

A unified interpretation of stress in molecular systems

Nikhil Chandra Admal · E. B. Tadmor

The authors would like to dedicate this article to Jack Irving, who passed away in 2008 at the age of 87. Irving, while a graduate student on leave from Princeton, worked with Prof. John Kirkwood at Caltech on the fundamental non-equilibrium statistical mechanics theory which serves as the basis for the present article.

Received: 11 January 2010/ Published online: 27 May 2010

Abstract The microscopic definition for the Cauchy stress tensor has been examined in the past from many different perspectives. This has led to different expressions for the stress tensor and consequently the “correct” definition has been a subject of debate and controversy. In this work, a unified framework is set up in which all existing definitions can be derived, thus establishing the connections between them. The framework is based on the non-equilibrium statistical mechanics procedure introduced by Irving, Kirkwood and Noll, followed by spatial averaging. The Irving–Kirkwood–Noll procedure is extended to multi-body potentials with continuously differentiable extensions and generalized to *non-straight bonds*, which may be important for particles with internal structure. Connections between this approach and the direct spatial averaging approach of Murdoch and Hardy are discussed and the Murdoch–Hardy procedure is systematized. Possible sources of non-uniqueness of the stress tensor, resulting separately from both procedures, are identified and addressed. Numerical experiments using molecular dynamics and lattice statics are conducted to examine the behavior of the resulting stress definitions including their convergence with the spatial averaging domain size and their symmetry properties.

Keywords Stress · Microscopic Interpretation · Statistical Mechanics · Irving–Kirkwood–Noll procedure · Spatial averaging

PACS 74A25

This work was partly supported through NSF (DMS-0757355). This article has drawn heavily upon material from Ellad Tadmor and Ronald Miller, *Modeling Materials: Continuum, Atomistic and Multiscale Techniques*, ©2010 Ellad Tadmor and Ronald Miller, forthcoming Cambridge University Press, reproduced with permission. The final publication is available at www.springerlink.com.

Nikhil Chandra Admal
Department of Aerospace Engineering and Mechanics, The University of Minnesota.
E-mail: admal002@umn.edu

Ellad B. Tadmor
Department of Aerospace Engineering and Mechanics, The University of Minnesota.
E-mail: tadmor@aem.umn.edu

1 Introduction

Continuum mechanics provides an efficient theoretical framework for modeling materials science phenomena. To characterize the behavior of materials, *constitutive relations* serve as an input to the continuum theory. These constitutive models have functional forms which must be consistent with material frame-indifference and the laws of thermodynamics and include parameters that are fitted to reproduce experimental observations. With the advent of modern computing power, atomistic simulations through “numerical experiments” offer the potential for studying different materials and arriving at their constitutive laws from first principles. This could make it possible to design new materials and to improve the properties of existing materials in a systematic fashion. To use the data obtained from an atomistic simulation to build a constitutive law, which is framed in the language of continuum mechanics, it is necessary to understand the connection between continuum fields and the underlying microscopic dynamics.

Another arena where the connection between continuum and atomistic concepts is important is the field of *multiscale modeling*. This discipline involves the development of computational tools for studying problems where two or more length and/or time scales play a major role in determining macroscopic behavior. A prototypical example is fracture mechanics where the behavior of a crack is controlled by atomistic-scale phenomena at the crack-tip, while at the same time long-range elastic stress fields are set up in the body. Many advances have been made in the area of multiscale modeling in recent years. Some common atomistic/continuum coupling methods are quasicontinuum [53, 59], coupling of lengthscales [51], cluster quasicontinuum [31], bridging domain [66], coupled atomistics and discrete dislocations [54], and heterogeneous multiscale methods [15], to name just a few. Refer to [58] for a comparison of some prominent atomistic/continuum coupling multiscale methods. In a multiscale method, a key issue involves the transfer of information between the discrete model and the continuum model. It is therefore of practical interest to understand how to construct definitions of continuum fields for an atomistic system, to ensure a smooth transfer of information between the discrete and continuum domains.

In this paper, we focus on just one aspect of the continuum-atomistic connection, namely the interpretation of the (Cauchy) stress tensor in a discrete system. This question has been explored from many different perspectives for nearly two hundred years and this has led to various definitions that do not appear to be consistent with each other. As a result, the “correct” definition for the stress tensor has been a subject of great debate and controversy. We begin with a brief historical survey.

A brief history of microscopic definitions for the stress tensor

Interest in microscopic definitions for the stress tensor dates back at least to Cauchy in the 1820s [5, 6] with his aim to define stress in a crystalline solid. Cauchy’s original definition emerges from the intuitive idea of identifying stress with the force per unit area carried by the bonds that cross a given surface. A comprehensive derivation of Cauchy’s approach is given in Note B of Love’s classic book on the theory of elasticity [35]. Since this approach is tied to the particular surface being considered,

it actually constitutes a definition for the *traction* (or *stress vector*) and not for the stress tensor. The first definition of stress as a tensorial quantity follows from the works of Clausius [8] and Maxwell [40, 41] in the form of *virial theorem*. Clausius states the virial theorem as

The mean vis viva of a system is equal to its virial.

By “vis viva” (literally “living force”), Clausius means kinetic energy, while the term “virial” comes from the Latin “vis” (pl. “vires”) meaning force. The virial theorem leads to a definition for pressure in a gas. Maxwell [40, 41] extended Clausius’ work and showed the existence of a tensorial version of the virial theorem (see Appendix A). The *virial stress* resulting from the virial theorem is widely used even today in many atomistic simulations due to its simple form and ease of computation. Unlike Cauchy’s original definition for stress, the virial stress includes a contribution due to the kinetic energy of the particles. This discrepancy was addressed by Tsai [63], who extended the definition given by Cauchy to finite temperature by taking into consideration the momentum flux passing through the surface. Let us refer to this stress vector as the *Tsai traction*.

An alternative approach for defining the stress tensor was pioneered in the landmark paper of Irving and Kirkwood [27]. Irving and Kirkwood derived the equations of hydrodynamics from the principles of non-equilibrium classical statistical mechanics and in the process established a pointwise definition for various continuum fields including the stress tensor. Although their work was indeed noteworthy, the stress tensor obtained involved a series expansion of the Dirac delta distribution which is not mathematically rigorous. Continuing their work, Noll [48] proved two lemmas, which allowed him to avoid the use of the Dirac delta distribution, and thus arrive at a closed-form expression for the stress tensor which does not involve a series expansion. We refer to the procedure introduced by Irving and Kirkwood and extended by Noll as the *Irving–Kirkwood–Noll procedure*. Schofield and Henderson [52] highlighted the non-uniqueness present in the stress tensor derived by Irving and Kirkwood, and pointed out that it could result in a non-symmetric stress tensor. There have been several attempts to improve on the Irving and Kirkwood procedure. In particular, Lutsko [36] reformulated this procedure in Fourier space. An error in Lutsko’s derivation was corrected by Cormier et al. [9].

Due to the stochastic nature of the Irving and Kirkwood stress, many difficulties arise when one tries to use their expression in atomistic simulations. To avoid these difficulties, Hardy and co-workers [23, 24] and independently Murdoch [43–47] developed a new approach that bypasses the mathematical complexity of the Irving and Kirkwood procedure. This is done by defining continuum fields as direct spatial averages of the discrete equations of motion using weighting functions with compact support. In particular, this approach leads to the so-called *Hardy stress* [23] often used in molecular dynamics simulations. Murdoch in [45] provides an excellent description of the spatial averaging approaches currently being used and discusses the non-uniqueness of the stress tensor resulting from the spatial averaging procedure. We refer to the direct spatial averaging approach as the *Murdoch–Hardy procedure*.

Another approach, which leads to a stress tensor very similar to that obtained by Irving and Kirkwood is the reformulation of elasticity theory using peridynam-

ics [55]. Lehoucq and Silling [33] have recently shown that Noll’s solution is a minimum solution in a variational sense. Morante et al. [42] proposed a new approach for defining the stress tensor using the invariance of partition function under infinitesimal canonical point transformations. However, their approach is limited to equilibrium statistical mechanics and involves taking derivatives of delta distributions.

We can summarize the “state of the art” for the microscopic definition of the stress tensor as follows. There are currently at least three definitions for the stress tensor which are commonly used in atomistic simulations: the virial stress, the Tsai traction, and the Hardy stress [68]. The importance of the Irving and Kirkwood formulation is recognized, however, it is not normally used in practice and its connection with the other stress definitions is not commonly understood. The difference between *pointwise* stress measures and temporal and/or spatially-averaged quantities is often not fully appreciated. The result is that the connection between the Cauchy stress tensor defined in continuum mechanics and its analogue, defined for a discrete system, remains controversial and continues to be a highly-debated problem.

A unified framework for the microscopic definition for the stress tensor

In this paper, a unified framework based on the Irving–Kirkwood–Noll procedure is established which *leads to all of the major stress definitions* discussed above and identifies additional possible definitions. Since all of the definitions are obtained from a common framework the connections between them can be explored and analyzed and the uniqueness of the stress tensor can be established. An overview of the approach and the organization of the paper are described below.

Before turning to the general framework, we begin in Section 2 with a derivation of the virial stress tensor within the framework of equilibrium statistical mechanics using the technique of canonical transformations. Although this derivation is quite different from the Irving–Kirkwood–Noll procedure, it provides insight into how the geometric ideas of mechanics can be used to derive the stress tensor. It also provides a limit to which the general non-equilibrium stress tensor must converge under equilibrium conditions in the thermodynamic limit. This is used later to establish the uniqueness of the stress tensor obtained from our general unified framework.

Next, we turn to the construction of the new unified framework. In Section 3, we extend the Irving–Kirkwood–Noll procedure [27, 48], originally derived for pair potential interactions, to multi-body potentials. Due to the invariance of the potential energy function with respect to the Euclidean group, it can be shown that any multi-body potential can be expressed as a function of distances between particles. When expressed in this form, we note that for a system of more than 4 particles, this function is only defined on a manifold since the $N(N - 1)/2$ distances between N particles in \mathbb{R}^3 are not independent for $N \geq 5$. To apply the Irving–Kirkwood–Noll procedure to multi-body potentials, we recognize that the potential energy function must be *extended* from its manifold to a higher-dimensional Euclidean space as a continuously differentiable function. We show that if such an extension exists, then an infinite number of equivalent extensions can be constructed using *Cayley-Menger determinants*, which describe the constraints that the distances between particles embedded in \mathbb{R}^3 must satisfy. Then for multi-body potentials that possess continuously

differentiable extensions (which is the case for most practical interatomic potentials), we establish the key result that due to the balance of linear and angular momentum, *the force on a particle in a discrete system can always be decomposed as a sum of central forces between particles*, i.e., forces that are parallel to the lines connecting the particles. In other words, the *strong law of action and reaction* is always satisfied for such multi-body potentials. We show, that although the net force on a particle calculated using *any* extension is the same, its decomposition into central forces is generally different for different extensions. Using this result we show that the pointwise stress tensor resulting from the Irving–Kirkwood–Noll procedure is non-unique and symmetric. We also show, that a generalization of Noll’s lemmas [48] to *non-straight bonds* gives a non-symmetric stress tensor that may be important for particles with internal structure, such as liquid crystals.

The *macroscopic* stress tensor corresponding to the pointwise stress tensor described above is obtained in Section 4 through a procedure of spatial averaging. The connection between this stress and the stress tensors obtained via the direct spatial averaging procedure introduced by Murdoch [43–47] and Hardy [23] is explored and in the process the Murdoch–Hardy procedure is systematized and generalized to multi-body potentials using the results of Section 3. The non-uniqueness of the stress tensor, inherent in the Murdoch–Hardy procedure is studied and a general class of possible definitions under this procedure are identified. The connection between the non-uniqueness in the Murdoch–Hardy procedure and the non-uniqueness mentioned in Section 3 is addressed.

In Section 5, various stress definitions including the Hardy stress, the Tsai traction and the virial stress are shown to be special cases of the macroscopic stress tensor derived from the extended Irving–Kirkwood–Noll procedure in Section 4. The original definitions for these measures are generalized in this manner to multi-body potentials. The existence of different extensions for the potential energy function, which led to non-uniqueness of the pointwise stress tensor discussed in Section 3, also result in the non-uniqueness of these definitions. However it is shown that the difference in the macroscopic stress tensor resulting from this non-uniqueness tends to zero in the *thermodynamic limit*¹. Another source of non-uniqueness explored in this section is that given any definition for the stress tensor, a new definition, which also satisfies the balance of linear momentum, can be obtained by adding to it an arbitrary tensor field with zero divergence. It is shown that in the thermodynamic limit the macroscopic stress tensor obtained in Section 4 converges to the virial stress derived in Section 2.

To address practical aspects of the different definitions obtained within the unified framework, Section 6 describes several “numerical experiments” involving molecular dynamics and lattice statics. These simulations are designed to examine the behavior of these stress definitions, including their convergence with averaging domain size and their symmetry properties. Our conclusions and directions for future research are presented in Section 7.

¹ The thermodynamic limit is the state obtained as the number of particles, N , and the volume, V , of the system tend to infinity in such a way that the ratio N/V is constant.

Notation

In this paper, vectors are denoted by lower case letters in bold font and tensors of higher order are denoted by capital letters in bold font. The tensor product of two vectors is denoted by the symbol “ \otimes ” and the inner product of two vectors is denoted by a dot “ \cdot ”. The inner product of two second-order tensors is denoted by “ \cdot ”. A second-order tensor operating on a vector is denoted by juxtaposition, e.g., $\mathbf{T}\mathbf{v}$. The gradient of a vector field, $\mathbf{v}(\mathbf{x})$, is denoted by $\nabla_{\mathbf{x}}\mathbf{v}(\mathbf{x})$, which in indicial notation is given by $[\nabla_{\mathbf{x}}\mathbf{v}]_{ij} = \partial v_i / \partial x_j$. The divergence of a tensor field, $\mathbf{T}(\mathbf{x})$, is denoted by $\text{div}_{\mathbf{x}}\mathbf{T}(\mathbf{x})$. The divergence of a vector field is defined as the trace of its gradient. The divergence of a second-order tensor field in indicial notation (with Einstein’s summation convention) is given by $[\text{div}_{\mathbf{x}}\mathbf{T}]_i = \partial T_{ij} / \partial x_j$. The notation described above is followed unless otherwise explicitly stated.

2 Stress in an equilibrium system

In this section, we obtain expressions for the Cauchy stress in an equilibrium system using the technique of canonical transformations. The basic philosophy behind canonical transformation is explained in the next section.

2.1 Canonical transformations

Consider a system consisting of N point masses whose behavior is governed by classical mechanics. Let $\mathbf{q}_{\alpha}(t)$ and $\mathbf{p}_{\alpha}(t)$ ($\alpha = 1, 2, \dots, N$) denote the generalized coordinates and momenta of the system.² For brevity, we sometimes use $\mathbf{q}(t)$ and $\mathbf{p}(t)$ to denote the vectors $(\mathbf{q}_1(t), \mathbf{q}_2(t), \dots, \mathbf{q}_N(t))$ and $(\mathbf{p}_1(t), \mathbf{p}_2(t), \dots, \mathbf{p}_N(t))$, respectively. The time evolution of the system can be studied through three well-known approaches, referred to as the *Newtonian formulation*, the *Lagrangian formulation*, and the *Hamiltonian formulation*. The first approach is used in molecular dynamics simulations, while the latter two approaches are more elegant and can sometimes be used to obtain useful information from systems in the absence of closed-form solutions.

In the Lagrangian formulation, a system is characterized by the vector $\mathbf{q}(t)$ and a Lagrangian function \mathcal{L} , given by

$$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}; t) = \mathcal{T}(\dot{\mathbf{q}}) - \mathcal{V}(\mathbf{q}), \quad (2.1)$$

where \mathcal{T} is the kinetic energy of the system, \mathcal{V} is the potential energy of the system, and $\dot{\mathbf{q}}(t)$ represents the time derivative of $\mathbf{q}(t)$. It is useful to think of \mathbf{q} as a point in a $3N$ -dimensional *configuration space*. The time evolution of $\mathbf{q}(t)$ in configuration space is described by a variational principle called *Hamilton’s principle*. Hamilton’s

²In a general theory of canonical transformations, \mathbf{q}_{α} and \mathbf{p}_{α} need not denote the actual position and momentum of particle α .

principle states that the time evolution of $\mathbf{q}(t)$ corresponds to the extremum of the action integral defined as a functional of \mathbf{q} by

$$\mathcal{A}[\mathbf{q}] = \int_{t_1}^{t_2} \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}; t) dt, \quad (2.2)$$

where t_1 , t_2 , $\mathbf{q}(t_1)$ and $\mathbf{q}(t_2)$ are held fixed with respect to the class of variations being considered [32, Section V.1]. In mathematical terms, we require that

$$\delta \mathcal{A} = 0, \quad (2.3)$$

while keeping the ends fixed as described above. The Euler–Lagrange equation arising from (2.3) is

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}_\alpha} \right) - \frac{\partial \mathcal{L}}{\partial \mathbf{q}_\alpha} = \mathbf{0}. \quad (2.4)$$

The Lagrangian formulation is commonly used as a calculation tool in solving simple problems.

Next, we note that the Lagrangian is the Legendre transform of the Hamiltonian \mathcal{H} , [32, Section VI.2],

$$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}; t) = \sup_{\mathbf{p}} [\mathbf{p} \cdot \dot{\mathbf{q}} - \mathcal{H}(\mathbf{p}, \mathbf{q}; t)]. \quad (2.5)$$

The Hamiltonian is the total energy of the system. Using the Hamiltonian, equation (2.3) can be rewritten as

$$\delta \int_{t_1}^{t_2} [\mathbf{p} \cdot \dot{\mathbf{q}} - \mathcal{H}(\mathbf{p}, \mathbf{q}; t)] dt = 0. \quad (2.6)$$

Note that in (2.3), the variation is only with respect to \mathbf{q} , whereas in (2.6), the functional depends on the functions \mathbf{q} and \mathbf{p} , and variations are taken with respect to both \mathbf{q} and \mathbf{p} independently. In both cases, t_1 , t_2 , $\mathbf{q}(t_1)$ and $\mathbf{q}(t_2)$ are held fixed. The variational principle given in (2.6) is commonly referred as the *modified Hamilton’s principle* [22] or simply as the “Hamiltonian formulation”. The advantage of the Hamiltonian formulation lies not in its use as a calculation tool, but rather in the deeper insight it affords into the formal structure of mechanics. The Euler–Lagrange equations associated with (2.6) are

$$\dot{\mathbf{q}}_\alpha = \nabla_{\mathbf{p}_\alpha} \mathcal{H}, \quad (2.7)$$

$$\dot{\mathbf{p}}_\alpha = -\nabla_{\mathbf{q}_\alpha} \mathcal{H}, \quad (2.8)$$

commonly called Hamilton’s equations. The above equations are also referred to as the *canonical equations of motion*³.

It is important to note that the Hamiltonian formulation is more general than the Lagrangian formulation, since it accords the coordinates and momenta independent status, thus providing the analyst with far greater freedom in selecting generalized

³The term “canonical” in this context has nothing to do with the canonical ensemble of statistical mechanics. The terminology was introduced by Jacobi to indicate that Hamilton’s equations constitute the simplest form of the equations of motion.

coordinates. We now think of (\mathbf{q}, \mathbf{p}) as a point in a $6N$ -dimensional *phase space*, as opposed to the $3N$ -dimensional configuration space of the Lagrangian formulation. The choice of \mathbf{q} and \mathbf{p} is not arbitrary, however, since the selected variables must satisfy the canonical equations of motion. For this reason \mathbf{q} and \mathbf{p} are called *canonical variables*.

The requirement that the generalized coordinates and momenta must be canonical means that new sets of generalized coordinates can be derived from a given set through a special kind of transformation defined below.

Definition 1 *Any transformation of generalized coordinates that preserves the canonical form of Hamilton's equations is said to be a canonical transformation.*⁴

The construction of canonical transformations is facilitated by the introduction of *generating functions* as explained below.

Generating functions

Consider two sets of canonical variables (\mathbf{Q}, \mathbf{P}) and (\mathbf{q}, \mathbf{p}) , related to each other through a canonical transformation given by

$$\mathbf{Q} = \mathbf{Q}(\mathbf{q}, \mathbf{p}, t), \quad \mathbf{P} = \mathbf{P}(\mathbf{q}, \mathbf{p}, t). \quad (2.9)$$

Since the variables are canonical, they satisfy the modified Hamilton's principle in (2.6),

$$\delta \int_{t_1}^{t_2} [\mathbf{p} \cdot \dot{\mathbf{q}} - \mathcal{H}(\mathbf{p}, \mathbf{q}; t)] dt = 0, \quad (2.10)$$

$$\delta \int_{t_1}^{t_2} [\mathbf{P} \cdot \dot{\mathbf{Q}} - \hat{\mathcal{H}}(\mathbf{P}, \mathbf{Q}; t)] dt = 0, \quad (2.11)$$

where $\hat{\mathcal{H}}$ is defined later as part of the canonical transformation. The integrands of (2.10) and (2.11) can therefore only differ by a quantity whose variation after integration is identically zero. A possible solution is

$$\delta \int_{t_1}^{t_2} [\mathbf{p} \cdot \dot{\mathbf{q}} - \mathbf{P} \cdot \dot{\mathbf{Q}} - (\mathcal{H} - \hat{\mathcal{H}})] dt = \delta \int_{t_1}^{t_2} \frac{dG}{dt} dt, \quad (2.12)$$

where G is an arbitrary scalar function of the canonical variables and time, with continuous second derivatives. The integral on the right is only evaluated at fixed integration bounds and its variation is zero. This is not obvious since there is no restriction on the variation of the momenta at the ends. We assume this to be true to avoid the introduction of differential forms. For a mathematically rigorous argument refer to [1, Section 45]⁵. The difference between the integrands of (2.10) and (2.11) therefore satisfies,

$$dG - \mathbf{p} \cdot d\mathbf{q} + \mathbf{P} \cdot d\mathbf{Q} + (\mathcal{H} - \hat{\mathcal{H}})dt = 0. \quad (2.13)$$

⁴This definition suffices for our purpose, but a more correct definition can be found in [1] using *differential forms*.

⁵Briefly the proof is based on the symmetry present in the geometry of any Hamiltonian system commonly called *symplectic geometry*.

Now, consider the case where $G = G_1(\mathbf{q}, \mathbf{Q}, t)$. The total differential of G is then

$$dG = \nabla_{\mathbf{q}} G_1 \cdot d\mathbf{q} + \nabla_{\mathbf{Q}} G_1 \cdot d\mathbf{Q} + \frac{\partial G_1}{\partial t} dt. \quad (2.14)$$

Substituting (2.14) into (2.13) gives

$$(\nabla_{\mathbf{q}} G_1 - \mathbf{p}) \cdot d\mathbf{q} + (\nabla_{\mathbf{Q}} G_1 + \mathbf{P}) \cdot d\mathbf{Q} + \left(\frac{\partial G_1}{\partial t} + \mathcal{H} - \hat{\mathcal{H}} \right) dt = 0. \quad (2.15)$$

Since \mathbf{q} , \mathbf{Q} and t are independent, the above equation is satisfied provided that

$$\mathbf{p}_\alpha = \frac{\partial G_1}{\partial \mathbf{q}_\alpha}, \quad \mathbf{P}_\alpha = -\frac{\partial G_1}{\partial \mathbf{Q}_\alpha}, \quad \hat{\mathcal{H}} = \mathcal{H} + \frac{\partial G_1}{\partial t}. \quad (2.16)$$

The above relations define the canonical transformation. Since G_1 generates the transformation, it is commonly called the *generating function* of the canonical transformation. Note that if G_1 does not depend on time t , then $\hat{\mathcal{H}} = \mathcal{H}$.

