Improving convergence in smoothed particle hydrodynamics simulations without pairing instability

Walter Dehnen* and Hossam Aly*

Department for Physics and Astronomy, University of Leicester, Leicester LE1 7RH

Accepted 2012 May 31. Received 2012 May 30; in original form 2012 April 11

ABSTRACT
The numerical convergence of smoothed particle hydrodynamics (SPH) can be severely restricted by random force errors induced by particle disorder, especially in shear flows, which are ubiquitous in astrophysics. The increase in the number $N_H$ of neighbours when switching to more extended smoothing kernels at fixed resolution (using an appropriate definition for the SPH resolution scale) is insufficient to combat these errors. Consequently, trading resolution for better convergence is necessary, but for traditional smoothing kernels this option is limited by the pairing (or clumping) instability. Therefore, we investigate the suitability of the Wendland functions as smoothing kernels and compare them with the traditional B-splines. Linear stability analysis in three dimensions and test simulations demonstrate that the Wendland kernels avoid the pairing instability for all $N_H$, despite having vanishing derivative at the origin (disproving traditional ideas about the origin of this instability; instead, we uncover a relation with the kernel Fourier transform and give an explanation in terms of the SPH density estimator). The Wendland kernels are computationally more convenient than the higher order B-splines, allowing large $N_H$ and hence better numerical convergence (note that computational costs rise sublinear with $N_H$). Our analysis also shows that at low $N_H$ the quartic spline kernel with $N_H \approx 60$ obtains much better convergence than the standard cubic spline.

Key words: hydrodynamics – methods: numerical.

1 INTRODUCTION
Smoothed particle hydrodynamics (SPH) is a particle-based numerical method, pioneered by Gingold & Monaghan (1977) and Lucy (1977), for solving the equations of hydrodynamics (recent reviews include Monaghan 2005, 2012; Rosswog 2009; Springel 2010; Price 2012). In SPH, the particles trace the flow and serve as interpolation points for their neighbours. This Lagrangian nature of SPH makes the method particularly useful for astrophysics, where typically open boundaries apply, though it becomes increasingly popular also in engineering (e.g. Monaghan 2012).

The core of SPH is the density estimator: the fluid density is estimated from the masses $m_i$ and positions $x_i$ of the particles via (the symbol $\approx$ denotes an SPH estimate)

$$\rho(x_i) \approx \hat{\rho}_i \equiv \sum_i m_j W(x_i-x_j,h_i), \quad (1)$$

where $W(x,h)$ is the smoothing kernel and $h_i$ the smoothing scale, which is adapted for each particle such that $h_i^3 \hat{\rho}_i =$constant (with $v$ the number of spatial dimensions). Similar estimates for the value of any field can be obtained, enabling discretization of the fluid equations. Instead, in conservative SPH, the equations of motion for the particles are derived, following Nelson & Papaloizou (1994), via a variational principle from the discretized Lagrangian:

$$L = \sum_i m_i \left[ \frac{1}{2} \dot{x}_i^2 - u(\rho_i,s_i) \right] \quad (2)$$

(Monaghan & Price 2001). Here, $u(\rho,s)$ is the internal energy as function of density and entropy $s$ (and possibly other gas properties), the precise functional form of which depends on the assumed equation of state. The Euler–Lagrange equations then yield

$$\dot{x}_i = \frac{1}{m_i} \frac{\partial L}{\partial \dot{x}_i} = \sum_j m_j \left[ \frac{\dot{\rho}_j}{\Omega_j \rho_j} \nabla W(x_{ij},h_i) + \frac{\dot{\rho}_j}{\Omega_j \rho_j} \nabla W(x_{ij},h_j) \right], \quad (3)$$

where $x_{ij} = x_i - x_j$ and $\dot{P}_j = \rho_j^2 \frac{\partial u}{\partial \rho_j}$, while the factors

$$\Omega_i = \frac{1}{v h_i^2 \rho_i} \frac{\partial (h_i^3 \hat{\rho}_i)}{\partial \ln h_i} \approx 1 \quad (4)$$

(Monaghan 2002; Springel & Hernquist 2002) arise from the adaptation of $h_i$ (Nelson & Papaloizou) such that $h_i^3 \hat{\rho}_i =$constant.

Equation (3) is a discretization of $\rho \dot{x} = -\nabla P$, and, because of its derivation from a variational principle, conserves mass, linear and angular momentum, energy, entropy and (approximately) circulation. However, its derivation from the Lagrangian is only valid if all fluid variables are smoothly variable. To ensure this, in particular for velocity and entropy, artificial dissipation terms have to be...
added to $\dot{x}_i$ and $\dot{u}_i$. Recent progress in restricting such dissipation
to regions of compressive flow (Cullen & Dehnen 2010; Read &
Hayfield 2012) has greatly improved the ability to model contact
continuities and their instabilities as well as near-inviscid flows.

SPH is not a Monte Carlo method, since the particles are not
randomly distributed, but typically follow a semiregular glass-like
distribution. Therefore, the density (and pressure) error is much
smaller than the $\gtrsim 15$ per cent expected from Poisson noise for $\sim 40$
neighbours and SPH obtains $O(h^2)$ convergence. However, some
level of particle disorder cannot be prevented, in particular in shear-
ing flows (as in turbulence), where the particles are constantly re-
arranged (even in the absence of any forces), but also after a shock,
where an initially isotropic particle distribution is squashed along
one direction to become anisotropic.

In such situations, the SPH force (3) in addition to the pressure gradient contains a random ‘$E_0$
error’ (Read, Hayfield & Agertz 2010), and SPH converges more
slowly than $O(h^2)$. Since shocks and shear flows are common in star-
and galaxy formation, the ‘$E_0$ errors’ may easily dominate the
overall performance of astrophysical simulations.

One can dodge the ‘$E_0$ error’ by using other discretizations of
$\rho \dot{x} = -\nabla \rho \rho$ (Morris 1996; Abel 2011). However, such approaches
unavoidably abandon momentum conservation and hence fail in
practice, in particular, for strong shocks (Morris 1996). Further-
more, with such modifications SPH no longer maintains particle
order, which it otherwise automatically achieves. Thus, the ‘$E_0$
error’ is SPH’s attempt to resurrect particle order (Price 2012), and
SPH converges more slowly than $O(h^2)$. Since shocks and shear flows are common in
star- and galaxy formation, the ‘$E_0$ errors’ may easily dominate the
overall performance of astrophysical simulations.

One can dodge the ‘$E_0$ error’ by using other discretizations of
$\rho \dot{x} = -\nabla \rho \rho$ (Morris 1996; Abel 2011). However, such approaches
unavoidably abandon momentum conservation and hence fail in
practice, in particular, for strong shocks (Morris 1996). Further-
more, with such modifications SPH no longer maintains particle
order, which it otherwise automatically achieves. Thus, the ‘$E_0$
error’ is SPH’s attempt to resurrect particle order (Price 2012), and
SPH converges more slowly than $O(h^2)$. Since shocks and shear flows are common in
star- and galaxy formation, the ‘$E_0$ errors’ may easily dominate the
overall performance of astrophysical simulations.

One can dodge the ‘$E_0$ error’ by using other discretizations of
$\rho \dot{x} = -\nabla \rho \rho$ (Morris 1996; Abel 2011). However, such approaches
unavoidably abandon momentum conservation and hence fail in
practice, in particular, for strong shocks (Morris 1996). Further-
more, with such modifications SPH no longer maintains particle
order, which it otherwise automatically achieves. Thus, the ‘$E_0$
error’ is SPH’s attempt to resurrect particle order (Price 2012), and
SPH converges more slowly than $O(h^2)$. Since shocks and shear flows are common in
star- and galaxy formation, the ‘$E_0$ errors’ may easily dominate the
overall performance of astrophysical simulations.

One can dodge the ‘$E_0$ error’ by using other discretizations of
$\rho \dot{x} = -\nabla \rho \rho$ (Morris 1996; Abel 2011). However, such approaches
unavoidably abandon momentum conservation and hence fail in
practice, in particular, for strong shocks (Morris 1996). Further-
more, with such modifications SPH no longer maintains particle
order, which it otherwise automatically achieves. Thus, the ‘$E_0$
error’ is SPH’s attempt to resurrect particle order (Price 2012), and
SPH converges more slowly than $O(h^2)$. Since shocks and shear flows are common in
star- and galaxy formation, the ‘$E_0$ errors’ may easily dominate the
overall performance of astrophysical simulations.

