial q_k}\right) + O(\delta t^4), \tag{5.54}$$

or, equivalently, the Hamiltonian

$$H = \sum_i \frac{1}{2} m_i v_i^2 + \Phi + \frac{\delta t^2}{12}\left(\sum_j m_j f_j^2 + \frac{1}{2}\sum_j \sum_k m_j \dot{q}_j \dot{q}_k \frac{\partial f_j}{\partial q_k}\right) + O(\delta t^4), \tag{5.55}$$

where Φ is the potential energy such that $f_i = \partial\Phi/\partial q_i$. In an SPH calculation, Φ is given by

$$\Phi = \sum_j m_j u_j + \Psi, \tag{5.56}$$

where Ψ is the potential of any body force. As a consequence, the Hamiltonian, and therefore the energy, will not show a secular increase or decrease with time. Note that the double summation term can be written

$$\sum_j \sum_k \dot{q}_j \dot{q}_k \frac{\partial f_j}{\partial q_k} = \sum_j v_j \frac{\mathrm{d} f_j}{\mathrm{d} t}. \tag{5.57}$$

Since $\mathrm{d}f_j/\mathrm{d}t$ can be estimated from f_j at two time steps, the contribution of this double summation can be computed at little cost.

The advantages of using symplectic integrators for molecular dynamics has been discussed by many authors (see, e.g. Leimkuhler *et al* (1997)).

6. Applications of the Euler equations

The most common application of the SPH equations without dissipation is to small oscillations. The simplest of these is the oscillation of an infinite, one-dimensional gas with constant initial density. The analysis in the case of constant h has been given by Monaghan (1989) and Morris (1996). However, we will give the dispersion relation appropriate for h a function of ρ. A more complicated example is the oscillation of a Toy star in one dimension. This case is important because it mimics the oscillations of a star and is more difficult because the eigen functions vary sharply near the surface where the density goes to zero.

6.1. Dispersion relation for an infinite one-dimensional gas

Consider an SPH system that consists of an infinite set of particles in one dimension with initial spacing Δ. They are perturbed by a velocity much less than the speed of sound c_s. Let the unperturbed quantities (shown by an over bar) and the space and time variation of all perturbed quantities be proportional to

$$\exp \mathrm{i}(k\bar{x}_a - \omega t), \tag{6.1}$$

where $\bar{x}_a = a\Delta$ is the unperturbed position of particle a. In an unpublished work, I have shown that the linearized one-dimensional SPH equations of motion, with $h \propto 1/\rho$ and $P = K\rho^\gamma$, give the dispersion relation

$$\omega^2 = \frac{c_s^2}{\gamma\bar{\Omega}}\left[\frac{(\gamma-2)\Phi^2}{\omega} + 2\Psi - \frac{h}{\bar{\Omega}}\frac{\partial\Phi^2}{\partial h} - \frac{\ell\Phi^2}{\bar{\Omega}^2}\right]. \tag{6.2}$$

The functions Φ, Ψ and ℓ are defined by

$$\Phi = \Delta\sum_c \sin(k\bar{x}_c)\frac{\partial W(\bar{x}_c, h)}{\partial \bar{x}_c}, \tag{6.3}$$

$$\Psi = \Delta\sum_c [1-\cos(k\bar{x}_c)]\frac{\partial^2 W(\bar{x}_c, h)}{\partial h^2} \tag{6.4}$$

and

$$\ell = -2h\Delta\sum_c \frac{\partial W(\bar{x}_c, h)}{\partial h} - h^2\Delta\sum_c \frac{\partial^2 W(\bar{x}_c, h)}{\partial h^2} \tag{6.5}$$

and c_s is the adiabatic sound speed.

If the wavelength is much larger than Δ (as in many simulations, where the wave length is typically 100Δ), the summation can be replaced by an integration. We find

$$\Phi = \int_\infty^\infty \sin(kx)\frac{\mathrm{d}W}{\mathrm{d}x}\mathrm{d}x = -k\tilde{W}(k,h), \tag{6.6}$$

where $\tilde{W}(k,h)$ is the Fourier Transform of the kernel, and

$$\Psi = \int_\infty^\infty (1-\cos(kx))\frac{\mathrm{d}^2 W}{\mathrm{d}x^2}\mathrm{d}x = k^2\tilde{W}(k,h). \tag{6.7}$$

In this limit the dispersion relation becomes

$$\omega^2 = \frac{c_s^2k^2}{\gamma\Omega}\left[\frac{(\gamma-2)\tilde{W}^2(k,h)}{\Omega} + 2\tilde{W}(k,h) - \frac{h}{\Omega}\frac{\partial\tilde{W}^2}{\partial h} - \frac{\ell\tilde{W}^2}{\Omega^2}\right]. \tag{6.8}$$

The Fourier transform of the one-dimensional Gaussian kernel and the spline kernels are

$$\mathrm{e}^{-(hk/2)^2} \tag{6.9}$$

and

$$\left(\frac{\sin(hk/2)}{hk/2}\right)^n, \tag{6.10}$$

respectively, where the latter is obtained from (2.4) (note that the cubic spline has $n = 4$). If $kh < 1$, we can approximate each of these by

$$1 - \beta h^2k^2, \tag{6.11}$$

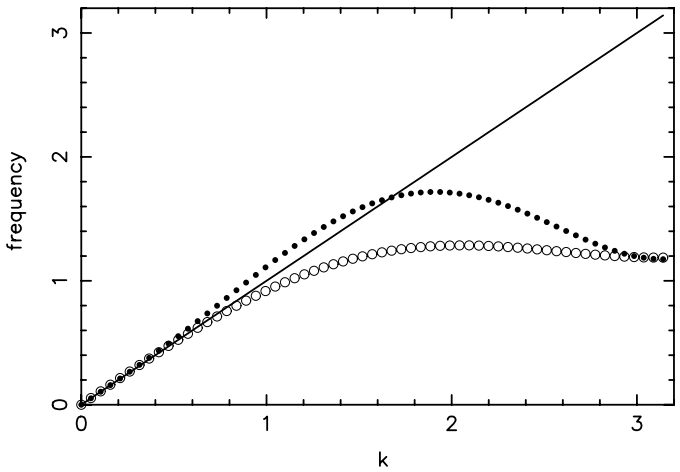

Figure 7. The SPH dispersion relation for sound waves in one dimension using the cubic spline and taking $\gamma = 5/3$ with full account of the variation of h with ρ. The black dots show the results when the variation of h with density is included. The open circles show the results when h is fixed.

with $\beta = 1/4$ for the Gaussian and $\beta = 1/6$ for the cubic spline. Using this approximation, and the further approximation $\Omega = 1$ and $\ell \sim 0$, the dispersion relation when h varies with ρ becomes

$$\omega^2 = k^2 c_s^2 \left[1 + \frac{2\beta h^2 k^2 (3-\gamma)}{\gamma} \right]. \tag{6.12}$$

If the contributions from the variation of h with ρ are neglected, the dispersion relation becomes

$$\omega^2 = k^2 c_s^2 \left[1 + \frac{2\beta h^2 k^2 (1-\gamma)}{\gamma} \right]. \tag{6.13}$$

This shows that, if h is constant, and γ lies in the normal range $1 \leqslant \gamma \leqslant 3$, the dispersion curve lies below the exact line $\omega = c_s k$, whereas if the variation of h is included the dispersion curve is above the exact line. The dispersion relation for the cubic spline in the case where, initially, $h = 1.2$ and $\gamma = 5/3$ is shown in figure 7. The variation of the dispersion relation is in agreement with the previous results when k is sufficiently small. In addition, we note that the dispersion relation for the case of varying h always lies above that for the case where h is constant. Both dispersion curves have the same limit when $k = \pi/\Delta$ because, for this k, $\Phi = 0$ and $\Psi = 8\Delta(\partial^2 W/\partial x^2)$ evaluated at $x = \Delta$. More accurate dispersion relations can be obtained if kernels which interpolate at a higher order are used. However, as mentioned earlier, these kernels are not satisfactory for shocks unless, in a dynamical calculation, there is a switch from lower order interpolation near shocks (e.g. where the cubic spline could be used) to a higher order interpolating kernel, elsewhere. An alternative is to use velocity smoothing (see later) with a suitable coefficient to cancel the error terms.

Figure 8 shows a velocity field after 4000 steps, for a one-dimensional gas with $\gamma = 1.4$, which was begun with density constant and velocity $0.05 c_s \sin(2\pi x)$. The velocity was reversed after 2000 steps. The total time of the simulation is equivalent to 13.7 periods. The SPH results were calculated using 100 particles with constant $h = 0.013$. The exact result (the reversed initial velocity) is shown by the continuous line, which is difficult to see because it passes through the points from the SPH simulation shown by filled symbols. Figure 9 shows the results for an initial velocity $0.05 c_s \sin(10\pi x)$. In this case the integration time is equivalent to 68.5 periods. The agreement between the SPH results and the exact values is excellent.

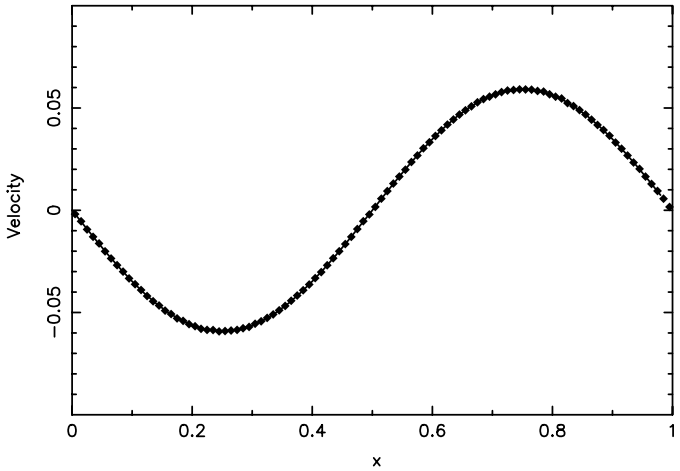

Figure 8. The velocity field for an oscillation in a one-dimensional gas at step 4000. The integration was performed using a Verlet symplectic integrator with 100 SPH points and the motion was reversed at step 2000. The SPH results are shown by filled symbols and the exact results by a continuous line (which is difficult to see because it passes through the SPH points).

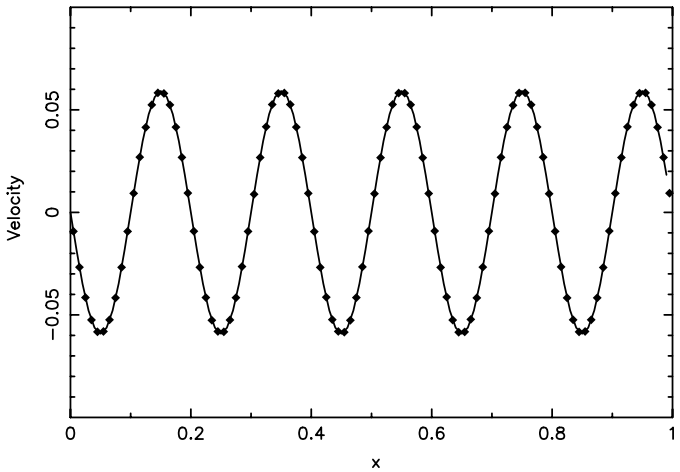

Figure 9. The velocity field for the conditions of the previous figure except that the initial velocity is $0.05c_s \sin(10\pi x)$.

6.2. Toy star oscillations

The usual tests in computational gas dynamics involve systems with rigid or periodic boundaries as in the previous test. These boundaries are quite useful for testing algorithms for industrial fluid dynamics. However, in astrophysics a more realistic test case is a finite mass of gas pulled together by gravity or a force which mimics gravity. The region outside the gas then has zero density. When finite difference methods are used for the dynamics of such a system they often give poor results because they do not handle the outer region of the gas moving into a vacuum. However, they do not present difficulties for particle methods such as SPH.

A useful class of such test problems are the Toy stars considered by Monaghan and Price (2004). The self-gravity is replaced by an attractive force proportional to the distance

and along the line of centres of any two particles. This force is the simplest many-body force. It was discovered by Newton, who pointed out that if two particles attract each other with a linear force then they move as if attracted to the centre of mass of the pair (see Chandrasekhar (1995) for a modern interpretation of Newton's Principia and, in particular, Newton's proposition LXIV, which discusses this force).

If there are N particles attracting each other with a force proportional to the separation, and directed along the line joining pairs of particles, then each particle moves as if it is independent of the others. The force appears as a linear force towards the centre of mass of the N particles (the particles, therefore, move in a common oscillator potential). In the case of two particles in three dimensions, the trajectories are closed Liassajous figures. A gaseous system with this force has a number of attractive features for testing algorithms for fluid dynamics. The linear modes of oscillation can be calculated easily and there is an exact non-linear solution where the velocity is a linear function of the coordinates but a non-linear function of time. This solution can be calculated very accurately by integrating a small number of ordinary differential equations and the results provide an excellent test of any computational fluid dynamics algorithm.

The simplest version of the Toy star assumes that pressure P is given in terms of the density ρ by $P = K\rho^2$, where K is a constant. This makes the problem analogous to the problem of shallow water motion in paraboloidal basins. There is extensive literature on this problem including the seminal papers of Ball (1963) and the general analysis by Holm (1991).