The generating functions of the form, $G = G_1(\mathbf{q}, \mathbf{Q}, t)$, does not generate all possible canonical transformations. In general, there are four primary classes of generating functions where the functional dependence is (\mathbf{q}, \mathbf{Q}) , (\mathbf{q}, \mathbf{P}) , (\mathbf{p}, \mathbf{Q}) and (\mathbf{p}, \mathbf{P}) .⁶ We have already encountered the first class, where $G = G_1(\mathbf{q}, \mathbf{Q}, t)$. The remaining classes can be obtained from the first through Legendre transformations. Consider for example, the following definition,

$$G = G_3(\mathbf{p}, \mathbf{Q}, t) + \mathbf{q} \cdot \mathbf{p}. \quad (2.17)$$

The total differential of this expression is

$$dG = \nabla_{\mathbf{p}} G_3 \cdot d\mathbf{p} + \nabla_{\mathbf{Q}} G_3 \cdot d\mathbf{Q} + \frac{\partial G_3}{\partial t} dt + \mathbf{q} \cdot d\mathbf{p} + \mathbf{p} \cdot d\mathbf{q}. \quad (2.18)$$

Substituting the above equation into (2.13) gives

$$(\nabla_{\mathbf{p}} G_3 + \mathbf{q}) \cdot d\mathbf{p} + (\nabla_{\mathbf{Q}} G_3 + \mathbf{P}) \cdot d\mathbf{Q} + \left(\frac{\partial G_3}{\partial t} + \mathcal{H} - \hat{\mathcal{H}} \right) dt = 0, \quad (2.19)$$

which leads to the following canonical transformation:

$$\mathbf{q}_\alpha = -\frac{\partial G_3}{\partial \mathbf{p}_\alpha}, \quad \mathbf{P}_\alpha = -\frac{\partial G_3}{\partial \mathbf{Q}_\alpha}, \quad \hat{\mathcal{H}} = \mathcal{H} + \frac{\partial G_3}{\partial t}. \quad (2.20)$$

The other two classes of transformation can be derived in a similar way.

Finally, an important property of a canonical transformation is that it preserves the volume of any element in phase space, i.e., $d\mathbf{q}d\mathbf{p} = d\mathbf{Q}d\mathbf{P}$ [22, page 402]. This means that for a change of variables between (\mathbf{p}, \mathbf{q}) and (\mathbf{P}, \mathbf{Q}) , the Jacobian of the transformation is unity.

⁶In addition to these four classes of transformation, it is possible to have a mixed dependence, where each degree of freedom can belong to a different class [22].

2.2 A derivation of the stress tensor under equilibrium conditions

In this section, we use the method of canonical transformations to derive an expression for the Cauchy stress tensor. In continuum mechanics, a body is identified with a regular region of Euclidean space \mathcal{E} referred to as the reference configuration. Any point $\mathbf{X} \in \mathcal{B}$ is referred to as a material point. The body \mathcal{B} is deformed via a smooth, one-to-one mapping $\varphi : \mathcal{E} \rightarrow \mathcal{E}$, which maps each $\mathbf{X} \in \mathcal{B}$ to a point,

$$\mathbf{x} = \varphi(\mathbf{X}), \quad (2.21)$$

in the deformed configuration,⁷ where we have assumed that the deformation is independent of time. The deformation gradient \mathbf{F} is defined as

$$\mathbf{F}(\mathbf{X}) = \nabla_{\mathbf{X}} \varphi. \quad (2.22)$$

The mapping φ is assumed to satisfy the condition that $\det \mathbf{F}$ is strictly positive. The Cauchy stress, $\boldsymbol{\sigma}$, is defined by [37]

$$\boldsymbol{\sigma}(\mathbf{T}, \mathbf{F}) = \frac{1}{\det \mathbf{F}} \nabla_{\mathbf{F}} \psi \mathbf{F}^T, \quad (2.23)$$

where $\psi(\mathbf{T}, \mathbf{F})$ is the Helmholtz free energy density function relative to the reference configuration. We are only focusing on a conservative elastic body.

A system in thermodynamic equilibrium⁸ can by definition only support a uniform state of deformation. Therefore, our material system is deformed via the affine mapping⁹

$$\mathbf{q}_{\alpha} = \mathbf{F} \mathbf{Q}_{\alpha}. \quad (2.24)$$

⁷We adopt the continuum mechanics convention of denoting variables in the reference configuration with upper-case letters, and variables in the deformed configuration with lower-case letters.

⁸A system is said to be in a state of thermodynamic equilibrium when all of its properties are independent of time and all of its intensive properties are independent of position [65]. To stress this, the term *uniform state of thermodynamic equilibrium* is sometimes used to describe this state.

⁹To understand this mapping, consider a system of N particles with positions \mathbf{q}_{α} ($\alpha = 1, 2, \dots, N$) confined to a parallelepiped container defined by the three linearly independent vectors \mathbf{l}_1 , \mathbf{l}_2 and \mathbf{l}_3 , which need not be orthogonal. This selection is done for convenience and does not limit the generality of the derivation as explained below. The position of a particle in the container can be expressed in terms of scaled coordinates $\xi_i^{\alpha} \in [0, 1]$ as

$$\mathbf{q}_{\alpha} = \xi_i^{\alpha} \mathbf{l}_i, \quad (*)$$

where Einstein's summation convention is applied to spatial indices. The deformation of the container is defined relative to a reference configuration where the cell vectors are \mathbf{L}_1 , \mathbf{L}_2 and \mathbf{L}_3 . The current and reference cell vectors are related through an affine mapping defined by \mathbf{F} ,

$$\mathbf{l}_i = \mathbf{F} \mathbf{L}_i. \quad (**)$$

Equations (*) and (**) can be combined to relate the position \mathbf{q}_{α} of particle α in the deformed configuration with its position in the reference configuration \mathbf{Q}_{α} .

$$\mathbf{q}_{\alpha} = \xi_i^{\alpha} (\mathbf{F} \mathbf{L}_i) = \mathbf{F} (\xi_i^{\alpha} \mathbf{L}_i) = \mathbf{F} \mathbf{Q}_{\alpha}. \quad (***)$$

This is exactly the mapping defined in (2.24). It provides a direct relationship between the positions of particles in the reference configuration and their position in the deformed configuration. Note that the assumed (parallelepiped) shape of the container does not enter into the relation, $\mathbf{q}_{\alpha} = \mathbf{F} \mathbf{Q}_{\alpha}$, which means that this relation holds for a container of any shape.

It is clear that if we enforce this mapping on our system, with no change in the momentum coordinates, then the newly obtained variables will not satisfy Hamilton's equations. Therefore any change of variables should be governed by a canonical transformation. The following generator function provides the desired canonical transformation

$$G_3(\mathbf{p}, \mathbf{Q}) = - \sum_{\alpha} \mathbf{p}_{\alpha} \cdot \mathbf{F} \mathbf{Q}_{\alpha}. \quad (2.25)$$

Substituting this generating function into (2.20) gives

$$\mathbf{q}_{\alpha} = - \frac{\partial G_3}{\partial \mathbf{p}_{\alpha}} = \mathbf{F} \mathbf{Q}_{\alpha}, \quad \mathbf{P}_{\alpha} = - \frac{\partial G_3}{\partial \mathbf{Q}_{\alpha}} = \mathbf{F}^{\top} \mathbf{p}_{\alpha}, \quad \hat{\mathcal{H}} = \mathcal{H}. \quad (2.26)$$

The first relation in the above equation is the desired transformation in (2.24). The second relation is the corresponding transformation that the momentum degrees of freedom must satisfy, so that the new set of coordinates (\mathbf{Q}, \mathbf{P}) are canonical. The third relation refers to the Hamiltonian of the system, which is assumed to be given by

$$\mathcal{H}(\mathbf{p}, \mathbf{q}) = \sum_{\alpha=1}^N \frac{\mathbf{p}_{\alpha} \cdot \mathbf{p}_{\alpha}}{2m_{\alpha}} + \mathcal{V}(\mathbf{q}_1, \dots, \mathbf{q}_N), \quad (2.27)$$

where \mathcal{V} denotes the potential energy of the system. Expressed in terms of the reference variables, (2.27) becomes

$$\begin{aligned} \hat{\mathcal{H}}(\mathbf{P}, \mathbf{Q}, \mathbf{F}) &= \mathcal{H}(\mathbf{p}(\mathbf{Q}, \mathbf{P}, \mathbf{F}), \mathbf{q}(\mathbf{Q}, \mathbf{P}, \mathbf{F})) \\ &= \sum_{\alpha=1}^N \frac{\mathbf{F}^{-\top} \mathbf{P}_{\alpha} \cdot \mathbf{F}^{-\top} \mathbf{P}_{\alpha}}{2m_{\alpha}} + \mathcal{V}(\mathbf{F} \mathbf{Q}_1, \dots, \mathbf{F} \mathbf{Q}_N). \end{aligned} \quad (2.28)$$

We now proceed to derive the expression for the Cauchy stress tensor using (2.23). The Helmholtz free energy density for the canonical ensemble is given by [26]

$$\psi(T, \mathbf{F}) = - \frac{k_B T \ln Z}{V_0}, \quad (2.29)$$

where k_B is the Boltzmann's constant, T is the absolute temperature, V_0 is the volume of the body in the reference configuration, and $Z(T, \mathbf{F})$ is the *partition function* defined as

$$Z(T, \mathbf{F}) := \frac{1}{N! h^{3N}} \int_{\Gamma_0} e^{-\hat{\mathcal{H}}/k_B T} d\mathbf{P} d\mathbf{Q}, \quad (2.30)$$

It is important to note that (***) does *not* impose a kinematic constraint that dictates the position of particle α in the deformed configuration based on its position in the reference configuration (as does the Cauchy–Born rule used in multiscale methods [58]). We will see later that this will merely be used as a change of variables, where, instead of integrating over the deformed configuration with the variables \mathbf{q} , the integration is carried out over a given reference configuration using the variables \mathbf{Q} . In both cases the same result is obtained. However, by using the referential variables the dependence on the deformation gradient is made explicit.

where h is Planck's constant and Γ_0 denotes the phase space in the reference configuration. With this definition, the statistical mechanics phase average of a function $A(\mathbf{P}, \mathbf{Q})$ in the canonical ensemble is

$$\langle A \rangle(T, \mathbf{F}) = \int_{\Gamma_0} A(\mathbf{P}, \mathbf{Q}) W_c(\mathbf{P}, \mathbf{Q}, T, \mathbf{F}) d\mathbf{P} d\mathbf{Q}, \quad (2.31)$$

where

$$W_c(\mathbf{P}, \mathbf{Q}, T, \mathbf{F}) = \frac{1}{N!h^{3N}Z} e^{-\hat{\mathcal{H}}(\mathbf{P}, \mathbf{Q}, \mathbf{F})/k_B T} \quad (2.32)$$

is the canonical distribution function. Substituting (2.29) and (2.30) into (2.23), we obtain

$$\boldsymbol{\sigma} = -\frac{k_B T}{(\det \mathbf{F}) V_0 Z} \nabla_{\mathbf{F}} Z \mathbf{F}^T = \frac{1}{V} \left\langle \nabla_{\mathbf{F}} \hat{\mathcal{H}} \right\rangle \mathbf{F}^T, \quad (2.33)$$

where in the last step we have used the identity $V = (\det \mathbf{F}) V_0$ and where

$$\begin{aligned} \nabla_{\mathbf{F}} Z &= \frac{\partial}{\partial \mathbf{F}} \left[\frac{1}{N!h^{3N}} \int_{\Gamma_0} e^{-\hat{\mathcal{H}}/k_B T} d\mathbf{Q} d\mathbf{P} \right] \\ &= -\frac{1}{k_B T N!h^{3N}} \int_{\Gamma_0} \nabla_{\mathbf{F}} \hat{\mathcal{H}} e^{-\hat{\mathcal{H}}/k_B T} d\mathbf{Q} d\mathbf{P}. \end{aligned} \quad (2.34)$$

Next, we compute $\nabla_{\mathbf{F}} \hat{\mathcal{H}}$. In our derivation, we make use of indicial notation and the Einstein summation rule. To accommodate for the spatial indices, we push α representing the particle to the superscript position. Following this adjustment, we have

$$\frac{\partial \hat{\mathcal{H}}}{\partial F_{iJ}} = \frac{\partial}{\partial F_{iJ}} \left[\sum_{\alpha} \frac{p_k^{\alpha} p_k^{\alpha}}{2m^{\alpha}} + \mathcal{V}(\mathbf{q}^1, \dots, \mathbf{q}^N) \right] = \sum_{\alpha} \left[\frac{1}{m^{\alpha}} \frac{\partial p_k^{\alpha}}{\partial F_{iJ}} p_k^{\alpha} + \frac{\partial \mathcal{V}}{\partial q_k^{\alpha}} \frac{\partial q_k^{\alpha}}{\partial F_{iJ}} \right]. \quad (2.35)$$

From (2.26), we have

$$\frac{\partial q_k^{\alpha}}{\partial F_{iJ}} = \frac{\partial}{\partial F_{iJ}} (F_{kL} Q_L^{\alpha}) = \delta_{ik} Q_J^{\alpha}, \quad (2.36)$$

$$\frac{\partial p_k^{\alpha}}{\partial F_{iJ}} = \frac{\partial}{\partial F_{iJ}} (F_{Lk}^{-1} P_L^{\alpha}) = -F_{Jk}^{-1} F_{Li}^{-1} P_L^{\alpha} = -F_{Jk}^{-1} p_i^{\alpha}, \quad (2.37)$$

where in (2.37), we have used the following identity:

$$\frac{\partial F_{Lk}^{-1}}{\partial F_{iJ}} = -F_{Li}^{-1} F_{Jk}^{-1}. \quad (2.38)$$

Substituting (2.36) and (2.37) into (2.35), we have

$$\frac{\partial \hat{\mathcal{H}}}{\partial F_{iJ}} = -\sum_{\alpha} \left[\frac{p_i^{\alpha} F_{Jk}^{-1} p_k^{\alpha}}{m^{\alpha}} + f_{\alpha,i}^{\text{int}} Q_J^{\alpha} \right], \quad (2.39)$$

where $\mathbf{f}_\alpha^{\text{int}} = -\partial\mathcal{V}/\partial\mathbf{r}_\alpha$ is the internal force, defined in the deformed configuration, on particle α .¹⁰ In direct notation, we have

$$\nabla_{\mathbf{F}} \hat{\mathcal{H}} = - \sum_{\alpha} \left[\frac{\mathbf{p}_\alpha \otimes \mathbf{F}^{-1} \mathbf{p}_\alpha}{m_\alpha} + \mathbf{f}_\alpha^{\text{int}} \otimes \mathbf{Q}_\alpha \right]. \quad (2.40)$$

Substituting the above equation into (2.33) and using (2.26), we obtain an expression for the Cauchy stress:

$$\boldsymbol{\sigma}(T, \mathbf{F}) = -\frac{1}{V} \sum_{\alpha} \left\langle \frac{\mathbf{p}_\alpha \otimes \mathbf{p}_\alpha}{m_\alpha} + \mathbf{f}_\alpha^{\text{int}} \otimes \mathbf{q}_\alpha \right\rangle, \quad (2.41)$$

where the phase averaging is now being performed with respect to the variables \mathbf{p} and \mathbf{q} . The switch from phase averaging over \mathbf{P} and \mathbf{Q} in (2.31) to \mathbf{p} and \mathbf{q} above can be made because canonical transformations preserve the volume element in phase space as explained at the end of Section 2.1.

The expression in (2.41) for the Cauchy stress tensor is called the *virial stress*. A simpler derivation of the virial stress, based on time averages, is given in Appendix A. Although, the derivation here made use of the canonical ensemble, it is expected to apply to any ensemble in the thermodynamic limit (see footnote 1 on page 5) where all ensembles are equivalent. Continuum mechanics also tells us that the Cauchy stress tensor is symmetric, something that is not evident from the above equation. Discussion of the symmetry of the stress tensor, which hinges on an important property of $\mathbf{f}_\alpha^{\text{int}}$, is postponed to Section 5.

The virial stress defined above corresponds to the macroscopic stress tensor only under conditions of thermodynamic equilibrium in the thermodynamic limit. We now show that this expression for the stress tensor, as well as all other expressions in common use, can be derived as limiting cases of a more general formulation which begins with the Irving–Kirkwood–Noll procedure. We refer to this as the “unified framework” for the stress tensor.

3 Continuum fields as phase averages

In this section, we discuss the Irving and Kirkwood procedure [27], which laid the foundation for the microscopic definition of continuum fields for non-equilibrium systems. This work was later extended by Walter Noll [48]¹¹, who showed how closed-form analytical solutions can be obtained for the definition of certain continuum fields, which otherwise involved a non-rigorous¹² series expansion of the Dirac

¹⁰There is a subtle point here. Since we are using the canonical ensemble, the Hamiltonian \mathcal{H} neglects the interaction term of the system with the surrounding “heat bath”. This means that the potential energy \mathcal{V} in \mathcal{H} only includes the *internal* energy of the system and, therefore, its derivative with respect to the position of particle α gives the force $\mathbf{f}_\alpha^{\text{int}}$ on this particle due to its interactions with other particles in the system.

¹¹An English translation of this article appears in the current issue of the *Journal of Elasticity*.

¹²The derivation is non-rigorous in the sense that expressing the stress tensor as a series expansion is only possible when the probability density function, which is used in the derivation, is an analytic function of the spatial variables [48].

delta distribution in the original procedure. We refer to the procedure proposed by Noll in [48] as the *Irving–Kirkwood–Noll procedure*. The derivation presented in this section largely follows that of Noll [48], but extends it to more general atomistic models.

Consider a system \mathcal{M} modeled as a collection of N point masses/particles, each particle referred to as α ($\alpha = 1, 2, \dots, N$). We use the terms “particle” and “atom” interchangeably. The position, mass and velocity of particle α are denoted by \mathbf{x}_α , m_α and \mathbf{v}_α , respectively. The complete microscopic state of the system is known, at any instant of time, from the knowledge of position and velocity of each particle in \mathbb{R}^3 . Hence, the state of the system at time t , may be represented by a point $\Xi(t)$ in a $6N$ -dimensional phase space¹³. Let Γ denote the phase space. Therefore any point $\Xi(t) \in \Gamma$, can be represented as,

$$\begin{aligned}\Xi(t) &= (\mathbf{x}_1(t), \mathbf{x}_2(t), \dots, \mathbf{x}_N(t); \mathbf{v}_1(t), \mathbf{v}_2(t), \dots, \mathbf{v}_N(t)) \\ &=: (\mathbf{x}(t); \mathbf{v}(t)).\end{aligned}\tag{3.1}$$

In reality, the microscopic state of the system is never known to us, and the only observables identified are the macroscopic fields as defined in continuum mechanics. We identify the continuum fields with macroscopic observables obtained in a two-step process: (1) a pointwise field is obtained as a statistical mechanics phase average; (2) a macroscopic field is obtained as a spatial average over the pointwise field. The phase averaging in step (1) is done with respect to a probability density function $W : \Gamma \times \mathbb{R}^+ \rightarrow \mathbb{R}^+$ of class C^1 defined on all phase space for all t (W_c , defined in (2.32), is an example of a stationary (time-independent) probability density function defined for the canonical ensemble). The explicit dependence of W on time t , indicates that our system need not be in thermodynamic equilibrium.

As discussed in Section 2, the evolution of $\Xi(t)$ in the phase space is given by the following set of $2N$ first-order equations (Hamilton’s equations of (2.7)–(2.8)):

$$\dot{\mathbf{p}} = -\nabla_{\mathbf{x}} \mathcal{H},\tag{3.2a}$$

$$\dot{\mathbf{x}} = \nabla_{\mathbf{p}} \mathcal{H},\tag{3.2b}$$

where $\mathbf{p} := (\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N)$, \mathbf{p}_α denotes the momentum of each particle, and $\mathcal{H}(\mathbf{p}, \mathbf{x})$ is the Hamiltonian of the system.

The basic idea behind the original Irving and Kirkwood procedure is to prescribe/derive microscopic definitions for continuum fields, such that they are consistent with the balance laws of mass, momentum and energy. To arrive at these definitions, we repeatedly use the following theorem, commonly referred to as *Liouville’s theorem*, which relates to the conservation of volume in phase space.

As a system evolves, the phase space Γ is mapped into itself at every instant of time, and this mapping is governed by (3.2). If g_t denotes this mapping, then Liouville’s theorem essentially says that for any subset U of Γ , the volume of U remains invariant under the mapping g_t . This can be formally stated as,

¹³The usual convention is to represent the phase space via positions and momenta of the particles. For convenience, in this section, we represent the phase space via positions and velocities of the particles.