Another possibility to reduce the ‘$E_0$ error’ is to subtract an
average pressure from each particle’s $\dot{P}$, in equation (3). Effectively,
this amounts to adding a negative pressure term, which can cause the
tensile instability (see Section 3.1.2). Moreover, this trick is only useful in situations with little pressure variations, perhaps in
simulations of near-incompressible flows (e.g. Monaghan 2011).

The only remaining option for reducing the ‘$E_0$ error’ appears an
increase of the number $N_{\text{eff}}$ of particles contributing to the density
and force estimates (contrary to naive expectation, the computa-
tional costs grow sublinear with $N_{\text{eff}}$). The traditional way to try
to do this is by switching to a smoother and more extended ker-
nel, enabling larger $N_{\text{eff}}$ at the same smoothing scale $h$ (e.g. Price 2012).
However, the degree to which this approach can reduce the
‘$E_0$ errors’ is limited and often insufficient, even with an infinitely
extended kernel, such as the Gaussian. Therefore, one must also
consider ‘stretching’ the smoothing kernel by increasing $h$. This
inevitably reduces the resolution, but that is still much better than
obtaining erroneous results. Of course, the best balance between
reducing the ‘$E_0$ error’ and resolution should be guided by results
for relevant test problems and by convergence studies.

Unfortunately, at large $N_{\text{eff}}$ the standard SPH smoothing kernels
become unstable to the pairing (or clumping) instability (a cousin of the
tensile instability), when particles form close pairs reducing the
effective neighbour number. The pairing instability (first mentioned
by Schüller & Schmitt 1981) has traditionally been attributed to
the diminution of the repulsive force between close neighbours
approaching each other (Schüller & Schmitt 1981; Thomas &
Couchman 1992; Herant 1994; Swegle, Hicks & Attaway 1995;
Springel 2010; Price 2012). Such a diminishing near-neighbour
force occurs for all kernels with an inflection point, a necessary
property of continuously differentiable kernels. Kernels without

1 Strictly speaking, the ‘$E_0$ error’ term of Read et al. is only the dominant
contribution to the force errors induced by particle discreteness.

SPH without pairing instability 1069

© 2012 The Authors, MNRAS 425, 1068–1082
Monthly Notices of the Royal Astronomical Society © 2012 RAS
2.2 Smoothing kernels

After these definitions, let us list the desirable properties of the smoothing kernel (cf. Fulk & Quinn 1996; Price 2012):

(i) equation (1) obtains an accurate density estimate;
(ii) \( W(x, h) \) is twice continuously differentiable;
(iii) SPH is stable against pairing at the desired \( N_H \);
(iv) \( W(x, h) \) and \( \nabla W(x, h) \) are computationally inexpensive.

Here, condition (i) implies that \( W(x, h) \to \delta(x) \) as \( h \to 0 \) but also that \( W(x, h) \geq 0 \) is monotonically declining with \( |x| \); condition (ii) guarantees smooth forces, but also implies \( \nabla W(0) = 0 \).

2.2.1 B-splines

The most used SPH kernel functions are the Schoenberg (1946) B-spline functions, generated as 1D Fourier transforms² (Monaghan & Lattanzio 1985):

\[
w(r) = C b_n(r), \quad b_n(r) \equiv \frac{\Gamma(\infty)}{\sqrt{2\pi}} \int_0^\infty \frac{\sin k/n}{k/n} \cos kr \, dk, \tag{11}
\]

with normalization constant \( C \). These kernels consist of \( n/2 \) piecewise polynomials of degree \( n - 1 \) (see Table 1) and are \( n - 2 \) times continuously differentiable. Thus, the cubic spline \( n = 4 \) is the first useful, but the quartic and quintic have also been used. For large \( n \), the B-splines approach the Gaussian: \( b_n \to N(0, H^2/3n) \) (this follows from footnote 2 and the central limit theorem).

Following Monaghan & Lattanzio, \( \hat{h} \equiv 2H/n \) is conventionally used as smoothing scale for the B-splines independent of \( v \). This is motivated by their original purpose to interpolate equidistant one-dimensional data with spacing \( h \), but cannot be expressed via a functional \( h \equiv h[|W(x)|] \). Moreover, the resulting ratios between \( \hat{h} \) for the \( b_n \) do not match any of the definitions discussed above.³

² By this definition they are the \( n \)-fold convolution (in one dimension) of \( b_1(r) \) with itself (modulo a scaling), and hence are identical to the Irwin (1927)–Hall (1927) probability density for the sum \( r \) of \( n \) independent random variables, each uniformly distributed between \(-1/n \) and \( 1/n \).

³ Fig. 2 of Price (2012) seems to suggest that with this scaling the B-splines approach the Gaussian with \( \sigma = \hat{h}/\sqrt{2} \). However, this is just a coincidence for \( n = 6 \) (quintic spline) since \( \sigma = \sqrt{\pi/12}\hat{h} \) for the B-splines in 1D.

Instead, we use the more appropriate \( h = 2\sigma \) also for the B-spline kernels, giving \( H \approx 1.826 h \) for the cubic spline in 3D, close to the conventional \( H = 2\hat{h} \) (see Table 1).

2.2.2 ‘Triangular’ kernels

At low-order \( n \) the B-splines are only stable against pairing for modest values of \( N_H \) (we will be more precise in Section 3), while at higher \( n \) they are computationally increasingly complex.

Therefore, alternative kernel functions which are stable for large \( N_H \) are desirable. As the pairing instability has traditionally been associated with the presence of an inflection point (minimum of \( w^\prime \)), functions \( w(r) \) without inflection point have been proposed. These have a triangular shape at \( r \approx 0 \) and necessarily violate point (ii) of our list, but avoid the pairing instability.⁴ For comparison we consider one of them, the ‘HOCT4’ kernel of Read et al. (2010).

2.2.3 Wendland functions

The linear stability analysis of the SPH algorithm, presented in the next section, shows that a necessary condition for stability against pairing is the non-negativity of the multidimensional Fourier transform of the kernel. The Gaussian has non-negative Fourier transform for any dimensionality and hence would give an ideal kernel were it not for its infinite support and computational costs.

Therefore, we look for kernel functions of compact support which have non-negative Fourier transform in \( v \) dimensions and are low-order polynomials⁵ in \( r \). This is precisely the defining property of the Wendland (1995) functions, which are given by

\[
w(r) = C \psi_{\ell k}(r), \quad \psi_{\ell k}(r) \equiv I^2 \left( 1 - r \right)_{+}^{\ell}, \tag{12}
\]

⁴ Thomas & Couchman (1992) proposed such kernels only for the force equation (3), but to keep a smooth kernel for the density estimate. However, such an approach cannot be derived from a Lagrangian and hence necessarily violates energy and/or entropy conservation (Price 2012).

⁵ Polynomials in \( r^2 \) would avoid the computation of a square root. However, it appears that such functions cannot possibly have non-negative Fourier transform (Wendland, private communication).
Figure 1. Kernels of Table 1, the Gaussian and the HOCT4 kernel of Read et al. (2010) scaled to a common resolution scale of $h = 2\sigma$ for 3D (top: linear plot, arrows indicating $|x| = H$; bottom: logarithmic plot).

with $(\cdot)_+=\max\{0,\cdot\}$ and the linear operator

$$I[f](r) = \int_0^\infty sf(x)dx.$$  

(13)

In $v$ spatial dimensions, the functions $\psi_{\ell k}(|x|)$ with $\ell = k + 1 + \lfloor v/2 \rfloor$ have positive Fourier transform and are $2k$ times continuously differentiable. In fact, they are the unique polynomials in $|x|$ of minimal degree with these properties (Wendland 1995, 2005). For large $k$, they approach the Gaussian, which is the only non-trivial eigenfunction of the operator $I$. We list the first few Wendland functions for one, two and three dimensions in Table 1, and plot them for $v = 3$ in Fig. 1.

2.3 Kernel comparison

Fig. 1 plots the kernel functions $w(r)$ of Table 1, the Gaussian and the HOCT4 kernel, all scaled to the same $h = 2\sigma$ for $v = 3$. Amongst the various scalings (ratios for $h/H$) discussed in Section 2.1 above, this gives by far the best match between the kernels. The B-splines and Wendland functions approach the Gaussian with increasing order. The most obvious difference between them in this scaling is their central value. The B-splines, in particular of lower order, put less emphasis on small $r$ than the Wendland functions or the Gaussian.

Obviously, the HOCT4 kernel, which has no inflection point, differs significantly from all the others and puts even more emphasis on the centre than the Gaussian (for this kernel $\sigma \approx 0.228343H$).