6.3. Toy stars in one dimension

Suppose that we have an isolated group of N particles in one dimension interacting with linear forces such that the force on particle j owing to particle k is $\nu m_j m_k (x_k - x_j)$. The potential energy is

$$\Phi = \frac{1}{4}\nu \sum_j \sum_k m_j m_k (x_j - x_k)^2, \tag{6.14}$$

The equation of motion of the jth particle is then

$$m_j \frac{\mathrm{d}^2 x_j}{\mathrm{d}t^2} = -\nu m_j \sum_k m_k (x_j - x_k). \tag{6.15}$$

However, the centre of mass can be chosen as the origin, so the equation of motion becomes

$$\frac{\mathrm{d}^2 x_j}{\mathrm{d}t^2} = -\nu M x_j, \tag{6.16}$$

where M is the total mass. The motion of the N-body system is therefore identical to the independent motion of each particle in a harmonic potential. In the following, we replace $M\nu$ by Ω^2. The acceleration equation for a one-dimensional gaseous toy star with velocity v, density ρ and pressure P is

$$\frac{\mathrm{d}v}{\mathrm{d}t} = -\frac{1}{\rho}\frac{\partial P}{\partial x} - \Omega^2 x. \tag{6.17}$$

Solutions can be found for $P = K\rho^\gamma$ with K constant and any $\gamma \geqslant 1$. If $\gamma = 2$, the equations are identical in form to those for the shallow water equations with density replacing the water depth. The equilibrium quantities can be easily calculated and, when the equilibrium is disturbed by velocities that are small compared with the speed of sound, the equations can be linearized. The errors in the linearization may, however, be large near the surface where the speed of sound and the pressure fall to zero. The velocity eigenfunctions are Gegenbauer

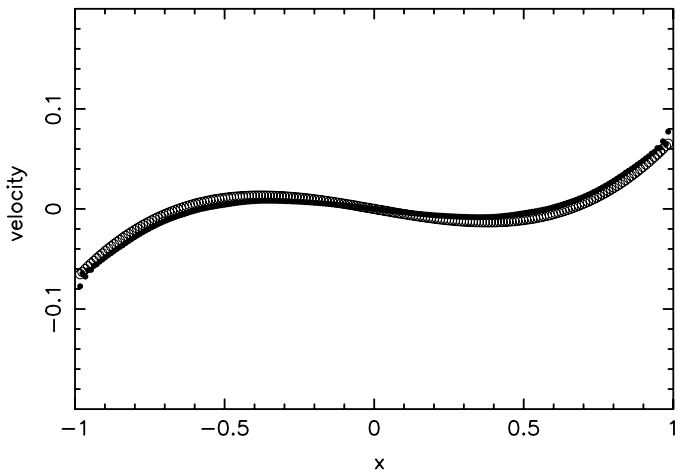

Figure 10. The velocity field for the Toy star oscillating with the velocity field in mode 3. The SPH results are shown by the filled symbols and the exact result by the circles.

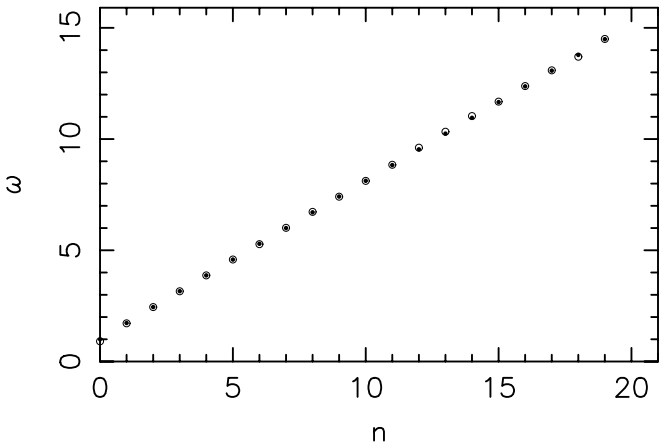

Figure 11. The SPH frequencies for Toy star eigenfunctions shown by filled circles compared with the exact results shown by open circles.

polynomials and the density eigenfunctions are Legendre polynomials. To simulate high order oscillations a large number of particles must be used to ensure that the resolution length is much smaller than the separation of the modes. If 400 particles are used then modes up to the twentieth can be simulated with high accuracy. The frequencies are always very accurate but the errors in the eigenfunctions are less accurate especially near the boundary. The velocity field for mode 3 is shown in figure 10 after 4 oscillation periods. The agreement with the perturbation solution is very good. A comparison between the SPH and the exact frequencies are shown in figure 11 for the first 20 modes. The agreement between theory and computation is excellent. These results show that the SPH method is able to accurately reproduce rather delicate and small oscillations in one dimension.

An attractive feature of the Toy stars is that exact non-linear solutions can be found. For the one-dimensional case with $P = K\rho^{\gamma}$ the solution has the form

$$v = A(t)x, \tag{6.18}$$

with

$$\rho^{(\gamma-1)} = H(t) - C(t)x^2, \tag{6.19}$$

so that the time-dependent radius of the toy star is $\sqrt{H/C}$. Substitution into the equations of motion and equating powers of x gives a set of ordinary differential equations for A, H and C. These can be integrated with high accuracy and compared with direct simulations by SPH. The agreement between the SPH results and the exact solution is again excellent.

The generalization of Toy stars to 2 and 3 dimensions is straightforward although the details become more complicated. Solutions can also be found with magnetic fields (see Monaghan and Price (2004)) who solve the one-dimensional MHD case).

7. Heat conduction and matter diffusion

The efficient solution of the heat conduction equation is fundamental for dissipative processes since similar techniques can be used for viscous dissipation or matter diffusion. An advantage of the SPH equations for these dissipative problems, as in the purely mechanical case, is that they can be written in such a way that they mimic fundamental properties of the system and allow complicated physics to be handled in a straightforward way. Appropriate forms of these equations have been derived (Brookshaw 1985, Cleary and Monaghan 1999, Monaghan *et al* 2005) and applied to a wide variety of heat conduction problems, including the Stefan problem and the freezing of alloy solutions (Monaghan *et al* 2005) and problems involving radiative transfer in the diffusion approximation (Whitehouse and Bate 2004).

7.1. The SPH heat conduction equation

A convenient form of the heat conduction equation without heat sources or sinks is

$$c_p \frac{\mathrm{d}T}{\mathrm{d}t} = \frac{1}{\rho}\nabla(\kappa\nabla T), \tag{7.1}$$

where T is the absolute temperature, c_p the heat capacity per unit mass at constant pressure, ρ the density, κ the coefficient of thermal conductivity and $\mathrm{d}/\mathrm{d}t$ the derivative following the motion. The spatial derivatives can be determined using the results of section 2.3, and the SPH form of (7.1) is

$$c_{p,a}\frac{\mathrm{d}T_a}{\mathrm{d}t} = \sum_b \frac{m_b}{\rho_a\rho_b}(\kappa_a + \kappa_b)(T_a - T_b)F_{ab}. \tag{7.2}$$

This equation shows that the contribution of particle b to the rate of change of T_a is positive if $T_b > T_a$ because $F_{ab} \leqslant 0$, i.e. the heat flows from the hotter element of the fluid to the cooler element as expected. As mentioned in section 2.3, this fundamental requirement could not be guaranteed if the second derivatives of the interpolation formula for T were calculated directly.

Equation (7.2) does not guarantee that the heat flux will be continuous when κ is discontinuous. Cleary and Monaghan (1999) show from an analysis of the finite difference case that this problem can be solved by replacing $(\kappa_a + \kappa_b)$ in (7.2) with

$$\frac{4\kappa_a\kappa_b}{(\kappa_a + \kappa_b)}. \tag{7.3}$$

The heat flux is then continuous even with jumps by a factor 10^3 in κ across 3 particle spacings. A slightly different κ term, based on similar ideas, gives satisfactory results for jumps in κ by a factor 10^9 (Parshikov and Medin 2002). However, because the very simple form (7.3) gives

excellent results for the normal range of material properties, the final SPH heat conduction equation is, therefore,

$$c_{p,a}\frac{\mathrm{d}T_a}{\mathrm{d}t} = \sum_b \frac{m_b}{\rho_a\rho_b}\frac{4\kappa_a\kappa_b}{(\kappa_a+\kappa_b)}(T_a - T_b)F_{ab}. \tag{7.4}$$

Cleary and Monaghan (1999) showed that this SPH form of the heat conduction equation had similar accuracy to finite difference methods and was not sensitive to the particle disorder that occurs in some SPH calculations. In addition, heat conduction problems with discontinuous κ, and with κ varying with T, were accurately integrated. Whitehouse and Bate (2004) studied heat conduction by radiation in the diffusion approximation obtaining accurate results for test problems.

If the particles are thermally isolated (so they can only exchange heat amongst themselves) then (7.2) shows (noting $F_{ab} = F_{ba}$), that the total heat content

$$\sum_a m_a c_{p,a} T_a \tag{7.5}$$

is constant.

7.2. Heat conduction with sources or sinks

When the system contains point sources or sinks, (7.1) becomes

$$\rho c_p \frac{\mathrm{d}T}{\mathrm{d}t} = \nabla(\kappa\nabla T) + \sum_k Q_k \delta(\boldsymbol{r} - \boldsymbol{R}_k), \tag{7.6}$$

where Q_k denotes the strength of the source or sink and is negative for a sink. $\boldsymbol{R}_k$ denotes the position of source/sink k and δ denotes a Dirac delta function. The SPH equation corresponding to (7.6) becomes

$$c_{p,a}\frac{\mathrm{d}T_a}{\mathrm{d}t} = \sum_b \frac{m_b}{\rho_a\rho_b}\frac{4\kappa_a\kappa_b}{(\kappa_a+\kappa_b)}(T_a - T_b)F_{ab} + \frac{1}{\rho_a}\sum_k Q_k\zeta_k W(\boldsymbol{r}_a - \boldsymbol{R}_k), \tag{7.7}$$

where the delta function has been replaced by a smoothing kernel, which is consistent with the smoothing of the original continuum equation and, to ensure that the rate of change of thermal energy owing to the source is correct, a normalizing factor ζ_k for source k defined by

$$\frac{1}{\zeta_k} = \sum_b \frac{m_b}{\rho_b} W(\boldsymbol{r}_b - \boldsymbol{R}_k, h), \tag{7.8}$$

has been introduced. The right-hand side is an SPH estimate of the constant 1 at the position of the source. From (7.7) the rate of change of thermal energy is

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_a m_a c_{p,a} T_a\right) = \sum_k Q_k, \tag{7.9}$$

as expected.

7.3. Salt diffusion

Denoting the mass fraction of salt by C so that the mass of salt in a mass M of liquid is CM, the diffusion of the salt is given by an equation similar in form to the heat conduction equation, namely,

$$\frac{\mathrm{d}C}{\mathrm{d}t} = \frac{1}{\rho}\nabla(D\nabla C), \tag{7.10}$$

where D is the coefficient of diffusion with dimensions of $\mathrm{ML}^{-1}\mathrm{T}^{-1}$. The SPH form of this equation follows in the same way as for the heat conduction equation. The SPH equation for the rate of change of the concentration C_a of particle a is given by

$$\frac{\mathrm{d}C_a}{\mathrm{d}t} = \sum_b \frac{m_b}{\rho_a \rho_b} \frac{4D_a D_b}{(D_a + D_b)} (C_a - C_b) F_{ab}. \tag{7.11}$$

The combination of D in the SPH equation ensures that the flux of material across an interface between two materials with different diffusion coefficients is constant. The total mass of salt is conserved by the SPH equation.

7.4. The increase of entropy

The SPH conduction equation results in entropy increasing in the absence of heat sinks. If S is the total entropy of the system then

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \sum_a m_a \frac{\mathrm{d}s_a}{\mathrm{d}t} = \sum_a \frac{m_a}{T_a} \frac{\mathrm{d}q_a}{\mathrm{d}t}, \tag{7.12}$$

where s_a is the entropy/mass of particle a, q_a is the heat content/mass of particle a and T is the absolute temperature. From equations (7.4) and (7.12), with an interchange of labels the change of entropy with time can be written as

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \frac{1}{2} \sum_a \sum_b \frac{m_a m_b}{\rho_a \rho_b} \frac{4\kappa_a \kappa_b}{(\kappa_a + \kappa_b)} \left(\frac{1}{T_a} - \frac{1}{T_b} \right) (T_a - T_b) F_{ab}. \tag{7.13}$$

Since $F_{ab} \leqslant 0$ we deduce that $\mathrm{d}S/\mathrm{d}t \geqslant 0$.

When the composition changes there is a further contribution to the entropy. To deduce this we first divide (7.11) by C_a. If the resulting equation is summed over a, and added to the same expression with the labels interchanged, the following positive definite quantity is obtained.

$$\frac{\mathrm{d}}{\mathrm{d}t} \sum_a m_a \ln C_a = \sum_a \sum_b m_a m_b \frac{4D_a D_b}{(D_a + D_b)} \left(\frac{1}{C_a} - \frac{1}{C_b} \right) \frac{(C_a - C_b)}{\rho_a \rho_b} F_{ab} \geqslant 0. \tag{7.14}$$

This quantity is the increase in entropy resulting from composition changes.

