<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Robustness and accuracy of SPH formulations for viscous flow</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Quinlan, Nathan J.; Lastiwka, Martin; Basa, Mihai</td>
</tr>
<tr>
<td><strong>Publication Date</strong></td>
<td>2009</td>
</tr>
<tr>
<td><strong>Link to publisher's version</strong></td>
<td><a href="http://dx.doi.org/10.1002/fld.1927">http://dx.doi.org/10.1002/fld.1927</a></td>
</tr>
<tr>
<td><strong>Item record</strong></td>
<td><a href="http://hdl.handle.net/10379/3832">http://hdl.handle.net/10379/3832</a></td>
</tr>
<tr>
<td><strong>DOI</strong></td>
<td><a href="http://dx.doi.org/10.1002/fld.1927">http://dx.doi.org/10.1002/fld.1927</a></td>
</tr>
</tbody>
</table>
Robustness and accuracy of SPH formulations for viscous flow

Mihai Basa, Nathan J. Quinlan*, Martin Lastiwka

Department of Mechanical and Biomedical Engineering, National University of Ireland, Galway

SUMMARY

Numerous methods are available for the modelling of viscous stress terms in Smoothed Particle Hydrodynamics (SPH). In this work, the existing methods are investigated systematically and evaluated for a range of Reynolds numbers using Poiseuille channel and lid-driven cavity test cases. The best results are obtained using two methods based on combinations of finite difference and SPH approximations, due to Morris et al. and Cleary. Gradients of high-valued functions are shown to be inaccurately estimated with standard SPH. A method that reduces the value of functions (in particular, pressure) before calculating gradients reduces this inaccuracy and is shown to improve performance. A mode of instability in Poiseuille channel flows, also reported in other works, is examined and a qualitative explanation is proposed. The choice of boundary implementation is shown to have a significant effect on transient velocity profiles in start-up of the flow. The use of at least linear extrapolation for in-wall velocities is shown to be preferable to mirroring of velocities. Consistency corrections to the kernel are also found to result in significant accuracy and stability improvements with most methods, although not in all cases. Copyright © 2008 John Wiley & Sons, Ltd.

KEY WORDS: mesh-free, instability, truncation error, boundary conditions, Poiseuille, Lid-driven cavity

1. INTRODUCTION

Smoothed Particle Hydrodynamics (SPH) [1] is a numerical technique that works with a model of fluid discretised into “particles” that move at fluid velocity, carrying values of mass, pressure, and all other fluid variables. Instead of using a mesh structure, information from local neighbourhoods of particles without permanent connectivity is used to interpolate values or calculate gradients of fluid variables. The mesh-free Lagrangian nature of SPH gives it unique advantages for modelling of astrophysical [2], multiphase [3] and free-surface flows [4]. In this paper we look at the accuracy of several existing approximations of viscous stress for Smoothed Particle Hydrodynamics, as well as different methods of imposing boundary conditions.

*Correspondence to: Nathan Quinlan, Department of Mechanical and Biomedical Engineering, National University of Ireland, Galway, IRELAND. Email: nathan.quinlan@nuigalway.ie

Contract/grant sponsor: Irish Research Council for Science, Engineering and Technology under the Embark Initiative, funded by the National Development Plan; contract/grant number: SC/2002/189

Copyright © 2008 John Wiley & Sons, Ltd.
Viscosity plays an essential role in many engineering fluid dynamics applications. Correct modelling of viscous phenomena is therefore a critical requirement of a general computational fluid dynamics method. We investigate methods of improving the models’ accuracy and robustness and reveal the ranges of Reynolds numbers at which they are usable. We also consider corrected forms of SPH and present a new approach to and a new analysis of corrections based on reduction of pressure.

In section 2, we review the basic SPH formulations used in this work. In section 3, we give a new analysis and formulation of corrections based on a reduction of the magnitude of pressure. Such corrections have previously been proposed, but not analysed (e.g. Morris et al. [5]). Section 4 reviews corrections which raise the order of consistency of the kernel and gradient operations of SPH. Section 5 contains definition and discussion of the various numerical models for viscous stress in SPH. Finally, in Section 6, numerical tests are presented to support the theoretical arguments of preceding sections and evaluate the performance of various formulations. A mechanism of instability in shear flow is observed and characterised. Section 6 includes a comprehensive exploration of the accuracy and robustness of SPH models of simple viscous flow for a range of Reynolds numbers, for the various formulations of viscous stress terms, consistency corrections and boundary conditions. We conclude with general recommendations for the modelling of viscous flow in SPH.

2. SPH FORMULATIONS

In SPH, to interpolate a function $F$ at a location $r_a$ from values at particles in the local neighbourhood, a form of weighted averaging is performed:

$$F_a \approx \sum_b F_b V_b W(r_b - r_a, h),$$  \hspace{1cm} (1)

where the index $b$ represents values at neighbour particles, and $V_b$ is a volume determined from particle $b$’s mass and density: $V_b = m_b / \rho_b$. The weighting kernel $W$ is usually a bell-shaped function with compact support that grants greater importance to particles closer to the point of interpolation, and satisfies certain conditions. The smoothing length $h$ characterises the width of the kernel’s compact support, playing a role similar to that of grid spacing in mesh-based methods. Particles within this compact support are the neighbours of $a$, including $a$ itself.

Similarly to interpolation, a partial derivative of the function $F$ along coordinate $x_i$, at $a$, can be obtained using:

$$\frac{\partial F}{\partial x_i} \bigg|_a \approx \left\langle \frac{\partial F}{\partial x_i} \right\rangle_a = \sum_b V_b F_b \frac{\partial W(r_a - r_b, h)}{\partial x_i}.$$  \hspace{1cm} (2)

We shall use the angle bracket notation $\langle \rangle_a$ as above, to denote that the quantity between the brackets is evaluated using an SPH approximation at the location of the “current” particle, $a$. In what follows we will write $W_{ab}$ as shorthand for $W(r_a - r_b, h)$. The notations $x_i$ and $x_j$ are used for coordinates, the presence of $i$ and $j$ indicating that the two coordinates can be different. The Einstein summation convention is not used in this work.
2.1. Governing equations

2.1.1. The momentum equation  Viscosity plays its most important role in the momentum equation of fluid flow, with a more minor role in the energy equation, and no direct role in mass conservation. The momentum equation in non-conservative form is written as:

\[
\rho \frac{D\mathbf{u}}{Dt} = -\nabla p + \nabla \cdot \tilde{\tau} + \mathbf{f}_e, \tag{3}
\]

where \( \mathbf{u} \) is flow velocity, \( p \) is pressure, \( \tilde{\tau} \) is the viscous stress tensor, and \( \mathbf{f}_e \) represents the contribution of external forces. For a fluid with Newtonian viscosity, the stress tensor is expressed as [6]:

\[
\tau_{ij} = \mu \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) - \lambda \nabla \cdot \mathbf{u} \delta_{ij}, \tag{4}
\]

where \( \lambda = \frac{2}{3} \mu \) represents the second viscosity coefficient.