2.4 Kernel Fourier transforms

For spherical kernels of the form (6), their Fourier transform only depends on the product $H|k|$, i.e. $\hat{W}(k) = \hat{w}(H|k|)$. In 3D ($F_v$ denotes the Fourier transform in $v$ dimensions)

$$\hat{w}(k) = F_v[w(r)](k) = 4\pi k^{-1} \int_0^\infty \sin(kr) w(r) r dr$$  

(14)

which is an even function and (up to a normalization constant) equals $-\kappa^{-1}df_3[w]/dk$. For the B-splines, which are defined via their 1D Fourier transform in equation (11), this gives immediately

$$F_v[B_{\nu}(r)](k) = 3 \left( \frac{n}{\kappa} \right)^{n+2} \sin^2 \frac{\kappa}{n} \left( 1 - \frac{x}{n} \cot \frac{\kappa}{n} \right)$$  

(15)

Figure 2. Fourier transforms $\hat{W}(k)$ for the Gaussian, the HOCT4 and the kernels of Table 1 scaled to the same common scale $h = 2\sigma$. Negative values are plotted with broken curves.

(which includes the normalization constant), while for the 3D Wendland kernels

$$F_3[\psi_{\ell k}(r)](k) = \left( -\frac{1}{\kappa} \frac{d}{dk} \right)^{\ell+1} F_3[(1 - r)^{\nu}](k)$$  

(16)

we abstain from giving individual functional forms).

All these are plotted in Fig. 2 after scaling them to a common $h = 2\sigma$. Notably, all the B-spline kernels obtain $\hat{W} < 0$ and oscillate about zero for large $|k|$ (which can also be verified directly from equation 15), whereas the Wendland kernels have $\hat{W}(k) > 0$ at all $k$, as does the HOCT4 kernel. As non-negativity of the Fourier transform is necessary (but not sufficient) for stability against pairing at large $N_H$ (see Section 3.1.2), in 3D the B-splines (of any order) fall prey to this instability for sufficiently large $N_H$, while, based solely on their Fourier transforms, the Wendland and HOCT4 kernels may well be stable for all neighbour numbers.

At large $|k|$ (small scales), the HOCT kernel has most power, caused by its central spike, while the other kernels have ever less small-scale power with increasing order, becoming ever smoother and approaching the Gaussian, which has least small-scale power.

The scaling to a common $h = 2\sigma$ in Fig. 2 has the effect that $\hat{W}(k)$ all overlap at small wave numbers, since their Taylor series

$$\hat{W}(k) = 1 - \frac{1}{2}\sigma^2 k^2 + O(|k|^4).$$  

(17)

2.5 Density estimation and correction

The SPH force (3) is inseparably related, owing to its derivation via a variational principle, to the derivative of the density estimate. Another important role of the SPH density estimator is to obtain accurate values for $\hat{P}$ in equation (3), and we will now assess the performance of the various kernels in this latter respect.

In Fig. 3, we plot the estimated density (1) versus neighbour number $N_H$ for the kernels of Table 1 and particles distributed in three-dimensional densest sphere packing (solid curves) or a glass (squares). While the standard cubic spline kernel underestimates the density (only values $N_H \lesssim 55$ are accessible for this kernel owing to the pairing instability), the Wendland kernels (and Gaussian, not shown) tend to overestimate it.

It is worthwhile to ponder about the origin of this density overestimation. If the particles were randomly rather than semiregularly
distributed, ñ obtained for an unoccupied position would be unbiased (e.g. Silverman 1986), while at a particle position the self-contribution \( mW(0, h) \) to \( ñ \) results in an overestimate. Of course, in SPH and in Fig. 3 particles are not randomly distributed, but at small \( N_h \) the self-contribution still induces some bias, as evident from the overestimation for all kernels at very small \( N_h \).

The HOCT4 kernel of Read et al. (2010, orange) with its central spike (cf. Fig. 1) shows by far the worst performance. However, this is not a peculiarity of the HOCT4 kernel, but a generic property of all kernels without inflection point.

These considerations suggest the corrected density estimate

\[
\hat{n}_{\text{corr}} = \hat{n} - \epsilon m W(0, h),
\]

which is simply the original estimate (1) with a fraction \( \epsilon \) of the self-contribution subtracted. The equations of motion obtained by replacing \( \hat{n} \) in the Lagrangian (2) with \( \hat{n}_{\text{corr}} \) are otherwise identical to equations (3) and (4) (note that \( \partial (\hat{n}_{\text{corr}})/\partial \ln h = \partial \hat{n}/\partial \ln h \), since \( \hat{n}_{\text{corr}} \) and \( \hat{n} \) differ only by a constant), in particular the conservation properties are unaffected. From the data of Fig. 3, we find that good results are obtained by a simple power law

\[
\epsilon = \epsilon_{100} \left( N_h/100 \right)^{-\alpha},
\]

with constants \( \epsilon_{100} \) and \( \alpha \) depending on the kernel. We use \( \epsilon_{100}, \alpha = (0.0294, 0.977), (0.01342, 1.579) \) and \( 0.0116, 2.236 \), respectively, for the Wendland kernels in 3 dimensions.

The dashed curves and triangles in Fig. 3 demonstrate that this approach obtains accurate density and hence pressure estimates.

### 3 Linear Stability and Sound Waves

The SPH linear stability analysis considers a plane-wave perturbation to an equilibrium configuration, i.e. the positions are perturbed according to

\[
x_i \rightarrow x_i + a \exp(i(k \cdot x_i - \omega t)),
\]

with displacement amplitude \( a \), wave vector \( k \) and angular frequency \( \omega \). Equating the forces generated by the perturbation to linear order in \( a \) to the acceleration of the perturbation yields a dispersion relation of the form

\[
a \cdot P(k) = \omega^2 a.
\]

This is an eigenvalue problem for the matrix \( P \) with eigenvector \( a \) and eigenvalue \( \omega^2 \). The exact (non-SPH) dispersion relation (with \( c^2 = \partial P/\partial \rho \), \( P = \rho^2 \partial u/\partial \rho \) at constant entropy)

\[
c^2 a \cdot k k = \omega^2 a
\]

has only one non-zero eigenvalue \( \omega^2 = c^2 k^2 \) with eigenvector \( a \parallel k \), corresponding to longitudinal sound waves propagating at speed \( c \).

The actual matrix \( P \) in equation (21) depends on the details of the SPH algorithm. For conservative SPH with equation of motion (3), Monaghan (2005) gives it for \( P \propto \rho^4 \) in one spatial dimension. We derive it in Appendix A for a general equation of state and any number \( v \) of spatial dimensions:

\[
P = c^2 u^{(2)} + \frac{2}{\rho} \left( U - u^{(2)} \right) + \frac{1}{2v \rho^2 \Omega^2 \ln \h} \partial \Omega^{(2)}
\]

where \( x^{(2)} \) is the outer product of a vector with itself, bars denote SPH estimates for the unperturbed equilibrium, \( t = \rho \Omega u \), and

\[
u(k) = \frac{1}{\rho(\Omega)} \sum_j m \sin h \cdot \hat{x}_j \nabla W(\hat{x}_j, \h),
\]

\[
U(k) = \frac{1}{\rho(\Omega)} \sum_j m (1 - \cos h \cdot \hat{x}_j) \nabla^{(2)} W(\hat{x}_j, \h),
\]

\[
\Xi = \frac{1}{v} \partial \ln \rho \hat{\Omega} \approx 0.
\]

Here and in the remainder of this section, curly brackets indicate terms not present in the case of a constant \( h = \bar{h} \), when our results reduce to relations given by Morris (1996) and Read et al. (2010).

Since \( P \) is real and symmetric, its eigenvalues are real and its eigenvectors mutually orthogonal. The SPH dispersion relation (21) can deviate from the true relation (22) in mainly two ways. First, the longitudinal eigenvalue \( \omega_{1,2}^2 \) (with eigenvector \( a \parallel k \)) may deviate from \( c^2 k^2 \) (wrong sound speed) or even be negative (pairing instability; Morris 1996; Monaghan 2000). Secondly, the other two eigenvalues \( \omega_{1,2}^2 \) may be significantly non-zero (transverse instability for \( \omega_{1,2}^2 < 0 \) or transverse sound waves for \( \omega_{1,2}^2 > 0 \)).

The matrix \( P \) in equation (23) is not accessible to simple interpretation. We will compute its eigenvalues for the various SPH kernels in Sections 3.2 and 3.3 and Figs 4–6, but first consider the limiting cases of the dispersion relation, allowing some analytic insight.

### 3.1 Limiting cases

There are three spatial scales: the wavelength \( \lambda = 2\pi/|k| \), the smoothing scale \( h \) and the nearest neighbour distance \( d_{nn} \). We will separately consider the limit \( \lambda \gg h \) of well-resolved waves, the continuum limit \( h \gg d_{nn} \) of large neighbour numbers and finally the combined limit \( \lambda \gg h \gg d_{nn} \).