7.5. Boundary and interface conditions

There is no need to place a special condition on the gradient of the temperature at the boundary to satisfy these conditions if SPH is used. If all the boundaries are adiabatic, then the particles interact amongst themselves and the symmetry of the SPH conduction equation ensures that the system conserves its thermal energy as shown earlier. If one or more boundary curves have fixed temperatures, the SPH particles on the boundaries are included in the heat conduction equation so that the heat transferred to the boundary during a time step can be calculated. After this is done the temperatures of the boundary particles are set back to the specified boundary temperatures for the next time step. The heat transferred to the boundary particle can be calculated from the temperature change. Cleary and Monaghan (1999) noted that near the boundaries, the SPH interpolation can give errors of a few per cent, and they made corrections to the density near the boundary to compensate for this. As noted earlier, SPH calculations do not need special interface conditions. The SPH particles exchange heat and material with neighbouring particles whether they are of the same or different phases.

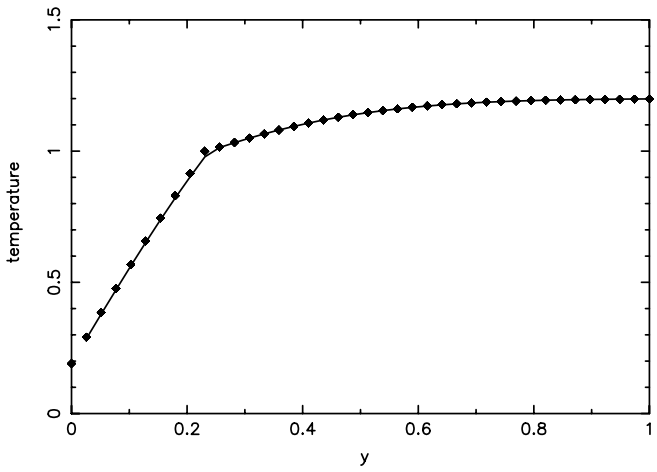

Figure 12. The temperature against the distance from the cooling boundary for a two-dimensional Stefan problem. The system is periodic in the x direction. The exact results are shown by the solid line and the SPH results by the solid diamonds. The change of slope shows the interface between solid and liquid.

7.6. The Stefan problem

An interesting application of SPH is to the Stefan problem where a pure substance is cooled sufficiently for it to freeze. In the standard treatment of this problem the following condition is required at the interface:

$$\kappa_1 \left(\frac{\mathrm{d}T}{\mathrm{d}y}\right)_1 - \kappa_2 \left(\frac{\mathrm{d}T}{\mathrm{d}y}\right)_2 = \rho L \frac{\mathrm{d}Y}{\mathrm{d}t}, \tag{7.15}$$

where L is the latent heat/mass and $\mathrm{d}Y/\mathrm{d}t$ is the rate of change of the position of the interface (Carslaw and Jaeger 1990). The condition expresses the fact that the difference in the heat flux on each side of the interface supplies the heat to change the phase. In this formulation the position of the interface is one of the unknowns.

The SPH treatment of the freezing is very simple. Initially, the SPH particles are assumed to be liquid and tagged with an integer to denote liquid particles and given the material properties of the liquid. As heat is conducted from the liquid some liquid SPH particles reach the solidification temperature T_m. The heat per unit mass q, lost by these particles after this time, is then stored and their temperatures are kept at T_m. If particle a is in this condition then, when q_a reaches L, the integer tag is changed to that for the solid phase, and the properties of this phase (thermal conductivity and heat capacity) are assigned to this particle. Between the solid particles and the liquid particles there is a region where the particles have reached the solidification temperature but have, not yet had their latent heat fully extracted. An example of the SPH solution of a one-dimensional Stefan problem is shown in figure 12, and for an axisymmetric Stefan problem with a heat sink in figure 13 (both taken from Monaghan *et al* (2005)). The agreement between the SPH results and the exact result (Carslaw and Jaeger 1990) is excellent.

8. Viscosity

The first use of viscosity in SPH equations was by Lucy (1977) who introduced an artificial bulk viscosity to prevent a slow build-up of acoustic energy from integration errors in an SPH simulation. A different, and more effective viscosity, which conserves linear and angular

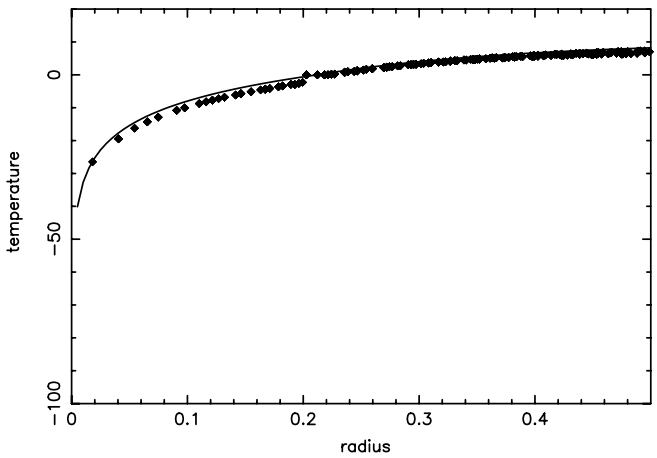

Figure 13. The temperature against the radius for the case of freezing is induced by a line sink in an axisymmetric system. The exact results are shown by the continuous line and the SPH results by symbols. Despite the particles being on a rectangular grid, the variation of the temperature is close to radial. Note that the small group of particles which have reached the freezing temperature but have not yet become ice particles, form a small horizontal line at a radius of approximately 0.22.

momentum was suggested and tested by Monaghan and Gingold (1983). The results obtained using this viscosity in a wide variety of shock problems involving gases, liquids and solids (Libersky and Petschek 1991) and, with a different version for a relativistic gas (Chow and Monaghan 1997) in one dimension, is in good agreement with the theory. With reference to shock problems, SPH does not give the widths of shock fronts as accurately as the methods based on Riemann solvers with similar resolution; however, no current method gives the width of a shock front accurately, since the width of real shock fronts is only a few molecular mean-free paths. Typical resolutions in numerical simulations are a factor 10^4 greater. The key is to get the pre- and post-shock values correct and SPH is capable of producing these to any degree of desired accuracy.

In two and more dimensions it is more difficult for SPH to match the accuracy of modern finite difference codes, but its advantage is that it is independant of the special properties of the ideal gas equation, which are built into the finite difference codes. Consequently, SPH can be used when the equation of state is complicated and Riemann solutions are unavailable (approximate linear solutions could be used but they are unreliable (Quirk (1994))).

The viscosity of real fluids can be implemented using ideas similar to those used for the artificial viscosity and for heat conduction (Cleary 1998, Cleary and Ha 2002). Applications have been made to a low Reynolds number flow (Morris *et al* 1997) and to systems involving more than one fluid in contact. An alternative approach is to calculate the velocity derivatives in the viscous term using SPH methods (Takeda *et al* 1994, Watkins *et al* 1996, Chaniotis *et al* 2002). These forms of the viscous stress tensor conserve linear momentum but not angular momentum. In many industrial fluid dynamics problems, the exact conservation of angular momentum is not an issue and the work of Chaniotis *et al* (2002) shows that SPH (together with a re-meshing strategy) gives excellent results.

8.1. Artificial viscosity

As its name suggests, artificial viscosity bears no relation to real viscosities, but is designed to allow shock phenomena to be simulated, or simply to stabilize a numerical algorithm. Artificial

viscosities are often constructed analogously to real gas viscosities, replacing the mean free path with the resolution length. The Navier–Stokes acceleration equation for viscous flow has the form

$$\frac{\mathrm{d}v_i}{\mathrm{d}t} = -\frac{1}{\rho}\frac{\partial P}{\partial x_i} + \frac{1}{\rho}\left[\frac{\partial}{\partial x_k}\left(\eta\left(\frac{\partial v_i}{\partial x_k} + \frac{\partial v_k}{\partial x_i} - \frac{2}{3}\delta_{ik}\nabla\cdot\boldsymbol{v}\right)\right) + \frac{\partial}{\partial x_i}(\zeta\nabla\cdot\boldsymbol{v})\right], \tag{8.1}$$

where η is the shear viscosity coefficient and ζ is the bulk viscosity, which is required when the internal degrees of freedom of the molecules of the fluid are activated during the flow. These viscosity coefficients are, in general, functions of temperature and density. For a monatomic gas $\eta \sim \frac{1}{3}\rho\lambda c_\mathrm{s}$, where λ is the mean free path and c_s is the speed of sound.

The viscous terms could be estimated directly using the SPH interpolation formula but, as in the case of heat conduction, this leads to equations, which do not conserve linear and angular momentum, and do not guarantee that the viscous dissipation will increase the entropy. Monaghan and Gingold (1983) devised a viscosity by simple arguments about its form and its relation to gas viscosity. The viscous term, denoted by Π_{ab} is added to the pressure terms in SPH equations to give

$$\frac{\mathrm{d}v_a}{\mathrm{d}t} = -\sum_b m_b\left(\frac{P_a}{\rho_a^2} + \frac{P_b}{\rho_b^2} + \Pi_{ab}\right)\nabla_a W_{ab}, \tag{8.2}$$

where

$$\Pi_{ab} = -\nu\left(\frac{\boldsymbol{v}_{ab}\cdot\boldsymbol{r}_{ab}}{r_{ab}^2 + \epsilon\bar{h}_{ab}^2}\right) \tag{8.3}$$

and $\epsilon \sim 0.01$ is introduced to prevent a singularity when $r_{ab} = 0$ and ν is defined by

$$\nu = \frac{\alpha\bar{h}_{ab}\bar{c}_{ab}}{\bar{\rho}_{ab}}, \tag{8.4}$$

where, for example, $\bar{h}_{ab} = (h_a + h_b)/2$. A further generalization, which has not been explored, but which gives a higher order viscosity, is to multiply the previous viscosity by any power of

$$\left|\frac{\boldsymbol{v}_{ab}\cdot\boldsymbol{r}_{ab}}{\bar{c}_{ab}}\right|. \tag{8.5}$$

The artificial viscosity term Π_{ab} is a Galilean invariant and vanishes for rigid rotation. When two particles approach each other, the artificial viscosity produces a repulsive force between the particles. When they recede from each other the force is attractive.

The SPH viscosity can be related to a continuum viscosity by converting the summation to integrals. The x component of the acceleration equation has the viscous contribution

$$f_x = \sum_b m_b\frac{\alpha\bar{c}_{ab}\bar{h}_{ab}}{\bar{\rho}_{ab}}\frac{\boldsymbol{v}_{ab}\cdot\boldsymbol{r}_{ab}}{r_{ab}^2 + \epsilon\bar{h}_{ab}^2}(x_a - x_b)F_{ab}. \tag{8.6}$$

If the $\epsilon\bar{h}_{ab}^2$ in the denominator is dropped this integral can be written as a sum of terms similar in form to those considered in section 2.3. If α, c, h and ρ are constant the continuum equivalent of f_x in two dimensions is

$$f_x = \alpha hc\left(\tfrac{3}{8}v^x_{xx} + \tfrac{1}{8}v^x_{yy} + \tfrac{1}{4}v^y_{xy}\right), \tag{8.7}$$

where v^y_{xy} denotes $\partial^2 v^x/\partial x\partial y$ and v^x denotes the x component of the velocity with a similar notation for the other terms. This shows that the shear viscosity coefficient $\eta = \rho\alpha hc/8$ and the bulk viscosity coefficient $\zeta = 5\eta/3$. Similar analysis in three dimensions shows that $\eta = \rho\alpha hc/10$ and $\zeta = 5\eta/3$.

If there are rapid changes in the parameters then the same argument used for the case of heat conduction with discontinuous thermal conductivity can be used. If we define (for two dimensions)

$$\mu_a = \tfrac{1}{8}\alpha_a h_a c_a \rho_a \tag{8.8}$$

and define a new Π_{ab} according to

$$\Pi_{ab} = -\frac{16\mu_a\mu_b}{\rho_a\rho_b(\mu_a+\mu_b)}\left(\frac{\boldsymbol{v}_{ab}\cdot\boldsymbol{r}_{ab}}{r_{ab}^2+\epsilon\bar{h}_{ab}^2}\right), \tag{8.9}$$

we produce a viscosity term which can be used for real viscosities and maintains the continuity of viscous stress accurately. This SPH viscosity was proposed by Cleary (1998) and applied to the simulation of viscous liquids in flow modelling for casting processes (see, e.g. Cleary and Ha (2002) and the references therein). Cleary determined the coefficient by numerical experiment and found that coefficient 16 should be replaced by 19.8. In three dimensions (where the factor 1/8 in (8.8) is replaced by 1/10) the analysis suggests a coefficient of 20.

In the case of shock tube problems, it is usual to turn the viscosity on for approaching particles and turn it off for receding particles. In this way, the viscosity is used for shocks and not rarefactions. Unfortunately, in astrophysical calculations, this rule means that the viscosity is turned on when the density increases in the shock-free regions, for example, when gravity pulls gas together.