In 2D, the viscous term can be written as:

\[
\nabla \cdot \tilde{\tau} = \left( \frac{\partial}{\partial x} \left[ \lambda \nabla \cdot \mathbf{u} + 2\mu \frac{\partial u_i}{\partial x_i} \right] + \frac{\partial}{\partial y} \left[ \mu \left( \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \right) \right] \right) \cdot \nabla W_{ab}. \tag{5}
\]

This form is useful for numerical implementations that use first derivatives exclusively.

Making the simplifying assumption that \( \mu \) does not vary in space, the 2D viscous term can also be written in the alternative form:

\[
\nabla \cdot \tilde{\tau} = \mu \left( \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 v}{\partial x \partial y} \right), \tag{6}
\]

in which second spatial derivatives of velocity components appear explicitly. In incompressible flow, since \( \nabla \cdot \mathbf{u} = 0 \), this further reduces to

\[
\nabla \cdot \tilde{\tau} = \mu \Delta \mathbf{u} = \mu (\nabla^2 u \mathbf{i} + \nabla^2 v \mathbf{j}). \tag{7}
\]

To achieve symmetry between particle interactions, and therefore conservation of linear momentum, one of the following symmetric expressions [1] for the momentum equation can be used, where \( \tilde{\sigma} = -\rho \mathbf{I} + \tilde{\tau} \) represents the total stress tensor:

\[
\frac{D\mathbf{u}_a}{Dt} = \sum_b m_b \left( \frac{\tilde{\sigma}_a}{\rho_a^2} + \frac{\tilde{\sigma}_b}{\rho_b^2} \right) \nabla W_{ab} + \mathbf{f}_e, \tag{8}
\]

or:

\[
\frac{D\mathbf{u}_a}{Dt} = \sum_b \frac{m_b}{\rho_a \rho_b} \left( \tilde{\sigma}_a + \tilde{\sigma}_b \right) \nabla W_{ab} + \mathbf{f}_e. \tag{9}
\]

2.1.2. Continuity equation  Two main choices exist in implementing a mass-conservation equation in SPH. The first choice is to use a true continuity equation based on the divergence of velocity, as in

\[
\frac{D\rho}{Dt} = -\rho \nabla \cdot \mathbf{u} \approx \sum_b m_b (\mathbf{u}_a - \mathbf{u}_b) \nabla W_{ab}. \tag{10}
\]

Copyright © 2008 John Wiley & Sons, Ltd.
Int. J. Numer. Meth. Fluids 2008; 00:1–24
Prepared using fidauth.cls
The second choice is to use the so-called summation form, an update of density based on local mass distribution:

\[ \rho \approx \langle \rho \rangle_a = \sum_b m_b W_{ab}. \tag{11} \]

Bonet and Lok [7] have shown, based on variational mechanics considerations, that there are links between the different forms of continuity and momentum equations. In order to maintain variational consistency, form (10) of continuity should be used together with form (9) of momentum, and similarly for forms (11) and (8). These combinations were used in this work as well.

Results obtained with the two combinations of equations were found to be very similar, and for simplicity only those obtained with the summation form are shown in this work.

2.1.3. Equation of state  We investigate two choices of equation of state, corresponding to a compressible and a weakly-compressible fluid. For compressible flows the equation of state for an ideal gas is used:

\[ p = \rho RT. \tag{12} \]

To approximate incompressible flow, we use a weakly-compressible formulation expressed through an equation of state introduced by Batchelor [8],

\[ p = \rho_0 c_0^2 \left[ \left( \frac{\rho}{\rho_0} \right)^\gamma - 1 \right], \tag{13} \]

with \( \gamma = 7 \), where \( c_0 \) is the speed of sound. This is a stiff equation of state, in the sense that small perturbations in density induce large changes in pressure.

2.1.4. Energy equation  In the implementation presented in this paper, the contribution of viscous dissipation to the energy equation is considered negligible at the low Mach numbers investigated, and is therefore ignored. When using the compressible fluid formulation (12), we use the following SPH energy equation in symmetrised form:

\[ \frac{De_a}{Dt} = \frac{1}{2} \sum_b m_b \left( \frac{p_a}{\rho_a^2} + \frac{p_b}{\rho_b^2} \right) (u_a - u_b) \nabla W_{ab}, \tag{14} \]

where \( e \) is internal energy.

When the weakly-compressible formulation (13) is used, internal energy and temperature become decoupled from other fluid properties, and therefore no equation of energy conservation is required.

3. PRESSURE REDUCTION

In an analysis of the standard 1D SPH derivative approximation, Quinlan et al. [9] expanded the truncation error series in powers of the smoothing length \( h \). The leading terms of the series are of the form

\[ \left\langle \frac{\partial F}{\partial x} \right\rangle_a - \frac{\partial F_a}{\partial x} = \frac{F_a}{h} \left[ \xi_0 \left( \frac{\Delta x}{h} \right)^3 + \frac{\xi^2}{2} O \left( \frac{\Delta x}{h} \right)^4 \right] + O \left( \frac{\Delta x}{h} \right)^5 + O \left( \frac{\Delta x}{h} \right)^{5+2} + \ldots \tag{15} \]
where $\Delta x$ is the mean one-dimensional particle volume, $\beta$ is a parameter related to kernel smoothness, and $\xi$ is the mean distance between a particle location and the centroid of its volume. For uniformly distributed particles, $\xi = 0$.