---

6 In equation (3) one omits the factors \( \Omega \), but still adapts \( h \) to obtain \( h' \hat{n} = \) constant, as some practitioners do, then the resulting dispersion relation has an asymmetric matrix \( P \) with potentially complex eigenvalues.
3.1.1 Resolved waves

If |k|h ≪ 1, the argument of the trigonometric functions in equations (24a,b) is always small and we can Taylor expand them. If we also assume a locally isotropic particle distribution, this gives to lowest order in |k| (I is the unit matrix; see also Section A3):

\[
P \rightarrow c^2 k^2 (I + \frac{\bar{\rho}}{\rho} \left[ \left( \frac{2\nu}{v + 2} - 1 \right) k^2 + \frac{\nu k^2}{v + 2} \right])\]

(25)

with the eigenvalues

\[
\omega_1^2 = \bar{c}^2 + 2 \frac{\bar{\rho}}{\rho} \left( \frac{3\nu}{v + 2} - 1 \right),
\]

(26a)

\[
\omega_{1,2}^2 = \frac{\nu k^2}{v + 2}.
\]

(26b)

The error of these relations is mostly dictated by the quality of the density estimate, either directly via \(\bar{\rho}, \bar{c}\) and \(P\), or indirectly via \(\Xi\). The density correction method of equation (18) can only help with the former, but not the latter. The difference between constant and adapted \(\tilde{h}\) is a factor 4/9 (for 3D) in favour of the latter.

3.1.2 Continuum limit

For large neighbour numbers \(N_H, H \gg d_{\text{nn}}, \Omega & \Pi \rightarrow 1, \bar{\rho} \rightarrow \rho\) and the sums in equations (24a,b) can be approximated by integrals\(^8\)

\[
\mu \rightarrow -k \hat{W}(k) \quad \text{and} \quad U \rightarrow k^2 \hat{W}(k),
\]

(27)

\(^7\) In his analysis of 1D SPH, Rasio (2000) also considers this simplification, but interprets it incorrectly as the limit \(|k|d_{\text{nn}} \ll 1\) regardless of \(h\).

\(^8\) Assuming a uniform particle distribution. A better approximation, which does not require \(d_{\text{nn}} \ll H\), is to assume some \textit{radial distribution function} \(g(r)\) (as in statistical mechanics of glasses) for the probability of any two particles having distance \(r\). Such a treatment may well be useful in the context of SPH, but it is beyond the scope of our study.

3.1.3 Resolved waves in the continuum limit

The combined limit of \(\lambda \gg H \gg d_{\text{nn}}\) is obtained by inserting the Taylor expansion (17) of \(\hat{W}\) into equation (28), giving

\[
\omega_{1,2}^2 = c^2 k^2 (1 + \sigma^2 k^2 \gamma^{-1}[2/v + 1 - \gamma] + \mathcal{O}(h^4 |k|^4)).
\]

(29)
Figure 5. As Fig. 4, but for the quartic and quintic spline and the HOCT4 kernel (left) and the Wendland $C^2$ to $C^6$ kernels (right). The fine details of these contours are specific to the densest sphere packing and will be different for more realistic glass-like particle distributions.
Monaghan (2005) gave an equivalent relation for \( \sigma \) proportional to the resolution scale, at least concerning sound waves. In Fig. 6 we plot the ratio between \( \omega \) and \( c_{\text{SPH}} \) for the kernel–\( N_H \) combinations of Table 2 and three different wave directions (as indicated) for particles in densest sphere packing with nearest-neighbour separation \( d_{\text{nn}} \). These curves are horizontal cuts through the top panels of Figs 4 and 5. The thin vertical lines indicate sound with wavelength \( \lambda = \frac{h}{d} \), corresponding to the dashed lines in Figs 4 and 5.

3.2 Linear stability of SPH kernels

We have evaluated the eigenvalues \( \omega^2 \) and \( \omega^2_{\perp,\parallel} \) of the matrix \( P \) in equation (23) for all kernels of Table 1, as well as the HOCT4 and Gaussian kernels, for unperturbed positions from densest sphere packing (face-centred cubic grid).\(^9\) In Figs 4 and 5, we plot the resulting contours of \( \omega^2c^2k^2 \) over wavenumber \( |k| \) and smoothing scale \( h \) (both normalized by the nearest-neighbour distance \( d_{\text{nn}} \) or \( N_H \) on the right axes (except for the Gaussian kernel when \( N_H \) is ill-defined and we give \( N_H \) instead) for two wave directions, one being a nearest-neighbour direction.

\(^9\) Avoiding an obviously unstable configuration, such as a cubic lattice, which may result in \( \omega^2 < 0 \) simply because the configuration itself was unstable, not the numerical scheme.

3.2.1 Linear stability against pairing

The top subpanels of Figs 4 and 5 refer to the longitudinal eigenvalue \( \omega^2_{\parallel} \), when green and red contours are for, respectively, \( \omega^2_{\parallel} \approx c_0^2k^2 \) and \( \omega^2_{\parallel} < 0 \), the latter indicative of the pairing instability. For the Gaussian kernel (truncated at 16\( \sigma \); Fig. 4) \( \omega^2_{\parallel} > 0 \) everywhere, proving its stability,\(^10\) similar to the HOCT4 and, in particular the higher degree, Wendland kernels. In contrast, all the B-spline kernels obtain \( \omega^2_{\parallel} < 0 \) at sufficiently large \( N_H \).

The quintic spline, Wendland \( C^2 \), and HOCT4 kernel each have a region of \( \omega^2_{\parallel} < 0 \) for \( |k| \) close to the Nyquist frequency and \( N_H \approx 100, \approx 40 \) and \( \approx 150 \), respectively. In numerical experiments similar to those described in Section 4.1, the corresponding instability for the quintic spline and Wendland \( C^2 \) kernels can be triggered by very small random perturbations to the grid equilibrium. However, such modes are absent in glass-like configurations, which naturally emerge by ‘cooling’ initially random distributions. This strongly suggests that these kernel–\( N_H \) combinations can be safely be used in practice. Whether this also applies to the HOCT4 kernel at \( N_H \approx 150 \) we cannot say, as we have not run test simulations for this kernel. Note that these islands of linear instability at small \( N_H \) are not in contradiction to the relation between kernel Fourier transform and stability and are quite different from the situation for the B-splines, which are only stable for sufficiently small \( N_H \).

3.2.2 Linear transverse instability?

The bottom subpanels of Figs 4 and 5 show \( \omega^2_{\perp,\parallel}c^2k^2 \), when both families of kernels have \( 0 \approx |\omega^2_{\perp,\parallel}c^2| < c^2k^2 \) with either sign occurring. \( \omega^2_{\perp,\parallel} < 0 \) implies growing transverse modes,\(^11\) which we indeed found in simulations starting from a slightly perturbed densest sphere packing. However, such modes are not present in glass-like configurations, which strongly suggests that transverse modes are not a problem in practice.

3.3 Numerical resolution of sound waves

The dashed lines in Figs 4 and 5 indicate sound with wavelength \( \lambda = \frac{8}{h} \). For \( h \gtrsim d_{\text{nn}} \), such sound waves are well resolved in the sense that the sound speed is accurate to \( \lesssim 1 \) per cent. This is similar to grid methods, which typically require about eight cells to resolve a wavelength.

The effective SPH sound speed can be defined as \( c_{\text{SPH}} \approx \omega_0/|k| \). In Fig. 6 we plot the ratio between \( c_{\text{SPH}} \) and the correct sound speed as function of wavenumber for three different wave directions and the 10 kernel–\( N_H \) combinations of Table 2 (which also gives their formal resolutions). The transition from \( c_{\text{SPH}} \approx c \) for long waves to \( c_{\text{SPH}} \ll c \) for short waves occurs at \( |k|d_{\text{nn}} \gtrsim 1 \), but towards longer waves for larger \( h d_{\text{nn}} \), as expected.

\(^10\) There is in fact \( \omega^2_{\parallel} < 0 \) at values for \( h \) larger than plotted. In agreement with our analysis in Section 3.1.2, this is caused by truncating the Gaussian, which (like any other modification to avoid infinite neighbour numbers) invalidates the non-negativity of its Fourier transform. These theoretical results are confirmed by numerical findings of D. Price (referee report), who reports pairing at large \( h d_{\text{nn}} \) for the truncated Gaussian.