Liouville's Theorem For any $U \subseteq \Gamma$, volume is preserved under the one-parameter group of transformations of phase space, $g_t : U \rightarrow \Gamma$, given by the mapping

$$(\mathbf{x}(0), \mathbf{p}(0)) \mapsto (\mathbf{x}(t), \mathbf{p}(t)),$$

where $\mathbf{x}(t)$ and $\mathbf{p}(t)$ are solutions of the Hamilton's system of equations (3.2), i.e.,

$$\text{vol}(U) = \text{vol}(g_t U). \quad (3.3)$$

Proof Let $\overline{\dot{\text{vol}}(g_t U)}$ denote the material time derivative of $\text{vol}(g_t U)$ in the sense that $\Xi(0)$ is held fixed while performing this differentiation. Then we have,

$$\overline{\dot{\text{vol}}(g_t U)} = \overline{\int_{g_t U} d\Xi(t)} = \int_U (\overline{\det \dot{\mathbf{F}}}) d\Xi_0,$$

where $\mathbf{F}(\Xi_0, t) := \nabla_{\Xi_0} \Xi(\Xi_0, t)$, $\Xi(\Xi_0, t) = g_t(\Xi_0)$ and $\Xi_0 = \Xi(0)$. Using the fact that

$$\overline{\det \dot{\mathbf{F}}} = (\det \mathbf{F}) \text{tr}(\dot{\mathbf{F}} \mathbf{F}^{-1}),$$

we obtain

$$\overline{\dot{\text{vol}}(g_t U)} = \int_U (\det \mathbf{F}) \text{tr}(\dot{\mathbf{F}} \mathbf{F}^{-1}) d\Xi_0. \quad (3.4)$$

Let

$$\dot{\Xi}(\Xi) := \left. \frac{d\Xi}{dt} \right|_{\Xi_0 = g_t^{-1}(\Xi)}. \quad (3.5)$$

From chain rule, we have

$$\nabla \dot{\Xi} = \left. \frac{d(\nabla \Xi)}{dt} \right|_{\Xi_0 = g_t^{-1}(\Xi)} \nabla_{\Xi} \Xi_0 = \dot{\mathbf{F}} \mathbf{F}^{-1}. \quad (3.6)$$

Therefore $\text{div} \dot{\Xi} = \text{tr}(\dot{\mathbf{F}} \mathbf{F}^{-1})$. Equation (3.4) can now be rewritten as

$$\overline{\dot{\text{vol}}(g_t U)} = \int_U (\det \mathbf{F}) (\text{div} \dot{\Xi})|_{\Xi(t)=g_t(\Xi_0)} d\Xi_0. \quad (3.7)$$

But from (3.1) and (3.2) we also have,

$$\text{div} \dot{\Xi} = \text{div}_{\mathbf{x}} \dot{\mathbf{x}} + \text{div}_{\mathbf{p}} \dot{\mathbf{p}} = \text{div}_{\mathbf{x}} (\nabla_{\mathbf{p}} \mathcal{H}) - \text{div}_{\mathbf{p}} (\nabla_{\mathbf{x}} \mathcal{H}) = 0.$$

Therefore $\overline{\dot{\text{vol}}(g_t U)} = 0$ for arbitrary t . Thus (3.3) holds. \square

Let $W(\Xi; t)$ denote the probability density function defined on $g_t(\Gamma)$. Hence, we have

$$\int_{g_t U} W(\Xi(t); t) d\Xi(t) = \int_U W(\Xi_0; 0) (\det \mathbf{F}) d\Xi_0. \quad (3.8)$$

As a consequence of Liouville's theorem, we have, $\det \mathbf{F} = 1$. Therefore

$$\frac{d}{dt} \int_{g_t U} W(\Xi(t); t) d\Xi(t) = 0. \quad (3.9)$$

Since (3.9) holds for all $U \subseteq \Gamma$, we have $\dot{W}(\Xi(t); t) = 0$. Hence, the time evolution of the probability density function is given by

$$\frac{\partial W}{\partial t} + \sum_{\alpha=1}^N [\mathbf{v}_\alpha \cdot \nabla_{\mathbf{x}_\alpha} W + \dot{\mathbf{v}}_\alpha \cdot \nabla_{\mathbf{v}_\alpha} W] = 0. \quad (3.10)$$

The above equation can be rewritten as

$$\frac{\partial W}{\partial t} + \sum_{\alpha=1}^N \left[\mathbf{v}_\alpha \cdot \nabla_{\mathbf{x}_\alpha} W - \frac{\nabla_{\mathbf{x}_\alpha} \mathcal{V}}{m_\alpha} \cdot \nabla_{\mathbf{v}_\alpha} W \right] = 0, \quad (3.11)$$

where, as before, $\mathcal{V}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$ denotes the potential energy of the system. Equation (3.11) is called *Liouville's equation*.

3.1 Phase averaging

Under the Irving–Kirkwood–Noll procedure, pointwise fields are defined as phase averages. This phase averaging is expressed via weighted marginal densities. For example, the pointwise mass density field is defined as

$$\rho(\mathbf{x}, t) := \sum_{\alpha} m_{\alpha} \int_{\mathbb{R}^{3N} \times \mathbb{R}^{3N}} W \delta(\mathbf{x}_{\alpha} - \mathbf{x}) d\mathbf{x} d\mathbf{v}, \quad (3.12)$$

where the integral represents a marginal density defined on \mathbb{R}^3 , δ denotes the Dirac delta distribution, and \sum_{α} denotes summation from $\alpha = 1$ to N . To avoid the Dirac delta distribution and for greater clarity we adopt Noll's notation as originally used in [48]. Hence (3.12) can be rewritten as

$$\begin{aligned} \rho(\mathbf{x}, t) &= \sum_{\alpha} m_{\alpha} \int W d\mathbf{x}_1 \dots d\mathbf{x}_{\alpha-1} d\mathbf{x}_{\alpha+1} \dots d\mathbf{x}_N d\mathbf{v} \\ &=: \sum_{\alpha} m_{\alpha} \langle W \mid \mathbf{x}_{\alpha} = \mathbf{x} \rangle, \end{aligned} \quad (3.13)$$

where $\langle W \mid \mathbf{x}_{\alpha} = \mathbf{x} \rangle$ denotes an integral of W over all its arguments except \mathbf{x}_{α} and \mathbf{x}_{α} is substituted with \mathbf{x} . Now consider the continuum velocity field. Unlike the definition of pointwise density field, which appears unambiguous, the pointwise velocity field can be defined in different ways. It may seem more natural to define the continuum velocity in an analogous fashion to the density field, i.e.,

$$\mathbf{v}(\mathbf{x}, t) = \frac{\sum_{\alpha} \langle W \mathbf{v}_{\alpha} \mid \mathbf{x}_{\alpha} = \mathbf{x} \rangle}{\sum_{\alpha} \langle W \mid \mathbf{x}_{\alpha} = \mathbf{x} \rangle}. \quad (3.14)$$

Alternatively, the pointwise velocity field can be defined via the momentum density field, $\mathbf{p}(\mathbf{x}, t)$, as follows:

$$\mathbf{p}(\mathbf{x}, t) := \sum_{\alpha} m_{\alpha} \langle W \mathbf{v}_{\alpha} \mid \mathbf{x}_{\alpha} = \mathbf{x} \rangle, \quad (3.15)$$

$$\mathbf{v}(\mathbf{x}, t) := \frac{\mathbf{p}(\mathbf{x}, t)}{\rho(\mathbf{x}, t)}. \quad (3.16)$$

Note that definitions (3.14) and (3.16) are equivalent for a single species material, but are not so in general. The definition given by (3.16) is the one used in practice. There are two reasons for this. First, the definition in (3.16) makes more physical sense since, following spatial averaging, it associates the continuum velocity with the velocity of the center of mass of the system of particles. Second, the definition in (3.16) satisfies the continuity equation as shown in Section 3.3, whereas (3.14) does not.¹⁴

3.2 Regularity assumptions for the probability density function

It is clear from the definitions in (3.13), (3.15) and (3.16) that the integrals in these equations converge under appropriate decay conditions on W . The following two conditions are sufficient for the convergence of all the integrals and the validity of the results in this section [48]:

1. There exists a $\delta > 0$ such that the function

$$W(\Xi; t) \prod_{\alpha=1}^N \|\mathbf{x}_\alpha\|^{3+\delta} \prod_{\beta=1}^N \|\mathbf{v}_\beta\|^{6+\delta} \quad (3.17)$$

and its first derivatives are bounded by a constant that only depends on time.

2. $\mathcal{V}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$ is a bounded C^1 function defined on the phase space, and having bounded first derivatives.¹⁵

Conditions (1) and (2) ensure the convergence of all the integrals considered in this section and swapping of integration and differentiation. Furthermore, let $\mathbf{G}(\Xi; t)$ be any vector or tensor-valued function of class C^1 defined on the phase space for all t , and which, for suitable functions $g(t)$ and $h(t)$, satisfies the condition

$$\sup_{\mathbf{x}_1 \in \mathbb{R}^3, \mathbf{x}_2 \in \mathbb{R}^3, \dots, \mathbf{x}_N \in \mathbb{R}^3} (\|\mathbf{G}\|, \|\operatorname{div}_{\mathbf{v}_\alpha} \mathbf{G}\|, \|\operatorname{div}_{\mathbf{x}_\alpha} \mathbf{G}\|) < g(t) \prod_{\beta=1}^N \|\mathbf{v}_\beta\|^3 + h(t), \quad (3.18)$$

where $\|\cdot\|$ refers to the norm defined through the inner product. Since the space of all tensors has a natural inner product defined as

$$\mathbf{S} : \mathbf{T} = \operatorname{tr}(\mathbf{S}^T \mathbf{T}), \quad (3.19)$$

¹⁴It would be interesting to explore how the equation of continuity fails for the definition in (3.14) by identifying the regions that act as sinks and sources. This is difficult to do for a general N particle system because the continuity equation quickly becomes unwieldy. Even for the much simpler case of a two-particle system, the answer is not trivial. A quick examination shows that the distribution of sinks and sources depends not only on the ratio of the masses but also on the probability density function W .

¹⁵If any two particles overlap, we would normally expect $\mathcal{V} \rightarrow \infty$. By specifying additional decay conditions for W , the case of unbounded \mathcal{V} can be handled. For simplicity, we assume \mathcal{V} to be bounded.

we have $\|\mathbf{S}\| = \sqrt{\mathbf{S} : \mathbf{S}}$. Under these conditions on $\mathbf{G}(\boldsymbol{\Xi}; t)$, we have¹⁶

$$\int_{\mathbb{R}^3} \mathbf{G} \cdot \nabla_{\mathbf{x}_\alpha} W d\mathbf{x}_\alpha = - \int_{\mathbb{R}^3} W \operatorname{div}_{\mathbf{x}_\alpha} \mathbf{G} d\mathbf{x}_\alpha, \quad (3.20a)$$

$$\int_{\mathbb{R}^3} \mathbf{G} \cdot \nabla_{\mathbf{v}_\alpha} W d\mathbf{v}_\alpha = - \int_{\mathbb{R}^3} W \operatorname{div}_{\mathbf{v}_\alpha} \mathbf{G} d\mathbf{v}_\alpha. \quad (3.20b)$$

The above identities are repeatedly used in deriving the equation of continuity and the equation of motion in the following sections.

3.3 Equation of continuity

Let us demonstrate that the pointwise fields defined in Section 3.1 satisfy the equation of continuity. The equation of continuity from continuum mechanics is given by [37]

$$\frac{\partial \rho}{\partial t} + \operatorname{div}_{\mathbf{x}}(\rho \mathbf{v}) = 0. \quad (3.21)$$

From (3.13) we have

$$\frac{\partial \rho}{\partial t}(\mathbf{x}, t) = \sum_{\alpha} m_{\alpha} \left\langle \frac{\partial W}{\partial t} \middle| \mathbf{x}_{\alpha} = \mathbf{x} \right\rangle.$$

Using Liouville's equation in (3.11), we have

$$\frac{\partial \rho}{\partial t}(\mathbf{x}, t) = \sum_{\alpha} m_{\alpha} \left\langle \sum_{\beta} \left(-\mathbf{v}_{\beta} \cdot \nabla_{\mathbf{x}_{\beta}} W + \frac{\nabla_{\mathbf{x}_{\beta}} \mathcal{V}}{m_{\beta}} \cdot \nabla_{\mathbf{v}_{\beta}} W \right) \middle| \mathbf{x}_{\alpha} = \mathbf{x} \right\rangle.$$

Now, consider the summand on the right-hand side of the above equation for a fixed α . From (3.20b), it is clear that $\left\langle \frac{\nabla_{\mathbf{x}_{\beta}} \mathcal{V}}{m_{\beta}} \cdot \nabla_{\mathbf{v}_{\beta}} W \middle| \mathbf{x}_{\alpha} = \mathbf{x} \right\rangle = 0$, for $\beta = 1, 2, \dots, N$, and from (3.20a), we also have $\left\langle \mathbf{v}_{\beta} \cdot \nabla_{\mathbf{x}_{\beta}} W \middle| \mathbf{x}_{\alpha} = \mathbf{x} \right\rangle = 0$, for $\beta \neq \alpha$. Therefore the above equation simplifies to

$$\frac{\partial \rho}{\partial t}(\mathbf{x}, t) = - \sum_{\alpha} m_{\alpha} \langle \mathbf{v}_{\alpha} \cdot \nabla_{\mathbf{x}_{\alpha}} W \mid \mathbf{x}_{\alpha} = \mathbf{x} \rangle.$$

Using the identity,

$$\operatorname{div}_{\mathbf{x}}(a \mathbf{w}) = \nabla_{\mathbf{x}} a \cdot \mathbf{w}, \quad (3.22)$$

where $a(\mathbf{x})$ is any C^1 scalar function of \mathbf{x} , and \mathbf{w} is any vector independent of \mathbf{x} , we obtain

$$\frac{\partial \rho}{\partial t}(\mathbf{x}, t) = - \sum_{\alpha} m_{\alpha} \operatorname{div}_{\mathbf{x}} \langle W \mathbf{v}_{\alpha} \mid \mathbf{x}_{\alpha} = \mathbf{x} \rangle.$$

¹⁶If \mathbf{G} is a second-order tensor or higher, then the dot product indicates tensor operating on a vector. Note that in (3.20), in the interest of brevity, we are breaking our notation of denoting a second-order tensor operating on a vector by juxtaposition.

Using (3.15) and (3.16) for the definition of the pointwise momentum density field, we have

$$\frac{\partial \rho}{\partial t}(\mathbf{x}, t) + \operatorname{div}_{\mathbf{x}}(\rho \mathbf{v}) = 0,$$

which is the continuity equation. We have established that the definitions given in (3.12) and (3.15) identically satisfy conservation of mass.

3.4 Equation of Motion

The equation of motion from continuum mechanics is given by [37]

$$\frac{\partial(\rho \mathbf{v})}{\partial t} + \operatorname{div}_{\mathbf{x}}(\rho \mathbf{v} \otimes \mathbf{v}) = \operatorname{div}_{\mathbf{x}} \boldsymbol{\sigma} + \mathbf{b}. \quad (3.23)$$

Here we identify $\boldsymbol{\sigma}$ with the pointwise stress tensor. From (3.15), we have

$$\frac{\partial \mathbf{p}}{\partial t}(\mathbf{x}, t) = \sum_{\alpha} m_{\alpha} \left\langle \mathbf{v}_{\alpha} \frac{\partial W}{\partial t} \middle| \mathbf{x}_{\alpha} = \mathbf{x} \right\rangle.$$

Again, using (3.11) we obtain,

$$\begin{aligned} \frac{\partial \mathbf{p}}{\partial t}(\mathbf{x}, t) &= \sum_{\alpha} m_{\alpha} \left\langle \mathbf{v}_{\alpha} \sum_{\beta} \left(-\nabla_{\mathbf{x}_{\beta}} W \cdot \mathbf{v}_{\beta} + \frac{\nabla_{\mathbf{x}_{\beta}} \mathcal{V}}{m_{\beta}} \cdot \nabla_{\mathbf{v}_{\beta}} W \right) \middle| \mathbf{x}_{\alpha} = \mathbf{x} \right\rangle \\ &= \sum_{\alpha} m_{\alpha} \sum_{\beta} \left\langle -(\mathbf{v}_{\alpha} \otimes \mathbf{v}_{\beta}) \nabla_{\mathbf{x}_{\beta}} W + \left(\mathbf{v}_{\alpha} \otimes \frac{\nabla_{\mathbf{x}_{\beta}} \mathcal{V}}{m_{\beta}} \right) \nabla_{\mathbf{v}_{\beta}} W \middle| \mathbf{x}_{\alpha} = \mathbf{x} \right\rangle. \end{aligned} \quad (3.24)$$

Now, consider the summand on the right-hand side of the above equation for fixed α and β . Using (3.20a), we have $\langle (\mathbf{v}_{\alpha} \otimes \mathbf{v}_{\beta}) \nabla_{\mathbf{x}_{\beta}} W \mid \mathbf{x}_{\alpha} = \mathbf{x} \rangle = \mathbf{0}$, for $\beta \neq \alpha$. From (3.20b), we have $\langle (\mathbf{v}_{\alpha} \otimes \nabla_{\mathbf{x}_{\beta}} \mathcal{V}) \nabla_{\mathbf{v}_{\beta}} W \mid \mathbf{x}_{\alpha} = \mathbf{x} \rangle = \mathbf{0}$, for $\beta \neq \alpha$, and for $\beta = \alpha$, we have

$$\langle (\mathbf{v}_{\alpha} \otimes \nabla_{\mathbf{x}_{\alpha}} \mathcal{V}) \nabla_{\mathbf{v}_{\alpha}} W \mid \mathbf{x}_{\alpha} = \mathbf{x} \rangle = -\langle \nabla_{\mathbf{x}_{\alpha}} \mathcal{V} W \mid \mathbf{x}_{\alpha} = \mathbf{x} \rangle,$$

using the fact that $\operatorname{div}_{\mathbf{u}}(\mathbf{u} \otimes \mathbf{w}) = \mathbf{w}$, for any vector \mathbf{u} and for any vector \mathbf{w} independent of \mathbf{u} . Therefore (3.24) simplifies to

$$\frac{\partial \mathbf{p}}{\partial t}(\mathbf{x}, t) = - \sum_{\alpha} m_{\alpha} \langle (\mathbf{v}_{\alpha} \otimes \mathbf{v}_{\alpha}) \nabla_{\mathbf{x}_{\alpha}} W \mid \mathbf{x}_{\alpha} = \mathbf{x} \rangle - \sum_{\alpha} \langle W \nabla_{\mathbf{x}_{\alpha}} \mathcal{V} \mid \mathbf{x}_{\alpha} = \mathbf{x} \rangle. \quad (3.25)$$

Using the identity,

$$\operatorname{div}_{\mathbf{x}}(a \mathbf{T}) = \mathbf{T} \nabla_{\mathbf{x}} a, \quad (3.26)$$

where $a(\mathbf{x})$ is any C^1 scalar function of \mathbf{x} , and \mathbf{T} is any tensor independent of \mathbf{x} , we can rewrite (3.25) as

$$\frac{\partial \mathbf{p}}{\partial t}(\mathbf{x}, t) = - \operatorname{div}_{\mathbf{x}} \sum_{\alpha} m_{\alpha} \langle (\mathbf{v}_{\alpha} \otimes \mathbf{v}_{\alpha}) W \mid \mathbf{x}_{\alpha} = \mathbf{x} \rangle - \sum_{\alpha} \langle W \nabla_{\mathbf{x}_{\alpha}} \mathcal{V} \mid \mathbf{x}_{\alpha} = \mathbf{x} \rangle. \quad (3.27)$$

Now, note that the term $\mathbf{v}_\alpha \otimes \mathbf{v}_\alpha$ can be written as

$$\begin{aligned}\mathbf{v}_\alpha \otimes \mathbf{v}_\alpha &= (\mathbf{v}_\alpha - \mathbf{v}) \otimes (\mathbf{v}_\alpha - \mathbf{v}) + \mathbf{v} \otimes \mathbf{v}_\alpha + \mathbf{v}_\alpha \otimes \mathbf{v} - \mathbf{v} \otimes \mathbf{v} \\ &= \mathbf{v}_\alpha^{\text{rel}} \otimes \mathbf{v}_\alpha^{\text{rel}} + \mathbf{v} \otimes \mathbf{v}_\alpha + \mathbf{v}_\alpha \otimes \mathbf{v} - \mathbf{v} \otimes \mathbf{v},\end{aligned}\quad (3.28)$$

where $\mathbf{v}_\alpha^{\text{rel}}$ is the velocity of particle α relative to the pointwise velocity field. Consider the first term on the right-hand side of (3.27). Substituting (3.28) into this expression we have,

$$\begin{aligned}& -\operatorname{div}_{\mathbf{x}} \sum_{\alpha} m_{\alpha} \langle (\mathbf{v}_\alpha \otimes \mathbf{v}_\alpha) W \mid \mathbf{x}_\alpha = \mathbf{x} \rangle \\ &= -\sum_{\alpha} m_{\alpha} \operatorname{div}_{\mathbf{x}} \langle (\mathbf{v}_\alpha^{\text{rel}} \otimes \mathbf{v}_\alpha^{\text{rel}}) W \mid \mathbf{x}_\alpha = \mathbf{x} \rangle - \operatorname{div}_{\mathbf{x}} \sum_{\alpha} [\mathbf{v} \otimes m_{\alpha} \langle \mathbf{v}_\alpha W \mid \mathbf{x}_\alpha = \mathbf{x} \rangle \\ &\quad + m_{\alpha} \langle \mathbf{v}_\alpha W \mid \mathbf{x}_\alpha = \mathbf{x} \rangle \otimes \mathbf{v} - m_{\alpha} \langle W \mid \mathbf{x}_\alpha = \mathbf{x} \rangle \mathbf{v} \otimes \mathbf{v}] \\ &= -\operatorname{div}_{\mathbf{x}} \sum_{\alpha} m_{\alpha} \langle (\mathbf{v}_\alpha^{\text{rel}} \otimes \mathbf{v}_\alpha^{\text{rel}}) W \mid \mathbf{x}_\alpha = \mathbf{x} \rangle - \operatorname{div}_{\mathbf{x}} (\rho \mathbf{v} \otimes \mathbf{v}),\end{aligned}\quad (3.29)$$

where we have used (3.13), (3.15) and (3.16) in the last step. Substituting (3.29) into (3.27), we obtain

$$\begin{aligned}\frac{\partial \mathbf{p}}{\partial t}(\mathbf{x}, t) + \operatorname{div}_{\mathbf{x}} (\rho \mathbf{v} \otimes \mathbf{v}) &= -\sum_{\alpha} m_{\alpha} \operatorname{div}_{\mathbf{x}} \langle (\mathbf{v}_\alpha^{\text{rel}} \otimes \mathbf{v}_\alpha^{\text{rel}}) W \mid \mathbf{x}_\alpha = \mathbf{x} \rangle \\ &\quad - \sum_{\alpha} \langle W \nabla_{\mathbf{x}_\alpha} \mathcal{V} \mid \mathbf{x}_\alpha = \mathbf{x} \rangle.\end{aligned}\quad (3.30)$$

The left-hand sides of (3.30) and (3.23) are identical. Therefore, the right-hand sides must also be equal. Hence

$$\operatorname{div}_{\mathbf{x}} \boldsymbol{\sigma} + \mathbf{b} = -\sum_{\alpha} m_{\alpha} \operatorname{div}_{\mathbf{x}} \langle (\mathbf{v}_\alpha^{\text{rel}} \otimes \mathbf{v}_\alpha^{\text{rel}}) W \mid \mathbf{x}_\alpha = \mathbf{x} \rangle - \sum_{\alpha} \langle W \nabla_{\mathbf{x}_\alpha} \mathcal{V} \mid \mathbf{x}_\alpha = \mathbf{x} \rangle. \quad (3.31)$$

To proceed, we divide the potential energy $\mathcal{V}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$ into two parts:

1. An *external* part, \mathcal{V}_{ext} , associated with long-range interactions such as gravity or electromagnetic fields.
2. An *internal* part, \mathcal{V}_{int} , associated with short-range particle interactions.