When the viscosity term Π_{ab} was first used (Monaghan and Gingold 1983) it was found to work well for shocks of moderate strength. However, in astrophysical calculations involving colliding gas clouds, where the Mach number can be very high, it was found that particles from one cloud could stream between the particles of the other cloud. Generally, this streaming is limited to a few particle spacings, and is, therefore, not a severe problem; however, it should not occur at all. To prevent it, an extra term was added to ν which then took the form (Monaghan 1992)

$$\nu = \frac{\bar{h}_{ab}}{\bar{\rho}_{ab}}\left(\alpha\bar{c}_{ab} - \beta\frac{\bar{h}_{ab}\boldsymbol{v}_{ab}\cdot\boldsymbol{r}_{ab}}{r_{ab}^2+\epsilon\bar{h}_{ab}^2}\right). \tag{8.10}$$

This form of ν, and hence Π_{ab}, evolved through various forms starting with the work of Lattanzio *et al* (1985) on interstellar cloud collisions. Good results have been obtained with the choice $\alpha = 1$ and $\beta = 2$. This form of the viscosity, though changed in details, is found naturally by considering aspects of the dissipative term in shock solutions based on Riemann solvers (Monaghan 1997). In this case

$$\Pi_{ab} = -\frac{K v_{\mathrm{sig}}(\boldsymbol{v}_{ab}\cdot\boldsymbol{r}_{ab})}{\bar{\rho}_{ab}|r_{ab}|}, \tag{8.11}$$

where $K \sim 0.5$. The signal velocity v_{sig} is defined by

$$v_{\mathrm{sig}} = c_a + c_b - \beta\boldsymbol{v}_{ab}\cdot\hat{\boldsymbol{r}} \tag{8.12}$$

where $\hat{\boldsymbol{r}} = \boldsymbol{r}_{ab}/|\boldsymbol{r}_{ab}|$ and $\beta \sim 4$. The signal velocity can be interpreted as follows. If the fluid is at rest we estimate the speed at which a sound wave from a approaches a sound wave from b as $(c_a + c_b)$. The extra term represents the change in speed if the fluids at a and b are moving relative to each other. The fact that this must be a Galilean invariant, and would vanish if they have the same velocity or rotate rigidly, leads directly to the form shown. This is discussed further by Monaghan (1997).

8.2. Viscous heating and the energy equations

Viscosity dissipates the flow and transfers energy from kinetic to thermal. The contribution to the thermal energy is always positive. Owing to the way in which the SPH viscosity was derived, viscous dissipation is best obtained directly from the SPH equations. By taking the scalar product of $\boldsymbol{v}_a$ and the acceleration equation, multiplying by m_a and summing over a, the viscous contribution to the rate of change of thermal energy can be identified (Monaghan and Gingold 1983, Monaghan 1997). The final result is the thermal energy equation

$$\frac{\mathrm{d}u_a}{\mathrm{d}t} = \frac{P_a}{\Omega_a \rho_a^2} \sum_b m_b \boldsymbol{v}_{ab} \cdot \nabla_a W_{ab} + \frac{1}{2} \sum_a m_a \sum_b m_b \Pi_{ab} \boldsymbol{v}_{ab} \cdot \nabla_a W_{ab}. \tag{8.13}$$

Referring now to the definition of Π_{ab}, for example to (8.3), the contribution to the viscous dissipation of particle a from b can be written as (using the same definition of $F_{ab} \leqslant 0$ as before)

$$-\left(\frac{\alpha \bar{h}_{ab} \bar{c}_{ab}}{\bar{\rho}_{ab}}\right) \frac{F_{ab} (\boldsymbol{v}_{ab} \cdot \boldsymbol{r}_{ab})^2}{r_{ab}^2 + \eta^2}, \tag{8.14}$$

which is $\geqslant 0$. This confirms that the SPH dissipation increases the thermal energy as it should. In addition to increasing the thermal energy, the viscous dissipation should increase the total entropy of the system. From the first law of thermodynamics

$$T \frac{\mathrm{d}s}{\mathrm{d}t} = \mathrm{d}u - \frac{P}{\rho^2} \mathrm{d}\rho. \tag{8.15}$$

In SPH form this becomes

$$T_a \frac{\mathrm{d}s_a}{\mathrm{d}t} = \frac{1}{2} \sum_b m_b \Pi_{ab} \boldsymbol{v}_{ab} \cdot \nabla_a W_{ab} \tag{8.16}$$

and from the previous results the change in the entropy of any particle owing to viscous dissipation is positive.

8.3. Dissipation and the thermokinetic energy equation

It was shown earlier (section 3) that the thermokinetic equation takes the form

$$\frac{\mathrm{d}\hat{e}_a}{\mathrm{d}t} = -\sum_b m_b \left(\frac{P_a \boldsymbol{v}_b}{\rho_a^2} + \frac{P_b \boldsymbol{v}_a}{\rho_b^2}\right) \cdot \nabla_a W_{ab}, \tag{8.17}$$

where $\hat{e}_a = \frac{1}{2} v_a^2 + u_a$. In order to use this equation for shock phenomena it is necessary to add dissipative terms. Although these could be deduced by beginning with the definition of $\hat{e}$ and SPH equations for the derivatives, it is more convenient to be guided by ideas from Riemann solvers (Monaghan 1997). Hence, we need to add a dissipative term Υ_{ab} to the pressure–velocity terms in (8.17) where

$$\Upsilon_{ab} = -\frac{K v_{\mathrm{sig}}(a,b)(e_a^* - e_b^*)\hat{\boldsymbol{r}}}{\bar{\rho}_{ab}} \tag{8.18}$$

and

$$e_a^* = \tfrac{1}{2}(\boldsymbol{v}_a \cdot \hat{\boldsymbol{r}})^2 + u_a, \tag{8.19}$$

where $\hat{\boldsymbol{r}} = \boldsymbol{r}_{ab}/|\boldsymbol{r}_{ab}|$. Replacing the actual kinetic energy with the kinetic term using the velocity along the line joining the particles a and b guarantees that the contribution to the thermal energy from viscous dissipation will be positive, and that the entropy will increase with time (Monaghan 1997). It is often assumed that the constant K and the signal velocity

$v_{\rm sig}$ are the same as in the dissipative term (8.11) in the acceleration equation but that is not necessary. Starting with the equation for the rate of change of $\hat{e}$

$$\frac{\mathrm{d}\hat{e}_a}{\mathrm{d}t} = -\sum_b m_b \left(\frac{P_a \boldsymbol{v}_b}{\rho_a^2} + \frac{P_b \boldsymbol{v}_a}{\rho_b^2} + \Upsilon_{ab} \right) \cdot \nabla_a W_{ab}, \tag{8.20}$$

it is possible to deduce the rate of change of thermal energy (Monaghan 1997). This takes the following form

$$\frac{\mathrm{d}u_a}{\mathrm{d}t} = \frac{P_a}{\rho_a^2} \sum_b m_b \boldsymbol{v}_{ab} \cdot \nabla_a W_{ab} + \text{dissipative term}, \tag{8.21}$$

where the dissipative term is

$$\sum_b m_b \frac{K v_{\rm sig}(a,b)}{\bar{\rho}_{ab}} \left(u_a - u_b - \frac{1}{2}(\boldsymbol{v} \cdot \hat{\boldsymbol{r}})^2 \right) |\boldsymbol{r}_{ab}| F_{ab}. \tag{8.22}$$

The terms involving u, namely,

$$\sum_b m_b \frac{K v_{\rm sig}(a,b)}{\bar{\rho}_{ab}} (u_a - u_b) |\boldsymbol{r}_{ab}| F_{ab} \tag{8.23}$$

give heat diffusion. This expression, is a variant on the heat diffusion term described in section 7.1 and has similar properties. The diffusion coefficient is proportional to $v_{\rm sig}(a,b)|\boldsymbol{r}_{ab}|$. A heat diffusion conduction term was used in the thermal energy by Lattanzio and Monaghan (1991) in their discussion of fragmenting molecular gas clouds.

8.4. Reducing artificial dissipation

Artificial dissipation is very successful for handling shocks but it can be too large in other parts of the flow. For example, artificial viscous dissipation increases the Reynolds number of a flow, artificially, with the result that, for example, the Kelvin–Helmoltz shear instabilities are heavily diffused. Balsara (1995) suggested reducing viscous dissipation by multiplying Π_{ab} by the factor

$$\frac{|\nabla \cdot \boldsymbol{v}|}{|\nabla \cdot \boldsymbol{v}| + |\nabla \times \boldsymbol{v}|}, \tag{8.24}$$

made symmetric, for example, by replacing $\nabla \cdot \boldsymbol{v}$ by the average for the interacting pair of particles. Colagrossi (2004) found that it is preferable to replace the previous factor by

$$\frac{|\nabla \cdot \boldsymbol{v}|_{ab}}{|\nabla \cdot \boldsymbol{v}|_{ab} + \sqrt{E^{ij}E^{ij}} + 10^{-4}\bar{c}_{ab}/h}, \tag{8.25}$$

where c is the speed of sound and the rate of strain tensor E^{ij} is defined by

$$E^{ij} = \frac{1}{2}\left(\frac{\partial v^i}{\partial x^j} + \frac{\partial v^j}{\partial x^i} \right). \tag{8.26}$$

The indices denote cartesian tensors and the summation convention is used in evaluating $E^{ij}E^{ij}$. Colagrossi (2004) found that the replacement of $\nabla \times \boldsymbol{v}$ with the term involving E^{ij} gave improved results for problems involving slightly compressible fluids. A particularly impressive example being the rotation in two dimensions of a rotating square of water in an otherwise empty space.

Another very useful approach is to note that the dissipation terms have the same coefficients K and $v_{\rm sig}$ in both Π_{ab} and Υ_{ab}. In general, different coefficients, or signal speeds could be used for the viscous and the thermal energy terms. Furthermore, each particle can have its own

coefficient determined by the conditions it encounters. Morris and Monaghan (1997) explored this idea for the artificial viscous terms in gas dynamics where it is desirable to reduce the viscosity away from shocks. In finite difference calculations this is achieved by switches based on the first and second spatial derivatives of physical variables such as the momentum flux. However, spatial derivatives sit uncomfortably with particle methods for which time derivatives are more natural. The basic problem is like trying to predict the onset of a stock market crash from the time variation of the market. For shock simulation the coefficient α in (8.4) or equivalently K in (8.11) should be different for each particle and should change with time according to the conditions the particle is in, becoming large at shocks, but relaxing back to a small value when the flow is calmer. A simple way to do this for a typical particle a is to determine its α_a from the equation

$$\frac{\mathrm{d}\alpha_a}{\mathrm{d}t} = -\frac{(\alpha_a - \alpha_0)}{\tau} + S_a, \tag{8.27}$$

where τ is a suitable time scale $\propto h/c_s$, $\alpha_0 \sim 0.1$ is the ambient value of α, and S_a is a source term. S should increase as the particle approaches a shock in such a way that α increases to approximately 1. Morris and Monaghan (1997) discuss a choice of $S \propto \nabla \cdot \boldsymbol{v}$ and show that it gives good results. Rosswog *et al* (2000) take

$$S = \mathrm{Max}(-\nabla \cdot \boldsymbol{v}, 0)(2 - \alpha). \tag{8.28}$$

Not only are the good results for the shocks retained, but elsewhere in the flow the viscosity is also reduced by approximately a factor of 10. However, in many astrophysical problems, where a collapse of gas clouds occurs, $-\nabla \cdot \boldsymbol{v}$ can increase without the occurrence of shocks. It would, therefore, be desirable to relate S to some other quantity related to the change of entropy.

Price (2004a) suggested that the artificial thermal conductivity should also vary with each particle and proposed a similar equation to that for α but using a source term proportional to $|\nabla\sqrt{u}|$. In terms of (8.18) it means splitting the dissipative term into a viscous part and a heat conduction part and using a different K for each. Price obtains improved results for shock tube phenomena especially near contact discontinuities. In the same way he tested a dissipation term for the magnetic fields in MHD simulations and found it improved his SPH simulations. Whether or not the onset of a shock could be predicted more satisfactorily with higher derivatives is an open question.

9. Applications to shock and rarefaction problems

There have been widespread applications of SPH to shocks in gases, liquids and metals (see, e.g. Libersky and Petschek (1991), Johnson *et al* (1996), Monaghan (1997)). There is only space here to describe some elementary examples from gas dynamics.

The first case we consider is the rarefaction wave. This can be set up by placing SPH particles in the region $-0.5 \leqslant x \leqslant 0.5$ with uniform separation Δx and the density $\rho = 1$. For this example, we use 200 particles and set $\gamma = 1.4$, the initial $h = 1.5\Delta x$, and the thermal energy/mass is 2. The SPH acceleration, continuity and thermal energy equation were integrated. In figure 14 the velocity field for $x \geqslant 0$ is shown. The exact velocity field is shown by the solid line and the SPH results are shown by solid diamonds. The agreement between the two is excellent.

We now consider the shock tube used by Sod (1978) as a test for numerical techniques. The system is one-dimensional with uniform conditions on each side of a diaphragm which breaks at $t = 0$. To the left of the diaphragm ($x < 0$) the conditions are $\rho, P, v, \gamma = 1.0$, 1.0, 0.0, 1.4 and to the right (0.125, 0.1, 0.0, 1.4). The evolved system consists of (from the left), the undisturbed original conditions, a rarefaction, a contact discontinuity and a shock.