\(^11\) Read et al. (2010) associate \( \omega^2_{\perp,\parallel} < 0 \) with a ‘banding instability’ which appeared near a contact discontinuity in some of their simulations. However, they fail to provide convincing arguments for this connection, as their stability analysis is compromised by the use of the unstable cubic lattice.
In order to test our theoretical predictions regarding the pairing in-
native, the quartic spline with \( \rho \approx 42 \) is the most common choice in astrophysical simulations, whereas for the cubic spline
is an indicator for the regularity of the particle distribution around

For resolved waves (\( \lambda \gtrsim 8 \ h \): left of the thin vertical lines in Fig. 6), \( \text{c}_{\text{SPH}} \) obtains a value close to \( c \), but with clear differences between the various kernel–\( N \) \( h \) combinations. Surprisingly, the standard cubic spline kernel, which is used almost exclusively in astrophysics, performs very poorly with errors of few per cent, for both \( N = 42 \) and 55. This is in stark contrast to the quartic spline with similar \( N = 60 \) but \( \text{c}_{\text{SPH}} \) accurate to \( < 1 \) per cent. Moreover, the quartic spline with \( N = 60 \) resolves shorter waves better than the cubic spline with a smaller \( N = 55 \), in agreement with Table 2.

We should note that these results for the numerical sound speed assume a perfectly smooth simulated flow. In practice, particle disorder degrades the performance, in particular for smaller \( N \), and the resolution of \( \text{SPH} \) is limited by the need to suppress this degradation via increasing \( h \) (and \( N \)).

### 4 TEST SIMULATIONS

In order to assess the Wendland kernels and compare them to the standard B-spline kernels in practice, we present some test simulations which emphasize the pairing, strong shear, and shocks. All these simulations are done in 3D using periodic boundary conditions, \( K = K_0^{1/3} \), conservative \( \text{SPH} \) (equation 3) and the Cullen & Dehnen (2010) artificial viscosity treatment, which invokes dissipation only for compressive flows, and an artificial conductivity similar to that of Read & Hayfield (2012). For some tests we used various values of \( N \) per kernel, but mostly those listed in Table 2.

#### 4.1 Pairing in practice

In order to test our theoretical predictions regarding the pairing instability, we evolve noisy initial conditions with 32,000 particles until equilibrium is reached. Initially, \( x_i = 0 \), while the initial \( x_i \) are generated from densest sphere packing by adding normally distributed offsets with (1D) standard deviation of one unperturbed nearest-neighbour distance \( d_{\text{min}} \). To enable a uniform-density equilibrium (a glass), we suppress viscous heating.

The typical outcome of these simulations is either a glass-like configuration (right-hand panel of Fig. 7) or a distribution with particle pairs (left-hand panel of Fig. 7). In order to quantify these outcomes, we compute for each particle the ratio

\[
r_{\text{min},i} = \min_{j \neq i} |x_i - x_j|/H_i
\]

between its actual nearest-neighbour distance and kernel-support radius. The maximum possible value for \( r_{\text{min}} \) occurs for densest sphere packing, when \( |x_i - x_j| \geq d_{\text{min}} = (n/\sqrt{3})^{1/3} \) with \( n \) the number density. Replacing \( \rho_i \) in equation (7) with \( m/n \), we obtain

\[
r_{\text{min}} \lesssim (3N_{\text{hi}}/2^{5/2} \pi)^{1/3}.
\]

Thus, the ratio

\[
q_{\text{min},i} = \frac{r_{\text{min},i}}{(3N_{\text{hi}}/2^{5/2} \pi)^{1/3}} \approx \min_{j \neq i} \left\{ \frac{|x_i - x_j|}{d_{\text{min,grid}}} \right\}
\]

is an indicator for the regularity of the particle distribution around particle \( i \). It obtains a value very close to one for perfect densest sphere packing and near zero for pairing, while a glass typically gives \( q_{\text{min},i} \sim 0.7 \).

Fig. 8 plots the final value for the overall minimum of \( q_{\text{min},i} \) for each of a set of simulations. For all values tested for \( N_{\text{hi}} \) (up to 700), the Wendland kernels show no indication of a single particle pair. This is in stark contrast to the B-spline kernels, all of which suffer from particle pairing. The pairing occurs at \( N_{\text{hi}} \gtrsim 67 \) and 190 for the quartic and quintic spline, respectively, whereas for the cubic spline \( \min_i \{q_{\text{min},i}\} \) approaches zero more gradually, with \( \min_i \{q_{\text{min},i}\} \lesssim 0.16 \) at \( N_{\text{hi}} \gtrsim 55 \). These thresholds match quite well the suggestions of the linear stability analysis in Figs 4 and 5 (except that the

| Kernel | \( N_{\text{hi}} \) | \( N_{\text{h}} \) | \( h/d_{\text{min}} \) | \( b(\rho/m)^{1/3} \) |
|--------|-------------|-------------|----------------|----------------|
| Cubic spline | 42 | 6.90 | 1.052 | 1.181 |
| Cubic spline | 55 | 9.04 | 1.151 | 1.292 |
| Quartic spline | 60 | 7.29 | 1.072 | 1.203 |
| Quintic spline | 180 | 17.00 | 1.421 | 1.595 |
| Wendland \( C^2 \) | 100 | 13.77 | 1.325 | 1.487 |
| Wendland \( C^3 \) | 200 | 18.58 | 1.464 | 1.643 |
| Wendland \( C^4 \) | 400 | 27.22 | 1.662 | 1.866 |
| HOCT4 | 442 | 42.10 | 1.923 | 2.158 |
| Gaussian | \( \infty \) | 10.00 | 1.191 | 1.337 |
| Gaussian | \( \infty \) | 20.00 | 1.500 | 1.684 |

![Figure 7](https://example.com/figure7.png)

**Figure 7.** Final \( x \) and \( y \) positions for particles at \(|z| < d_{\text{min}}/2\) for the tests of Section 4.1. Symbol size is linear in \( z \); two overlapping symbols of the same size indicate particle pairing.

![Figure 8](https://example.com/figure8.png)

**Figure 8.** Final value of the overall minimum of \( q_{\text{min},i} \) (equation 32) for simulations starting from noisy initial conditions for all kernels of Table 1 (same colour coding as in Fig. 3; using the density correction of Section 2.5 for the Wendland kernels) as function of \( N_{\text{hi}} \). For densest sphere packing \( q_{\text{min},i} \approx 1 \), for a glass \( 0 < q_{\text{min},i} < 1 \), while \( q_{\text{min},i} \approx 0 \) indicates particle pairing.
indicates of instability of the quintic spline at \( N_H \approx 100 \) and the Wendland \( C^2 \) kernel at \( N_H \approx 40 \) are not reflected in our tests here. The quintic (and higher order) splines are the only option amongst the B-spline kernels for \( N_H \) appreciably larger than \( \sim 50 \).

We also note that min. \{\( q_{\text{min.}} \} \) grows substantially faster, in particular early on, for the Wendland kernels than for the B-splines, especially when operating close to the stability boundary.

### 4.2 The Gresho–Chan vortex test

As discussed in the Introduction, particle disorder is unavoidably generated in shearing flows, inducing ‘\( E_0 \) errors’ in the forces and causing modelling errors. A critical test of this situation consists of a differentially rotating fluid of uniform density in centrifugal balance (Gresho & Chan 1990, see also Liska & Wendroff 2003; Springel 2010; Read & Hayfield 2012). The pressure and azimuthal velocity are

\[
P = \begin{cases} 
  P_0 + 12.5 R^2 & \text{for } 0 \leq R < 0.2, \\
  P_0 + 12.5 R^2 + 4 - 20 R + 4 \ln(5R) & \text{for } 0.2 \leq R < 0.4, \\
  P_0 + 2(2 \ln 2 - 1) & \text{for } 0.4 \leq R; 
\end{cases} \tag{33a} 
\]

\[
v_\phi = \begin{cases} 
  5 R & \text{for } 0 \leq R < 0.2, \\
  2 - 5 R & \text{for } 0.2 \leq R < 0.4, \\
  0 & \text{for } 0.4 \leq R, 
\end{cases} \tag{33b} 
\]

where \( P_0 = 5 \) and \( R \) is the cylindrical radius. We start our simulations from densest sphere packing with effective one-dimensional particle numbers \( N_{1D} = 51, 102, 203 \) or 406. The initial velocities and pressure are set as in equations (33).

There are three different causes for errors in this test. First, an overly viscous method reduces the differential rotation, as shown by Springel (2010); this effect is absent from our simulations owing to the usage of the Cullen & Dehnen (2010) dissipation switch. Secondly, the ‘\( E_0 \) error’ generates noise in the velocities which in turn triggers some viscosity. Finally, finite resolution implies that the sharp velocity kinks at \( R = 0.2 \) and 0.4 cannot be fully resolved (in fact, the initial conditions are not in SPH equilibrium because the pressure gradient at these points is smoothed such that the SPH acceleration is not exactly balanced with the centrifugal force).