The first error term scales with the absolute value $F_a$ of the function whose gradient is to be evaluated, and grows with decreasing smoothing length. This $1/h$ effect has been verified in numerical experiments in both 1D and 3D. Thus, error in SPH gradient evaluations is sensitive to the function values, and this sensitivity is strongest when $h$ is small and particle distribution non-uniformity (quantified by $\xi$) is high.

This suggests that the accuracy of SPH discretisations could be improved if the absolute values of differentiated field variables (pressure and velocity) in governing equations can be reduced by some means. In fact, several widely used formulations of SPH employ various methods which effectively reduce function values, although the rationale for this has never previously been explained in terms of truncation error. These methods are sometimes introduced by the authors under other reasonings, such as to introduce density into the formulation, although forms not including density yet retaining the reduction process are reported to work as well or better [10]. For example, Monaghan [1] argued that gradients of velocity are more accurate when written in the following form, because it includes density:

$$\nabla \cdot \mathbf{u} = (\nabla \cdot (\rho \mathbf{u}) - \mathbf{u} \cdot \nabla \rho) / \rho ,$$

which, when discretised leads to

$$\nabla \cdot \mathbf{u}_a \approx \frac{1}{\rho_a} \sum_b m_b (\mathbf{u}_b - \mathbf{u}_a) \cdot \nabla W_{ab} .$$

Watkins et al. [10], however, report better results than those obtained with (17) are produced when replacing $\rho_a$ with $\rho_b$ in the above equation:

$$\nabla \cdot \mathbf{u}_a \approx \sum_b m_b \rho_b (\mathbf{u}_b - \mathbf{u}_a) \cdot \nabla W_{ab} .$$

This last expression, although not presented as such by Watkins et al., can be considered an SPH discretisation of the following identity, which does not include density:

$$\nabla \cdot \mathbf{u} = \nabla \cdot \mathbf{u} - \mathbf{u} \cdot \nabla 1 .$$

The $(\mathbf{u}_b - \mathbf{u}_a)$ term (which reduces the value of the function whose gradient is taken) is present in both resulting equations (17) and (18), although density is not. It may be that the truncation error term discussed above is partly responsible for the popularity of such value-reducing formulations.

Morris et al. [5] (and others, such as [11, 12]) also found that better results could be obtained for Poiseuille flow by subtracting hydrostatic pressure from absolute pressure in the momentum equation. This has the effect of replacing the pressure gradient (which is linear, in the example used by Morris et al.) with a uniform body force, but it also greatly reduces the absolute value of pressure, thus reducing the truncation error according to Equation (15).

The truncation error term has its most important impact in the evaluation of pressure gradients at low Mach number and low Reynolds number, where pressure gradients are relatively small and are balanced largely by viscous stresses. In these cases, absolute pressure divided by smoothing length ($p/h$, or $F_a/h$ in the notation of Equation (15)) is numerically
much larger than the physical pressure gradient. In compressible and shocked flows, by contrast, the physical pressure gradient $|\nabla p|$ is more significant in comparison with $p/h$. In typical SPH applications such as unsteady free-surface flows with wave-breaking and sloshing, inertial forces are more important. Therefore, among engineering fluid dynamics applications of SPH, low-speed viscous flows may be most vulnerable to the $F_a/h$ error in pressure gradient.

In order to reduce the effects of this truncation error term, we introduce a form different from that of Morris et al., in that we subtract a local average pressure based on neighbour particles, instead of a steady hydrostatic pressure field. The momentum equation becomes:

$$\frac{D\mathbf{u}_a}{Dt} = -\sum_b m_b \left( \frac{p_a - p_{\text{avg}}}{\rho_a^2} + \frac{p_b - p_{\text{avg}}}{\rho_b^2} \right) \nabla W_{ab} + \nabla \cdot \mathbf{\tilde{\tau}} + f_e,$$

(20)

with $p_{\text{avg}} = \sum_{b=1}^{N_b} p_b/N_b$, where $N_b$ is the number of neighbour particles. Because the constant $p_{\text{avg}}$ is subtracted only when (effectively) $\nabla p$ is computed for the momentum equation, it is a purely numerical modification. This is in contrast with XSPH [18], in which a similar modification is made to velocity, but the modified velocity is then used throughout the algorithm.

The localised pressure reduction in our method breaks the symmetry in $a$ and $b$ (as does the hydrostatic pressure method [5]). As a result, momentum is no longer exactly conserved, especially in situations where strong pressure gradients exist. However, the method brings pressure values closer to zero than a global method would.

Results comparing solutions obtained with and without pressure reduction are presented in section 6.1.1. We have found that pressure reduction can bring strong (and sometimes critical) improvements to robustness, with little computational effort, and have therefore used it in all other simulations presented here.

4. CONSISTENCY CORRECTIONS TO SPH OPERATIONS

The order of consistency of an operation represents the order of the polynomial function on which the operation evaluates exactly, when using values at discrete particle positions. It can be shown that for some operations traditional SPH is not zero-order consistent, even for regularly distributed particles [16]. Several correction methods exist which improve SPH’s order of consistency and ability to deal with disordered particle spacing. In this section we present two such correction methods which are used in the test cases presented in section 6.

4.1. Zero-order correction

Shepard interpolation can be used to improve SPH’s consistency properties [13, 7]. The method can be viewed as a re-normalisation of the standard SPH summation procedure:

$$\langle F_i \rangle_{a}^{[0]} = \frac{\sum_b F_b V_b W_{ab}}{\sum_b V_b W_{ab}} = \frac{\langle F \rangle_a}{\langle 1 \rangle_a},$$

(21)

where $\langle 1 \rangle_a$ denotes the SPH approximation of a uniform unit-valued field, i.e. $\sum_b V_b W_{ab}$. The superscript “$[0]$” denotes the order of consistency achieved with the corrected method.