It is natural to associate \mathcal{V}_{ext} with the body force field \mathbf{b} in (3.31). We therefore define $\mathbf{b}(\mathbf{x}, t)$ as

$$\mathbf{b}(\mathbf{x}, t) := -\sum_{\alpha} \langle W \nabla_{\mathbf{x}_\alpha} \mathcal{V}_{\text{ext}} \mid \mathbf{x}_\alpha = \mathbf{x} \rangle. \quad (3.32)$$

Substituting (3.32) into (3.31), we have

$$\operatorname{div}_{\mathbf{x}} \boldsymbol{\sigma} = -\sum_{\alpha} m_{\alpha} \operatorname{div}_{\mathbf{x}} \langle (\mathbf{v}_\alpha^{\text{rel}} \otimes \mathbf{v}_\alpha^{\text{rel}}) W \mid \mathbf{x}_\alpha = \mathbf{x} \rangle - \sum_{\alpha} \langle W \nabla_{\mathbf{x}_\alpha} \mathcal{V}_{\text{int}} \mid \mathbf{x}_\alpha = \mathbf{x} \rangle. \quad (3.33)$$

From (3.33), we see that the pointwise stress tensor has two contributions:

$$\boldsymbol{\sigma}(\mathbf{x}, t) = \boldsymbol{\sigma}_k(\mathbf{x}, t) + \boldsymbol{\sigma}_v(\mathbf{x}, t), \quad (3.34)$$

where $\boldsymbol{\sigma}_k$ and $\boldsymbol{\sigma}_v$ are, respectively, the *kinetic* and *potential* parts of the pointwise stress. The kinetic part is given by

$$\boldsymbol{\sigma}_k(\mathbf{x}, t) = - \sum_{\alpha} m_{\alpha} \langle (\mathbf{v}_{\alpha}^{\text{rel}} \otimes \mathbf{v}_{\alpha}^{\text{rel}}) W \mid \mathbf{x}_{\alpha} = \mathbf{x} \rangle. \quad (3.35)$$

It is evident that the kinetic part of the stress tensor is symmetric. The presence of a kinetic contribution to the stress tensor appears at odds with the continuum definition of stress that is stated solely in terms of the forces acting between different parts of the body. This discrepancy has led to controversy in the past about whether the kinetic term belongs in the stress definition [67]. The confusion is related to the difference between absolute velocity and relative velocity defined in (3.28) [63]. The kinetic stress reflects the momentum flux associated with the vibrational kinetic energy portion of the internal energy.

Continuing with (3.33), the potential part of the stress must satisfy the following differential equation:

$$\text{div}_{\mathbf{x}} \boldsymbol{\sigma}_v(\mathbf{x}, t) = \sum_{\alpha} \langle W \mathbf{f}_{\alpha}^{\text{int}} \mid \mathbf{x}_{\alpha} = \mathbf{x} \rangle, \quad (3.36)$$

where

$$\mathbf{f}_{\alpha}^{\text{int}} := -\nabla_{\mathbf{x}_{\alpha}} \mathcal{V}_{\text{int}}, \quad (3.37)$$

is the force on particle α due to internal interactions. Equation (3.36) needs to be solved in order to obtain an explicit form for $\boldsymbol{\sigma}_v$. In the original paper of Irving and Kirkwood [27], this was done by applying a Taylor expansion to the Dirac delta distribution appearing in the right-hand side of the equation. In contrast, Noll showed that a closed-form solution for $\boldsymbol{\sigma}_v$ can be obtained by recasting the right-hand side in a different form and applying a lemma proved in [48]. We proceed with Noll's approach, except we place no restriction on the nature of the interatomic potential energy \mathcal{V}_{int} . The potential energy considered in [27] and [48] is limited to pair potentials.

General interatomic potentials

In general, the internal part of the potential energy, also called the *interatomic potential energy*, depends on the positions of all particles in the system:

$$\mathcal{V}_{\text{int}} = \widehat{\mathcal{V}}_{\text{int}}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N), \quad (3.38)$$

where the “hat” indicates that the functional dependence is on absolute particle positions (as opposed to distances later on). We assume that $\widehat{\mathcal{V}}_{\text{int}} : \mathbb{R}^{3N} \rightarrow \mathbb{R}$ is a continuously differentiable function.¹⁷ This function must satisfy the following invariance principle:

¹⁷Note that this assumption may fail in systems undergoing first-order magnetic or electronic phase transformations.

The internal energy of a material system is invariant with respect to the Euclidean group $\mathcal{G} := \{\mathbf{x} \mapsto \mathbf{Q}\mathbf{x} + \mathbf{c} \mid \mathbf{x} \in \mathbb{R}^3, \mathbf{Q} \in O(3), \mathbf{c} \in \mathbb{R}^3\}$, where $O(3)$ denotes the full orthogonal group.

To exploit this invariance, let us consider the action of \mathcal{G} on \mathbb{R}^{3N} , i.e., the action of any combination of translation and rotation (proper or improper), which is represented by an element $g : \mathbf{x} \mapsto \mathbf{Q}\mathbf{x} + \mathbf{c}$ in \mathcal{G} , on any configuration of N particles represented by a vector $(\mathbf{x}_1, \dots, \mathbf{x}_N) \in \mathbb{R}^{3N}$:

$$g \cdot (\mathbf{x}_1, \dots, \mathbf{x}_N) = (\mathbf{Q}\mathbf{x}_1 + \mathbf{c}, \dots, \mathbf{Q}\mathbf{x}_N + \mathbf{c}). \quad (3.39)$$

This action splits \mathbb{R}^{3N} into disjoint sets of equivalence classes [14], which we now describe. For any $\mathbf{u} = (\mathbf{x}_1, \dots, \mathbf{x}_N) \in \mathbb{R}^{3N}$, let $\mathcal{O}_{\mathbf{u}} \subset \mathbb{R}^{3N}$ denote an equivalence class which is defined as¹⁸

$$\mathcal{O}_{\mathbf{u}} := \{g \cdot \mathbf{u} \mid g \in \mathcal{G}\}, \quad (3.40)$$

where $g \cdot \mathbf{u}$ denotes the action of g on \mathbf{u} defined in (3.39). In other words, $\mathcal{O}_{\mathbf{u}}$ represents the set of all configurations which are related to the configuration \mathbf{u} by a rigid body motion and/or reflection. Due to the invariance of the potential energy, we can view the function \mathcal{V}_{int} as a function on the set of equivalence classes, i.e.,

$$\bar{\mathcal{V}}_{\text{int}}(\mathcal{O}_{\mathbf{u}}) = \hat{\mathcal{V}}_{\text{int}}(\mathbf{u}), \quad (3.41)$$

because

$$\hat{\mathcal{V}}_{\text{int}}(\mathbf{v}) = \hat{\mathcal{V}}_{\text{int}}(\mathbf{u}) \quad \forall \mathbf{v} \in \mathcal{O}_{\mathbf{u}}. \quad (3.42)$$

Now, consider a set $\mathcal{S} \subset \mathbb{R}^{N(N-1)/2}$, defined as

$$\mathcal{S} := \{(r_{12}, r_{13}, \dots, r_{1N}, r_{23}, \dots, r_{(N-1)N}) \mid r_{\alpha\beta} = \|\mathbf{x}_{\alpha} - \mathbf{x}_{\beta}\|, (\mathbf{x}_1, \dots, \mathbf{x}_N) \in \mathbb{R}^{3N}\}. \quad (3.43)$$

In other words, the set \mathcal{S} consists of all possible $N(N-1)/2$ -tuples of real numbers which correspond to the distances between N particles in \mathbb{R}^3 .¹⁹ In technical terms, the coordinates of any point in \mathcal{S} are said to be *embeddable* in \mathbb{R}^3 . Note that \mathcal{S} is a proper subset of $\mathbb{R}^{N(N-1)/2}$ as it consists of only those $N(N-1)/2$ -tuple distances which satisfy certain geometric constraints. In fact, the set \mathcal{S} represents a $(3N-6)$ -dimensional manifold in $\mathbb{R}^{N(N-1)/2}$, commonly referred to as the *shape space*.

Let ϕ be the mapping taking a point in configuration space to the corresponding set of distances in \mathcal{S} , i.e., $\phi : \mathbb{R}^{3N} \rightarrow \mathcal{S} : (\mathbf{x}_1, \dots, \mathbf{x}_N) \mapsto (r_{12}, \dots, r_{(N-1)N})$, where $r_{\alpha\beta}$ from here onwards is used to denote $\|\mathbf{x}_{\alpha} - \mathbf{x}_{\beta}\|$. Since the Euclidean group preserves distances, it immediately follows that the map

$$\bar{\phi} : \{\text{Equivalence classes}\} \rightarrow \mathcal{S}, \quad (3.44)$$

¹⁸The notation “ $\{g \cdot \mathbf{u} \mid g \in \mathcal{G}\}$ ” should be read as “the set of all $g \cdot \mathbf{u}$, such that g is in the Euclidean group \mathcal{G} ”.

¹⁹The key here is that not all $N(N-1)/2$ combinations of real numbers constitute a valid set of physical distances. The distances must satisfy certain geometric constraints in order to be physically meaningful as explained below.

defined as $\bar{\phi}(\mathcal{O}_u) = \phi(u)$, is a bijection (one-to-one and onto mapping) from the set of equivalence classes to the set \mathcal{S} .²⁰ This essentially means that for every set of equivalent configurations, i.e., configurations related to each other by a rigid body motion and/or reflection, there exists a unique $N(N-1)/2$ -tuple of distances and vice versa. From (3.41) and (3.44), it immediately follows that the potential energy of the system can be completely described by a function $\check{\mathcal{V}}_{\text{int}} : \mathcal{S} \rightarrow \mathbb{R}$, defined as

$$\check{\mathcal{V}}_{\text{int}}(s) := \bar{\mathcal{V}}_{\text{int}}(\bar{\phi}^{-1}(s)) \quad \forall s \in \mathcal{S}. \quad (3.45)$$

We now restrict our discussion to those systems for which there exists a continuously differentiable extension of $\check{\mathcal{V}}_{\text{int}}$, defined on the shape space, to $\mathbb{R}^{N(N-1)/2}$.²¹ This is justifiable because of the fact that all interatomic potentials used in practice, for a system of N particles, are either continuously differentiable functions on $\mathbb{R}^{N(N-1)/2}$, or can easily be extended to one. For example, the pair potential and the embedded-atom method (EAM) potential [12] are continuously differentiable functions on $\mathbb{R}^{N(N-1)/2}$, while the Stillinger-Weber [57] and the Tersoff [60] potentials can be easily extended to $\mathbb{R}^{N(N-1)/2}$ by expressing the angles appearing in them as a function of distances between particles. Therefore, we assume that there exist a continuously differentiable function $\mathcal{V}_{\text{int}} : \mathbb{R}^{N(N-1)/2} \rightarrow \mathbb{R}$, such that the restriction of \mathcal{V}_{int} to \mathcal{S} is equal to $\check{\mathcal{V}}_{\text{int}}$:

$$\mathcal{V}_{\text{int}}(s) = \check{\mathcal{V}}_{\text{int}}(s) \quad \forall s = (r_{12}, \dots, r_{(N-1)N}) \in \mathcal{S}. \quad (3.46)$$

An immediate question that arises is whether this extension is unique in a neighborhood of $s \in \mathcal{S}$. Note that for $N \leq 4$, $3N-6 = N(N-1)/2$. Therefore, for $N \leq 4$, for every point $s \in \mathcal{S}$, there exists a neighborhood in $\mathbb{R}^{N(N-1)/2}$ which lies in \mathcal{S} . However, for $N > 4$, there may be multiple extensions of $\check{\mathcal{V}}_{\text{int}}$.

As noted above, the reason we are considering an extension is to define the partial derivative of the potential energy with respect to each coordinate of a point in $\mathbb{R}^{N(N-1)/2}$. This will be used later to define the stress tensor. For example, the partial derivative of $\mathcal{V}_{\text{int}}(\zeta_{12}, \dots, \zeta_{(N-1)N})$ with respect to ζ_{12} at any point $s = (r_{12}, \dots, r_{(N-1)N}) \in \mathcal{S}$, defined as

$$\frac{d\mathcal{V}_{\text{int}}}{d\zeta_{12}}(s) = \lim_{\epsilon \rightarrow 0} \frac{\mathcal{V}_{\text{int}}(r_{12} + \epsilon, \dots, r_{N(N-1)/2}) - \mathcal{V}_{\text{int}}(r_{12}, \dots, r_{N(N-1)/2})}{\epsilon}, \quad (3.47)$$

requires us to evaluate the function at non-embeddable points.

²⁰ $\bar{\phi}$ is surjective (onto) by the definition of \mathcal{S} . The proof that it is injective (one-to-one) is similar to the proof of the *basic invariance theorem* for the simultaneous invariants of vectors due to Cauchy, which can be found in [62, Section 11].

²¹The extension is necessary since $\check{\mathcal{V}}_{\text{int}}$ is defined in (3.45) only on the set \mathcal{S} . We need to extend the definition to *all* points in $\mathbb{R}^{N(N-1)/2}$, whether they correspond to a set of physical distances or not, in order to be able to compute derivatives as explained later in the text. This issue has been overlooked in the past (see for example [13]), which leads to the conclusion that the stress tensor is always symmetric. It turns out that this conclusion is correct (at least for point masses without internal structure), but the reasoning is more involved as we show later.

It will be shown later that the quantity evaluated in (3.47) may differ for different extensions. On the other hand, $\nabla_{\mathbf{x}_\alpha} \mathcal{V}_{\text{int}}$ is uniquely defined for any extension. This is because

$$\begin{aligned}\nabla_{\mathbf{x}_\alpha} \mathcal{V}_{\text{int}}(\mathbf{s}) &= \nabla_{\mathbf{x}_\alpha} \check{\mathcal{V}}_{\text{int}}(\mathbf{s}) \\ &= \nabla_{\mathbf{x}_\alpha} \bar{\mathcal{V}}_{\text{int}}(\bar{\phi}^{-1}(\mathbf{s})) \\ &= \nabla_{\mathbf{x}_\alpha} \hat{\mathcal{V}}_{\text{int}}(\mathbf{u}),\end{aligned}\tag{3.48}$$

where $\bar{\phi}^{-1}(\mathbf{s}) = \mathcal{O}_{\mathbf{u}}$, which implies $\phi(\mathbf{u}) = \mathbf{s}$,²² and we have used (3.46), (3.45) and (3.41) in the first, second and the last equality respectively.

We next address the possibility of having multiple extensions for the potential energy by studying the various constraints that the distances between particles have to satisfy in order to be embeddable in \mathbb{R}^3 . We demonstrate, through a simple example, how multiple extensions for the potential energy can lead to a non-unique decomposition of the force on a particle, which in turn leads to a non-unique pointwise stress tensor.

Central-force decomposition and the possibility of alternate extensions

We will now show that the force on a particle can always be decomposed as a sum of central forces. The force on a particle due to internal interactions is defined in (3.37). Therefore, for any configuration $\mathbf{u} \in \mathbb{R}^{3N}$, we have

$$\mathbf{f}_\alpha^{\text{int}}(\mathbf{u}) = -\nabla_{\mathbf{x}_\alpha} \hat{\mathcal{V}}_{\text{int}}(\mathbf{u}).\tag{3.49}$$

Using (3.48), (3.49) takes the form

$$\begin{aligned}\mathbf{f}_\alpha^{\text{int}}(\mathbf{u}) &= -\nabla_{\mathbf{x}_\alpha} \mathcal{V}_{\text{int}}(\mathbf{s})|_{\mathbf{s}=\phi(\mathbf{u})} \\ &= \sum_{\substack{\beta \\ \beta \neq \alpha}} \mathbf{f}_{\alpha\beta}(\mathbf{u}),\end{aligned}\tag{3.50}$$

where $\mathbf{s} = \phi(\mathbf{u}) = (r_{12}, \dots, r_{(N-1)N})$ and

$$\mathbf{f}_{\alpha\beta}(\mathbf{u}) := \begin{cases} \frac{\partial \mathcal{V}_{\text{int}}}{\partial \zeta_{\alpha\beta}}(\phi(\mathbf{u})) \frac{\mathbf{x}_\beta - \mathbf{x}_\alpha}{r_{\alpha\beta}} & \text{if } \alpha < \beta, \\ \frac{\partial \mathcal{V}_{\text{int}}}{\partial \zeta_{\beta\alpha}}(\phi(\mathbf{u})) \frac{\mathbf{x}_\beta - \mathbf{x}_\alpha}{r_{\alpha\beta}} & \text{if } \alpha > \beta, \end{cases}\tag{3.51}$$

is the contribution to the force on particle α due to the presence of particle β .

Note that $\mathbf{f}_{\alpha\beta}$ is parallel to the direction $\mathbf{x}_\beta - \mathbf{x}_\alpha$ and satisfies $\mathbf{f}_{\alpha\beta} = -\mathbf{f}_{\beta\alpha}$. We therefore note the important result that the *internal force on a particle, for any interatomic potential that has a continuously differentiable extension, can always be decomposed as a sum of central forces, i.e., forces parallel to directions connecting*

²²Note that the vector \mathbf{u} appearing in (3.48) can be replaced by any $\mathbf{v} \in \mathcal{O}_{\mathbf{u}}$.

the particle to its neighbors.²³ We will see later in Section 3.5 that the central-force decomposition is the only physically-meaningful partitioning of the force.

The remaining question is how different potential energy extensions affect the force decomposition in (3.50). We have already established in (3.48) and (3.49) that the force $\mathbf{f}_\alpha^{\text{int}}$ is independent of the particular extension used. However, we show below that the individual terms in the decomposition, $\mathbf{f}_{\alpha\beta}$, are *not* unique. These terms depend on the manner in which the potential energy, defined on the shape space, is extended to its neighborhood in the higher-dimensional Euclidean space.

In order to construct different extensions, we use the geometric constraints that the distances have to satisfy in order for them to be embeddable in \mathbb{R}^3 .²⁴ The nature of these constraints is studied in the field of *distance geometry*, which describes the geometry of sets of points in terms of the distances between them (see Appendix B). One of the main results of this theory, is that the constraints are given by *Cayley-Menger determinants*, which are related to the volume of a simplex formed by N points in an $N - 1$ dimensional space.

For simplicity let us restrict our discussion to one dimension. It is easy to see that in one dimension the number of independent coordinates are $N - 1$ and for $N > 2$ the number of interatomic distances exceeds the number of independent coordinates. Therefore, let the material system \mathcal{M} consist of three point masses interacting in one dimension. The standard pair potential representation for this system, which is an extension of the potential energy to the higher-dimensional Euclidean space, is given by

$$\mathcal{V}_{\text{int}}(\zeta_{12}, \zeta_{13}, \zeta_{23}) = \mathcal{V}_{12}(\zeta_{12}) + \mathcal{V}_{13}(\zeta_{13}) + \mathcal{V}_{23}(\zeta_{23}). \quad (3.52)$$

Since the calculation gets unwieldy, let us consider the special case when the particles are arranged to satisfy $x_1 < x_2 < x_3$, for which $r_{13} = r_{12} + r_{23}$. Using (3.50), the internal force, f_1^{int} , evaluated at this configuration, is decomposed as

$$\begin{aligned} f_1^{\text{int}}(r_{12}, r_{13}, r_{23}) &= -\frac{d\mathcal{V}_{\text{int}}}{dx_1} = -\frac{d\mathcal{V}_{12}}{dx_1} - \frac{d\mathcal{V}_{13}}{dx_1} \\ &= \mathcal{V}'_{12}(r_{12}) + \mathcal{V}'_{13}(r_{13}) \\ &=: f_{12} + f_{13}. \end{aligned} \quad (3.53)$$

²³ The result that the force on a particle, modeled using any interatomic potential with a continuously differentiable extension, can be decomposed as sum of central forces may seem strange to some readers. This may be due to the common confusion in the literature of using the term “central-force models” to refer to simple pair potentials. In fact, we see that due to the invariance requirement stated on Page 21, *all* interatomic potentials (including those with explicit bond angle dependence) that can be expressed as a continuously differentiable function as described in the text, are central-force models. By this we mean that the force on any particle (say α) can be decomposed as a sum of terms, $\mathbf{f}_{\alpha\beta}$, aligned with the vectors joining particle α with its neighbors and satisfying action and reaction.

The difference between the general case and that of a pair potential is that for a pair potential, $\|\mathbf{f}_{\alpha\beta}\|$ depends *only* on the distance $r_{\alpha\beta}$ between the particles, whereas for a general potential, the dependence is on a larger set of distances, $\|\mathbf{f}_{\alpha\beta}\| = \frac{\partial \mathcal{V}_{\text{int}}}{\partial \zeta_{\alpha\beta}}(r_{12}, r_{13}, \dots, r_{(N-1), N})$, i.e., $\|\mathbf{f}_{\alpha\beta}\|$ depends on the *environment* of the “bond” between α and β . For this reason, $\mathbf{f}_{\alpha\beta}$ for a pair potential is a property of particles α and β alone and can be physically interpreted as the “force exerted on particle α by particle β ”. Whereas, in the more general case of arbitrary interatomic potentials, the physical significance of the interatomic force is less clear and at best we can say that $\mathbf{f}_{\alpha\beta}$ is the “contribution to the force on particle α due to the presence of particle β ”.

²⁴ We thank Ryan Elliott for suggesting this line of thinking.

We now provide an alternate extension to the standard pair potential representation given in (3.52). The Cayley-Menger determinant corresponding to a cluster of three points (see (B.6)) is identically equal to zero at every point on the shape space. This is because the shape space corresponds to a configuration of three collinear points, and the area of the triangle formed by three collinear points is zero. Thus, we have

$$\begin{aligned}\chi(r_{12}, r_{13}, r_{23}) &= (r_{12} - r_{13} - r_{23})(r_{23} - r_{12} - r_{13}) \\ &\quad \times (r_{13} - r_{23} - r_{12})(r_{12} + r_{13} + r_{23}) \\ &= 0.\end{aligned}\tag{3.54}$$

Using the identity in (3.54), an alternate extension $\mathcal{V}_{\text{int}}^A$ is constructed:

$$\mathcal{V}_{\text{int}}^A(\zeta_{12}, \zeta_{13}, \zeta_{23}) = \mathcal{V}_{\text{int}}(\zeta_{12}, \zeta_{13}, \zeta_{23}) + \chi(\zeta_{12}, \zeta_{13}, \zeta_{23}).\tag{3.55}$$

Note that $\mathcal{V}_{\text{int}}^A$ is indeed an extension because from (3.54) it is clear that $\mathcal{V}_{\text{int}}^A$ is equal to \mathcal{V}_{int} at every point on the shape space of the system and it is continuously differentiable because $\chi(\zeta_{12}, \zeta_{13}, \zeta_{23})$, being a polynomial, is infinitely differentiable. Let us now see how the internal force, f_1^{int} , for the special configuration considered in this example, is decomposed using the new extension:

$$\begin{aligned}f_1^{\text{int}} &= -\frac{d\mathcal{V}_{\text{int}}^A}{dx_1} = -\frac{d\mathcal{V}_{\text{int}}}{dx_1} - \frac{d\chi}{dx_1} \\ &= \left(\mathcal{V}'_{12} - \frac{\partial\chi}{\partial\zeta_{12}}(s) \frac{\partial\zeta_{12}}{\partial x_1}(s) \right) + \left(\mathcal{V}'_{13} - \frac{\partial\chi}{\partial\zeta_{13}}(s) \frac{\partial\zeta_{13}}{\partial x_1}(s) \right) \\ &= (f_{12} - 8r_{12}r_{23}(r_{12} + r_{23})) + (f_{13} + 8r_{12}r_{23}(r_{12} + r_{23})) \\ &=: \tilde{f}_{12} + \tilde{f}_{13},\end{aligned}\tag{3.56}$$

It is clear from (3.53) and (3.56) that the central-force decomposition is not the same for the two representations, i.e., $f_{12} \neq \tilde{f}_{12}$ and $f_{13} \neq \tilde{f}_{13}$, however the force on particle 1, f_1^{int} , is the same in both cases as expected.