In Fig. 9 we plot the azimuthal velocity at time \( t = 1 \) for a subset of all particles at our lowest resolution of \( N_{1D} = 51 \) for four different kernel–\( N_H \) combinations. The leftmost is the standard cubic spline with \( N_H = 55 \), which considerably suffers from particle disorder and hence ‘\( E_0 \) errors’ (but also obtains too low \( v_\phi \) at \( R < 0.2 \)).

The second is the Wendland \( C^2 \) kernel with \( N_H = 100 \), which still suffers from the ‘\( E_0 \) error’. The last two are for the Wendland \( C^0 \) kernel with \( N_H = 400 \) and the Wendland \( C^2 \) kernel with \( N_H = 100 \) but with \( P_0 = 0 \) in equation (33a). In both cases, the ‘\( E_0 \) error’ is much reduced (and the accuracy limited by resolution) either because of large neighbour number or because of a reduced pressure.

In Fig. 10, we plot the convergence of the \( L_1 \) velocity error for the Gresho–Chan test with increasing number of particles (for a cubic lattice \( N_{1D} \) equals the number of cells along one side of the computational domain) for various kernel–\( N_H \) combinations (those with filled squares are shown in Fig. 9). A comparison with fig. 6 of Read & Hayfield (2012) shows that the Wendland \( C^0 \) kernel with \( N_H = 400 \) performs better than the HOCT4 kernel with \( N_H = 442 \).

**Figure 9.** Velocity profiles the Gresho–Chan vortex test at time \( t = 1 \) with the lowest resolution of \( N_{1D} = 50.8 \) for three different kernel–\( N_H \) pairs (as indicated) and one simulation (right-hand panel) with \( P_0 = 0 \) in equation (33a), when the ‘\( E_0 \) error’ is naturally much weaker. Only particles at \( |z| < 0.05 \) are plotted.

**Figure 10.** Convergence of the \( L_1 \) velocity error for the Gresho–Chan test with increasing number of particles (for a cubic lattice \( N_{1D} \) equals the number of cells along one side of the computational domain) for various kernel–\( N_H \) combinations (those with filled squares are shown in Fig. 9). A comparison with fig. 6 of Read & Hayfield (2012) shows that the Wendland \( C^0 \) kernel with \( N_H = 400 \) performs better than the HOCT4 kernel with \( N_H = 442 \).
lower for the cubic spline than any other kernel. This is caused by systematically too low \( v_\varphi \) in the rigidly rotating part at \( R < 0.2 \) (see leftmost panel in Fig. 9) at all resolutions. The good performance of the quartic spline is quite surprising, in particular given the rather low \( N_H \). The quintic spline at \( N_H = 180 \) and the Wendland \( C^6 \) kernel at \( N_H = 200 \) obtain very similar convergence, but are clearly topped by the Wendland \( C^6 \) kernel at \( N_H = 400 \), demonstrating that high neighbour number is really helpful in strong shear flows.

### 4.3 Shocks

Our final test is the classical Sod (1978) shock tube, a 1D Riemann problem, corresponding to an initial discontinuity in density and pressure. Unlike most published applications of this test, we perform 3D simulations with glass-like initial conditions. Our objective here is (1) to verify the ‘\( E_0 \) error’ reductions at larger \( N_H \) and (2) the resulting trade-off with the resolution across the shock and contact discontinuities. Other than for the vortex tests of Section 4.2, we only consider one value for the number \( N \) of particles but the same six kernel–\( N_H \) combinations as in Fig. 10. The resulting profiles of velocity, density and thermal energy are plotted in Fig. 11 together with the exact solutions.

Note that the usual overshooting of the thermal energy near the contact discontinuity (at \( x = 0.17 \)) is prevented by our artificial conductivity treatment. This is not optimized and likely over smooths the thermal energy (and with it the density). However, here we concentrate on the velocity.

For the cubic spline with \( N_H = 55 \), there is significant velocity noise in the post-shock region. This is caused by the re-ordering of the particle positions after the particle distribution becomes anisotropically compressed in the shock. This type of noise is a well-known issue with multidimensional SPH simulations of shocks (e.g. Springel 2010; Price 2012). With increasing \( N_H \) the velocity noise is reduced, but because of the smoothing of the velocity jump at the shock (at \( x = 0.378 \)) the \( L_1 \) velocity error does not approach zero for large \( N_H \).

Instead, for sufficiently large \( N_H \) (\( > 200 \) in this test), the \( L_1 \) velocity error saturates: any ‘\( E_0 \) error’ reduction for larger \( N_H \) is balanced by a loss of resolution. The only disadvantage of larger \( N_H \) is an increased computational cost (by a factor \( \approx 1.5 \) when moving from the quintic spline with \( N_H = 180 \) to the Wendland kernel \( C^6 \) with \( N_H = 400 \), see Fig. 12).

### 5 DISCUSSION

#### 5.1 What causes the pairing instability?

The Wendland kernels have an inflection point and yet show no signs of the pairing instability. This clearly demonstrates that the traditional ideas for the origin of this instability (à la Swegle et al. 1995, see the Introduction) were incorrect. Instead, our linear stability analysis shows that in the limit of large \( N_H \) pairing is caused by a negative kernel Fourier transform \( \tilde{W} \), whereas the related tensile instability with the same symptoms is caused by an (effective) negative pressure. While it is intuitively clear that negative pressure causes pairing, the effect of \( \tilde{W} < 0 \) is less obvious. Therefore, we now provide another explanation, not restricted to large \( N_H \).

#### 5.1.1 The pairing instability as artefact of the density estimator

By their derivation from the Lagrangian (2), the SPH forces \( m_i \dot{x}_i = \partial L / \partial x_i = -\partial U / \partial x_i \) tend to reduce the estimated total thermal energy \( \tilde{U} = \sum m_i u(\hat{\rho}_i, s_i) \) at fixed entropy.\(^{12}\) Thus, hydrostatic equilibrium corresponds to an extremum of \( \tilde{U} \), and stable equilibrium to a minimum when small positional changes meet opposing forces. Minimal \( \tilde{U} \) is obtained for uniform \( \hat{\rho}_i \), since a redistribution of the particles in the same volume but with a spread of \( \hat{\rho}_i \) gives larger \( \tilde{U} \) (assuming uniform \( s_i \)). An equilibrium is metastable, if \( \tilde{U} \) is only a local (but not the global) minimum. Several extrema can occur if different particle distributions, each obtaining (near-)uniform \( \hat{\rho}_i \), have different average \( \hat{\rho}_i \). Consider, for example, particles in densest sphere packing, replace each by a pair and increase the spacing by \( 2^{1/3} \), so that the average density \( \rho \) (but not \( \hat{\rho} \)) remains unchanged. This fully paired distribution is in equilibrium with uniform \( \hat{\rho}_i \), but the effective neighbour number is reduced by a factor of 2 (for the same smoothing scale). Now

\(^{12}\) This holds, of course, also for an isothermal gas, when \( u = \) constant, but not the entropy \( s \), so that \( (\partial s / \partial \rho)_u \neq 0 \).
the global minimum of \( \hat{U} \) is highly degenerate. This method has neither a pairing instability, nor ‘\( E_0 \) errors’, nor the re-ordering capacity of SPH, but requires additional terms for that latter purpose.

5.2 Are there more useful kernel functions?

 Neither the B-splines nor the Wendland functions have been designed with SPH or the task of density estimation in mind, but derive from interpolation of function values \( y \) for given points \( x_i \).

The B-splines were constructed to exactly interpolate polynomials on a regular 1D grid. However, this for itself is not a desirable property in the context of SPH, in particular for 2D and 3D.

The Wendland functions were designed for interpolation of scattered multidimensional data, viz

\[
\hat{s}(x) = \sum_j \alpha_j w((x - x_j)).
\]

The coefficients \( \alpha_j \) are determined by matching the interpolant \( s(x) \) to the function values, resulting in the linear equations

\[
y_i = \sum_j \alpha_j w(|x_i - x_j|), \quad i = 1, \ldots, n.
\]

If the matrix \( W \) has compact support, then it is sparse, which greatly reduces the complexity of the problem. The Wendland functions were designed to fit this bill. As a side effect they have non-negative Fourier transform (according to Bochner 1933), which together with their compact support, smoothness and computational simplicity makes them ideal for SPH with large \( N_H \).