The Shepard correction can extended to evaluations of first- and second-order derivatives by applying these to both sides of Equation (21). Using the shorthand notation defined at
the beginning of section 2, and for simplicity omitting the subscript \( a \) from all angle brackets, these can be written as:

\[
\begin{align*}
\left\langle \frac{\partial F}{\partial x_i} \right\rangle^{[0]} & = \frac{\partial}{\partial x_i} \left( \left\langle F \right\rangle^{(1)} \right) = \left( \left\langle \frac{\partial F}{\partial x_i} \right\rangle^{(1)} - \left\langle F \right\rangle \left\langle \frac{\partial}{\partial x_i} 1 \right\rangle \right) / (1)^2, \\
\left\langle \frac{\partial^2 F}{\partial x_i \partial x_j} \right\rangle^{[0]} & = \frac{\partial}{\partial x_i} \left( \frac{\partial}{\partial x_j} \left( \left\langle F \right\rangle^{(1)} \right) \right) = \\
& \left( \left\langle \frac{\partial^2 F}{\partial x_i \partial x_j} \right\rangle^{(1)} - \left\langle \frac{\partial F}{\partial x_j} \right\rangle - \left\langle \frac{\partial F}{\partial x_i} \right\rangle \left\langle \frac{\partial}{\partial x_j} 1 \right\rangle - \left\langle F \right\rangle \left\langle \frac{\partial^2}{\partial x_i \partial x_j} 1 \right\rangle \right) / (1)^2 + \\
& + 2\left\langle F \right\rangle \left\langle \frac{\partial}{\partial x_i} 1 \right\rangle \left\langle \frac{\partial}{\partial x_j} 1 \right\rangle / (1)^3.
\end{align*}
\]

(22)

(23)

4.2. First-order correction

The benefits of a first-order mixed kernel-and-gradient consistency correction, as developed by Bonet and Lok [7], are investigated in this work. The method consists of applying a corrected gradient operation to a Shepard-corrected kernel:

\[
\left\langle \frac{\partial F}{\partial x_i} \right\rangle^{[1]} = \sum_b V_b F_b \otimes \nabla^{[1]} W^{[0]}_{ab},
\]

(24)

with the corrected kernel being:

\[
W^{[0]}_{ab} = W_{ab} \langle 1 \rangle_a,
\]

(25)

and its corrected gradient:

\[
\nabla^{[1]} W^{[0]}_{ab} = L_a \nabla W^{[0]}_{ab}, \text{ where } L_a = \left( \sum_b V_b \nabla W^{[0]}_{ab} \otimes r_b \right)^{-1}.
\]

(26)

5. DISCRETISATIONS OF THE VISCOUS TERM IN SPH

The five viscosity approximations investigated in this work differ principally in the method used for computing second derivatives of velocity, which appear in the viscous term. The “direct derivatives” model (referred to as model A here) is based upon direct computation of second derivatives using a second-order kernel derivative in the SPH sums. This method has been found to be prone to problems such as instability and inaccuracy [11], which are sometimes alleviated by the use of higher-order kernels, such as Gaussian ones [14], or alternatively by re-meshing strategies [12]. The “two-pass” model, B, obtains second derivatives by nesting two first derivative operations, needing two passes over all particles (hence the name). This method has been found to produce good results, with the only drawback being increased computational cost [10, 15, 2]. Models C, D and E are based on a combination of finite-difference derivatives with standard SPH derivatives. This is a robust technique first used in artificial viscosity [17] and heat conduction [19] and then adapted for modelling physical viscous stresses [20, 5].
5.1. Direct derivatives: model A

An intuitive method of modelling viscous stresses is to write the viscous term in the form shown in equation (6), in which second derivatives appear explicitly. These are then approximated directly, using

\[ \left\langle \frac{\partial^2 F}{\partial x_i \partial x_j} \right\rangle_a = \sum_b V_b F_b \frac{\partial^2 W_{ba}}{\partial x_i \partial x_j}. \]  

(27)

After a change of coordinates in the derivatives, from Cartesian coordinates to the kernel’s native variable of distance \( r = \sqrt{x^2 + y^2 + z^2} \), the second derivative can be expressed as:

\[ \frac{\partial^2 W}{\partial x_i \partial x_j} = \frac{x_i x_j}{r^2} \left( \frac{\partial^2 W}{\partial r^2} - \frac{1}{r} \frac{\partial W}{\partial r} \right) + \delta_{ij} \frac{1}{r} \frac{\partial W}{\partial r}. \]  

(28)

Due to the presence of the Kronecker delta \( \delta_{ij} \), the kernel derivative takes two different forms depending on whether or not \( i \neq j \) (i.e. whether or not the coordinates in the second derivative are the same). A plot of the kernel in the two different cases is shown in Figure 1, for a 2D B-spline kernel.

A few conclusions regarding the order of consistency of direct second derivatives in special situations can be drawn from the different types of symmetries present in these two kernel shapes. We will show in the next paragraph that the \( i \neq j \) case has special consistency properties which do not hold for \( i = j \). These properties emerge in particle distributions in which a given particle’s neighbours are arranged in symmetric pairs across either of the coordinate planes passing through \( a \). This symmetry of particle distribution is a feature of Cartesian and staggered Cartesian distributions most often used to initialise SPH simulations.

The second derivative of a constant or a linear function is zero. Exact evaluation of a zero second derivative would be guaranteed if the contributions of all neighbour particles in the SPH sum (27) cancel out. The kernel for the \( i \neq j \) case exhibits antisymmetry across the coordinate planes passing through \( a \) (as shown in Figure 1-b). If neighbour particles are symmetrically distributed on either side of any of the two coordinate planes, they will contribute in equal magnitude but with opposite signs to the second derivative of \( F \) at \( a \), when \( F \) is constant or linear. When summing over all particles in a regular arrangement that is aligned with the
coordinate axes, the contributions from neighbours will therefore either cancel out two-by-two, or be zero (for particles on the symmetry planes).

In conclusion, in an appropriately ordered particle distribution (such as Cartesian, staggered grid, etc.) the SPH approximation to derivatives \( \partial^2 F / \partial x_i \partial x_j \) with \( i \neq j \) will exhibit zero-order consistency by virtue of the antisymmetry of the kernel. However, for the \( i = j \) case, the kernel exhibits symmetry across the coordinate planes, which does not lead to an intrinsic cancellation. In this last case there is no regular particle distribution for which the derivative (27) will be guaranteed to exhibit zero-order consistency.