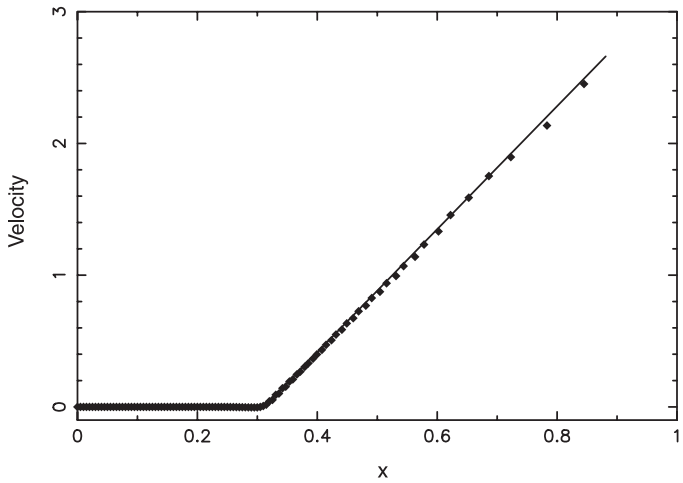

Figure 14. The velocity field for the one-dimensional rarefaction waves from the expansion of uniform gas initially in the region $-0.5 \leqslant x \leqslant 0.5$ is shown. The results for the right half $x \geqslant 0$ of the domain are also shown. The exact velocity field is shown by the solid line and the SPH results by the solid diamonds.

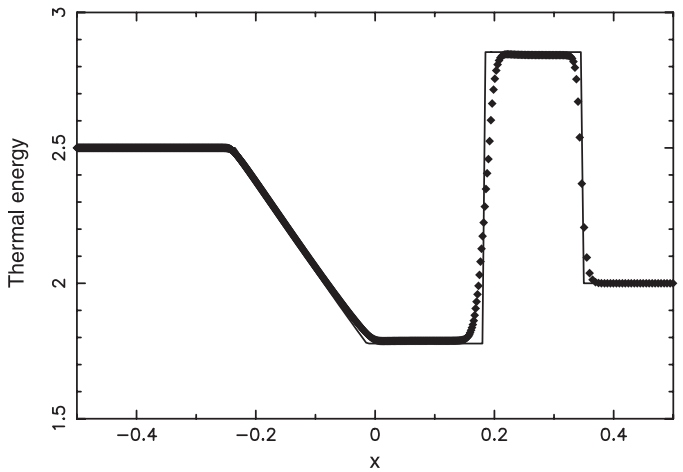

Figure 15. The thermal energy for a shock tube where the initial density ratio is 8 : 1. The particles have equal mass but the spacing is a factor of 8 smaller to the left of the initial diaphragm. The simulation uses 100 particles to the right of the initial diaphragm and 800 particles to the left. The exact results are shown by the continuous line and the SPH results by solid diamonds.

Between the shock and the rarefaction the pressure and velocity are constant. The density and thermal energy change discontinuously at the contact discontinuity. In this simulation the viscosity (8.3) with (8.10) was used with the coefficients: $\alpha = 1$ and $\beta = 2$. The particles have equal mass. Since there is an initial discontinuity in all the properties other than the initial velocity, the density and thermal energy are smoothed at the interface (as in Monaghan (1997)). Hence to be consistent with the particles having constant mass and the density being smooth, we must smooth the spacing. Price (2004a) finds that using h, calculated consistently with the density, according to (4.2) and (4.3) gives better results. The thermal energy is shown in figure 15 and the velocity field in figure 16. The solid lines show the exact results. The

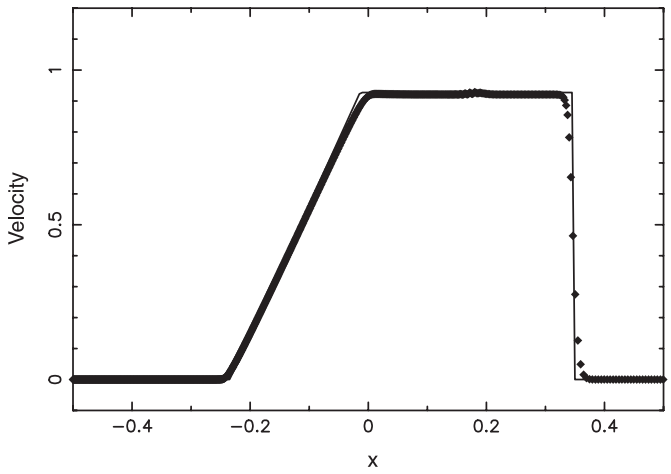

Figure 16. The velocity field for a shock tube. The exact results are shown by the continuous line and the SPH results by solid diamonds. The value of h and the particle spacing becomes smaller as the gas passes through the shock. The width is, therefore, ∼3 initial particle spacings. Note the variation in the velocity owing to a slight jump in the pressure at the contact discontinuity.

post-shock values are accurate to within 1%, though the shock fronts are broader than the comparable Riemann solver shocks. The actual broadening is smaller than the number of particles across the shock would suggest because, on entering the shock, the particles move closer together, and h becomes smaller.

10. Applications of SPH to liquids

A liquid such as water is slightly compressible but, for many fluid dynamical problems, it can be approximated by an artificial incompressible fluid, and this is the basis of most of the finite difference numerical algorithms for liquids. An alternative approach, better suited to SPH, is to approximate the liquid by an artificial fluid which is slightly compressible. All that is required is that the speed of sound be large enough for the density fluctuations to be negligible (Monaghan 1994). The equation of state most frequently used is due to Cole (1948), (see also Batchelor (1974)) which, when atmospheric pressure is negligible, has the form

$$P = B\left(\left(\frac{\rho}{\rho_0}\right)^{\gamma} - 1\right), \tag{10.1}$$

where ρ_0 is a reference density, $\gamma \sim 7$ and B is chosen so that the speed of sound is large enough to keep the relative density fluctuation $|\delta\rho|/\rho$ small. Since

$$\frac{|\delta\rho|}{\rho} \sim \frac{v^2}{c_s^2}, \tag{10.2}$$

where v is the maximum speed of the fluid we can ensure $|\delta\rho|/\rho \sim 0.01$ if $v/c_s < 0.1$. The speed of sound at the reference density is

$$c_s^2 = \frac{\gamma B}{\rho_0}. \tag{10.3}$$

Therefore, if $B = 100\rho_0 v^2/\gamma$, the relative density fluctuations should be ∼0.01. This requires an estimate of the maximum speed to be found which is often very easy to do. An example

of the results that can be achieved with SPH is shown in figure 5, in section 2, where the SPH calculations of Colagrossi (2004) are compared with those from a combination of level set and finite difference methods. This figure shows the evolution of liquid which is initially in the shape of a square and is distorted by a velocity field with spatially varying viscosity.

10.1. Boundaries

Most problems involving liquids also involve boundaries which may be fixed or moving or they might represent the surfaces of rigid bodies, wholly or partially, within the fluid. These boundaries may be handled easily by replacing the boundary with particles which interact with the fluid with prescribed forces. In this way, complicated problems involving fluids interacting with rigid bodies (which may float) and contained within an arbitrarily moving rigid body can be treated easily. An example would be the dynamics of a damaged car ferry with water pouring into the decks containing the cars.

Let $\boldsymbol{f}_{ka}$ be the force per unit mass on boundary particle k due to fluid particle a. To ensure that linear and angular momentum of the entire system is conserved in the absence of external forces, the force on a due to k must be equal and opposite to the force on k due to a. The most obvious way to specify the forces would be to use a Lennard–Jones force acting between the centres of the particles (Monaghan 1994). However, the large variation in the force on a particle moving parallel to the boundary causes large disturbance to flow near a boundary. A better procedure is the following (Monaghan *et al* 2004).

Consider the interaction between a fluid particle a and a boundary particle k where the local unit normal to the boundary is $\boldsymbol{n}_k$. If the distance measured normal to the boundary, from the boundary particle to the fluid particle, is denoted by y and the tangential distance by x, then a suitable form for the force per unit mass on boundary particle k due to fluid particle a is

$$\boldsymbol{f}_{ka} = -\frac{m_a}{m_a + m_k} B(x, y)\boldsymbol{n}_k, \tag{10.4}$$

where $B(x, y)$ is chosen to ensure that B rapidly increases as y decreases towards zero (to prevent penetration of the walls) and the variation with x ensures that the force on a particle moving parallel to the wall is constant. The total force/unit mass on boundary particle k due to all fluid particles is then $\boldsymbol{f}_k = \sum_a \boldsymbol{f}_{ka}$. The force per unit mass on fluid particle a due to boundary particle k is

$$\boldsymbol{f}_{ak} = \frac{m_k}{m_a + m_k} B(x, y)\boldsymbol{n}_k, \tag{10.5}$$

so that the forces $m_k\boldsymbol{f}_{ka} = -m_a\boldsymbol{f}_{ak}$ are equal and opposite. The total force per unit mass on fluid particle a from all boundary particles is $\boldsymbol{f}_a = \sum_k \boldsymbol{f}_{ak}$.

The equation of motion of a fluid particle a is then

$$\frac{\mathrm{d}v_a}{\mathrm{d}t} = -\sum_b m_b \left(\frac{P_a}{\rho_a^2} + \frac{P_b}{\rho_b^2} + \Pi_{ab}\right) \nabla_a W_{ab} + \boldsymbol{f}_a. \tag{10.6}$$

10.2. Motion of a rigid body interacting with a liquid

If the system consists of a liquid containing a rigid body with centre of mass $\boldsymbol{R}$ and centre of mass velocity $\boldsymbol{V}$, the equations of motion of this body are first, the equation for the motion of the centre of mass

$$M\frac{\mathrm{d}\boldsymbol{V}}{\mathrm{d}t} = \sum_k m_k \boldsymbol{f}_k, \tag{10.7}$$

where the summation over k only refers to the boundary particles on the surface of the rigid body. Second, the equation for the angular velocity Ω about the centre of mass which, in the case of 2D motion, is

$$I\frac{\mathrm{d}\Omega}{\mathrm{d}t} = \tau, \tag{10.8}$$

where I is the moment of inertia (a scalar for the present case) and τ is the total torque about the centre of mass. The torque can be calculated from the forces on the boundary particles of the rigid body. We then get

$$I\frac{\mathrm{d}\boldsymbol{\Omega}}{\mathrm{d}t} = \sum_k m_k(\boldsymbol{r}_k - \boldsymbol{R}) \times \boldsymbol{f}_k, \tag{10.9}$$

where the direction of $\boldsymbol{\Omega}$ is perpendicular to the plane of the motion and $\boldsymbol{R}$ is the position of the centre of mass. The rigid body boundary particles move as part of the rigid body so that the change in position of boundary particle k is given by

$$\frac{\mathrm{d}\boldsymbol{r}_k}{\mathrm{d}t} = \boldsymbol{V} + \boldsymbol{\Omega} \times (\boldsymbol{r}_k - \boldsymbol{R}). \tag{10.10}$$

From (10.6) and (10.7) we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_a m_a \boldsymbol{v}_a + M\boldsymbol{V}\right) = 0, \tag{10.11}$$

because the pair forces in each term cancel. Linear momentum is, therefore, conserved. To prove the conservation of angular momentum is a little more complicated. Since it has not been given in the literature, I now give it here.

The rate of change of the angular momentum of the rigid body about a fixed origin, when the motion takes place in a plane, is

$$\frac{\mathrm{d}\boldsymbol{J}}{\mathrm{d}t} = M\boldsymbol{R} \times \frac{\mathrm{d}\boldsymbol{V}}{\mathrm{d}t} + I\frac{\mathrm{d}\Omega}{\mathrm{d}t}. \tag{10.12}$$

Using the previous equations the right-hand side becomes

$$\boldsymbol{R} \times \sum_k m_k \boldsymbol{f}_k + \sum_k m_k(\boldsymbol{r}_k - \boldsymbol{R}) \times \boldsymbol{f}_k = \sum_k m_k \boldsymbol{r}_k \times \boldsymbol{f}_k \tag{10.13}$$

$$= \sum_k \sum_a m_k \boldsymbol{r}_k \times \boldsymbol{f}_{ka}. \tag{10.14}$$

The rate of change of the angular momentum of the liquid is

$$\sum_a m_a \boldsymbol{r}_a \times \frac{\mathrm{d}\boldsymbol{v}_a}{\mathrm{d}t} = \sum_a \sum_k m_a \boldsymbol{r}_a \times \boldsymbol{f}_{ak}, \tag{10.15}$$

because the pressure forces give zero net contribution to the total angular momentum of the fluid as we showed earlier. Adding the rate of change of angular momentum of the rigid body and the liquid and recalling that $m_k \boldsymbol{f}_{ka} = -m_a \boldsymbol{f}_{ak}$ gives

$$\sum_k \sum_a m_a(\boldsymbol{r}_a - \boldsymbol{r}_k) \times \boldsymbol{f}_{ka} = \sum_k \sum_a \frac{m_a m_k}{m_a + m_k}(\boldsymbol{r}_a - \boldsymbol{r}_k) \times \hat{\boldsymbol{n}}_k B(x, y). \tag{10.16}$$

Now consider the contribution to the rate of change of angular momentum from a liquid particle a. Suppose this particle lies between two boundary particles k and $(k+1)$ and suppose the tangential distance to k is x and to $(k+1)$ is $(1-x)$ assuming the unit of length is the separation of the boundary particles. The contribution from the previous summation is then

$$xB(x, y) - (1-x)B(1-x, y) = 0 \tag{10.17}$$

provided $B(x, y) = (1 - x)\Gamma(y)$ which is the choice made in the next section. The rate of change of total angular momentum is, therefore, zero. This proof requires that the masses of the boundary particles are equal. If they are not equal then the forces must be scaled so that the torque from neighbouring boundary particles vanishes.