So far, the Wendland functions are the only kernels which are stable against pairing for all \( N_H \) and satisfy all other desirable properties from the list in Section 2.2.

5.3 What is the SPH resolution scale?

In smooth flows, i.e. in the absence of particle disorder, the only error of the SPH estimates is the bias induced by the smoothing operation. For example, assuming a smooth density field

\[
\hat{\rho_i} \approx \rho(x_i) + \frac{1}{2} \sigma^2 \nabla^2 \rho(x_i) + O(h^4)
\]

(e.g. Monaghan 1985; Silverman 1986) with \( \sigma \) defined in equation (8). Since \( \sigma \) also sets the resolution of sound waves (Section 3.1.3), our definition (10), \( h = 2\sigma \), of the SPH resolution scale is appropriate for smooth flows. The result (34) is the basis for the traditional claim of \( O(h^2) \) convergence for smooth flows. True flow discontinuities are smeared out over a length scale comparable to \( h \) (though we have not made a detailed investigation of this).

In practice, however, particle disorder affects the performance and, as our test simulations demonstrated, the actual resolution of SPH can be much worse than the smooth-flow limit suggests.

5.4 Are large neighbour numbers sensible?

There is no consensus about the best neighbour number in SPH: traditionally the cubic spline kernel is used with \( N_H \approx 42 \), while Price (2012) favours \( N_H = 57 \) (at or even beyond the pairing-instability limit) and Read et al. (2010) use their HOCT4 kernel with even \( N_H = 442 \) (corresponding to a 1.7 times larger \( h_0 \)). From a pragmatic point of view, the number \( N \) of particles, the neighbour number \( N_H \), and the smoothing kernel (and between them the numerical

---

13 The density correction of Section 2.5 does not affect these arguments, because during a simulation \( \epsilon \) in equation (18) is fixed and in terms of our considerations here the solid curves in Fig. 3 are simply lowered by a constant.
resolution) are numerical parameters which can be chosen to optimize the efficiency of the simulation. The critical question therefore is

Which combination of $N$ and $N_{H}$ (and kernel) most efficiently models a given problem at a desired fidelity?

Clearly, this will depend on the problem at hand as well as the desired fidelity. However, if the problem contains any chaotic or turbulent flows, as is common in star- and galaxy formation, then the situation exemplified in the Gresho–Chan vortex test of Section 4.2 is not atypical and large $N_{H}$ may be required for sufficient accuracy.

But are high neighbour numbers affordable? In Fig. 12, we plot the computational cost versus $N_{H}$ for different kernels. At $N_{H} \lesssim 400$ the costs rise sublinearly with $N_{H}$ (because at low $N_{H}$ SPH is data- rather than computation dominated) and high $N_{H}$ are well affordable. In the case of the vortex test, they are absolutely necessary as Fig. 13 demonstrates: for a given numerical accuracy, our highest $N_{H}$ makes optimal use of the computational resources (in our code memory usage does not significantly depend on $N_{H}$, so CPU time is the only relevant resource).

6 SUMMARY

Particle disorder is unavoidable in strong shear (ubiquitous in astrophysical flows) and causes random errors of the SPH force estimator. The good news is that particle disorder is less severe than Poissonian shot noise and the resulting force errors (which are dominated by the $E_{0}$ term of Read et al. 2010) are not catastrophic. The bad news, however, is that these errors are still significant enough to spoil the convergence of SPH.

In this study we investigated the option to reduce the $E_{0}$ errors by increasing the neighbour number in conjunction with a change of the smoothing kernel. Switching from the cubic to the quintic spline at fixed resolution $h$ increases the neighbour number $N_{H}$ only by a factor of 1.74, hardly enough to combat $E_{0}$ errors. For a significant reduction of these errors one has to trade resolution and significantly increase $N_{H}$ beyond conventional values.

The main obstacle with this approach is the pairing instability, which occurs for large $N_{H}$ with the traditional SPH smoothing kernels. In Section 3 and appendix A, we have performed (it appears for the first time) a complete linear stability analysis for conservative SPH in any number of spatial dimensions. This analysis shows that SPH smoothing kernels whose Fourier transform is negative for some wave vector $k$ will inevitably trigger the SPH pairing instability at sufficiently large neighbour number $N_{H}$. Such kernels therefore require $N_{H}$ to not exceed a certain threshold in order to avoid the pairing instability (not to be confused with the tensile instability, which has the same symptoms but is caused by a negative effective pressure independent of the kernel properties).

Intuitively, the pairing instability can be understood in terms of the SPH density estimator: if a paired particle distribution obtains a lower average estimated density, its estimated total thermal energy $\tilde{U}$ is smaller and hence favourable. Otherwise, the smallest $\tilde{U}$ occurs for a regular distribution, driving the automatic maintenance of particle order, a fundamental ingredient of SPH.

The Wendland (1995) functions, presented in Section 2.2.3, have been constructed, albeit for different reasons, to possess a non-negative Fourier transform, and be of compact support with simple functional form. The first property and the findings from our tests in Section 4.1 demonstrate the remarkable fact that these kernels are stable against pairing for all neighbour numbers (this disproves the long-cultivated myth that the pairing instability was caused by a maximum in the magnitude of the kernel gradient). Our 3D test simulations show that the cubic, quartic and quintic spline kernels become unstable to pairing for $N_{H} > 55, 67$ and 190, respectively (see Fig. 8), but operating close to these thresholds cannot be recommended.

A drawback of the Wendland kernels is a comparably large density error at low $N_{H}$. As we argue in Section 5.1.1, this error is directly related to the stability against pairing. However, in Section 2.5 we present a simple method to correct for this error without affecting the stability properties and without any other adverse effects.

We conclude, therefore, that the Wendland functions are ideal candidates for SPH smoothing kernels, in particular when large $N_{H}$ are desired, since they are computationally superior to the high-order B-splines. All other alternative kernels proposed in the literature are computationally more demanding and are either centrally spiked, like the HOCT4 kernel of Read et al. (2010), or susceptible to pairing like the B-splines (e.g. Cabezón, García-Senz & Relaño 2008).

Our tests of Section 4 show that simulations of both strong shear flows and shocks benefit from large $N_{H}$. These tests suggest that for $N_{H} \sim 200$ and 400, respectively, the Wendland $C^{4}$ and $C^{0}$ kernels are most suitable. Compared to $N_{H} = 55$ with the standard cubic spline kernel, these kernel--$N_{H}$ combinations have lower resolution ($h$ increased by factors of 1.27 and 1.44, respectively), but obtain much better convergence in our tests.

For small neighbour numbers, however, these tests and our linear stability analysis unexpectedly show that the quartic B-spline kernel with $N_{H} = 60$ is clearly superior to the traditional cubic spline and can compete with the Wendland $C^{4}$ kernel with $N_{H} = 100$. The reason for this astonishing performance of the quartic spline is unclear, perhaps the fact that near $x = 0$ this spline is more than three times continuously differentiable plays a role.

We note that, while the higher degree Wendland functions are new to SPH, the Wendland $C^{4}$ kernel has already been used (Monaghan 2011, for example, employs it for 2D simulations). However, while its immunity to the pairing instability has been noted (e.g. Robinson...
REFERENCES

Abel T., 2011, MNRAS, 413, 271
Bochner S., 1933, Math. Ann., 108, 378
Cabezón R. M., García-Senz D., Relaño A., 2008, J. Comput. Phys., 227, 8523
Cullen L., Dhenen W., 2010, MNRAS, 408, 669
Fulk D. A., Quinn D. W., 1996, J. Comput. Phys., 126, 165
Gingold R. A., Monaghan J. J., 1977, MNRAS, 181, 375
Gray J. P., Monaghan J. J., Swift R. P., 2001, Comput. Methods Applied Mech. Eng., 190, 6641
Gresho P. M., Chan S. T., 1990, Int. J. Numer. Methods Fluids, 11, 621
Hall P., 1927, Biometrika, 19, 232
Herant M., 1994, Mem. Soc. Astron. Ital., 65, 1013
Heß S., Springel V., 2010, MNRAS, 406, 2289
Irwin J. O., 1927, Biometrika, 19, 225
Irwin J. O., 1927, Biometrika, 19, 240
Liska R., Wendroff B., 2003, SIAM J. Sci. Comput., 25, 995
Lucy L. B., 1977, AJ, 82, 1013
Monaghan J. J., 1985, MNRAS, 216, 883
Price D. J., 2012, J. Comput. Phys., 231, 759
Rasio F. A., 2000, Progress Theor. Phys. Suppl., 138, 609
Read J. I., Hayfield T., Agertz O., 2010, MNRAS, 405, 1513
Read J. I., Hayfield T., 2012, MNRAS, 422, 3037
Rasio F. A., 2000, Progress Theor. Phys. Suppl., 138, 609
Read J. I., Hayfield T., 2012, MNRAS, 422, 3037
Read J. I., Hayfield T., Agertz O., 2010, MNRAS, 405, 1513
Robinson M. J., 2009, PhD thesis, Monash University, Australia
Rosswog S., 2009, New Astron. Rev., 53, 78
Schoenrich J. I., 1974, Q. Applied Math., 4, 45
Schüller L., Schmitt D., 1981, A&A, 97, 373
Silverman B. W., 1986, Density Estimation for Statistics and Data Analysis.
Springer Berlin Heidelberg, London
Sod G. A., 1978, J. Comput. Phys., 27, 1
Springel V., 2010, ARA&A, 48, 391