It is very interesting to note that $\mathcal{V}_{\text{int}}^A$ is *not* a pair potential (based on the definition of a pair potential), but it is equivalent to a pair potential, i.e., it agrees with a pair potential on the shape space. Thus, the set of continuously differentiable extensions of a given interatomic potential function form an equivalence class. It is not clear at this stage if these equivalence classes can be fully expressed in terms of the Cayley-Menger determinant constraints.

Although the above example is quite elementary, this process can be extended to any arbitrary number of particles in three dimensions. Any given potential can be altered to an equivalent potential by adding a function of the Cayley-Menger determinants corresponding to any cluster of 5 or 6 particles (see Appendix B). This function must be continuously differentiable and equal to zero when all of its arguments are zero. For example, a new representation in three dimensions can be constructed by adding a linear combination of the Cayley-Menger determinants:

$$\mathcal{V}_{\text{int}}^* = \mathcal{V}_{\text{int}}(\zeta_{12}, \dots, \zeta_{(N-1)N}) + \sum_{k=1}^m \lambda_k \chi_k,\tag{3.57}$$

where there are m constraints defined by the Cayley-Menger determinants χ_k , and λ_k are constants.²⁵

From this point on, we abuse our notation slightly, and write for any $\mathbf{s} = (r_{12}, \dots, r_{(N-1)N}) \in \mathcal{S}$:

$$\frac{\partial \mathcal{V}_{\text{int}}}{\partial r_{\alpha\beta}} \quad \text{for} \quad \frac{\partial \mathcal{V}_{\text{int}}}{\partial \zeta_{\alpha\beta}}(\mathbf{s}). \quad (3.58)$$

Also, we assume that there exists a continuously differentiable extension whenever we write $\mathbf{f}_{\alpha\beta}$ and sometimes refer to a continuously differentiable extension as an extension.

Derivation of the pointwise stress tensor

We now return to the differential equation in (3.36) for the potential part of the pointwise stress tensor. Substituting the force decomposition given in (3.50) corresponding to a continuously differentiable extension, into (3.36), we obtain

$$\text{div}_{\mathbf{x}} \boldsymbol{\sigma}_{\text{v}}(\mathbf{x}, t) = \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \langle W \mathbf{f}_{\alpha\beta} \mid \mathbf{x}_{\alpha} = \mathbf{x} \rangle. \quad (3.59)$$

On using the identity

$$\langle \mathbf{f}_{\alpha\beta} W \mid \mathbf{x}_{\alpha} = \mathbf{x} \rangle = \int_{\mathbb{R}^3} \langle \mathbf{f}_{\alpha\beta} W \mid \mathbf{x}_{\alpha} = \mathbf{x}, \mathbf{x}_{\beta} = \mathbf{y} \rangle d\mathbf{y}, \quad (3.60)$$

equation (3.59) takes the form

$$\text{div}_{\mathbf{x}} \boldsymbol{\sigma}_{\text{v}}(\mathbf{x}, t) = \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \int_{\mathbb{R}^3} \langle W \mathbf{f}_{\alpha\beta} \mid \mathbf{x}_{\alpha} = \mathbf{x}, \mathbf{x}_{\beta} = \mathbf{y} \rangle d\mathbf{y}. \quad (3.61)$$

We now note that, being anti-symmetric, the integrand in the right-hand side of the above equation satisfies all the necessary conditions for the application of Lemma C.1 given in Appendix A. Conditions (1) and (2) in Appendix A are satisfied through the regularity conditions on W . Therefore, using Lemma C.1, which was proved by

²⁵Note that (3.57) has the same form as a Lagrangian with the λ terms playing the role of Lagrange multipliers. For a static minimization problem, we seek to minimize $\mathcal{V}_{\text{int}}^*$, without violating the physical constraints relating the distances to each other. (This is equivalent to minimizing $\hat{\mathcal{V}}_{\text{int}}$ with respect to the positions of particles.) Thus, the original constrained minimization of \mathcal{V}_{int} is replaced by the problem of finding the saddle points of $\mathcal{V}_{\text{int}}^*$.

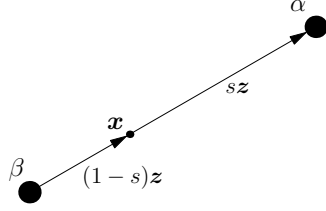


Fig. 3.1 A schematic diagram helping to explain the vectors appearing in the pointwise potential stress expression in (3.62). The bond α - β is defined by the vector \mathbf{z} . When $s = 0$, atom α is located at point \mathbf{x} , and when $s = 1$, atom β is located at \mathbf{x} .

Noll in [48], we have

$$\sigma_v(\mathbf{x}, t) \quad (3.62)$$

$$\begin{aligned} &= \frac{1}{2} \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \int_{\mathbb{R}^3} \int_{s=0}^1 \langle -\mathbf{f}_{\alpha\beta} W \mid \mathbf{x}_\alpha = \mathbf{x} + s\mathbf{z}, \mathbf{x}_\beta = \mathbf{x} - (1-s)\mathbf{z} \rangle ds \otimes \mathbf{z} d\mathbf{z} \\ &= \frac{1}{2} \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \int_{\mathbb{R}^3} \frac{\mathbf{z} \otimes \mathbf{z}}{\|\mathbf{z}\|} \int_{s=0}^1 \left\langle \frac{\partial \mathcal{V}_{\text{int}}}{\partial r_{\alpha\beta}} W \mid \mathbf{x}_\alpha = \mathbf{x} + s\mathbf{z}, \mathbf{x}_\beta = \mathbf{x} - (1-s)\mathbf{z} \right\rangle ds d\mathbf{z}, \end{aligned} \quad (3.63)$$

where in passing to the second line, we have used (3.51) and the identity $\mathbf{x}_\alpha - \mathbf{x}_\beta = \mathbf{x} + s\mathbf{z} - [\mathbf{x} - (1-s)\mathbf{z}] = \mathbf{z}$. For the special case of a pair potential, $\partial \mathcal{V}_{\text{int}} / \partial r_{\alpha\beta} = \mathcal{V}'_{\alpha\beta}(r_{\alpha\beta})$, and (3.62) reduces to the expression originally given in [48].

The expression for the potential part of the pointwise stress tensor in (3.62) is a general result applicable to all interatomic potentials. We make some important observations regarding this expressions below:

1. Although the expression for σ_v appears complex, it is actually conceptually quite simple. σ_v at a point \mathbf{x} is the superposition of the expectation values of the forces in all possible bonds passing through \mathbf{x} . The variable \mathbf{z} selects a bond length and direction and the variable s slides the bond through \mathbf{x} from end to end (see Fig. 3.1).
2. σ_v is symmetric. This is clear because the term $\mathbf{z} \otimes \mathbf{z}$ is symmetric. Since the kinetic part of the stress in (3.35) is also symmetric, the conclusion is that the *pointwise stress tensor is symmetric for all interatomic potentials*.
3. Since σ_v depends on the nature of the force decomposition and different extensions of a given potential energy can result in different force decompositions, we conclude that the pointwise stress tensor is *non-unique* for all interatomic potentials (including the pair potential). We show in Section 5.5 that the difference due to any two pointwise stress tensors, resulting from different extensions for the interatomic potential energy, tends to zero as the volume of the domain over which these pointwise quantities are spatially averaged tends to infinity. Therefore, as expected, the macroscopic stress tensor, which is defined in the thermodynamic

limit (see footnote 1 on page 5), is always unique and is independent of the potential energy extension.

4. Another source of non-uniqueness is that any expression of the form, $\sigma_v + \tilde{\sigma}$, where $\text{div}_x \tilde{\sigma} = \mathbf{0}$, also satisfies the balance of linear momentum and is therefore also a solution. We address this issue in Section 5.1, where we show that in the thermodynamic limit under equilibrium conditions, the spatially averaged counterpart to σ_v converges to the virial stress derived in Section 2.

The above results hinge on the use of the central-force decomposition in (3.50). One may wonder whether other *non-central* decompositions exist, and if yes, why are these discarded. This is discussed in the next section.

3.5 Non-central-force decompositions and the strong law of action and reaction

In the previous section, we showed that as a consequence of the invariance of the potential energy with respect to the Euclidean group, for any interatomic potential with a continuously differentiable extension, the force on a particle can always be represented as a sum of central forces. In this section, we show that other *non-central-force decompositions* are possible, however that these violate the *strong law of action and reaction*, which we prove below, and therefore do not constitute physically-meaningful force decompositions.

A proposal for a non-symmetric stress tensor for a three-body potential

As an example, let us now consider the case of a three-body potential. For simplicity, we assume that the potential only has three-body terms and all particles are identical. Under these conditions, the internal potential energy is

$$\mathcal{V}_{\text{int}} = \sum_{\substack{\alpha, \beta, \gamma \\ \alpha < \beta < \gamma}} \hat{\mathcal{V}}(\mathbf{x}_\alpha, \mathbf{x}_\beta, \mathbf{x}_\gamma), \quad (3.64)$$

where $\hat{\mathcal{V}}(\mathbf{x}_\alpha, \mathbf{x}_\beta, \mathbf{x}_\gamma)$ is the potential energy of an isolated cluster, $\{\alpha, \beta, \gamma\}$, and $\sum_{\substack{\alpha, \beta, \gamma \\ \alpha < \beta < \gamma}}$ represents a triple sum. We know that a central-force decomposition can be obtained by following the procedure outlined in the previous section and that this leads to a symmetric pointwise stress tensor in (3.62). Alternatively, a *non-symmetric* three-body stress tensor is derived as follows. To keep things simple, we derive the stress tensor for a system containing only three particles. Rewriting (3.64) for this case, we have

$$\mathcal{V}_{\text{int}} = \hat{\mathcal{V}}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \sum_{\alpha=1}^3 \phi_\alpha, \quad (3.65)$$

where

$$\phi_\alpha = \frac{1}{3} \hat{\mathcal{V}}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \quad (3.66)$$

is the potential energy assigned to particle α , equal to one-third of the total potential energy. Substituting (3.65) into (3.36), we obtain

$$\begin{aligned} \operatorname{div}_{\mathbf{x}} \boldsymbol{\sigma}_{\mathbf{v}}(\mathbf{x}, t) &= - \sum_{\alpha, \beta} \langle W \nabla_{\mathbf{x}_{\alpha}} \phi_{\beta} \mid \mathbf{x}_{\alpha} = \mathbf{x} \rangle \\ &= - \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \langle W \nabla_{\mathbf{x}_{\alpha}} \phi_{\beta} \mid \mathbf{x}_{\alpha} = \mathbf{x} \rangle - \sum_{\alpha} \langle W \nabla_{\mathbf{x}_{\alpha}} \phi_{\alpha} \mid \mathbf{x}_{\alpha} = \mathbf{x} \rangle. \end{aligned} \quad (3.67)$$

Since the cluster of three particles is isolated, the net force on the cluster due to internal interactions is zero. Therefore, from (3.65), we have

$$\nabla_{\mathbf{x}_{\alpha}} \phi_{\alpha} = - \sum_{\beta \neq \alpha} \nabla_{\mathbf{x}_{\beta}} \phi_{\alpha}. \quad (3.68)$$

Using this relation, equation (3.67) simplifies to

$$\operatorname{div}_{\mathbf{x}} \boldsymbol{\sigma}_{\mathbf{v}}(\mathbf{x}, t) = - \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \langle W (\nabla_{\mathbf{x}_{\alpha}} \phi_{\beta} - \nabla_{\mathbf{x}_{\beta}} \phi_{\alpha}) \mid \mathbf{x}_{\alpha} = \mathbf{x} \rangle. \quad (3.69)$$

Let

$$\bar{\mathbf{f}}_{\alpha\beta} := \nabla_{\mathbf{x}_{\beta}} \phi_{\alpha} - \nabla_{\mathbf{x}_{\alpha}} \phi_{\beta}. \quad (3.70)$$

Now, using the identity

$$\langle \bar{\mathbf{f}}_{\alpha\beta} W \mid \mathbf{x}_{\alpha} = \mathbf{x} \rangle = \int_{\mathbb{R}^3} \langle \bar{\mathbf{f}}_{\alpha\beta} W \mid \mathbf{x}_{\alpha} = \mathbf{x}, \mathbf{x}_{\beta} = \mathbf{y} \rangle d\mathbf{y},$$

and the definition given in (3.70), equation (3.69) takes the form

$$\operatorname{div}_{\mathbf{x}} \boldsymbol{\sigma}_{\mathbf{v}}(\mathbf{x}, t) = \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \int_{\mathbb{R}^3} \langle W \bar{\mathbf{f}}_{\alpha\beta} \mid \mathbf{x}_{\alpha} = \mathbf{x}, \mathbf{x}_{\beta} = \mathbf{y} \rangle d\mathbf{y}. \quad (3.71)$$

Let us now study the definition of $\bar{\mathbf{f}}_{\alpha\beta}$ given in (3.70). From (3.66) we have

$$\bar{\mathbf{f}}_{\alpha\beta} = -\nabla_{\mathbf{x}_{\alpha}} \phi_{\beta} + \nabla_{\mathbf{x}_{\beta}} \phi_{\alpha} = \frac{1}{3} \left[-\frac{\partial \hat{\mathcal{V}}}{\partial \mathbf{x}_{\alpha}} + \frac{\partial \hat{\mathcal{V}}}{\partial \mathbf{x}_{\beta}} \right] = \frac{1}{3} (\mathbf{f}_{\alpha}^{\text{int}} - \mathbf{f}_{\beta}^{\text{int}}). \quad (3.72)$$

The above equation suggests how the force $\mathbf{f}_{\alpha}^{\text{int}}$ is decomposed. For example, $\mathbf{f}_1^{\text{int}}$ is decomposed as

$$\mathbf{f}_1^{\text{int}} = \bar{\mathbf{f}}_{12} + \bar{\mathbf{f}}_{13} = \frac{1}{3} (\mathbf{f}_1^{\text{int}} - \mathbf{f}_2^{\text{int}}) + \frac{1}{3} (\mathbf{f}_1^{\text{int}} - \mathbf{f}_3^{\text{int}}). \quad (3.73)$$

Rearranging this relation gives

$$\mathbf{f}_1^{\text{int}} + \mathbf{f}_2^{\text{int}} + \mathbf{f}_3^{\text{int}} = \mathbf{0}, \quad (3.74)$$

which is true since the cluster $\{1, 2, 3\}$ is isolated.

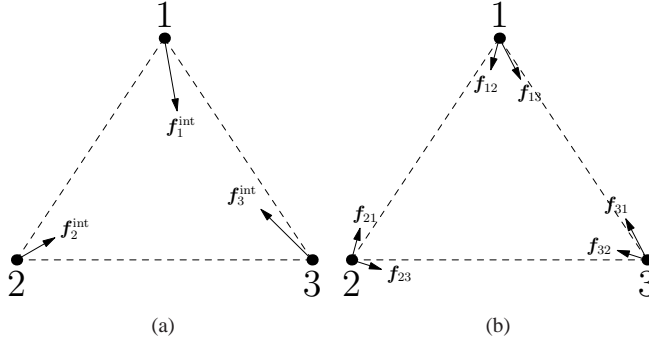


Fig. 3.2 (a) shows the force on each particle in a system consisting of 3 particles which interact through a 3-body potential given in (3.65). Since the potential is derived from an energy decomposition, we have $f_1^{\text{int}} + f_2^{\text{int}} + f_3^{\text{int}} = 0$. (b) shows the force decomposition of each f_α^{int} such that $f_{\alpha\beta} = -f_{\beta\alpha}$, but not necessarily parallel to the line joining particles α and β .

From (3.72) it is clear that $\bar{f}_{\alpha\beta}$ is anti-symmetric with respect to its arguments. Therefore, the integrand on the right-hand side of (3.71) satisfies all the necessary conditions for the application of Lemma C.1 given in Appendix C. Conditions (1) and (2) in Appendix C are satisfied through the regularity conditions on W . Therefore, using Lemma C.1, we have

$$\bar{\sigma}_v(x, t) = \frac{1}{2} \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \int_{\mathbb{R}^3} \int_{s=0}^1 \langle -(\nabla_{x_\beta} \phi_\alpha - \nabla_{x_\alpha} \phi_\beta) W \mid x_\alpha = x + sz, x_\beta = x - (1-s)z \rangle ds \otimes z dz. \quad (3.75)$$

The stress $\bar{\sigma}_v$ is non-symmetric in general because $\bar{f}_{\alpha\beta}$, defined in (3.72), need not be parallel to the line joining particles α and β as shown in Fig. 3.2. We therefore have two expressions for the stress for the same three-body potential. The symmetric expression in (3.62) and the non-symmetric expression in (3.75). We show next that the non-central-force decomposition that led to the non-symmetric stress tensor is not physically meaningful since it violates the strong law of action and reaction.

Weak and strong laws of action and reaction²⁶

The following derivation hinges on the fact that in a material system the balance laws of linear and angular momentum must be satisfied for any part of the body.

Consider a system of N particles with masses m_α ($\alpha = 1, \dots, N$). The total force on particle α is

$$f_\alpha = f_\alpha^{\text{ext}} + \sum_{\substack{\beta \\ \beta \neq \alpha}} f_{\alpha\beta}, \quad (3.76)$$

²⁶This derivation is due to Roger Fosdick [20].

where $\mathbf{f}_\alpha^{\text{ext}}$ is the external force on particle α , and, as above, $\mathbf{f}_{\alpha\beta}$ is the contribution to the force on particle α due to the presence of particle β . No assumptions are made regarding the terms $\mathbf{f}_{\alpha\beta}$ or the interatomic potential from which they are derived.

A “part” \wp_t of the system consists of $K \leq N$ particles. We suppose \mathbf{x}_0 is a fixed point in space. Let $\mathbf{F}^{\text{ext}}(\wp_t)$ denote the total force on the part \wp_t external to the part. Let $\mathbf{M}^{\text{ext}}(\wp_t; \mathbf{x}_0)$ denote the total external moment on \wp_t about \mathbf{x}_0 . Let $\mathbf{L}(\wp_t)$ be the linear momentum of the part \wp_t and $\mathbf{H}(\wp_t; \mathbf{x}_0)$ be the angular momentum of \wp_t about \mathbf{x}_0 .

We adopt the following balance laws, valid for all parts of the system:²⁷

$$\mathbf{F}^{\text{ext}}(\wp_t) = \frac{d\mathbf{L}}{dt}(\wp_t), \quad (3.77)$$

$$\mathbf{M}^{\text{ext}}(\wp_t; \mathbf{x}_0) = \frac{d\mathbf{H}}{dt}(\wp_t; \mathbf{x}_0). \quad (3.78)$$

We now show that by applying these balance laws to particular parts of the system, that the strong law of action and reaction can be established. As a first observation, let \wp_t consist of the single particle α . The external force and linear momentum for $\wp_t = \{\alpha\}$ is

$$\mathbf{F}^{\text{ext}}(\{\alpha\}) = \mathbf{f}_\alpha^{\text{ext}}(t) + \sum_{\substack{\gamma \\ \gamma \neq \alpha}} \mathbf{f}_{\alpha\gamma}(t), \quad (3.79)$$

$$\mathbf{L}(\{\alpha\}) = m_\alpha \dot{\mathbf{x}}_\alpha(t) \quad (\text{no sum}). \quad (3.80)$$

The balance of linear momentum in (3.77) requires

$$\mathbf{f}_\alpha^{\text{ext}} + \sum_{\substack{\gamma \\ \gamma \neq \alpha}} \mathbf{f}_{\alpha\gamma} = m_\alpha \ddot{\mathbf{x}}_\alpha. \quad (3.81)$$

The external moment and angular momentum of \wp_t is

$$\mathbf{M}^{\text{ext}}(\{\alpha\}; \mathbf{x}_0) = (\mathbf{x}_\alpha(t) - \mathbf{x}_0) \times (\mathbf{f}_\alpha^{\text{ext}} + \sum_{\substack{\gamma \\ \gamma \neq \alpha}} \mathbf{f}_{\alpha\gamma}) = m_\alpha (\mathbf{x}_\alpha(t) - \mathbf{x}_0) \times \ddot{\mathbf{x}}_\alpha(t), \quad (3.82)$$

where we have used (3.81), and

$$\mathbf{H}(\{\alpha\}; \mathbf{x}_0) = (\mathbf{x}_\alpha - \mathbf{x}_0) \times m_\alpha \dot{\mathbf{x}}_\alpha(t). \quad (3.83)$$

The balance of angular momentum in (3.78) is satisfied identically, since

$$\begin{aligned} m_\alpha (\mathbf{x}_\alpha(t) - \mathbf{x}_0) \times \ddot{\mathbf{x}}_\alpha(t) &= \frac{d}{dt} [(\mathbf{x}_\alpha(t) - \mathbf{x}_0) \times m_\alpha \dot{\mathbf{x}}_\alpha(t)] \\ &= \dot{\mathbf{x}}_\alpha(t) \times m_\alpha \dot{\mathbf{x}}_\alpha(t) + (\mathbf{x}_\alpha(t) - \mathbf{x}_0) \times m_\alpha \ddot{\mathbf{x}}_\alpha(t) \\ &= m_\alpha (\mathbf{x}_\alpha(t) - \mathbf{x}_0) \times \ddot{\mathbf{x}}_\alpha(t). \end{aligned}$$

²⁷The view that the balance of linear momentum and the balance of angular momentum are fundamental laws of mechanics lies at the basis of continuum mechanics. See, for example, Truesdell’s article “Whence the Law of Moment and Momentum?” in [61].

As a second observation, let \wp_t consist of the union of the two particles α and β . The external force and linear momentum are

$$\mathbf{F}^{\text{ext}}(\{\alpha, \beta\}) = \mathbf{f}_\alpha^{\text{ext}} + \mathbf{f}_\beta^{\text{ext}} + \sum_{\substack{\gamma \\ \gamma \neq \alpha \neq \beta}} (\mathbf{f}_{\alpha\gamma} + \mathbf{f}_{\beta\gamma}), \quad (3.84)$$

$$\mathbf{L}(\{\alpha, \beta\}) = m_\alpha \dot{\mathbf{x}}_\alpha + m_\beta \dot{\mathbf{x}}_\beta. \quad (3.85)$$

The balance of linear momentum in (3.77) requires

$$\mathbf{f}_\alpha^{\text{ext}} + \mathbf{f}_\beta^{\text{ext}} + \sum_{\substack{\gamma \\ \gamma \neq \alpha \neq \beta}} (\mathbf{f}_{\alpha\gamma} + \mathbf{f}_{\beta\gamma}) = m_\alpha \ddot{\mathbf{x}}_\alpha + m_\beta \ddot{\mathbf{x}}_\beta. \quad (3.86)$$

Subtracting (3.81) for particles α and β gives

$$\sum_{\substack{\gamma \\ \gamma \neq \alpha \neq \beta}} (\mathbf{f}_{\alpha\gamma} + \mathbf{f}_{\beta\gamma}) - \sum_{\substack{\gamma \\ \gamma \neq \alpha}} \mathbf{f}_{\alpha\gamma} - \sum_{\substack{\gamma \\ \gamma \neq \beta}} \mathbf{f}_{\beta\gamma} = \mathbf{0}, \quad (3.87)$$

from which

$$\mathbf{f}_{\alpha\beta} + \mathbf{f}_{\beta\alpha} = \mathbf{0}. \quad (3.88)$$

This relation is called the *weak law of action and reaction* [22]. It shows that $\mathbf{f}_{\alpha\beta} = -\mathbf{f}_{\beta\alpha}$, but does not guarantee that $\mathbf{f}_{\alpha\beta}$ lies along the line connecting particles α and β .