Similar analyses can be made for other SPH operations. The first derivative, for example, uses a radially antisymmetric kernel. Therefore it exhibits zero-order consistency with regular particle distributions, even when these are not aligned with the coordinate axes. The interpolation operation, on the other hand, uses a symmetric kernel, and therefore has no consistency bonus.

This illustrates one reason why accuracy obtained with direct second derivatives (or the interpolation operation) can be poorer than that obtained with first derivatives, for regular particle distributions. It also suggests that methods that correct consistency are expected to be of particular benefit when direct second derivatives are used. This is indeed shown to be the case in section 6.

5.2. Two-pass derivatives: model B

The numerical viscosity model referred to as B in this work was introduced by Flebbe et al. [15] and Watkins et al. [10]. The method consists of nesting two first derivatives in order to obtain second derivatives. Derivatives of velocity are calculated and stored at particle locations in a first pass; derivatives of these are then calculated in a second pass, and used in the viscous term (5). Using the notation defined in section 2, with \( c \) used to denote neighbours of \( b \), this process can be written as:

\[
\left\langle \frac{\partial^2 F}{\partial x_i \partial x_j} \right\rangle^B_a = \left\langle \frac{\partial}{\partial x_i} \left\langle \frac{\partial F}{\partial x_j} \right\rangle_b \right\rangle_a = \sum_b V_b \left( \sum_c V_c F_c \frac{\partial W_{bc}}{\partial x_j} \right) \frac{\partial W_{ab}}{\partial x_i}. \tag{29}
\]

The two-pass method is a simpler approach than the direct method, as no new SPH operations have to be implemented, and any corrections for first derivatives can be used without modification.

Certain side-issues appear when using the two-pass method. Due to the two-pass process, information is effectively “gathered” from an area twice as large as a normal compact support. The second derivatives use first derivatives from particles up to one compact-support radius away – these in turn needed information up to another radius away from there, when the first pass was performed. This larger area of dependence affects boundary implementations that use ghost particles – boundaries need to be twice as thick as normal.

5.3. Mixed finite difference/SPH approximations: models C, D, E

We look at three different models which combine a finite-difference and an SPH derivative to express second derivatives. The models implement the following simplified viscous term, which is valid for incompressible flows (see section 2.1.1):

\[
\nabla \cdot \tilde{\tau} = \mu \Delta u = \mu (\nabla^2 u \textbf{i} + \nabla^2 v \textbf{j} + \nabla^2 w \textbf{k}). \tag{30}
\]
A method of computing second derivatives which is used in two of the three models was first employed for artificial viscosity by Monaghan and Gingold [17]. The artificial viscosity term was introduced to stabilise computation of flows with shocks, and was not meant to model a physical viscosity. It is given by:

\[
\Pi_{ab} = \left\{ \begin{array}{ll}
-\sum_b m_b \frac{-\alpha (c_a + c_b) \mu_{ab} + \beta \mu_{ab}^2}{0.5(\rho_a + \rho_b)} \nabla W_{ab} & \text{if } \mathbf{u}_{ab} \cdot \mathbf{r}_{ab} < 0 \\
0 & \text{if } \mathbf{u}_{ab} \cdot \mathbf{r}_{ab} > 0
\end{array} \right. ,
\]

where the \( \mu_{ab} \) term is a combination of a finite-difference first derivative with an SPH derivative of velocity:

\[
\mu_{ab} = h \frac{\mathbf{u}_{ab} \cdot \mathbf{r}_{ab}}{r_{ab}^2 + \eta^2}.
\]

Here \( \mathbf{u}_{ab} = u_a - u_b, \mathbf{r}_{ab} = r_a - r_b, \) and \( \eta = (0.1)h \) is a small term that keeps the denominator from approaching zero. The constants \( \alpha \) and \( \beta \) take different values depending on the type of problem solved (the values \( \alpha = 1 \) and \( \beta = 2 \) being common [21]). One of the benefits of this formulation is that it conserves linear and angular momentum, due to its symmetry in \( a \) and \( b \).

Model C was derived by Violeau et al. [4] from the work of Monaghan, and is given by

\[
\frac{1}{\rho} \frac{\mathbf{\nabla}}{\cdot} \tilde{\tau} \bigg|_a \approx \sum_b m_b \frac{8}{\rho_a + \rho_b} \left( \frac{\mu_a}{\rho_a} + \frac{\mu_b}{\rho_b} \right) \frac{\mathbf{u}_{ab} \cdot \mathbf{r}_{ab}}{r_{ab}^2 + \eta^2} \nabla W_{ab} ,
\]

Cleary [20] introduced the form (model D):

\[
\frac{1}{\rho} \frac{\mathbf{\nabla}}{\cdot} \tilde{\tau} \bigg|_a \approx \sum_b m_b \frac{\xi}{\rho_a \rho_b} \frac{4\mu_a \mu_b}{(\mu_a + \mu_b)} \frac{\mathbf{u}_{ab} \cdot \mathbf{r}_{ab}}{r_{ab}^2 + \eta^2} \nabla W_{ab} ,
\]

where the parameter \( \xi \) takes the recommended value of 4.96333, independently of \( \mu \), as obtained through calibration against known solutions in a Couette flow. In Poiseuille test runs performed by the current authors, however, this value of \( \xi \) was found to give highly inaccurate velocities. The value 4.24 was found to provide much better results. The lack of an analytically derived value for \( \xi \) is troubling, and the 17% discrepancy between the calibrated values for the simplest cases raises serious doubts about its universality. This is a serious disadvantage of model D.

A third model based on a different approach is that of Morris, Fox and Zhu [5] (model E):

\[
\frac{1}{\rho} \frac{\mathbf{\nabla}}{\cdot} \tilde{\tau} \bigg|_a \approx \sum_b m_b \frac{(\mu_a + \mu_b)}{\rho_a \rho_b} \frac{\mathbf{r}_{ab} \cdot \nabla W_{ab}}{r_{ab}^2} \mathbf{u}_{ab} .
\]

A significant difference between this and the previous two models is that the viscous force is aligned with the relative velocity vector \( \mathbf{u}_{ab} \) instead of radially as given by the kernel gradient. This formulation therefore conserves linear momentum exactly, while preserving angular momentum only approximately, since the inter-particle viscous forces are equal and opposite, but not collinear [5].