APPENDIX A: LINEAR STABILITY ANALYSIS

We start from an equilibrium with particles of equal mass $m$ on a regular grid and impose a plane-wave perturbation to the unperturbed positions $x_i$, (a bar denotes a quantity obtained for the unperturbed equilibrium):

$$x_i = \bar{x}_i + \xi_i, \quad \xi_i = a \Phi_i(x, t), \quad \Phi_i(x, t) = e^{ik \cdot x - \omega t}.$$  

(A1)

as in equation (30). We derive the dispersion relation $\omega(k, a)$ by equating the SPH force imposed by the perturbation (to linear order) to its acceleration

$$\dot{\xi}_i = -a \omega^2 \Phi_i(x, t).$$  

(A2)

To obtain the perturbed SPH forces to linear order, we develop the internal energy of the system, and hence the SPH density estimate, to second order in $\xi_i$. If $\bar{\rho}_i = \rho + \rho_1 + \rho_2$ with $\rho_1$ and $\rho_2$ the first- and second-order density corrections, respectively, then

$$\dot{\xi}_i = \frac{1}{m} \frac{\partial L}{\partial \xi_i} = -\frac{P}{\rho} \sum_k \frac{1}{\rho} \frac{\partial \rho_{12}}{\partial \xi_i} \left( \frac{\xi^2}{2 \rho} - \frac{2P}{\rho} \right) \frac{1}{2} \sum_k \frac{\partial \rho_{12}}{\partial \xi_i}.$$  

(A3)

\textbf{A1 Fixed $h$}

Let us first consider the simple case of constant $h = \bar{h}$ which remains unchanged during the perturbation. Then, $\rho_i = \rho_{in}/n_i$ with $\rho_{in} = m \sum_j (\xi_{ij} \cdot \nabla W(\bar{x}_{ij}, \bar{h})$.  

(A4)

Inserting

$$\dot{\xi}_i = a(1 - e^{-ik \cdot x_i}) \Phi_i,$$  

(A5)

into (A4) gives

$$\partial_{a1} = \Phi_i \sum_1 m (1 - e^{-ik \cdot x_i}) a \nabla W(\bar{x}_{ij}, \bar{h}) = i \Phi_i \cdot a \cdot t,$$  

(A6)

where (assuming a symmetric particle distribution)

$$t(k) = \sum_j m \sin k \cdot \bar{x}_j \cdot \nabla W(\bar{x}_{ij}, \bar{h}).$$  

(A7)

We can then derive

$$\sum_k \phi_{11} \partial_{a11}/\partial \xi_i = \sum_k \phi_{11} \sum_j m \nabla W(\bar{x}_{ij}, \bar{h}) (\delta_{ik} - \delta_{ij}) = \sum_j m (\phi_{11} + \phi_{1}) \nabla W(\bar{x}_{ij}, \bar{h}) = i \Phi_i \cdot a \cdot t,$$  

(A8)

$$\sum_k \phi_{12} \partial_{a12}/\partial \xi_i = \sum_k m (\xi_{ij} \cdot \nabla W(\bar{x}_{ij}, \bar{h}) (\delta_{ik} - \delta_{ij}) = 2 \sum_j m (\xi_{ij} \cdot \nabla W(\bar{x}_{ij}, \bar{h}) = 2 \Phi_i \sum_j m (1 - e^{-ik \cdot x_i}) a \cdot \nabla W(\bar{x}_{ij}, \bar{h}) = 2 \Phi_i \cdot a \cdot T,$$  

(A9)

with

$$T(k) = \sum_j m (1 - \cos k \cdot \bar{x}_j) W(\bar{x}_{ij}, \bar{h}).$$  

(A10)
Inserting these results into (A3), we get
\[ \dot{\xi}_i = -\frac{2\tilde{P}}{\rho} \frac{a T}{\rho} \Phi_i - \left( \dot{\xi}^2 - \frac{2\tilde{P}}{\rho} \right) \frac{a T}{\rho^2} \Phi_i, \] (A11)

A2 Adapting smoothing
If the \( h_i \) are adapted such that \( M_t \equiv h_t \dot{\rho}_t \) remains a global constant \( M \), the estimated density is simply \( \dot{\rho}_t = M h_t^{-1} \). We start by expanding \( M_t \) to second order in both \( a \) and \( \eta_i \equiv \ln(h_i/h) \). Using a prime to denote differentiation with respect to \( \ln h \), we have
\[ M_t = \dot{M} + h_t \dot{\rho}_t \left[ \dot{\eta}_i + \eta_i \dot{\rho}_i + \frac{1}{2} \eta_i \dot{\rho}_t + \frac{1}{2} \eta_i \dot{\rho}_i \right], \] (A12a)
with \( \dot{\Omega} \equiv 1 \) as defined in equation (4) and
\[ \dot{\Pi} = \frac{1}{v \dot{h}^2 \dot{\rho} \dot{\rho} / \partial (\ln h)} = \frac{1}{v \dot{h} h^2 \dot{\rho} / \partial (\ln h)} \approx 1. \] (A13)

By demanding \( M_t = \dot{M} \), we obtain for the first- and second-order contributions to \( \eta_i \):
\[ -v \dot{\rho} \eta_{i1} = \frac{\dot{\eta}_i}{\Omega}, \] (A14a)
\[ -v \dot{\rho} \eta_{i2} = \frac{\dot{\eta}_i}{2\Omega} + \frac{\eta_i \dot{\rho}_i}{\Omega} + \frac{\eta_i \dot{\rho}_i}{\Omega} + \frac{2 \dot{\rho} \dot{\rho}}{2\Omega} \] (A14b)
\[ \hat{\Pi} = \frac{\eta_i \dot{\rho}_i}{2 \Omega}, \] (A14c)
From these expressions and \( \dot{\rho}_t = \rho e^{-\Delta h} \) we obtain the first- and second-order density corrections:
\[ \rho_{i1} = -v \dot{\rho} \eta_{i1} = \frac{\dot{\eta}_i}{\Omega}, \] (A15a)
\[ \rho_{i2} = -v \dot{\rho} \eta_{i2} + \frac{1}{2} \dot{\rho} \rho \dot{\rho} \eta_{i1} = \frac{\dot{\eta}_i}{2 \Omega} \rho e^{2 \rho} \dot{\rho} \rho \dot{\rho} + \frac{\dot{\rho} \dot{\rho} / \partial (\ln h)}{2 \rho} \] (A15b)

A3 The limit \( k \rightarrow 0 \)
From equation (5) \( h^2 \cdot \nabla W(x, h) = -\partial (h^2 \cdot \nabla W(x, h)) \), such that
\[ v^2 \dot{\rho} \hat{\Pi}_i = \sum_j \dot{m}_j (x_j \cdot \nabla W(x_j, h)) \] (A18a)
Assuming local spatial isotropy we then get in the limit \( k \rightarrow 0 \):
\[ t \rightarrow -\Omega \dot{\rho} \cdot k, \] (A19a)
\[ t' \rightarrow -v (\hat{\Pi} - \hat{\Pi}) \dot{\rho} \cdot k, \] (A19b)
\[ T \rightarrow \frac{(2 \Omega + v \hat{\Pi}) \hat{\rho} \cdot k^2}{2 + v} + \frac{v \hat{\Pi} - \hat{\Pi}}{2(2 + v)} |k|^2 I. \] (A19c)
Inserting these into (A16) gives
\[ \dot{\xi} = -\dot{\rho} \cdot \frac{a T}{\rho} \Phi_i - \frac{2}{\rho} \frac{a T}{\rho^2} \Phi_i \left[ \frac{2 v}{2 + v} - 1 \right] a \cdot k \Phi_i + \frac{v k^2 a}{v + 2} \Phi_i. \] (A20)

This paper has been typeset from a \LaTeX\ file prepared by the author.