10.3. The boundary force

Monaghan *et al* (2004) write $B(x, y)$ as a product $\Gamma(y)\chi(x)$ where the function $\chi(x)$ is defined by

$$\chi(x) = \begin{cases} \left(1 - \dfrac{x}{\Delta p}\right), & \text{if } 0 < x < \Delta p, \\ 0, & \text{otherwise,} \end{cases}$$

where Δp is the boundary particle spacing. This factor ensures that a fluid particle moving parallel to the wall will always feel the same force because, when it is between any two boundary particles, the total force from them is constant, regardless of where it lies between them.

The essential condition on the function $\Gamma(y)$ is that it increases as y decreases to prevent penetration of the wall. Monaghan *et al* (2004) choose a form related to the gradient of the cubic spline with the argument $q = y/h$. The gradient of the cubic spline has a maximum at $q = 2/3$. For $0 < q < 2/3$ they replace the value of the gradient by its maximum. Thus,

$$\Gamma(y) = \beta \begin{cases} \frac{2}{3}, & \text{if } 0 < q < \frac{2}{3}, \\ (2q - \frac{3}{2}q^2), & \text{if } \frac{2}{3} < q < 1, \\ \frac{1}{2}(2 - q)^2, & \text{if } 1 < q < 2, \\ 0, & \text{otherwise,} \end{cases}$$

where β is $0.02\, c_s^2/y$. This term is an estimate of the maximum force/mass necessary to stop a particle moving at the estimated maximum speed. The factor $1/y$ ensures that a faster moving particle can be stopped. Other details concerning boundaries, including the treatment of corners, are given by Monaghan *et al* (2004). If there is more than one rigid body interacting with the fluid, then the same methods can be used but now there may be an interaction between the bodies. There are, therefore, two types of boundary forces, namely, boundary–fluid and boundary–boundary. The best choice of boundary force is not known.

Other authors (e.g. Colagrossi and Landrini 2003) prefer to replace the rigid boundaries by ghost particles. This has advantages when the geometry is simple because the use of ghost particles gives less disturbance to the fluid. However, for complicated geometries, for example, those describing engine body parts in liquid metal moulding (Cleary and Ha 2002), or in geophysical flows, boundary particles are easier to use and can be more accurate. A simple generalization is to allow the boundary particles to have a different interaction with different fluids. For example, in the case of a dusty gas, the dust and gas SPH particles could interact with the boundary particles with different forces.

10.4. Applications to rigid bodies in water

The early applications were to bores, dam collapse, wave makers and breaking waves, though not to a high accuracy because only a small number of particles were used (Monaghan 1994). Further applications, with comparisons between SPH and experimental results for waves on beaches were made by Monaghan and Kos (1999), who also studied the generation of solitary waves by dropping boxes (Monaghan and Kos 2000) or by sliding boxes down ramps

Monaghan *et al* (2004). The mechanism by which a rising bubble could sink a ship was studied using experiments and SPH simulations by May and Monaghan (2003). Colagrossi and Landrini (2003) describe applications to more than one fluid in dam break (see also Colicchio *et al* (2002)). They also considered the rise of bubbles in water and the effect of air on wave breaking. Their work incorporates a number of improvements for these problems including a periodic re-initialization of the density field based on moving least squares interpolation (Belytschko *et al* 1998), and a generalized Balsara correction (discussed in section 9). Because of the contrast in density they use the acceleration equation (5.35) with $\Phi = 1$, and the convergence equation (5.26).

Sloshing tanks have been studied by Colagrossi (2004) and Colagrossi *et al* (2003) who found that the SPH simulations revealed an aspect of the sloshing not noticed previously. Subsequent specially designed experiments confirmed this prediction.

10.5. Turbulence

There have been limited studies of turbulence using SPH. Studies of wave breaking by Colagrossi (2004), and Landrini *et al* (2003) show that detailed properties of the complex vortices resulting from wave breaking can be recovered using SPH. A fully Lagrangian turbulence model based on the Lagrangian averaged alpha model (Holm 1999), Mohseni *et al* 2003) has been worked out (Monaghan 2002, 2004), but no comparisons have been made with other, more traditional, methods. In the SPH Lagrangian averaged model a typical particle a is moved with the XSPH smoothed velocity $\hat{v}_a$ (Monaghan 1989). This was originally defined by

$$\hat{\boldsymbol{v}}_a = \boldsymbol{v}_a + \epsilon \sum_b m_b \frac{(\boldsymbol{v}_b - \boldsymbol{v}_a)}{\bar{\rho}_{ab}} W_{ab}, \tag{10.18}$$

where $\epsilon \sim 0.5$ is constant and the kernel need not be the same as the kernel used in calculating the density. This smoothed velocity brings the particle velocity closer to the average velocity in its neighbourhood and reduces the particle disorder. Moving the particles with the new velocity does not change the linear or angular momentum. However, if the particles are moving with the smoothed velocity, energy is not conserved. To conserve it, and bring the algorithm into agreement with the alpha model of turbulence, we replace (10.18) with

$$\hat{\boldsymbol{v}}_a = \boldsymbol{v}_a + \epsilon \sum_b m_b \frac{(\hat{\boldsymbol{v}}_b - \hat{\boldsymbol{v}}_a)}{\bar{\rho}_{ab}} W_{ab}. \tag{10.19}$$

Moving particles with this velocity still conserves linear and angular momentum. In the continuum limit the previous equation becomes

$$\hat{\boldsymbol{v}}_a = \boldsymbol{v}_a + \frac{1}{2}\frac{\epsilon}{\rho}\nabla^j \left(\rho \nabla^j \hat{\boldsymbol{v}}_a\right) \left(\int q^2 W(q,h)\, \mathrm{d}\boldsymbol{q} \right), \tag{10.20}$$

where $\nabla^j = \partial/\partial x^j$. To compare with the continuum α model define

$$\alpha_{\mathrm{turb}} = \frac{1}{2}\epsilon \int q^2 W(q,h)\, \mathrm{d}\boldsymbol{q} \sim \epsilon \boldsymbol{h}^2, \tag{10.21}$$

so that (10.20) agrees, when ρ is constant, with Holm (1999).

The smoothing algorithm is similar to a discrete time-stepping of a diffusion equation. For example, the diffusion equation

$$\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}\tau} = \frac{\kappa}{\rho}\nabla^j(\rho \nabla^j \boldsymbol{v}) \tag{10.22}$$

can be approximated by the implicit, discrete time stepping

$$\boldsymbol{v}^{n+1} = \boldsymbol{v}^n + \delta\tau \frac{\kappa}{\rho} \nabla^j (\rho \nabla^j \boldsymbol{v}^{n+1}). \tag{10.23}$$

If $\hat{\boldsymbol{v}}$ is identified with $\boldsymbol{v}^{n+1}$, $\boldsymbol{v}$ with $\boldsymbol{v}^n$, the two equations (10.20) and (10.23) become equivalent. Because the implicit smoothing is stable it can be used for any value of $\epsilon > 0$. However, in practice, this implicit equation must be solved by iteration and several iterations may be needed if $\epsilon > 1$.

To complete the dynamics the most convenient approach is to use a Lagrangian (Monaghan 2002).

$$L = \sum_b m_b \left(\frac{1}{2} \hat{\boldsymbol{v}}_b \cdot \boldsymbol{v}_b - u_b - \Phi_b \right), \tag{10.24}$$

where Φ is a potential energy. The kinetic energy term can be written as

$$\frac{1}{2} \sum_b m_b \hat{\boldsymbol{v}}_b \cdot \hat{\boldsymbol{v}}_b + \frac{\epsilon}{4} \sum_a \sum_b m_a m_b \frac{\hat{\boldsymbol{v}}_{ab}^2}{\bar{\rho}_{ab}} W_{ab}, \tag{10.25}$$

from which the canonical momentum of particle a can be calculated. Remarkably, it is just $m_a \boldsymbol{v}_a$.

The smoothing of the velocity makes the Lagrangian averaged model similar to the large Eddy simulation method. However, the Lagrangian leads to a different set of equations from those used in LES simulations and variable resolution length is built into the equations. An interesting feature of these equations is that, in the absence of any dissipation, they result in the energy being redistributed so that the energy transfer to short length scales is impeded (Mohseni *et al* 2003, Monaghan 2004). Various aspects of these equations are discussed by Monaghan (2002, 2004). There is a need to apply the SPH turbulence model to standard problems such as turbulence decay in two- and three-dimensional boxes. An interesting astrophysical example to study would be the turbulence in toy stars.

Particle methods lead naturally to the idea of studying turbulence along the lines of statistical mechanics, that is in terms of the velocity and spatial distributions of the particles. No detailed work on this has appeared in the literature though there has been some analysis of probability distributions using SPH (Welton 1998, Welton and Pope 1997).

10.6. Multiphase flow

It is straightforward to include more than one fluid in SPH simulations. Each fluid has its own set of SPH particles with an appropriate equation of state. All the SPH particles are used in the summations. If the fluids are incompressible, the technique described earlier, where the speed of sound is artificial, and sufficiently large to make density fluctuations negligible, can be used. Gravity currents flowing into a stratified fluid have been studied using both experiment and simulation (Monaghan *et al* 1999) and air–water interactions have been simulated successfully by Colagrossi and Landrini (2003). Dusty gas occurs in both astrophysics and in volcanic outbursts. A formulation of SPH suitable for dusty gas (Monaghan and Kocharyan 1995) is available but no applications have appeared in the literature, yet.

11. Elasticity and fracture

The equations of elastic dynamics are the acceleration equation

$$\frac{\mathrm{d}v^i}{\mathrm{d}t} = \frac{1}{\rho} \frac{\partial \sigma^{ij}}{\partial x^j} + g^i, \tag{11.1}$$

where the stress tensor is given by

$$\sigma^{ij} = -P\delta^{ij} + S^{ij}, \tag{11.2}$$

S^{ij} is the deviatoric stress and g^i denotes the ith component of a body force per unit mass. In linear elastic theory, the deviatoric stress can be obtained from Hooke's law with shear modulus μ

$$\frac{\mathrm{d}S^{ij}}{\mathrm{d}t} = 2\mu\left(\dot{\epsilon}^{ij} - \frac{1}{3}\delta^{ij}\dot{\epsilon}^{kk}\right) + S^{ik}R^{jk} + R^{ik}S^{kj}, \tag{11.3}$$

where

$$\dot{\epsilon}^{ij} = \frac{1}{2}\left(\frac{\partial v^i}{\partial x^j} + \frac{\partial v^j}{\partial x^i}\right) \tag{11.4}$$

and

$$R^{ij} = \frac{1}{2}\left(\frac{\partial v^i}{\partial x^j} - \frac{\partial v^j}{\partial x^i}\right). \tag{11.5}$$

Alternative laws for the time change of the deviatoric stress (Ellero *et al* 2002) can be used without any change in the formalism. The pressure P is normally obtained from the Tillotson or Mie Gruniessen equation of state. The elastic equations can be converted into SPH form following the principles already established. In particular, the acceleration equation becomes

$$\frac{\mathrm{d}v_a^i}{\mathrm{d}t} = \sum_b m_b\left(\frac{\sigma_a^{ij}}{\rho_a^2} + \frac{\sigma_b^{ij}}{\rho_b^2} + \Pi_{ab}\right)\frac{\partial W_{ab}}{\partial x_a^j} + g^i \tag{11.6}$$

and the velocity derivatives in the equation for the deviatoric stress and the tensor R^{ij} can be estimated using the methods in section 2.2.

The elastic dynamical equations were first studied by Libersky and Petschek (1991). A comprehensive discussion by Randles and Libersky (1996) covers many aspects of elastic SPH. The elastic equations were combined with an elastic fracture model (Grady and Kipp 1987, Benz and Asphaug 1994, 1995) in order to study asteroid/comet and planetesimal collisions (Michel *et al* (2004) give a comprehensive review of this work). The brittle fracture model of Grady and Kipp (1987) is based on Griffiths theory of fracture. A set of flaws (cracks) is assigned to the SPH particles at random, according to the Weibull distribution. Depending on the flaw, tension may or may not cause it to grow. The growth is associated with the local damage quantified by a damage parameter D. When D is zero it means that the material is perfectly elastic and when D increases to 1 the material is completely damaged and the contribution of the deviatoric stress is zero. The precise way in which the flaws are assigned, and the equation for D, are discussed in detail by Benz and Aspaugh.

Since material carries its damage with it, Lagrangian models like SPH are uniquely designed to model fracture. SPH, in particular, gives a good description of the fragments and provides a natural transition from the continuum to the fragmented state. This method has also been used to study fracture in and around volcanoes (Gray and Monaghan 2004).

In the initial application of SPH to elastic problems it was noticed that, under tension, particles tended to clump in pairs. This instability was first analysed by Phillips and Monaghan (1985) in the context of magneto-hydrodynamics. They showed that the tension which always exists in magnetic fields can cause an instability. The instability was re-discovered by Swegle *et al* (1995) in the context of elastic simulations and called the tensile instability. Various methods have been proposed to eliminate it from SPH simulations. The most successful has been the artificial stress method (Monaghan 2000) and Gray *et al* (2001) which also includes using the XSPH (Monaghan 1989) smoothed velocity. Others include

additional stress points (Dyka *et al* 1997), and the correction of the kernels to give exact linear interpolation (Dilts 1999, Bonet and Lok 1999, Bonet and Kulasegaram 2000, 2001). One skeleton in the SPH closet is that the normal SPH elastic equations do not conserve angular momentum. A spectacular example is a rotating elastic wheel which loses its rotation after one rotation or less. Hoover *et al* (2004) show that by using strong XSPH smoothing the loss of angular momentum could be reduced. Atomic models of elasticity conserve angular momentum exactly and it would be worth investigating whether an SPH elastic model can be based on the atomic models with the SPH particles mimicking atoms.