Next, the external moment and angular momentum of \wp_t is

$$\begin{aligned} \mathbf{M}^{\text{ext}}(\{\alpha, \beta\}; \mathbf{x}_0) &= (\mathbf{x}_\alpha - \mathbf{x}_0) \times (\mathbf{f}_\alpha^{\text{ext}} + \sum_{\substack{\gamma \\ \gamma \neq \alpha \neq \beta}} \mathbf{f}_{\alpha\gamma}) + (\mathbf{x}_\beta - \mathbf{x}_0) \times (\mathbf{f}_\beta^{\text{ext}} + \sum_{\substack{\gamma \\ \gamma \neq \beta \neq \alpha}} \mathbf{f}_{\beta\gamma}) \\ &= (\mathbf{x}_\alpha - \mathbf{x}_0) \times (m_\alpha \ddot{\mathbf{x}}_\alpha - \mathbf{f}_{\alpha\beta}) + (\mathbf{x}_\beta - \mathbf{x}_0) \times (m_\beta \ddot{\mathbf{x}}_\beta - \mathbf{f}_{\beta\alpha}), \end{aligned} \quad (3.89)$$

where we have used (3.81), and

$$\mathbf{H}(\{\alpha, \beta\}; \mathbf{x}_0) = (\mathbf{x}_\alpha - \mathbf{x}_0) \times m_\alpha \dot{\mathbf{x}}_\alpha + (\mathbf{x}_\beta - \mathbf{x}_0) \times m_\beta \dot{\mathbf{x}}_\beta. \quad (3.90)$$

The balance of angular momentum in (3.78) requires

$$\begin{aligned} &(\mathbf{x}_\alpha - \mathbf{x}_0) \times (m_\alpha \ddot{\mathbf{x}}_\alpha - \mathbf{f}_{\alpha\beta}) + (\mathbf{x}_\beta - \mathbf{x}_0) \times (m_\beta \ddot{\mathbf{x}}_\beta - \mathbf{f}_{\beta\alpha}) \\ &= \frac{d}{dt} [(\mathbf{x}_\alpha - \mathbf{x}_0) \times m_\alpha \dot{\mathbf{x}}_\alpha + (\mathbf{x}_\beta - \mathbf{x}_0) \times m_\beta \dot{\mathbf{x}}_\beta] \\ &= \dot{\mathbf{x}}_\alpha \times m_\alpha \dot{\mathbf{x}}_\alpha + (\mathbf{x}_\alpha - \mathbf{x}_0) \times m_\alpha \ddot{\mathbf{x}}_\alpha + \dot{\mathbf{x}}_\beta \times m_\beta \dot{\mathbf{x}}_\beta + (\mathbf{x}_\beta - \mathbf{x}_0) \times m_\beta \ddot{\mathbf{x}}_\beta \\ &= (\mathbf{x}_\alpha - \mathbf{x}_0) \times m_\alpha \ddot{\mathbf{x}}_\alpha + (\mathbf{x}_\beta - \mathbf{x}_0) \times m_\beta \ddot{\mathbf{x}}_\beta, \end{aligned} \quad (3.91)$$

which simplifies to

$$(\mathbf{x}_\alpha - \mathbf{x}_0) \times \mathbf{f}_{\alpha\beta} + (\mathbf{x}_\beta - \mathbf{x}_0) \times \mathbf{f}_{\beta\alpha} = \mathbf{0},$$

and, after using (3.88), we obtain

$$(\mathbf{x}_\alpha - \mathbf{x}_\beta) \times \mathbf{f}_{\alpha\beta} = \mathbf{0}. \quad (3.92)$$

This shows that $\mathbf{f}_{\alpha\beta}$ must be *parallel* to the line joining particles α and β . This is the *strong law of action and reaction*. We have shown that this law must hold for any force decomposition, in order for the balance of linear and angular momentum to hold for any subset of a system of particles.

The possibility of non-symmetric stress

Based on the proof given above for the strong law of action and reaction, we argue that only force decompositions that satisfy the strong law of action and reaction provide a physically-meaningful definition for $\mathbf{f}_{\alpha\beta}$. For example, the definition in (3.70) is not physical because if it were used to compute the external moment acting on a subsystem of particles, as is done above, the balance of angular momentum would be violated. For this reason, this decomposition and the corresponding non-symmetric stress in (3.75) are discarded. The conclusion is that *a pointwise stress tensor for a discrete system of point masses without internal structure has to be symmetric*.

In the next section, we discuss the possibility of expanding the class of solutions resulting from Irving–Kirkwood–Noll procedure in a way that makes it possible to obtain non-symmetric stress tensors for systems where the point particles have internal structure. This involves a relaxation of the assumption that the “bonds” connecting two particles are necessarily straight.

3.6 Generalized non-symmetric pointwise stress for particles with internal structure

In Section 3.4, we saw that the Irving–Kirkwood–Noll procedure, when applied to multi-body potentials, results in a symmetric non-unique pointwise stress tensor. We now seek to find additional solutions to (3.61), which are not obtained using the standard Irving–Kirkwood–Noll procedure. In arriving at (3.62) using Lemma C.1, we can see that the contribution to the potential part of the stress at position \mathbf{x} is due to all possible bonds, *assumed to be straight lines*, that pass through \mathbf{x} . The question that naturally arises is to what extent can this assumption be weakened. In other words, can Lemma C.1 be generalized in a suitable manner so that non-straight bonds can be accommodated? Such a possibility was first discussed by Schofield and Henderson [52], who used the Irving and Kirkwood approach with a series expansion of the Dirac-delta distribution. It will be shown in this section, using Noll’s more rigorous approach, that solutions giving rise to non-straight bonds are possible.

From a physical standpoint, non-straight bonds are possible for systems with internal degrees of freedom. An example would be the dipole-dipole interactions between water molecules resulting from the electrical dipole of each molecule. The possibility of internal degrees of freedom was already raised by Kirkwood in his 1946 paper [29]. The idea is to relate the shape of the non-straight bonds to the additional physics associated with the internal degrees of freedom. This issue will be further explored in future work. For now, we only investigate the possible existence

of additional solutions other than that given by (3.62). We begin by describing the shape of a bond in a more precise way through the following definition.

Definition 2 *The “path of interaction” between any two interacting particles α and β is the unique contour that connects α and β , such that there is a non-zero contribution to the potential part of the pointwise stress, σ_v , at any point on this contour.*

In this section, the terms *bond* and *path of interaction* are used synonymously. Therefore, for the case of the pointwise stress tensor in (3.62), this path of interaction is given by the straight line joining α and β . It is shown in Appendix C, that under certain restrictions on the path of interaction, Lemma C.1 can be generalized to Lemma C.2 given in Appendix C. Roughly speaking these restrictions are given by the following conditions:

1. The shape of the bond connecting particles α and β only depends on the distance between α and β .²⁸
2. For any two pairs of particles (α, β) and (γ, δ) separated by the same distance, the bonds $\alpha - \beta$ and $\gamma - \delta$ are related by a rigid body motion. In addition, if $\mathbf{x}_\alpha - \mathbf{x}_\beta = \mathbf{x}_\gamma - \mathbf{x}_\delta$ then this rigid body motion involves only translation.

From condition 1, it is clear that the shape of the bonds can be described by contours $\Upsilon_l : [0, 1] \rightarrow \mathbb{R}^3$, for $l > 0$, with $\Upsilon_l(0) = (0, 0, 0)$, $\Upsilon_l(1) = (l, 0, 0)$ along with some mild restrictions. Hence, defining the contours Υ_l for $l > 0$ is equivalent to defining the paths of interaction between any two points in \mathbb{R}^3 . For a precise definition of Υ_l and the paths of interaction see Appendix C. Since all the necessary conditions for the application of Lemma C.2 are same as those for Lemma C.1, we can use Lemma C.2 in the Irving–Kirkwood–Noll procedure instead of Lemma C.1. In doing so, we arrive at a definition for the generalized pointwise stress tensor σ_v^G , for given paths of interaction with the above mentioned properties. This is given by

$$\sigma_v^G(\mathbf{x}, t) = \frac{1}{2} \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \int_{\mathbb{R}^3} \int_{s=0}^1 \langle \mathbf{f}_{\alpha\beta} W \mid \mathbf{x}_\alpha = \mathbf{x}_\perp + s\mathbf{z}, \mathbf{x}_\beta = \mathbf{x}_\perp - (1-s)\mathbf{z} \rangle \otimes \mathbf{Q}_z \Upsilon'_{\|\mathbf{z}\|}(s) ds d\mathbf{z}, \quad (3.93)$$

where $\mathbf{f}_{\alpha\beta}$ is defined in (3.51), and $\mathbf{x}_\perp(s, \mathbf{x}, \mathbf{z}) = \mathbf{x} - s\mathbf{z} - \mathbf{Q}_z \Upsilon_{\|\mathbf{z}\|}(s)$, $\mathbf{Q}_z \in SO(3)$. Here, \mathbf{x}_\perp represents the projection of \mathbf{x} onto the line joining the endpoints, $\mathbf{x}_\alpha = \mathbf{x}_\perp + s\mathbf{z}$ and $\mathbf{x}_\beta = \mathbf{x}_\perp - (1-s)\mathbf{z}$, of the path of interaction being considered. \mathbf{Q}_z represents the rotation part of the rigid body motion described in condition 2 that maps the contour $\Upsilon_{\|\mathbf{x}_\alpha - \mathbf{x}_\beta\|}$ to the path of interaction that connects \mathbf{x}_α and \mathbf{x}_β .

Equation (3.93) is a general expression for the potential part of the pointwise stress tensor, of which (3.62) is a special case. We discuss several key features of this expression below:

²⁸ This is in essence a constitutive postulate similar to the assumption in pair potentials that the energy depends on only the distance between particles. A more general dependence of the shape of the path of interaction on the environment of a pair of particles might be possible, but is not pursued here. See also Definition 3 in Appendix C and following discussion.

1. Equation (3.93) is unique only up to given paths of interaction for a given potential energy extension. It is a more general result than (3.62), since (3.62) can be obtained from (3.93) by assuming that the path of interaction between any two points is the straight line connecting them. For this special case it is easy to see that $\mathbf{x}_\perp = \mathbf{x}$ and $\mathbf{Q}_z \mathcal{R}'_{\|z\|}(s) = -\mathbf{z}$.
2. σ_v^G is in general non-symmetric, whereas the stress obtained through the standard Irving–Kirkwood–Noll procedure is always symmetric for any multi-body potential with an extension. Since the kinetic part of the stress tensor σ_k (see (3.35)) is symmetric, it follows that the total pointwise stress tensor obtained from the generalized stress tensor is usually non-symmetric. Therefore, under the present setting, the balance of angular momentum is satisfied only through the presence of couple stresses. This suggests that non-straight bonds might correspond to systems with particles having internal degrees of freedom.
3. Since both (3.62) and (3.93) are valid definitions for the potential part of the pointwise stress, the question of which one to choose depends on the presence of internal degrees of freedom in each particle. In the absence of internal degrees of freedom only straight bonds are possible due to symmetry. The issue of the pointwise stress being unique only up to a divergence-free tensor-valued function is partially addressed here, since the expression given by the difference between the two definitions is divergence-free.
4. The expression in (3.93) is very similar to (3.62). The pointwise stress at \mathbf{x} is a superposition of the expectations of the force of all possible *bonds/paths of interaction* passing through \mathbf{x} . The vector \mathbf{z} selects an orientation and the size of the vector connecting the two ends of the bond, and s slides it through \mathbf{x} from end to end as shown in Fig. 3.3.

3.7 Definition of the pointwise traction vector

In this section, we derive the formula for the pointwise traction vector $\mathbf{t}(\mathbf{x}, \mathbf{n}; t)$ defined on the surface passing through \mathbf{x} , with normal \mathbf{n} at time t . The following derivation is based on [48] and it can be easily extended to curved paths of interaction. As usual, let \mathcal{M} denote our material system. Let $\Omega \subset \mathbb{R}^3$ be a domain in three-dimensional space with continuously differentiable surface \mathcal{S} , representing a part of the body. By this definition, each of the N point masses described by \mathcal{M} either belong to Ω or in the space surrounding Ω , denoted by Ω^c . Let \mathbf{f} denote the force exerted by the particles in Ω^c on particles in Ω . We note that in continuum mechanics \mathbf{f} is related to \mathbf{t} by

$$\mathbf{f}(t) = \int_{\mathcal{S}} \mathbf{t}(\mathbf{x}, \mathbf{n}, t) dS(\mathbf{x}), \quad (3.94)$$

where $\mathbf{n}(\mathbf{x})$ is the outer normal at $\mathbf{x} \in \mathcal{S}$. Using the Cauchy relation, $\mathbf{t}(\mathbf{x}, \mathbf{n}, t) = \boldsymbol{\sigma}(\mathbf{x}, t)\mathbf{n}$, we obtain

$$\mathbf{f}(t) = \int_{\mathcal{S}} \boldsymbol{\sigma} \mathbf{n} dS(\mathbf{x}). \quad (3.95)$$

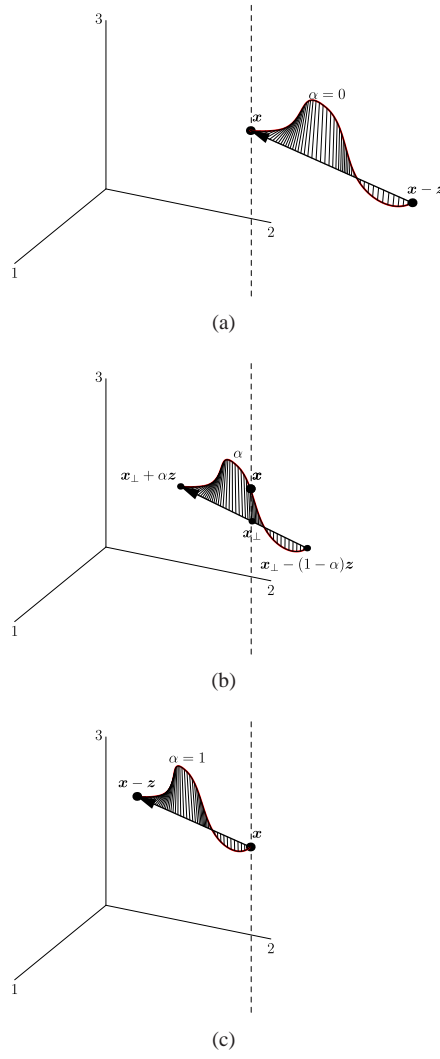


Fig. 3.3 A schematic diagram helping to explain the vectors appearing in the inner integral of (3.93) for a given point \mathbf{x} . The integral in (3.93) is an integral over all possible paths of interaction that pass through the point \mathbf{x} . The inner integral with respect to s , with \mathbf{z} fixed, is an integral over those paths, where \mathbf{z} is the vector joining its endpoints. Frame (a) shows a path of interaction when $s = 0$. As s is increased the path “slides” through \mathbf{x} . Frame (b) shows the path for an arbitrary s in the interval $[0, 1]$. The end points are represented by $\mathbf{x}_{\perp} + s\mathbf{z}$ and $\mathbf{x}_{\perp} - (1-s)\mathbf{z}$. Frame (c) shows the position of the path for $s = 1$.

Now, note that the net force exerted by Ω^c on Ω due to particle interaction, denoted by $\mathbf{f}_v(t)$, is given by

$$\mathbf{f}_v(t) = \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \int_{\mathbf{u} \in \Omega} \int_{\mathbf{v} \in \Omega^c} \langle \mathbf{f}_{\alpha\beta} W \mid \mathbf{x}_{\alpha} = \mathbf{u}, \mathbf{x}_{\beta} = \mathbf{v} \rangle d\mathbf{u} d\mathbf{v}, \quad (3.96)$$

where $\mathbf{f}_{\alpha\beta}$ is defined in (3.51). Since the integrand in (3.96) satisfies all the conditions for the application of the lemmas in Appendix C, we can now use a special case of Lemma C.3 by restricting to straight bonds.²⁹ We therefore have,

$$\begin{aligned} \mathbf{f}_v(t) &= \frac{1}{2} \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \\ &\int_S \int_{\mathbb{R}^3} \int_{s=0}^1 \langle -\mathbf{f}_{\alpha\beta} W \mid \mathbf{x}_\alpha = \mathbf{x} + s\mathbf{z}, \mathbf{x}_\beta = \mathbf{x} - (1-s)\mathbf{z} \rangle (\mathbf{z} \cdot \mathbf{n}) ds d\mathbf{z} dS(\mathbf{x}). \end{aligned} \quad (3.97)$$

We now note that \mathbf{f}_v in (3.97) exactly satisfies

$$\mathbf{f}_v(t) = \int_S \boldsymbol{\sigma}_v \mathbf{n} dS(\mathbf{x}), \quad (3.98)$$

where $\boldsymbol{\sigma}_v$ is given by (3.62). It is therefore clear that \mathbf{f}_v describes the potential part of the interaction force \mathbf{f} . Hence, it is natural to assign a potential part of the pointwise traction vector, \mathbf{t}_v to \mathbf{f}_v , given by

$$\begin{aligned} \mathbf{t}_v(\mathbf{x}, \mathbf{n}; t) &:= \boldsymbol{\sigma}_v \mathbf{n} \\ &= \frac{1}{2} \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \int_{\mathbb{R}^3} \int_{s=0}^1 \langle -\mathbf{f}_{\alpha\beta} W \mid \mathbf{x}_\alpha = \mathbf{x} + s\mathbf{z}, \mathbf{x}_\beta = \mathbf{x} - (1-s)\mathbf{z} \rangle (\mathbf{z} \cdot \mathbf{n}) ds d\mathbf{z}. \end{aligned} \quad (3.99)$$

The above formula is conceptually quite simple. *It gives the measure of the force per unit area of all the bonds that cross the surface, where the force is calculated with respect to a surface measure (see (3.97)).* Using this viewpoint, we motivate the definitions for the macroscopic traction vector and the stress tensor, when we incorporate spatial averaging in the next section.

It is now natural to assign the kinetic contribution to the force across the surface to the kinetic part of the pointwise stress tensor. Subtracting (3.98) from (3.95), we obtain the kinetic contribution to the force across a surface,

$$\mathbf{f}_k(t) := \int_S (\boldsymbol{\sigma} - \boldsymbol{\sigma}_v) \mathbf{n} dS(\mathbf{x}) = \int_S \boldsymbol{\sigma}_k \mathbf{n} dS(\mathbf{x}).$$

Therefore the kinetic contribution to the pointwise traction vector \mathbf{t}_k is given by

$$\begin{aligned} \mathbf{t}_k(\mathbf{x}, \mathbf{n}; t) &:= \boldsymbol{\sigma}_k \mathbf{n} \\ &= - \sum_{\alpha} m_{\alpha} \langle \mathbf{v}_{\alpha}^{\text{rel}} (\mathbf{v}_{\alpha}^{\text{rel}} \cdot \mathbf{n}) W \mid \mathbf{x}_{\alpha} = \mathbf{x} \rangle. \end{aligned} \quad (3.100)$$

Finally, we note that the definitions of \mathbf{t}_v and \mathbf{t}_k are functions of \mathbf{x} and \mathbf{n} alone. Hence, this result is related to the work of Fosdick and Virga [21], who give a variational proof for the stress theorem of Cauchy in the continuum version. In that work

²⁹Specifically, for straight bonds, we set: $\mathbf{x}_{\perp} = \mathbf{x}$ and $\mathbf{Q}_z \mathcal{I}'_{\|\mathbf{z}\|}(s) = -\mathbf{z}$.

the traction vector is allowed to depend on the unit normal and the surface gradient and is shown to be independent of the surface gradient.

The fields defined and derived in this section are pointwise quantities. In the next section, expressions for macroscopic fields are obtained by spatially averaging the pointwise fields over an appropriate macroscopic domain.

4 Spatial averaging

In the previous section, the Irving–Kirkwood–Noll procedure was used to construct pointwise fields from the underlying discrete microscopic system using the principles of classical statistical mechanics. Although the resulting fields resemble the continuum mechanics fields and satisfy the continuum conservation equations, they are not macroscopic continuum fields. For example, the pointwise stress field in (3.62), at sufficiently low temperature, will be highly non-uniform, exhibiting a criss-cross pattern with higher stresses along bond directions, even when macroscopically the material is nominally under uniform or even zero stress.

To measure the fields derived in the previous section in an experiment, one needs a probe which can extract data only from a single point of interest in space. Since this is not possible practically, there is no way we can correlate the experimental data with theoretical predictions. Therefore a true macroscopic quantity is by necessity an average over some spatial region surrounding the continuum point where it is nominally defined.³⁰ Thus, if $f(\mathbf{x}, t)$ is an Irving–Kirkwood–Noll pointwise field, such as density or stress, the corresponding macroscopic field $f_w(\mathbf{x}, t)$ is given by

$$f_w(\mathbf{x}, t) = \int_{\mathbb{R}^3} w(\mathbf{y} - \mathbf{x}) f(\mathbf{y}, t) d\mathbf{y}, \quad (4.1)$$

where $w(\mathbf{r})$ is a weighting function representing the properties of the probe and its lengthscale.

The important thing to note is that due to the linearity of the phase averaging in the Irving–Kirkwood–Noll procedure, the averaged macroscopic function $f_w(\mathbf{x}, t)$ satisfies the same balance equations as does the pointwise measure $f(\mathbf{x}, t)$.

Weighting function

The weighting function $w(\mathbf{r})$ is an \mathbb{R}^+ -valued function with compact support so that $w(\mathbf{r}) = 0$ for $\|\mathbf{r}\| > \lambda$, where λ is a microstructural lengthscale. The weighting function has units of volume⁻¹ and must satisfy the normalization condition

$$\int_{\mathbb{R}^3} w(\mathbf{r}) d\mathbf{r} = 1. \quad (4.2)$$

³⁰We do not include time averaging, because this is indirectly performed due to the presence of W . The reasoning for this comes from the *frequentist's* interpretation of probability, wherein the probability of a state is equal to the fraction of the total time spent by the system in that state.

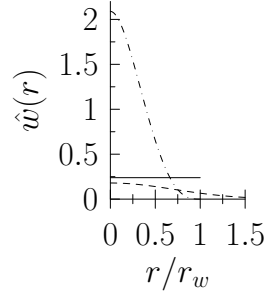


Fig. 4.1 Three weighting functions for spatial averaging: uniform weighting (solid line) in (4.3); Gaussian weighting (dashed line) in (4.4); Quartic spline weighting (dash-dot line) in (4.5). Note that the areas under the curves are not equal because the normalization in (4.2) is according to volume.