Graham and Hughes [22] analysed the effect of the number of neighbours per particle on the accuracy and convergence of model E in a Poiseuille channel. It was shown that convergence of the viscous stress approximation with increasing particle numbers is not always achieved.
6. TESTS

Two test cases are used to investigate the accuracy and robustness characteristics of the five viscous term approximations: a Poiseuille channel and a lid-driven cavity. Varying levels of consistency corrections are applied. Results are also compared for the compressible and weakly compressible formulations given by Equations (12) and (13) respectively.

Time stepping in these runs was performed with a second-order Runge-Kutta method. Smoothing lengths were set such that each particle initially had approximately 30 neighbours, corresponding to a $\Delta x/h = 0.637$, where $\Delta x$ is the inter-particle spacing and $h$ the smoothing length. The kernel used was the B-spline kernel of Monaghan and Lattanzio [23].

6.1. Test case 1: Poiseuille channel

The Poiseuille channel test consists of 2D viscous flow through a straight long channel under the influence of an external uniform body force per unit mass, $f_e$, acting in the direction of flow. If flow starts from a zero-velocity initial condition, the unsteady velocity profile in nondimensional form is given by the series solution

$$\frac{U(y, t)}{U_0} = 1 - \frac{32}{\pi} \left( \frac{y}{d} \right)^2 \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)^3} \cos \left( \frac{(2n+1)\pi y}{2d} \right) \exp \left( -\frac{(2n+1)^2\pi^2\mu t}{4d^2\rho} \right),$$

where $y$ is the channel’s transverse coordinate, with the origin placed at the centreline, $2d$ is the channel width, and $U_0$ is the steady-state centreline velocity. The steady-state velocity profile, $U(y, \infty)$, is parabolic, defined by

$$U(y, \infty) = \frac{|f_e| \rho}{2\mu} (d^2 - y^2).$$

In the initial state, particles are distributed on a Cartesian grid with 20 particles across the width and length of the channel, and zero velocity. In dimensional terms, all runs were initialised with $\rho \approx 1$ kg/m$^3$, $d = 0.5$ m, and $\mu = 1.0$ Pa·s. $|f_e|$ was adjusted to set the Reynolds number (based on $U_0$ and $d$) from approximately 0.1 to 10,000 as required. Initial pressure was 125 kPa for the compressible case, and 0 for the weakly compressible case.

6.1.1. Pressure reduction results

The method of pressure reduction, described in section 3, was found to have a very significant effect on solution stability when the compressible formulation is used and no consistency corrections are applied, irrespective of viscosity approximation. If pressure reduction is not used, solutions develop critical instabilities. These initiate in regions where particle arrangement is disturbed most, due to fluid shear. An example is shown in Figure 2 for $Re = 0.3$, using the compressible equation of state (12).

In the weakly compressible formulation pressure values are usually much lower than in the compressible one, and therefore pressure-induced instabilities are not as pronounced.

6.1.2. Boundary conditions

The boundaries in the Poiseuille channel problem are the inlet, outlet and no-slip impermeable walls. Because the channel is assumed to be infinite in length, the inlet and outlet can be simulated using periodic boundaries. Walls are modelled using mirror particles: for each fluid particle within a compact support radius of the wall, a virtual
particle is placed at a symmetric position on the inside of the wall [24], with opposite normal velocity.

Several approaches for setting the axial velocity of the virtual wall particles were tested: mirroring of fluid particles, linear extrapolation, quadratic extrapolation based on particles closest to the wall, and a globalised quadratic extrapolation based on the width of the channel. The last method, proposed by Watkins et al. [10], requires some a priori knowledge of the global flowfield. A further alternative investigated with all methods was the use of particles on the wall itself. For every fluid particle within a certain distance of the wall, a zero-velocity particle was placed at the nearest point on the wall. This approach was tested in conjunction with each of the methods of setting axial velocities mentioned above, using viscosity model B and first-order consistency corrections in a compressible formulation.

The largest differences between results obtained with different no-slip boundary treatments are found at the start of the simulation. At that time, fluid velocity is lowest and small variations are significant. Once the flow is established the methods usually produce quite similar results. A comparison of results obtained at a short time after initiation of the simulation is presented in Figure 4.

The worst results at flow startup are obtained when using the axial velocity mirroring method. At the very start of the simulation, this method creates only a thin band of high shear rate near the wall (see Figure 3). This band can be sufficiently small relative to particle compact supports that particles “see” through it, into the mirrored zone behind it where viscous forces are inverted. This can erroneously reduce the viscous forces acting on particles near the wall. By contrast, the linear and quadratic extrapolation methods project a shear field into the wall that is constant or grows linearly (respectively). Consequently, fluid particles react more strongly (and realistically) to the presence of the wall with extrapolation methods.

For the Poiseuille channel results presented in the following sections, we have used linear extrapolation of velocities, with no particles on the edges of walls themselves.
6.1.3. Accuracy and robustness  In this section we wish to quantify the accuracy and robustness of the five viscous stress models. The results presented span an array of variations on the SPH method, with two main variables: choice of equation of state and consistency correction level. Two equations of state are tried, corresponding to ideal gas and weakly compressible formulations (as detailed in section 2.1.3). Three correction levels are investigated for both formulations: no correction, zero-order and first-order (as detailed in section 4), for a total of six test configurations.

Each of the six cases have also been run using the two forms of the continuity equation and corresponding momentum equation formulations presented in section 2.1.2. However, differences between the two forms were found to be negligible, and therefore only the summation-form results (which are marginally better) are shown.

In order to quantify the effect of Reynolds number on flow robustness and accuracy, each of the six cases is run for a range of Reynolds numbers from 0.1 to 10,000. The RMS error is
defined as:

\[
\varepsilon(t) = \sqrt{\frac{1}{N} \sum_{a=1}^{N} \left( \frac{u_a(t) - U(y_a, t)}{U(y_a, t)} \right)^2},
\]

(38)

where \( N \) is the total number of fluid particles, \( u_a(t) \) are particle velocities and \( U(y, t) \) are analytical velocities at particle locations, as defined in Equation (36). The simulation is considered to have failed when the RMS error \( \varepsilon \geq 12\% \). The simulation usually fails with a large and sharp increase in error, and therefore results are not very sensitive to this threshold value.