12. Special and general relativistic SPH

The continuum equations for special relativity can be derived from the derivative of the energy momentum tensor

$$\frac{\partial T^{\mu\nu}}{\partial x^{\nu}} = 0. \tag{12.1}$$

For a non-dissipative gas of baryons each with rest mass m_0, $T^{\mu\nu}$

$$T^{\mu\nu} = (nm_0c^2 + nu + P)U^{\mu}U^{\nu} + P\eta^{\mu\nu}, \tag{12.2}$$

where U^{μ} is a 4-velocity and c is the speed of light. In the following, the velocity unit is c and the energy unit is m_0c^2. In (12.1) n and u are the baryon number density and the energy/baryon in the rest frame of the element of fluid they refer to. The metric tensor $\eta^{\mu\nu}$ has the signature $(-1, 1, 1, 1)$. The resulting equations can be solved using a variety of computational algorithms. When the gas is ideal (i.e. $P = (\Gamma - 1)nu$) excellent results have been obtained using Riemann methods (Marti and Mueller 2003). For more complicated equations of state, for example, those that are used to mimic heavy ion collisions (Amsden *et al* 1978), particle in cell (PIC) methods have been used (Amsden *et al* 1978). In this latter case, after the collision, the rapidly expanding pion gas ceases to behave like a continuum fluid and behaves more like a set of particles in a process called 'freeze-out'. This situation would be very easy to simulate using SPH because it handles the transition from continuum fluid to particles seamlessly.

SPH equations for special relativity can be derived either from the continuum equations (Mann 1991, Laguna *et al* 1993, Chow and Monaghan 1997) or from a Lagrangian (Monaghan and Price 2001). The Lagrangian is

$$L = -\int T^{\mu\nu}U_{\mu}U_{\nu}\,\mathrm{d}\boldsymbol{r}, \tag{12.3}$$

or

$$L = -\int n(1+u)\,\mathrm{d}\boldsymbol{r}. \tag{12.4}$$

The SPH formalism can be set up in a selected frame, conveniently called the computing frame. In this frame the baryon number density is

$$N = nU^0 = n\gamma = n/\sqrt{(1-v^2)}. \tag{12.5}$$

Using standard SPH interpolation but replacing the mass m_b for SPH particle b by the number of baryons ν_b, Monaghan and Price (2001) show that the Lagrangian becomes

$$L = -\sum_b \nu_b\sqrt{(1-v_b^2)}(1+u_b), \tag{12.6}$$

with

$$N(\boldsymbol{r}) = \sum_b \nu_b W(|\boldsymbol{r} - \boldsymbol{r}_b|). \tag{12.7}$$

The Lagrangian equations then give the acceleration equation

$$\frac{\mathrm{d}\boldsymbol{p}_a}{\mathrm{d}t} = -\nu_a \sum_b \nu_b \left(\frac{P_a}{N_a^2} + \frac{P_b}{N_b^2} \right) \nabla_a W_{ab}, \tag{12.8}$$

where the canonical momentum $\boldsymbol{p}_a$ is given by

$$\boldsymbol{p}_a = \nu_a \left(1 + u_a + \frac{P_a}{n_a} \right). \tag{12.9}$$

This equation is identical to that obtained from the continuum equations by Chow and Monaghan (1997). To apply the SPH equations to strong shocks it is necessary to add dissipative terms. The dissipative terms for the non-relativistic case can be chosen by analogy with the actual viscosity and heat conduction of a gas. However, in the relativistic case there are no accurate relativistic dissipative terms (if they existed they would involve relativistic fields) to act as a guide The approach of Eckart (1940), and Landau and Lifshitz (1993) gives dissipative terms which are unstable (Hiscock and Lindblom 1985), while Carter's approach fails to correctly describe the non-relativistic gases (Olson and Hiscock 1990). Chow and Monaghan (1997), therefore, based their dissipation terms on those chosen for Riemann problems which, in the SPH form, are similar to those worked out by Amsden *et al* (1978) by considering baryon scattering. These dissipation terms are very effective and give a degree of accuracy comparable to methods based on Riemann solvers. The disappointing early SPH calculations of Mann (1991) and Laguna *et al* (1993) can be attributed to their poor choice of artificial viscosity. No attempt has yet been made to solve these problems with h and N, calculated consistently as described earlier for h and ρ.

The general relativistic equations for fluid dynamics in a specified metric can also be obtained from a Lagrangian. The resulting SPH equations (Monaghan and Price 2001) differ from those of Siegler and Riffert (2000) which do not conserve momentum. The shock calculations of Siegler and Riffert (2000) show unphysical jumps at the contact discontinuity. These are due to the lack of heat conduction in the dissipative terms. At present no satisfactory dissipative terms have appeared in the literature. One obvious approach would be to use the signal velocities found for Riemann solvers, then construct dissipative terms along the lines of those used by Chow and Monaghan (1997).

13. Prospects for the future

The features of SPH which make it an effective computational algorithm are ultimately due to the fact that it can be derived from a Lagrangian and has the conservation properties of a Lagrangian system. As a result, the conservation of momentum and energy together with the approximate invariant of the circulation follow naturally. However, SPH also conserves composition, that is, each particle carries its composition unchanged unless the material the particle represents undergoes chemical transformations. This property of carrying composition unchanged has not been fully exploited despite its importance in both industry and astronomy. In the latter case the extent to which elements are fully mixed in clusters of stars is known from observation, but has not been studied with simulations.

Another attractive feature of SPH is that the resolution adjusts smoothly to changes in the density, but there is no reason, other than computational efficiency, why the resolution could

not be changed in response to steep local gradients in other quantities, for example, temperature gradients. Preliminary steps have already been taken. These are:

- Direct splitting Kitsionas and Whitworth (2002).
- Adjustment of h and particle position Børve *et al* (2001).
- Regridding of the particles Chaniotis *et al* (2002).

All of these have defects which may be overcome. The first has not been tested for splitting and merging in problems where the split–merge rule depends, for example, on the temperature gradient, nor has it been tested for liquids with a stiff equation of state where density perturbations may lead to large pressure changes. The second is computationally intensive, however, that may be the price one has to pay for a better algorithm. The third has not been extensively tested with the split–merge rule based on a variety of gradients and, in current formulations, leads to excess diffusion, but this can be expected to be greatly reduced in the future.

To achieve a robust SPH algorithm for splitting and merging, it might be useful to reflect on what happens in nature. For example, when a gas moves into a region of high temperature the atoms smoothly ionize, producing more particles, then smoothly recombine if they enter a cooler state. This is exactly the process that would be natural for SPH and it could be implemented by allowing a particle to split, but placing the new particles close together so the effect on the flow will initially be negligible. The original particle, now less massive, could be tagged to provide a nucleus for merging. It mimics the role of the ion in the ionized gas example. The flow would gradually spread the new particles because of their slightly different velocities. The merging could occur by allowing the split particles to be attracted to the tagged particles. This would be a continuous process similar to the way ions and electrons in an ionized gas combine when cooled.

The next class of software advances would be the development of more efficient strategies for handling very low Mach number flows. These are required for industrial, geological and oceanographic hydrodynamic problems, and for simulating the dynamics of elastic materials. Recent work (Hu and Adams, preprint (2005)) has produced accurate and robust SPH algorithms for multi-phase flow including surface tension effects involving three fluids, but the maximum density ratio of the fluids is 100, which is an order of magnitude less than the air–water density ratio. This work improves on that of Colagrossi and Landrini described in this review. If the efficient strategies can be found then the low Mach number flows in geology could be handled efficiently. At present, the most straightforward application of SPH is in the volcanic outbursts as these are often close to the speed of sound. An implicit code would enable SPH codes to be devised for plate tectonics. In marine engineering we could look forward to simulations of water–metal impact leading to breakage and providing information about the stability of ships, especially those containing dynamically significant moving parts, as in a car ferry.

Within the category of software development we mention two, the first being concerned with MHD problems. We can expect significant advances in the next few years as the technique of Børve *et al* (2001) is made more efficient, and that of Price and Monaghan (2004a) is improved. The second is concerned with the multi-scaling problems, where calculations at the atomic level are linked to macroscopic dynamics. Many researchers have noted that SPH allows a seamless transition from the continuum to the fragments in problems involving fracture and splashing fluids. It is natural, therefore, to predict that SPH will provide an effective approach to multi-scaling simulations.

Finally, it is worth noting that apart from the advances in software there have been significant advances in hardware as well. New chips (FPGA), in particular, can be programmed

to implement the SPH summations in hardware. This will lead to SPH simulations which are extremely fast and will make previously difficult problems trivial.

References

Amsden A A, Goldhaber A S, Harlow F H and Nix J R 1978 Relativistic two-fluid model of nucleus–nucleus collsions *Phys. Rev.* C **17** 2080–96

Ball F K 1963 Some general theorems concerning the finite motion of a shallow liquid lying on a paraboloid *J. Fluid Mech.* **17** 240–56

Balsara D S 1995 von Neumann stability analysis of smooth particle hydrodynamics—suggestions for optimal algorithms *J. Comput. Phys.* **121** 357–72

Batchelor G K 1974 *An Introduction to Fluid Mechanics* (Cambridge: Cambridge University Press)

Bate M R, Bonnell I A and Bromm V 2003 The formation of a star cluster predicting the properties of stars and Brown Dwarfs *Mon. Not. R. Astron. Soc.* **399** 577–99

Bate M R, Bonnell I A and Price N M 1995 Modelling accretion in protobinary systems *Mon. Not. R. Astron. Soc.* **277** 362–76

Belytschko T, Krongauz Y, Organ D and Gerlack G 1998 On the completeness of meshfree particle methods *Int. J. Numer. Methods Eng.* **43** 785–819

Ben Moussa B, Lanson N and Vila J P 1999 Convergence of meshless methods for conservations laws: applications to Euler equations *Int. Ser. Numer. Math.* **129** 31–40

Benz W 1990 Smoothed particle hydrodynamics—a review *The Numerical Modelling of Nonlinear Stellar Pulsations* ed J R Buchler (Dordrecht: Kluwer) pp 269–88

Benz W and Asphaug E 1994 Impact simulations with fracture: I. Method and tests *Icarus* **1233** 98–116

Benz W and Asphaug E 1995 Simulations of brittle solids using smoothed particle hydrodynamics *Comput. Phys. Commun.* **87** 253–65

Benz W, Slattery W L and Cameron A G W 1986 The origin of the moon and the single impact hypothesis *Icarus* **66** 515–35

Bonet J and Kulasegaram S 2000 Correction and stabilization of smooth particle hydrodynamics methods with applications in metal forming simulations *Int. J. Numer. Methods Eng.* **47** 1189–214

Bonet J and Kulasegaram S 2001 Remarks on tension instability of Eulerian and Lagrangian corrected smooth particle hydrodynamics (CSPH) methods *Int. J. Numer. Methods Eng.* **52** 1203–20

Bonet J and Lok T-S L 1999 Variational and momentum preservation aspects of smooth particle hydrodynamic formulations *Comput. Methods Appl. Mech. Eng.* **180** 97–115

Boneva L I, Kendall D and Stepanov I 1971 Spline transformations: three new diagnostic aids for statistical data analysis *J. R. Stat. Soc.* B **33** 1–37

Børve S, Omang M and Trulsen J 2001 Regularized smoothed particle hydrodynamics: a new approach to simulating magnetohydrodynamic shocks *Astrophys. J* **561** 82–93

Brookshaw L 1985 A method of calculating radiative heat diffusion in particle simulations *Proc. Astron. Soc. Aust.* **6** 207–10

Carslaw H S and Jaeger J C 1990 *Conduction of Heat in Solids* (Oxford: Oxford University Press)

Chaniotis A K, Poulikakos D and Kououtsakos P 2002 Remeshed smoothed particle hydrodynamics for the simulations of viscous and heat conducting flows *J. Comput. Phys.* **182** 67–90

Chandrasekhar S 1995 *Newton's Principia for the Common Reader* (Oxford: Clarendon)

Chow E and Monaghan J J 1997 Ultrarelativistic SPH *J. Comput. Phys.* **134** 296–305

Cleary P W 1998 Modelling confined multi-material heat and mass flows using SPH *Appl. Math. Modelling* **22** 981–93

Cleary P W and Ha J 2002 Flow modelling in casting processes *Appl. Math. Modelling* **26** 171–90

Cleary P W and Monaghan J J 1999 Conduction modelling using smoothed particle hydrodynamics *J. Comput. Phys.* **148** 227–64

Colagrossi A 2004 Dottorato di Ricerca in Meccanica Teorica ed Applicata XVI CICLO A meshless Lagrangian method for free-surface and interface flows with fragmentation *PhD Thesis* Universita di Roma, La Sapienza

Colagrossi A and Landrini M 2003 Numerical simulation of interfacial flows by smoothed particle hydrodynamics *J. Comput. Phys.* **191** 448–75

Colagrossi A, Lugni C, Douset V, Bertram V and Faltinsen O 2003 Numerical and experimental study of sloshing in partially filled rectangular tanks *6th Numerical Towing Tank Symp. (Rome, Italy)*

Cole R H 1948 *Underwater Explosions* (Princeton, NJ: Princeton University Press)

Colicchio G, Colagrossi A, Greco M and Landrini M 2002 Free surface flow after a dam break *Ship Technol. Res.* **49** 95–104