This condition ensures that the correct macroscopic stress is obtained when the point-wise stress is uniform. For a spherically-symmetric distribution, $w(\mathbf{r}) = \hat{w}(r)$, where $r = \|\mathbf{r}\|$. The normalization condition in this case is

$$\int_0^\infty \hat{w}(r) 4\pi r^2 dr = 1.$$

The simplest choice for $\hat{w}(r)$ is a spherically-symmetric uniform distribution over a specified radius r_w , given by

$$\hat{w}(r) = \begin{cases} 1/V_w & \text{if } r \leq r_w, \\ 0 & \text{otherwise,} \end{cases} \quad (4.3)$$

where $V_w = \frac{4}{3}\pi r_w^3$ is the volume of the sphere. This function is discontinuous at $r = r_w$. If this is a concern, a "mollifying" function that smoothly takes $w(r)$ to zero at r_w over some desired range can be added [45].³¹ Another possible choice for $\hat{w}(r)$ is a Gaussian function [23]

$$\hat{w}(r) = \pi^{-\frac{3}{2}} r_w^{-3} \exp \left[-r^2/r_w^2 \right]. \quad (4.4)$$

This function does not have compact support. However it decays rapidly with distance so that a numerical cutoff can be imposed where its value drops below a specified tolerance. Another possibility is a quartic spline used in meshless method applications (where it is called a *kernel function* [2]),

$$\hat{w}(r) = \begin{cases} \frac{105}{16\pi r_w^3} \left(1 + 3\frac{r}{r_w}\right) \left(1 - \frac{r}{r_w}\right)^3 & \text{if } r \leq r_w, \\ 0 & \text{otherwise.} \end{cases} \quad (4.5)$$

This spline has the advantage that it goes smoothly to zero at $r = r_w$, i.e., $\hat{w}(r_w) = 0$, $\hat{w}'(r_w) = 0$, and $\hat{w}''(r_w) = 0$. Fig. 4.1 shows the plots of the three weighting functions given above.

³¹An example of a mollifying function is given later in equation (6.13).

4.1 Spatial averaging and macroscopic fields

Continuum fields such as density and momentum density fields are defined using (4.1) as the ensemble average via the probability density function W , followed by a spatial average via the weight function w as follows:

$$\rho_w(\mathbf{x}, t) := \sum_{\alpha} m_{\alpha} \int_{\mathbb{R}^3} w(\mathbf{y} - \mathbf{x}) \langle W \mid \mathbf{x}_{\alpha} = \mathbf{y} \rangle d\mathbf{y}, \quad (4.6)$$

$$\mathbf{p}_w(\mathbf{x}, t) := \sum_{\alpha} m_{\alpha} \int_{\mathbb{R}^3} w(\mathbf{y} - \mathbf{x}) \langle W \mathbf{v}_{\alpha} \mid \mathbf{x}_{\alpha} = \mathbf{y} \rangle d\mathbf{y}. \quad (4.7)$$

It is straightforward to show that using definitions, (4.6) and (4.7), the macroscopic version of the generalized pointwise stress tensor given by (3.93) divides into potential and kinetic parts as,

$$\begin{aligned} \sigma_{w,v}(\mathbf{x}, t) &= \frac{1}{2} \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \int_{\mathbb{R}^3} w(\mathbf{y} - \mathbf{x}) \int_{s=0}^1 \int_{\mathbb{R}^3} \langle \mathbf{f}_{\alpha\beta} W \mid \mathbf{x}_{\alpha} = \mathbf{y}_{\perp} + s\mathbf{z}, \mathbf{x}_{\beta} = \mathbf{y}_{\perp} - (1-s)\mathbf{z} \rangle \otimes \mathbf{Q}_z \mathbf{r}'_{\parallel z}(s) ds d\mathbf{z} d\mathbf{y}, \\ &\quad (4.8) \end{aligned}$$

where $\mathbf{f}_{\alpha\beta}$ is defined in (3.51), $\mathbf{y}_{\perp} = \mathbf{y} - s\mathbf{z} - \mathbf{Q}_z \mathbf{r}_{\parallel z}(s)$, $\mathbf{Q}_z \in SO(3)$, and

$$\sigma_{w,k}(\mathbf{x}, t) = - \sum_{\alpha} \int_{\mathbb{R}^3} w(\mathbf{y} - \mathbf{x}) m_{\alpha} \langle (\mathbf{v}_{\alpha}^{\text{rel}} \otimes \mathbf{v}_{\alpha}^{\text{rel}}) W \mid \mathbf{x}_{\alpha} = \mathbf{y} \rangle d\mathbf{y}. \quad (4.9)$$

We now intend to express the potential part of stress in a more convenient form. This is done by two consecutive changes of variables. Under the assumption that \mathbf{Q}_z and $\mathbf{r}_{\parallel z}$ are differentiable with respect to \mathbf{z} and $\|\mathbf{z}\|$, respectively, the Jacobian of the transformation $(s, \mathbf{y}, \mathbf{z}) \mapsto (s, \mathbf{y}_{\perp}, \mathbf{z})$ is unity. Therefore,

$$\begin{aligned} \sigma_{w,v}(\mathbf{x}, t) &= \frac{1}{2} \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \int_{\mathbb{R}^3} w(\mathbf{y} - \mathbf{x}) \int_{s=0}^1 \int_{\mathbb{R}^3} \langle \mathbf{f}_{\alpha\beta} W \mid \mathbf{x}_{\alpha} = \mathbf{y}_{\perp} + s\mathbf{z}, \mathbf{x}_{\beta} = \mathbf{y}_{\perp} - (1-s)\mathbf{z} \rangle \otimes \mathbf{Q}_z \mathbf{r}'_{\parallel z}(s) ds d\mathbf{z} d\mathbf{y}_{\perp}, \\ &\quad (4.10) \end{aligned}$$

where $\mathbf{y} = \mathbf{y}(s, \mathbf{y}_{\perp}, \mathbf{z})$. A second change of variables is introduced as follows

$$\mathbf{y}_{\perp} + s\mathbf{z} = \mathbf{u}, \quad \mathbf{y}_{\perp} - (1-s)\mathbf{z} = \mathbf{v}, \quad (4.11)$$

which implies,

$$\mathbf{z} = \mathbf{u} - \mathbf{v}, \quad \mathbf{y}_{\perp} = (1-s)\mathbf{u} + s\mathbf{v}. \quad (4.12)$$

The Jacobian of the transformation is

$$J = \det \begin{bmatrix} \nabla_{\mathbf{u}} \mathbf{z} & \nabla_{\mathbf{v}} \mathbf{z} \\ \nabla_{\mathbf{u}} \mathbf{y}_{\perp} & \nabla_{\mathbf{v}} \mathbf{y}_{\perp} \end{bmatrix} = \det \begin{bmatrix} \mathbf{I} & -\mathbf{I} \\ (1-s)\mathbf{I} & s\mathbf{I} \end{bmatrix} = 1. \quad (4.13)$$

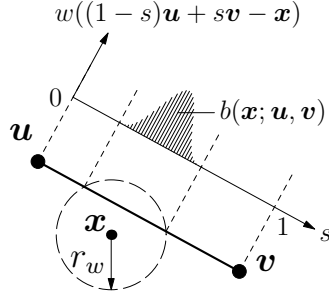


Fig. 4.2 The bond function $b(\mathbf{x}; \mathbf{u}, \mathbf{v})$ is the integral of the weighting function centered at \mathbf{x} along the line connecting points \mathbf{u} and \mathbf{v} . The graph shows the result for a quartic spline weighting function. The bond function is the area under the curve.

Using (4.11), (4.12) and (4.13) to rewrite (4.10), we obtain

$$\sigma_{w,v}(\mathbf{x}, t) = \frac{1}{2} \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \int_{\mathbb{R}^3 \times \mathbb{R}^3} \langle -\mathbf{f}_{\alpha\beta} W \mid \mathbf{x}_\alpha = \mathbf{u}, \mathbf{x}_\beta = \mathbf{v} \rangle \otimes \mathbf{b}(\mathbf{x}; \mathbf{u}, \mathbf{v}) d\mathbf{u} d\mathbf{v}, \quad (4.14)$$

where

$$\mathbf{b}(\mathbf{x}; \mathbf{u}, \mathbf{v}) := - \int_{s=0}^1 w(\hat{\mathbf{y}} - \mathbf{x}) \mathbf{Q}_{\mathbf{u}-\mathbf{v}} \mathbf{\Upsilon}'_{\|\mathbf{u}-\mathbf{v}\|} ds \quad (4.15)$$

is called the *bond vector*, with

$$\hat{\mathbf{y}}(s, \mathbf{u}, \mathbf{v}) = \mathbf{y}(s, \mathbf{y}_\perp(s, \mathbf{u}, \mathbf{v}), \mathbf{z}(\mathbf{u}, \mathbf{v})).$$

For the special case of straight bonds, we have

$$\hat{\mathbf{y}} = (1-s)\mathbf{u} + s\mathbf{v} \quad \text{and} \quad \mathbf{Q}_{\mathbf{u}-\mathbf{v}} \mathbf{\Upsilon}'_{\|\mathbf{u}-\mathbf{v}\|}(s) = -(\mathbf{u} - \mathbf{v}).$$

Therefore the bond vector simplifies to

$$\begin{aligned} \mathbf{b}(\mathbf{x}; \mathbf{u}, \mathbf{v}) &= (\mathbf{u} - \mathbf{v}) \int_{s=0}^1 w((1-s)\mathbf{u} + s\mathbf{v} - \mathbf{x}) ds \\ &= (\mathbf{u} - \mathbf{v}) b(\mathbf{x}; \mathbf{u}, \mathbf{v}), \end{aligned} \quad (4.16)$$

where $b(\mathbf{x}; \mathbf{u}, \mathbf{v})$ is commonly referred to as the *bond function*. The geometrical significance of the bond function is explained in Fig. 4.2.

For the special case of straight bonds, equation (4.14) simplifies to

$$\sigma_{w,v}(\mathbf{x}, t) = \frac{1}{2} \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \int_{\mathbb{R}^3 \times \mathbb{R}^3} \langle -\mathbf{f}_{\alpha\beta} W \mid \mathbf{x}_\alpha = \mathbf{u}, \mathbf{x}_\beta = \mathbf{v} \rangle \otimes (\mathbf{u} - \mathbf{v}) b(\mathbf{x}; \mathbf{u}, \mathbf{v}) d\mathbf{u} d\mathbf{v}. \quad (4.17)$$

The expressions for the potential and kinetic parts of the spatially-averaged stress tensor in equations (4.9) and (4.17) are our main result and constitute the general definitions for the macroscopic stress computed for a discrete system. It will be shown in Section 5 that these relations reduce to the Hardy stress tensor [23] under suitable approximations. The issue of the uniqueness of the stress tensor (in the sense that any divergence-free field can be added to it) is deferred to Section 5.5.

4.2 Comparison with the Murdoch–Hardy procedure

An alternative procedure of defining continuum fields to the one described above, due to Murdoch [43, 46, 47] and Hardy [23], only involves spatial averaging. We refer to this approach as the *Murdoch–Hardy procedure*. Under the Murdoch–Hardy procedure, continuum fields are defined as direct spatial averages of microscopic variables without incorporating statistical mechanics ideas. Therefore, the Murdoch–Hardy procedure is purely deterministic in nature. For example, the density and momentum density fields at a particular instant of time, corresponding to a given weighting function w , are defined as

$$\tilde{\rho}_w(\mathbf{x}, t) := \sum_{\alpha} m_{\alpha} w(\mathbf{x}_{\alpha}(t) - \mathbf{x}), \quad (4.18a)$$

$$\tilde{\mathbf{p}}_w(\mathbf{x}, t) := \sum_{\alpha} m_{\alpha} \mathbf{v}_{\alpha}(t) w(\mathbf{x}_{\alpha}(t) - \mathbf{x}), \quad (4.18b)$$

respectively, where \mathbf{x}_{α} and \mathbf{v}_{α} are deterministic quantities. We denote spatially-averaged variables obtained from the Murdoch–Hardy procedure with a superposed tilde to distinguish them from quantities obtained in Section 4.1. Equation (4.18) is used to “smear” a discrete system to form a continuum. The reasoning for abandoning statistical mechanics is the lack of knowledge of the ensemble of the system as explained by Murdoch and Bedeaux in [46]:

Physical interpretations of any given ensemble average clearly depends on the definition of the ensemble . . . for example, if a container is filled to a given level with water and then poured onto a surface, the lack of precision with which the pouring is effected may result in many different macroscopic flows. Here no single description is available within deterministic continuum mechanics: in this case the ensemble (defined in terms of the water molecules and limited knowledge of how the pouring takes place) relations involve averages associated with all possible flows. Clearly relations involving ensemble averages are associated with a much greater variety of behavior than is describable in terms of deterministic continuum mechanics.

We share the same concern regarding the ambiguity in the definition of an ensemble. For example, in an experiment where an austenite-martensite phase transformation occurs, the resulting micro-structure consists of a complex spatial configuration of martensitic variants, and this depends largely on the microscopic details of the system, such as cracks, lattice defects, etc. Therefore, in this case, macroscopic variables cannot completely describe the ensemble of interest. To avoid this difficulty, Murdoch proposes a time average in place of ensemble average. Nevertheless, it should be noted that from classical statistical mechanics, the ensemble of interest and its corresponding distribution exists *in principle*. Therefore the framework described in Section 4.1 is a correct framework in which to phrase the problem. A practical calculation can then be performed, for example, by replacing the ensemble averages with time averages in a molecular dynamics calculation (see Section 5). We stress the importance of writing a continuum field variable as an ensemble average followed by spatial average, rather than a spatial average followed by a time average, as is done

in the Murdoch–Hardy procedure, because it helps to give a unified picture of all the previous definitions for continuum fields and stress in particular. This is discussed in the next section.

It is interesting to note that by relaxing the connection with statistical mechanics, the Murdoch–Hardy procedure allows for a much wider class of definitions for the stress tensor [45] in addition to the non-uniqueness characterized so far, due to the presence of multiple extensions for the potential energy and allowing non-straight bonds. In this section we intend to systematize this procedure. The source of non-uniqueness resulting from multiple definitions of the stress tensor is studied, thus helping us to identify a much larger class of possible definitions. In this new systematic approach, the steps involved in the Murdoch–Hardy procedure are as follows:

1. Develop a continuum system by smearing out the discrete system using (4.18).
2. Introduce a *non-local* constitutive law for the continuum that is consistent with the discrete version of force balance given later in (4.19).
3. For each constitutive law, define a stress tensor, which satisfies the equation of motion for the continuum.

To understand the above three steps, we explore the Murdoch–Hardy procedure in more detail. The continuity equation is satisfied in a trivial way [44]. We now look at the equation of motion.

Equation of motion

The motion of particle α is governed by Newton’s second law,

$$\sum_{\substack{\beta \\ \beta \neq \alpha}} \mathbf{f}_{\alpha\beta}^d(t) + \mathbf{b}_{\alpha}^d(t) = m_{\alpha} \dot{\mathbf{v}}_{\alpha}(t), \quad (4.19)$$

where $\mathbf{f}_{\alpha\beta}^d(t) := \mathbf{f}_{\alpha\beta}(\mathbf{u}(t))$, $\mathbf{f}_{\alpha\beta}(\mathbf{u})$ are the terms in the central-force decomposition obtained from a multi-body potential with an extension (see Section 3.4) and $\mathbf{b}_{\alpha}^d(t)$ is defined as

$$\mathbf{b}_{\alpha}^d(t) := -\nabla_{\mathbf{x}_{\alpha}} \mathcal{V}_{\text{ext}}(\mathbf{x}_1(t), \dots, \mathbf{x}_N(t)).$$

The superscript “d” in (4.19) and the above equation are used to stress the fact that the quantities are deterministic in nature. Equation (4.19) is a force balance equation for the discrete system. We now design an analogous force balance equation for the smeared continuum defined by (4.18), such that (4.19) always holds.

For the sake of simplicity in notation, from here onwards we use $\mathbf{f}_{\alpha\beta}$ to denote both $\mathbf{f}_{\alpha\beta}(\mathbf{u})$ and $\mathbf{f}_{\alpha\beta}^d(t)$, whenever it is clear from the context. The same goes with the usage of $\mathbf{b}_{\alpha}(t)$ for $\mathbf{b}_{\alpha}^d(t)$.

Force balance for the smeared continuum

Multiplying (4.19) by $w(\mathbf{x}_i - \mathbf{x})$ and summing over all particles, we have

$$\tilde{\mathbf{f}}_w + \tilde{\mathbf{b}}_w = \sum_{\alpha} m_{\alpha} \dot{\mathbf{v}}_{\alpha} w(\mathbf{x}_{\alpha} - \mathbf{x}), \quad (4.20)$$

where

$$\tilde{\mathbf{f}}_w(\mathbf{x}, t) := \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \mathbf{f}_{\alpha\beta}(t) w(\mathbf{x}_\alpha(t) - \mathbf{x}), \quad (4.21)$$

$$\tilde{\mathbf{b}}_w(\mathbf{x}, t) := \sum_{\beta} \mathbf{b}_\beta(t) w(\mathbf{x}_\beta(t) - \mathbf{x}). \quad (4.22)$$

To arrive at a form similar to the equation of motion of continuum mechanics given in (3.23), equation (4.20) is rewritten as

$$\begin{aligned} \tilde{\mathbf{f}}_w + \tilde{\mathbf{b}}_w &= \frac{\partial}{\partial t} \sum_{\alpha} m_{\alpha} w(\mathbf{x}_{\alpha} - \mathbf{x}) \mathbf{v}_{\alpha} - \sum_{\alpha} m_{\alpha} \mathbf{v}_{\alpha} (\nabla w(\mathbf{x}_{\alpha} - \mathbf{x}) \cdot \mathbf{v}_{\alpha}) \\ &= \frac{\partial \tilde{\mathbf{p}}_w}{\partial t} + \operatorname{div}_{\mathbf{x}} \sum_{\alpha} m_{\alpha} w(\mathbf{x}_{\alpha} - \mathbf{x}) \mathbf{v}_{\alpha} \otimes \mathbf{v}_{\alpha}. \end{aligned} \quad (4.23)$$

Similar to (3.16), we define the continuum velocity as

$$\tilde{\mathbf{v}}_w(\mathbf{x}, t) := \frac{\tilde{\mathbf{p}}_w(\mathbf{x}, t)}{\tilde{\rho}_w(\mathbf{x}, t)}, \quad (4.24)$$

and the relative velocity of a particle with respect to the continuum velocity as

$$\tilde{\mathbf{v}}_{\alpha}^{\text{rel}}(\mathbf{x}, t) := \mathbf{v}_{\alpha}(t) - \tilde{\mathbf{v}}_w(\mathbf{x}, t). \quad (4.25)$$

Using (4.18) and (4.25), we obtain

$$\begin{aligned} \sum_{\alpha} m_{\alpha} \tilde{\mathbf{v}}_{\alpha}^{\text{rel}} w(\mathbf{x}_{\alpha}(t) - \mathbf{x}) &= \tilde{\mathbf{p}}_w(\mathbf{x}, t) - \tilde{\rho}_w(\mathbf{x}, t) \tilde{\mathbf{v}}_w(\mathbf{x}, t) \\ &= \mathbf{0}, \end{aligned}$$

the last equality being true by the definition (4.24). From (4.25) and the above equation, it follows that

$$\sum_{\alpha} m_{\alpha} w(\mathbf{x}_{\alpha} - \mathbf{x}) \mathbf{v}_{\alpha} \otimes \mathbf{v}_{\alpha} = \sum_{\alpha} m_{\alpha} w(\mathbf{x}_{\alpha} - \mathbf{x}) \tilde{\mathbf{v}}_{\alpha}^{\text{rel}} \otimes \tilde{\mathbf{v}}_{\alpha}^{\text{rel}} + \tilde{\rho}_w \tilde{\mathbf{v}}_w \otimes \tilde{\mathbf{v}}_w.$$

Substituting this into (4.23) and rearranging, we have

$$\tilde{\mathbf{f}}_w - \operatorname{div}_{\mathbf{x}} \sum_{\alpha} m_{\alpha} (\tilde{\mathbf{v}}_{\alpha}^{\text{rel}} \otimes \tilde{\mathbf{v}}_{\alpha}^{\text{rel}}) w(\mathbf{x}_{\alpha} - \mathbf{x}) + \tilde{\mathbf{b}}_w = \frac{\partial \tilde{\mathbf{p}}_w}{\partial t} + \operatorname{div}_{\mathbf{x}} (\tilde{\rho}_w \tilde{\mathbf{v}}_w \otimes \tilde{\mathbf{v}}_w).$$

Comparing the above equation with the equation of motion, (3.23), we have

$$\operatorname{div}_{\mathbf{x}} \tilde{\boldsymbol{\sigma}}_w(\mathbf{x}, t) = \tilde{\mathbf{f}}_w(\mathbf{x}, t) - \operatorname{div}_{\mathbf{x}} \sum_{\alpha} m_{\alpha} (\tilde{\mathbf{v}}_{\alpha}^{\text{rel}} \otimes \tilde{\mathbf{v}}_{\alpha}^{\text{rel}}) w(\mathbf{x}_{\alpha} - \mathbf{x}), \quad (4.26)$$

where $\tilde{\sigma}_w$ is the stress tensor corresponding to the weighting function w . From (4.26) it is clear that the kinetic part and the potential part of the stress tensor, $\tilde{\sigma}_{w,k}$ and $\tilde{\sigma}_{w,v}$, respectively, are given by

$$\tilde{\sigma}_{w,k}(\mathbf{x}, t) = - \sum_{\alpha} m_{\alpha} (\tilde{\mathbf{v}}_{\alpha}^{\text{rel}} \otimes \tilde{\mathbf{v}}_{\alpha}^{\text{rel}}) w(\mathbf{x}_{\alpha} - \mathbf{x}), \quad (4.27a)$$

$$\text{div}_{\mathbf{x}} \tilde{\sigma}_{w,v}(\mathbf{x}, t) = \tilde{\mathbf{f}}_w(\mathbf{x}, t). \quad (4.27b)$$

Any solution to (4.27b) is a valid candidate for the definition of $\tilde{\sigma}_{w,v}$. Murdoch [45] proposes several possible candidates, and highlights the possibility of having multiple definitions. To understand the connection between the different possible definitions, we look back at (4.20) and (4.21). Equation (4.20) is a force balance equation for any “continuum particle” at \mathbf{x} , and $\tilde{\mathbf{f}}_w$, defined in (4.21), is the force per unit volume acting on it. It is not immediately clear from (4.21) how two continuum particles at positions \mathbf{x} and \mathbf{y} interact with each other. This interaction can be given by a non-local constitutive law. The main idea is to recast (4.21) as

$$\tilde{\mathbf{f}}_w(\mathbf{x}, t) = \int_{\mathbb{R}^3} \mathbf{g}(\mathbf{x}, \mathbf{y}, t) d\mathbf{y}, \quad (4.28)$$

for some $\mathbf{g}(\mathbf{x}, \mathbf{y}, t)$, which we call the *generator of the non-local constitutive law*. This function describes the interaction between the continuum particles at \mathbf{x} and \mathbf{y} . To satisfy Newton’s third law, we also need \mathbf{g} to be anti-symmetric with respect to its arguments \mathbf{x} and \mathbf{y} . Unfortunately the representation given in (4.28) is not unique and, since every choice of \mathbf{g} leads to a different stress definition, this is one of the sources of non-uniqueness in the definition for the stress tensor in the Murdoch–Hardy procedure. We describe two different constitutive laws, which lead to the Hardy stress and the doubly-averaged stress (DA stress)³² [47].