The results which follow are presented in terms of nondimensionalised time \( t^* \):

\[
t^* = t \frac{\pi^2 \mu}{4 \rho d^2}.
\]

(39)

We nondimensionalise time using the time constant of the exponential term for \( n = 0 \) in the series (36). This is the slowest-growing and largest term in the series, and therefore it gives a meaningful time scale for evolution of the flow to steady state.

The results are grouped into two graphs for each level of consistency correction. Plots on the left in Figure 5 quantify solution robustness as a function of Reynolds number, by indicating the nondimensionalised time \( t^* \) reached at simulation failure (higher values of \( t^* \) at failure indicate better long-term robustness). Regions to the left and below the curves in these graphs can be considered stable. Plots in the right column quantify accuracy in terms of RMS velocity error as a function of \( t^* \) at \( Re=0.5 \). Plots on the right are aligned such that the \( t^* \) axes of the left and right plots are the same.

The search for simulation failure was stopped when either the corresponding analytical solution reached practically steady flow, defined as \( U(0, t)/U_0 > 1 - 10^{-12} \), or the number of timesteps rose above approximately 500,000.

Model A is found to be unstable when only standard SPH operations are used, which is expected, as this model is often declared in the literature as being unstable without the use of Gaussian kernels [14] or remeshing procedures [12]. With the zero-order correction introduced in section 4.1, however, the direct second-derivative model produces more stable and accurate results. This is consistent with the theory presented in section 5.1 above. We believe this is the first time the significant benefits of zero-order corrections have been shown for the direct second derivatives model, A.

Model B provides a mixed result, showing the least robustness, although accuracy is very good. Like A, model B benefits significantly from the zero-order correction. The best results for both robustness and accuracy are obtained using the mixed finite-difference/SPH models D and E. Model C is almost as robust as D and E, but gives some of the least accurate results of all the models. The benefits of consistency correction methods are small for the mixed models, and in the case of model D, first-order corrections actually appear to slightly degrade the accuracy of the solution.

An important caveat regarding the presented results is that model D’s empirical coefficient \( \xi \) was hand-calibrated, a tuning process from which none of the other models benefited. The increase in error with model D when high-order corrections are applied could also suggest that the constant may have been over-fitted for the uncorrected case, effectively compensating for inaccuracies in the SPH operations in that context.
Figure 5. Plots on left: Nondimensionalised time $t^*$ at which the simulation fails, plotted against Reynolds number. Plots on right: evolution of error (horizontal axis) plotted against $t^*$, at $Re = 0.5$.

The vertical axis, $t^*$, is aligned in left and right plots.
6.1.4. A description of the instability

According to the results in the previous section, true long-term simulations of Poiseuille flow poses serious challenges at all but the lowest Reynolds numbers. Similar difficulties as those presented here have been reported by other authors [10, 11, 25], or in cases where such difficulties are not explicitly discussed, studies take into consideration only lower Reynolds numbers [5] at which the failure appears much later.

All the models studied appear to exhibit the same two-stage behaviour with respect to long-term robustness at the Reynolds numbers studied. The first stage consists of a period of stable and uniform flow, in which particles maintain movement in ordered rows along the channel. The second stage consists of agitated particle movement and strong pressure and density fluctuations which often lead to solution failure, sometimes very quickly. The transition to agitated flow is rapid, and begins near the walls, quickly spreading inwards to the rest of the fluid. For sufficiently low Reynolds numbers the first stage can last well into effective steady state by the time the second stage initiates.

Using dynamical systems theory, it has been shown [26] that Poiseuille flow is potentially unstable for \( \text{Re} \geq 2645 \), when flow can separate into a pattern of recirculation bubbles on the sides of the channel and a weaving central jet. The present long-term failure may sometimes briefly resemble this flow pattern. However, the numerical failure can appear at \( \text{Re} \) as low as 1 (see Figure 5), and it therefore it does not simulate the physical instability. This is also contrary to an explanation based on turbulence, made by Watkins et al. regarding an instability obtained at a Reynolds number of 20 [10].

The failure has been observed to be at least in part due to a sensitivity of particles to transverse displacements. The instability does not appear if the transverse coordinates of particles are artificially fixed, for example.

The initial Cartesian particle distribution and the predominance of longitudinal forces in the Poiseuille channel leads particles to flow in ordered lengthwise rows. Particle positions in these rows are potentially unstable. If, for example, a particle shifts out of its row and towards the wall, viscous forces will decrease its velocity and bring it towards the particle
immediately behind. As the two particles approach, the inter-particle pressure forces will push them further out of their row and into a “swinging” movement past each other (see Figure 6). Particles ahead and behind the illustrated pair will have a weaker and weaker balancing effect on them, as the two approach, due to increasing distance. Adjacent rows of particles are progressively displaced by the particles’ movement, leading to a highly disturbed flow with strong noise in density, pressure and velocities. This motion is nonphysical.

Robustness does not appear to significantly or consistently depend on the number of particles or timestep size. Close examination reveals that transverse particle motion consists of two distinct behaviours. The first is a transverse oscillation that takes place with all particles in the row moving together. The amplitude of this motion usually remains below a uniform maximum amplitude of 0.1-0.5% of inter-particle distance, without growing over time. The second component is a slow relative drift of particles out of their rows. This secondary movement is initially of much lower amplitude but has an exponential rate of growth, and is a true indicator for the failure. In Figure 7 a plot of total kinetic energy evolution with time is shown alongside a plot of the transverse coordinate drift (i.e. relative difference in transverse coordinates) between two adjacent particles in the same row. The kinetic energy reaches and maintains an almost constant value for a long time, which at first suggests a stable state has been reached. However, there is a continuous and accelerated increase in relative drift between the two particles. This begins at very small scales but eventually reaches the macroscopic scale, leading to solution failure.