Couchman H M P, Thomas P A and Pearce F R 1995 HYDRA: an adaptive mesh implementation of P^3M—*SPH Astrophys. J.* **452** 797–813

Dilts G A 1999 Moving least squares hydrodynamics: consistency and stability *Int. J. Numer. Methods* **44** 1115–55

Dyka C T, Randles P W and Ingel R P 1997 Stress points for tension instability in SPH *Int. J. Numer. Methods Eng.* **40** 2325–41

Eckart C 1940 The thermodynamics of irreversible processes: III. Relativistic theory of the simple fluid *Phys. Rev.* **58** 919–24

Eckart C 1960 Variation principles of hydrodynamics *Phys. Fluids* **3** 421–7

Ellero M, Kroger M and Hess S 2002 Viscoelastic flows studied by smoothed particle dynamics *J. Non-Newtonian Fluid Mech.* **105** 35–51

Español P and Revenga M 2003 Smoothed dissipative particle dynamics *Phys. Rev.* E **67** 026705

Feynman R P 1957 Applications of quantum mechanics to liquid helium *Low Temp. Phys.* **1** 17–53

Frank J and Reich S 2003 Conservation properties of smoothed particle hydrodynamics applied to the shallow-water equations *BIT* **43** 40–54

Fulk D A and Quinn D W 1996 An analysis of 1-D smoothed particle hydrodynamics kernels *J. Comput. Phys.* **126** 165–80

Gingold R A and Monaghan J J 1977 Smoothed particle hydrodynamics: theory and application to non-spherical stars *Mon. Not. R. Astron. Soc.* **181** 375–89

Gingold R A and Monaghan J J 1978 Binary fission in damped rotating polytropes *Mon. Not. R. Astron. Soc.* **184** 481–99

Gingold R A and Monaghan J J 1979 A numerical study of the Roche and Darwin problems for polytropic stars *Mon. Not. R. Astron. Soc.* **188** 45–58

Gingold R A and Monaghan J J 1980 The Roche problem for polytropes in central orbits *Mon. Not. R. Astron. Soc.* **191** 897–924

Gingold R A and Monaghan J J 1982 Kernel estimates as a basis for general particle methods in hydrodynamics *J. Comput. Phys.* **46** 429–53

Grady D E and Kipp M E 1987 Dynamic rock fragmentation *Fracture Mechanics of Rock* (New York: Academic) chapter 10, pp 429–73

Gray J, Monaghan J J and Swift R P 2001 SPH elastic dynamics *Comput. Methods Appl. Mech. Eng.* **190** 6641–62

Gray J A and Monaghan J J 2004 Numerical modelling of stress fields and fracture around magma chambers *Volcanology Geothermal Res.* **135** 259–83

Hernquist L and Katz N 1989 TREESPH—A unification of SPH with the hierarchical tree method *Astrophys. J.* (Suppl.) **70** 419–46

Hiscock W A and Lindblom L 1985 Generic instabilities in first order dissipative relativistic fluid theories *Phys. Rev.* **31** 725–33

Hockney R W and Eastwood J W 1988 *Computer Simulation Using Particles* (Bristol: Hilger)

Holm D 1991 Elliptical vortices and integrable Hamiltonian dynamics of the rotating shallow-water equations *J. Fluid Mech.* **227** 393–406

Holm D 1999 Fluctuation effects on 3D Lagrangian mean and Eulerian mean notion *Physica* D **133** 215–69

Hoover W G 1998 Isomorphism linking smooth particles and embedded atoms *Physica* A **260** 244–54

Hoover W G, Hoover C C and Merritt E C 2004 Smooth particle applied mechanics: conservation of angular momentum with tensile stability and velocity averaging *Phys. Rev.* E **69** 016702-1-10

Hu X Y and Adams N A 2005 A multi-phase SPH method for macroscopic and mesoscopic flows *J. Comput. Phys.* at press

Johnson G R and Beissel S R 1996 Normalized smoothing functions for impact computations *Int. J. Numer. Methods Eng.* **569** 501–18

Johnson G R, Stryk R A and Beissel S R 1996 SPH for high velocity impact computations *Comput. Methods Appl. Mech. Eng.* **139** 347–73

Kahan W H 1980 IEEE floating point standard ed K L E Nickel *Interval Mathematics* (New York: Academic)

Kitsionas S and Whitworth A P 2002 Smoothed particle hydrodynamics with particle splitting applied to self-gravitating flows *Mon. Not. R. Astron. Soc.* **330** 129–36

Laguna L D, Miller W A and Zureck W H 1993 Smooth particle hydrodynamics near a black hole *Astrophys. J* **404** 678–85

Lamb H 1932 *Hydrodynamics* (Cambridge: Cambridge University Press)

Landau L D and Lifshitz E M 1993 *Fluid Mechanics* 2nd edn (Oxford: Pergamon) p 512

Landrini M, Colagrossi L and Faltinsen O 2003 Sloshing in 2D flows by the SPH method *Proc. 8th Int. Conf. on Numerical Ship Hydrodynamics (Pusan, Korea)*

Lattanzio J C and Monaghan J J 1991 A simulation of the collapse and fragmentation of cooling molecular clouds *Mon. Not. R. Astron. Soc.* **375** 177–89
Lattanzio J C, Monaghan J J, Pongracic H and Schwarz M P 1985 Interstellar cloud collisions *Mon. Not. R. Astron. Soc.* **215** 125–47
Leimkuhler B J, Reich S and Skeel R D 1997 Integration methods for molecular dynamics *IMA Volume in Mathematics and its Applications* Vol 82 ed K Schulten and J Mesirov (Springer) pp 161–86
Libersky L and Petschek A G 1991 Smooth particle hydrodynamics with strength of materials *Advances in Free Lagrange Methods* ed H E Trease and M J Fritts (Springer)
Lucy L B 1977 A numerical approach to the testing of the fission hypothesis *Astron. J.* **82** 1013–24
Mann P J 1991 A relativistic smoothed particle hydrodynamics code tested with the shock tube *Comput. Phys. Commun.* **67** 245–60
Marri S and White S D M 2003 Smoothed particle hydrodynamics for galaxy-formation simulations: improved treatments of multiphase gas, of star formation and of supernovae feedback *Mon. Not. R. Astron. Soc.* **345** 561–74
Marti J M and Mueller E 2003 Numerical hydrodynamics in special relativity *Living Rev. (Electronic)* **7**
May D and Monaghan J J 2003 Can a single bubble sink a ship *Am. J. Phys.* **71** 842–9
Michel P, Benz W and Richardson D 2004 Catastrophic disruption of asteroids and family formation: a review of numerical simulations, including both fragmentation and gravitational reaccumulation *Planet. Space Sci.* **52** 1109–17
Mohseni K, Kosovic B, Shkoller S and Marsden J E 2003 Numerical simulations of the Lagrangian averaged Navier–Stokes equations for homogeneous isotropic turbulence *Phys. Fluids* A **15** 524–44
Monaghan J J 1985a Extrapolating B splines for interpolation *J. Comput. Phys.* **60** 253–62
Monaghan J J 1985b Particle methods for hydrodynamics *Comput. Phys. Rep.* **3** 71–124
Monaghan J J 1989 On the problem of penetration in particle methods *J. Comput. Phys.* **82** 1–15
Monaghan J J 1992 Smoothed particle hydrodynamics *Ann. Rev. Astron. Astrophys.* **30** 543–74
Monaghan J J 1994 Simulating free surface flows with SPH *J. Comput. Phys.* **110** 399–406
Monaghan J J 1997 SPH and Riemann solvers *J. Comput. Phys.* **136** 298–307
Monaghan J J 2000 SPH without a tensile instability *J. Comput. Phys.* **159** 290–311
Monaghan J J 2002 SPH compressible turbulence *Mon. Not. R. Astron. Soc.* **335** 843–52
Monaghan J J 2004 Energy distribution in a particle alpha model *J. Turbul.* **12** 1
Monaghan J J, Cas R F, Kos A and Hallworth M 1999 Gravity currents descending a ramp in a stratified tank *J. Fluid Mech.* **379** 36–9
Monaghan J J and Gingold R A 1983 Shock simulation by the particle method SPH *J. Comput. Phys.* **52** 374–89
Monaghan J J, Huppert H E and Worster M G 2005 Solidification using smoothed particle hydrodynamics *J. Comput. Phys.* **206** 684–705
Monaghan J J and Kocharyan A 1995 SPH simulation of multi phase flow *Comput. Phys. Commun.* **87** 225–35
Monaghan J J and Kos A 1999 Solitary waves on a Cretan beach *J. Waterways Port Coastal Ocean Eng.* **1111** 145–54
Monaghan J J and Kos A 2000 Scott Russell's wave generator *Phys. Fluids* A **12** 622–30
Monaghan J J, Kos A and Issa N 2004 Fluid motion generated by impact *J. Waterway Port Coastal Ocean Eng.* **129** 250–9
Monaghan J J and Price D J 2001 Variational principles for relativistic smoothed particle hydrodynamics *Mon. Not. R. Astron. Soc.* **328** 381–92
Monaghan J J and Price D J 2004 Toy stars in one dimension *Mon. Not. R. Astron. Soc.* **350** 1449–56
Morris J P 1996 Analysis of smoothed particle hydrodynamics with applications *PhD Thesis* Monash University, Melbourne, Australia
Morris J P, Fox P J and Zhu Y J 1997 Modelling low Reynolds number incompressible flows using SPH *J. Comput. Phys.* **136** 214–26
Morris J P and Monaghan J J 1997 A switch to reduce SPH viscosity *J. Comput. Phys.* **136** 41–50
Olson T S and Hiscock W A 1990 Stability, causality and hyperbolicity in Carter's 'regular' theory of relativistic heat conducting fluids *Phys. Rev.* D **41** 3687–95
Parshikov A N and Medin S A 2002 Smoothed particle hydrodynamics using interparticle contact algorithms *J. Comput. Phys.* **180** 358–82
Parzen E 1962 On estimations of a probability density and mode *Ann. Math. Stat.* **33** 1065–76
Phillips G J and Monaghan J J 1985 A numerical method for three-dimensional simulations of collapsing, isothermal, magnetic gas clouds *Mon. Not. R. Astron. Soc.* **216** 883–95
Price D J and Monaghan J J 2004a Smoothed particle magnetohydrodynamics: I. Algorithms and tests in one dimension *Mon. Not. R. Astron. Soc.* **348** 123–38

Price D J and Monaghan J J 2004b Smoothed particle magnetohydrodynamics: II. Variational principles and variable smoothing length terms *Mon. Not. R. Astron. Soc.* **348** 139–52

Price D J 2004a Magnetic fields in astrophysics *PhD Thesis* University of Cambridge, Cambridge, UK www.astro.ex.ac.uk/people/dprice

Price D J 2004b Private communication

Quirk J J 1994 A contribution to the great Riemann solver debate *Int. J. Numer. Methods Fluids* **18** 555–74

Randles P W and Libersky L 1996 Smoothed particle hydrodynamics some recent improvements and applications *Comput. Methods Appl. Mech. Eng.* **139** 375–408

Rosenblatt M 1956 Remarks on some nonparametric estimates of a density function *Ann. Math. Stat.* **27** 832–7

Rosswog S and Davies M B 2002 High resolution calculations of merging neutron stars-I. Model description and hydrodynamic evolution *Mon. Not. R. Astron. Soc.* **334** 481–97

Rosswog S, Davies M B, Thielemann F-K and Piran T 2000 Merging neutron stars: asymmetric systems *Astron. Astrophys.* **360** 171–84

Schoenberg I J 1946 Contributions to the problem of approximation of equidistant data by analytic functions: part A *Q. Appl. Math.* **IV** 45–99

Siegler S and Riffert H 2000 Smoothed particle hydrodynamics simulations of ultra-relativistic shocks with artificial viscosity *Astrophys. J.* **531** 1053–66

Sod G A 1978 A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws *J. Comput. Phys.* **27** 1–31

Springel V and Hernquist L 2002 Cosmological smoothed particle hydrodynamics simulations: the entropy equation *Mon. Not. R. Astron. Soc.* **333** 649–64

Steinmetz M and Mueller E 1993 On the capabilities and limits of smoothed particle hydrodynamics *Astron. Astrophys.* **268** 391–410

Swegle J, Hicks J and Attaway S 1995 Smoothed particle hydrodynamics stability analysis *J. Comput. Phys.* **116** 123–34

Takeda H, Miyama S and Sekiya M 1994 Numerical simulation of viscosity using SPH *Prog. Theor. Phys.* **92** 939–60

Von Neumann J 1944 Proposal and analysis of a new numerical method for the treatment of hydrodynamical shock problems *Von Neumann Collected Works* ed A Taub (Oxford: Pergamon)

Watkins S J, Bhattal A S, Francis N, Turner J A and Whitworth A P 1996 A new prescription for viscosity in smoothed particle hydrodynamics *Astron. Astrophys.* (Suppl.) **119** 177–87

Welton W 1998 Two-dimensional PDF/SPH simulations of compressible turbulent flows *J. Comput. Phys.* **139** 410–43

Welton W and Pope S 1997 PDF models of compressible turbulence using SPH *J. Comput. Phys.* **134** 150–68

Whitehouse S C and Bate M R 2004 Smoothed particle hydrodynamics with radiative transfer in the flux limited approximation *Mon. Not. R. Astron. Soc.* **353** 1078–94