6.2. Test case 2: Lid-driven cavity

The lid-driven cavity test case involves viscous flow inside a square cavity in which one of the walls slides at a constant speed [28, 29], and is defined as a SPHERIC benchmark case for SPH
Figure 8. Velocity vectors for the lid driven cavity, at Re=100 (left) and 1000 (right) using model E with zero-order consistency corrections.

The test case presents a number of challenges, such as a strong initial shear discontinuity, along with corner singularities. No analytical solutions exist, but high-resolution simulation results are available.

Walls are modelled using stationary particles, with tangential velocities extrapolated normally to the wall, using information gathered by the wall particles closest to the fluid. Density and pressure of wall particles are allowed to evolve according to the continuity and state equations. For this test case only the weakly-compressible formulation (with the summation form of the continuity equation) is shown. The test was performed at Reynolds numbers of 100 and 1000, using 50 × 50 fluid particles. The number of neighbours per particle is 30 (i.e. ∆x/h = 0.637), as in the Poiseuille channel test.

This is quite a challenging test for SPH at Re=1000, and all the stress term models underpredict the total kinetic energy. In Figure 9 total kinetic energy in the cavity is compared against reference FVM solutions, for the two Reynolds numbers and three levels of consistency correction. The reference solutions in this figure were obtained using the OpenFOAM 1.4.1 CFD package [27], using a PISO incompressible algorithm on a 200 × 200 uniform mesh for Re=100, and 300 × 300 for Re=1000. Time and energy are nondimensionalised with respect to the lid velocity, length of cavity sides and fluid density.

The lid-driven cavity is a more demanding test case than the Poiseuille channel, and some of the viscosity models exhibit different behaviours between the two tests. Model B, which produced good results in the Poiseuille channel test, has difficulties handling the lid-driven cavity. Although at Re = 100 model B appears to have have a better kinetic energy evolution, this is revealed to be due to non-physical particle agitation which leads to failure for Re = 1000 when corrections are applied. Model A is stable when consistency corrections are used, but is still among the poorer performing methods. Model E, which achieved excellent results in the first test, is somewhat less robust here, showing spikes in kinetic energy, especially in the uncorrected case. Model C, which achieved relatively poor accuracy in the Poiseuille channel is nearly as accurate as model D, which continues to produce the best results, by a very small
Figure 9. Kinetic energy evolution in the lid-driven cavity problem, for Re=100 (left plots) and Re=1000 (right plots), using a low 50 × 50 particle resolution.
When centreline velocities are compared against reference solutions (Figure 10), almost all of the viscosity models perform quite similarly at Re=100, with the only outlier being model B, which exhibits severe clumping and agitation of particles. Model A is robust when the 0-order correction is applied, although it has a noisier velocity field than the other methods, especially at the higher Reynolds number. The change in test problem does not appear to significantly affect the results of model D – the empirical constant $\xi$ determined for uncorrected Poiseuille flow at Re=0.5 seems adequate and produces good accuracy at Re=100. However, it should be noted that the lid-driven cavity problem is not as sensitive to Reynolds number as the Poiseuille test, and errors in the $\xi$ factor (which is effectively a scaling factor on the viscosity coefficient $\mu$, and therefore Re) would have a smaller effect.

Overall, the lid-driven cavity test confirms the general trends for accuracy observed in the Poiseuille channel test of section 6.1. Robustness for a given Reynolds number is found to be better, mostly, in the lid-driven cavity than in the Poiseuille channel, possibly due to the existence of more isotropic shear stresses and strains which do not allow a failure mechanism such as in the channel test to form. The differences in accuracy and robustness of the various models between the two test cases, however, suggest that comparative results are case-dependent. It should also be noted that the accuracy of all models would be influenced by the choice of ratio $\Delta x/h$, as well as the total number of particles $N$, parameters which are not analysed in this work.

7. CONCLUSIONS

Differences in accuracy and robustness between five different viscosity models have been quantified in a variety of test situations with the Poiseuille channel and lid-driven cavity problems. The best overall choices for robustness and accuracy, under the test conditions,
appear to be the finite-difference-based methods of Cleary [20], model D, or that of Morris et al. [5], model E. It should be noted, however, that model D benefited from a tuning process that is only possible for flows whose solution is known. This is a serious disadvantage for this model.

The influence of the ratio $\Delta x/h$ and the total number of particles on the relative performance of the models was not analysed and so the conclusions of the comparative testing may not be universally definitive. However, the choices for these parameters are representative values for the benchmark cases studied, giving a reasonable practical trade-off of accuracy and computational cost. The effect of these parameters on one SPH viscous stress formulation has previously been studied in detail by Graham and Hughes [22].

Previous researchers have reported success with the technique of subtracting a hydrostatic pressure from the pressure used in the momentum equation. However, the rationale for this correction has not been explained. Here, we have shown that this correction also acts to reduce a truncation error term which scales with the absolute value of pressure. We have introduced the subtraction of a local average pressure, instead of a global hydrostatic pressure distribution, to minimise absolute values of pressure. Similar methods of reducing function values before taking the gradient are found throughout SPH in implementations of the continuity equation and the two-pass method, for example, but have been incompletely explained in the past.

Consistency corrections have been shown to be beneficial for the long-term robustness and accuracy of viscous calculations. A zero-order consistent direct second-derivative method has been developed and implemented with encouragingly stable results for both the Poiseuille and lid-driven cavity tests. The use of zero-order correction for the direct derivative method allowed it to be used as a viable viscosity model. To the current authors’ knowledge this is the first time such a correction has been used successfully with the direct second derivative method. Significant improvement in accuracy with the use of consistency correction methods were also quantified for the other models studied.

An explanation has been presented for a long-term instability in the Poiseuille channel, based on observed sensitivity to lateral displacement of particles.

An analysis of various choices of boundary velocity imposition for no-slip walls shows that the use of mirroring of tangential velocities is not a good choice, and at least linear extrapolation of velocities should be used.

ACKNOWLEDGEMENTS

The authors would like to thank Patrick Dooley for his help in obtaining a reference FVM solution for the lid-driven cavity test case.

This work was supported by Basic Research Grant SC/2002/189 from the Irish Research Council for Science, Engineering and Technology under the Embark Initiative, funded by the National Development Plan.

REFERENCES