For the case of Hardy stress, the generator \mathbf{g}^H is given by the equation

$$\mathbf{g}^H(\mathbf{x}, \mathbf{y}, t) = \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \mathbf{f}_{\alpha\beta} w(\mathbf{x}_{\alpha} - \mathbf{x}) \delta(\mathbf{x}_{\beta} - \mathbf{x}_{\alpha} + \mathbf{x} - \mathbf{y}), \quad (4.29)$$

where δ denotes the Dirac delta distribution.

The generator \mathbf{g}^D for the DA stress is given by

$$\mathbf{g}^D(\mathbf{x}, \mathbf{y}, t) = \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \mathbf{f}_{\alpha\beta} w(\mathbf{x}_{\alpha} - \mathbf{x}) w(\mathbf{x}_{\beta} - \mathbf{y}). \quad (4.30)$$

Fig. 4.3 shows the interaction between two continuum particles, with positions \mathbf{x} and \mathbf{y} , that are in a neighborhood of two interacting particles α and β respectively and not in a neighborhood of any other particle in the system. In this setup, it is clear from the generator for the Hardy stress, given in (4.29), that two continuum particles at \mathbf{x} and \mathbf{y} interact only when $\mathbf{y} - \mathbf{x}_{\beta} = \mathbf{x} - \mathbf{x}_{\alpha}$, as shown in Fig. 4.3(a).

³²Murdoch [45] refers to this stress as “Noll’s choice”. To avoid confusion with the stress derived through the Irving–Kirkwood–Noll procedure in Section 3, we name it the “DA stress”.

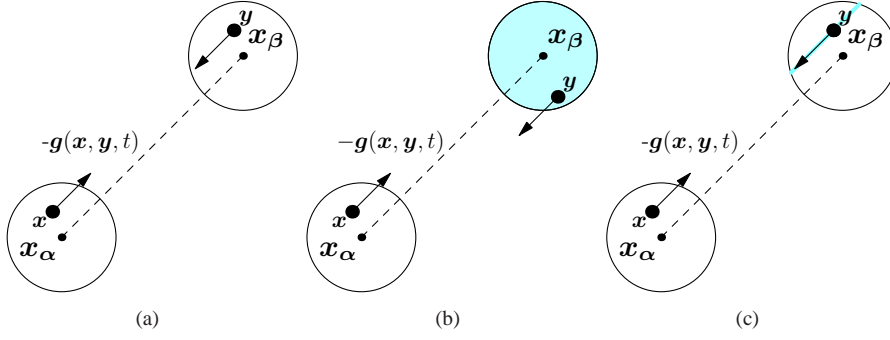


Fig. 4.3 A continuum particle x interacts with: (a) only that continuum particle at y , which is identically oriented to x_β as x is oriented to x_α , when the interaction is given by g^H ; (b) any continuum particle in the shaded region, when the interaction is given by g^D ; (c) any continuum particle on the shaded line, when the interaction is given by g^{HD} .

On the other hand, there is no such restriction on the generator for the DA stress, described by (4.30) (see Fig. 4.3(b)).³³ Although at this point there is no systematic way of suggesting additional possible generators, we can suggest a third generator, g^{HD} , which has properties that lie in between g^H and g^D . As shown in Fig. 4.3(c), when interaction is governed by g^{HD} , a continuum particle x interacts with y only when y lies on the line passing through x and parallel to $x_\alpha - x_\beta$. In all the three cases, the interaction force is always directed along the vector $x_\alpha - x_\beta$. Therefore by (4.28) we have three different integral representations for \tilde{f}_w with generators g^D , g^H , g^{HD} .

Now, in each of the integral representations of \tilde{f}_w given by (4.29) and (4.30), the integrand satisfies all the necessary conditions for the application of Lemma C.1 or Lemma C.2 in Appendix C.³⁴ For instance, using Lemma C.1 we obtain an expression for the potential part of the stress tensor, given by

$$\tilde{\sigma}_{w,v}(x, t) = -\frac{1}{2} \int_{\mathbb{R}^3} \left[\int_{s=0}^1 g(x + sz, x - (1-s)z, t) ds \right] \otimes z dz. \quad (4.31)$$

³³Note that for the DA stress, the force between two continuum particles at x and y is *not* parallel to $x - y$ in general. This is not a violation of the strong law of action and reaction, because the strong law only applies to discrete systems. It has been used in this derivation by requiring that $f_{\alpha\beta} = -f_{\beta\alpha}$ and $f_{\alpha\beta} \times (x_\alpha - x_\beta) = 0$.

³⁴The integrand should be continuously differentiable for Noll's lemma to be applicable. Although g^H is not continuously differentiable due to the presence of the Dirac delta distribution, this does not hinder us from applying the lemma since we can replace the Dirac delta distribution by an appropriate infinitely differentiable delta sequence and take a limit. See Appendix D for a rigorous derivation of this.

Substituting (4.29) into (4.31), we have the potential part of the Hardy stress:

$$\begin{aligned}\tilde{\sigma}_{w,v}^H &= -\frac{1}{2} \sum_{\substack{\alpha,\beta \\ \alpha \neq \beta}} \int_{\mathbb{R}^3} \left[\int_{s=0}^1 \mathbf{f}_{\alpha\beta} w(\mathbf{x}_\alpha - \mathbf{x} - s\mathbf{z}) \delta(\mathbf{x}_\beta - \mathbf{x}_\alpha + \mathbf{z}) ds \right] \otimes \mathbf{z} d\mathbf{z} \\ &= \frac{1}{2} \sum_{\substack{\alpha,\beta \\ \alpha \neq \beta}} \int_{s=0}^1 [-\mathbf{f}_{\alpha\beta} w((1-s)\mathbf{x}_\alpha + s\mathbf{x}_\beta - \mathbf{x}) \otimes (\mathbf{x}_\alpha - \mathbf{x}_\beta)] ds. \quad (4.32)\end{aligned}$$

Substituting (4.30) into (4.31) we have the potential part of the DA stress:

$$\tilde{\sigma}_{w,v}^{\text{DA}} = \frac{1}{2} \sum_{\substack{\alpha,\beta \\ \alpha \neq \beta}} \int_{\mathbf{z} \in \mathbb{R}^3} \int_{s=0}^1 [-\mathbf{f}_{\alpha\beta} w(\mathbf{x}_\alpha - \mathbf{x} - s\mathbf{z}) w(\mathbf{x}_\beta - \mathbf{x} + (1-s)\mathbf{z}) \otimes \mathbf{z}] ds d\mathbf{z}, \quad (4.33)$$

which was derived by Murdoch [47]. The conclusion is that the non-uniqueness of the generator in the systematized Murdoch–Hardy procedure leads to a non-unique definition for the stress tensor. Further sources of non-uniqueness can be introduced by having a different force decomposition corresponding to a different potential extension, or using curved paths of interaction instead of straight bonds and applying Lemma C.2, which is a generalization of Lemma C.1 in Appendix C. We do not pursue this generalization further here.

It is important to point out that the systematized Murdoch–Hardy procedure presented here does *not* describe all possible solutions to (4.27b). An example of a solution that cannot be obtained via our systematized Murdoch–Hardy procedure is the following definition suggested in [45]:

$$\tilde{\sigma}_{w,v}^*(\mathbf{x}, t) := \sum_{\substack{\alpha,\beta \\ \alpha \neq \beta}} \mathbf{f}_{\alpha\beta} \otimes (\mathbf{x} - \mathbf{x}_\alpha) \hat{a}(\|\mathbf{x} - \mathbf{x}_\alpha\|), \quad (4.34)$$

where $\hat{a}(u) := \frac{1}{u^3} \int_0^u s^2 \hat{w}(s) ds$. As is pointed out in [45], the expansion in (4.34) is not a physically-relevant definition for stress due to the following test case. Consider a stationary deformed body at zero temperature (i.e., where the particles occupy fixed positions without vibrating). In this case, the net force acting on any particle is zero. Since $\mathbf{f}_{\alpha\beta}$ is the only term in the summand of (4.34) which depends on β , (4.34) is equivalent to

$$\tilde{\sigma}_{w,v}^*(\mathbf{x}, t) := \sum_{\alpha} \left(\sum_{\substack{\beta \\ \beta \neq \alpha}} \mathbf{f}_{\alpha\beta} \right) \otimes (\mathbf{x} - \mathbf{x}_\alpha) \hat{a}(\|\mathbf{x} - \mathbf{x}_\alpha\|). \quad (4.35)$$

In our case, $\sum_{\beta} \mathbf{f}_{\alpha\beta} = \mathbf{f}_\alpha = \mathbf{0}$, for each particle α in the *interior* of the body which is considerably away from the surface compared to the interatomic distance. Hence, the only non-zero contribution to the stress is due to those particles close to the surface on which the net force due to other particles is non-zero. Moreover, although $\hat{a}(u)$ decays to zero as u increases, $\hat{a}(u) \neq 0$ for all $u \neq 0$. Thus, there is a non-zero stress at every point $\mathbf{x} \in \mathbb{R}^3$ (even outside the body!) due to particles close

to the surface of the body, which is obviously not physically reasonable. Nevertheless, $\tilde{\sigma}_{w,v}^*$ is mathematically still a valid definition since it satisfies the force balance equation [45]. However, it cannot be derived using the systematized Murdoch–Hardy procedure proposed here. Thus, the systematized Murdoch–Hardy procedure does not lead to all possible definitions that satisfy the force balance equation.

Finally, it is also worth noting that the fact that all balance laws are satisfied under the Murdoch–Hardy procedure should not come as a surprise, since w in (4.18) serves the same purpose as W in (3.13). In this view, w is seen as a function defined on a phase space, although one that does not evolve according to a flow described by Hamilton’s equations of motion, but still satisfies (3.10).³⁵ The corresponding flow in phase space is described as follows. Continuing with our notation introduced (3.1), let $\Xi(0) = (\mathbf{x}^s(0), \mathbf{v}^s(0)) = (\mathbf{x}_1^s(0), \dots, \mathbf{x}_N^s(0), \mathbf{v}_1^s(0), \dots, \mathbf{v}_N^s(0))$ denote any arbitrary point in phase space. We add a superscript “s” to stress the fact that an element in phase space is stochastic in nature. Consider the flow in phase space given by the mapping

$$\begin{aligned} \Xi(0) = (\mathbf{x}^s(0), \mathbf{v}^s(0)) &\mapsto (\mathbf{x}^s(0) + \mathbf{x}(t) - \mathbf{x}(0), \mathbf{v}^s(0) + \mathbf{v}(t) - \mathbf{v}(0)) \\ &= (\mathbf{x}^s(t), \mathbf{v}^s(t)) = \Xi(t), \end{aligned} \quad (4.36)$$

where the quantities $\mathbf{x}(t) = (\mathbf{x}_1(t), \dots, \mathbf{x}_N(t))$, $\mathbf{v}(t) = (\mathbf{v}_1(t), \dots, \mathbf{v}_N(t))$ denote the position and velocity of the particle and these are assumed to be known. (Typically these quantities are obtained from a molecular dynamics simulation.) Therefore the Murdoch–Hardy procedure can be interpreted as a probabilistic model constructed from the data, $\mathbf{x}(t)$ and $\mathbf{v}(t)$, obtained from a deterministic model – a molecular dynamics simulation. Note that \mathbf{x}^s and \mathbf{v}^s in (4.36) denote the positions and velocities of the particles in the probabilistic model. Then it is easy to see that if $W(\Xi; t)$ is given by

$$W(\Xi; t) = w(\mathbf{x}_1(t) - \mathbf{x}_1^s) \cdots w(\mathbf{x}_N(t) - \mathbf{x}_N^s), \quad (4.37)$$

then the definitions given by (3.13), (3.15) and (4.18) are consistent and W given by the above formula satisfies Liouville’s equation (see (3.10)), which was used in deriving the balance equations in Section 3. Note that unlike Section 3, $W(\Xi; t)$ defined in (4.37) is *not* a probability density function. (Its integral over phase space diverges, since it is independent of \mathbf{v} .) The key difference between the two approaches is that all quantities in the Irving–Kirkwood–Noll procedure are probabilistic, while this is not true for the Murdoch–Hardy procedure, if the above probabilistic interpretation is adopted. For example, $\mathbf{f}_{\alpha\beta}$ in the Murdoch–Hardy procedure is deterministic. Therefore the structure inherent in (3.61) through the marginal densities is absent in the Murdoch–Hardy procedure, thus giving additional non-uniqueness. It is shown in Section 5 that the Hardy stress can be derived using both approaches, while the DA stress is a result of the Murdoch–Hardy procedure alone.

³⁵This is only a mathematical argument. No physical significance should be drawn from this analogy.

4.3 Definition of the spatially-averaged traction vector

We close this section by defining the spatially-averaged traction vector, $\mathbf{t}_w(\mathbf{x}, \mathbf{n}; t)$, for a weighting function w , at a point \mathbf{x} relative to a plane with normal $\mathbf{n}(\mathbf{x})$. One possibility is to adopt the Cauchy relation using the spatially-averaged stress tensor,

$$\mathbf{t}_w(\mathbf{x}, \mathbf{n}; t) := \boldsymbol{\sigma}_w(\mathbf{x}, t)\mathbf{n}. \quad (4.38)$$

However, since $\boldsymbol{\sigma}_w$ is defined as a volume average, we immediately see that with this definition \mathbf{t}_w depends not only on the bonds that cross the surface, but also on nearby bonds that do not cross it [44]. Hence, equation (4.38) does not appear to be consistent with Cauchy's definition of traction.

We therefore seek an alternative definition for the spatially-averaged traction vector. In Section 3.7, we showed that the pointwise traction vector at a point on a surface is the expectation of the force per unit area of all the bonds that cross the surface, making it a property of the surface. We would like the spatially-averaged traction to have the same property. We therefore define it as an average over a *surface* rather than over a volume as for the stress. For simplicity, we consider the weighting function w_h , defined to be constant on the averaging domain, which is taken to be a generalized cylinder of height h , with its axis parallel to \mathbf{n} and enclosing \mathbf{x} . The traction $\mathbf{t}_w(\mathbf{x}, \mathbf{n}; t)$ is defined as

$$\mathbf{t}_w(\mathbf{x}, \mathbf{n}; t) := \lim_{h \rightarrow 0} \boldsymbol{\sigma}_{w_h}(\mathbf{x}, t)\mathbf{n}, \quad (4.39)$$

where $\boldsymbol{\sigma}_{w_h}$ is the stress associated with the weighting function, w_h . In a more general case, an arbitrary averaging domain can be collapsed onto a surface passing through \mathbf{x} , in many ways. Although this can be made mathematically more precise, we do not pursue that in this work. Definition (4.39) has a two-fold advantage over the definition in (4.38):

1. The traction vector is defined to be non-local on a surface, thus making it a property of the surface. This is physically more meaningful, and closer to the continuum definition.
2. The above definition differs from the traction definition in (4.38), because only the bonds which cross the surface contribute to the traction field.

In Section 5.2, we use definition (4.39) to define the Tsai traction starting with the spatial averaging discussed in Section 4.1 and in this way establish a link between the Tsai traction and the Irving–Kirkwood–Noll procedure.

5 Derivation of different stress definitions and the issue of uniqueness

In this section, we systematically derive various stress tensors commonly found in the literature from the methods developed in Section 3 and Section 4. The stress tensors discussed in this section are the Hardy, virial and DA stress tensors and the Tsai traction.

5.1 Hardy stress tensor

The Murdoch–Hardy procedure described in Section 4 was independently developed by Murdoch [43] and Hardy [23]. The motivation for Hardy’s study was to test the validity of the continuum description of phenomena in shock waves. The formulas suggested by Irving and Kirkwood were not useful due to the lack of knowledge regarding the probability density function and the infinite series expansion in the definition of the stress. As an alternative, Hardy used what we now term as the “Murdoch–Hardy procedure” to propose an instantaneous definition for stress, for the special case of pair potential, given by

$$\sigma_v^H(\mathbf{x}, t) = \frac{1}{2} \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \frac{(\mathbf{x}_\alpha(t) - \mathbf{x}_\beta(t)) \otimes (\mathbf{x}_\alpha(t) - \mathbf{x}_\beta(t))}{\|\mathbf{x}_\alpha - \mathbf{x}_\beta\|} \mathcal{V}'_{\alpha\beta} b(\mathbf{x}; \mathbf{x}_\alpha, \mathbf{x}_\beta), \quad (5.1a)$$

$$\sigma_k^H(\mathbf{x}, t) = - \sum_{\alpha} w(\mathbf{x}_\alpha(t) - \mathbf{x}) m_{\alpha} \mathbf{v}_{\alpha}^{\text{rel}}(t) \otimes \mathbf{v}_{\alpha}^{\text{rel}}(t), \quad (5.1b)$$

where b is the bond function defined in (4.16) and $\mathbf{v}_{\alpha}^{\text{rel}}$ is the velocity of particle α with respect to the continuum velocity, as defined in (4.24). To simplify the notation, the explicit dependence of \mathbf{x}_{α} and $\mathbf{v}_{\alpha}^{\text{rel}}$ on time is dropped from here onwards. Equations (5.1a) and (5.1b) may look familiar. They are similar to the spatially-averaged generalized stress in (4.9) and (4.17) (for the special case of a pair potential). If in these relations, the ensemble average is replaced by a time average, we obtain a time-averaged Hardy stress. However, in performing such an operation, we must note the following:

1. Under conditions of thermodynamic equilibrium (see footnote 8 on page 10), ensemble averages can be replaced by time averages provided that the system is assumed to be ergodic. Strictly speaking this time average should be done for infinite time, but for practical reasons we are restricted to finite time.
2. The Hardy stress tensor is valid under non-equilibrium conditions assuming that the system is in *local thermodynamic equilibrium*³⁶ at all points at every instant of time. This is plausible only when there is a clear separation of time scales between the microscopic equilibration time scale τ and macroscopic times. Here, τ is not being defined rigorously. Roughly speaking, τ must be sufficiently small so that macroscopic observables do not vary appreciably over it.

³⁶Local thermodynamic equilibrium is a weaker condition than uniform thermodynamic equilibrium (see footnote 8). The assumption is that the microscopic domain associated with each continuum particle is locally in a state of uniform thermodynamic (or at least metastable) equilibrium. This is the reason why concepts like temperature can be defined as field variables in continuum mechanics. See for example [17].

Under these assumptions, we may replace ensemble averages with time averages in (4.9) and (4.17) to obtain

$$\sigma_{w,k}(\mathbf{x}, t) = -\frac{1}{\tau} \sum_{\alpha} \int_t^{t+\tau} w(\mathbf{x}_{\alpha} - \mathbf{x}) m_{\alpha} \mathbf{v}_{\alpha}^{\text{rel}} \otimes \mathbf{v}_{\alpha}^{\text{rel}} dt, \quad (5.2a)$$

$$\sigma_{w,v}(\mathbf{x}, t) = \frac{1}{2\tau} \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \int_t^{t+\tau} [-\mathbf{f}_{\alpha\beta} \otimes (\mathbf{x}_{\alpha} - \mathbf{x}_{\beta}) b(\mathbf{x}; \mathbf{x}_{\alpha}, \mathbf{x}_{\beta})] dt, \quad (5.2b)$$

where $\mathbf{f}_{\alpha\beta}$, corresponding to a given potential extension, is defined in (3.51), and τ represents a microscopic time scale. We see that the Hardy stress is obtained through a rigorous process beginning with the statistical mechanics concepts introduced in Section 3. From here on, we will denote the stress in (5.2) as the “Hardy stress”, although we note that this definition constitutes a generalization of the original Hardy stress to arbitrary potentials and includes time averaging. Incidentally, the Hardy stress can also be derived from the systematized Murdoch–Hardy procedure described in Section 4.2. The kinetic part of the Hardy stress is the same as that obtained in the Murdoch–Hardy procedure. The potential part of the Hardy stress is derived using the generator \mathbf{g}^{H} given in (4.29). This was done in Section 4 (see (4.32)).

Note that the Hardy stress tensor is symmetric. One could modify this to a general form by choosing an arbitrary path of interaction, thus leading to a non-symmetric form (see Section 4.2). Also note that the stress tensor resulting from the generator \mathbf{g}^{HD} (see Fig. 4.3) would be symmetric, because the interaction force between two continuum particles is always aligned with the line connecting them. It is very important to observe that under non-equilibrium conditions, where we assume a local thermodynamic equilibrium at every instant of macroscopic time, we may assume that the averaging domain centered at a position \mathbf{x} moves with the continuum velocity $\mathbf{v}(\mathbf{x}, t)$. This fact will be used in Section 5.2.

5.2 Tsai traction

Cauchy’s original definition of stress emerges from the concept of traction acting across the internal surfaces of a solid via the bonds that cross the surface. It is therefore natural to attempt to define traction at the atomic level in a similar vein in terms of the force in bonds intersecting a given plane. This approach actually goes back to Cauchy himself as part of his effort in the 1820s to define the stress in crystalline systems [5, 6], which is described in detail in Note B in Love’s classical book on the theory of elasticity [35]. Cauchy’s derivation is limited to zero temperature equilibrium where the atoms are stationary. This approach was extended by Tsai [63] to the dynamical setting by also accounting for the momentum flux of atoms moving across the plane. The expression for the traction given in [63] appears to be based on intuition.

In this section, we show how the Tsai traction can be systematically derived from the Hardy stress tensor, which itself was derived from the generalized stress tensor defined in Section 4.1. We will see that the potential part of Tsai’s original definition

agrees with the results of our unified framework. However, Tsai's expression for the kinetic part of the traction depends on the absolute velocity of the particles and therefore is not invariant with respect to Galilean transformations. We show below that the correct expression for the Tsai traction vector $\mathbf{t}(\mathbf{x}, \mathbf{n}; t)$ across a plane P with normal \mathbf{n} is

$$\begin{aligned} \mathbf{t}_w(\mathbf{x}, \mathbf{n}; t) = & \frac{1}{A\tau} \int_t^{t+\tau} \sum_{\alpha\beta \cap P} \mathbf{f}_{\alpha\beta} \frac{(\mathbf{x}_\alpha - \mathbf{x}_\beta) \cdot \mathbf{n}}{|(\mathbf{x}_\alpha - \mathbf{x}_\beta) \cdot \mathbf{n}|} dt \\ & - \frac{1}{A\tau} \sum_{\alpha \leftrightarrow P} \frac{m_\alpha \mathbf{v}_\alpha^{\text{rel}}(t_{\leftrightarrow}) (\mathbf{v}_\alpha^{\text{rel}}(t_{\leftrightarrow}) \cdot \mathbf{n})}{|\mathbf{v}_\alpha^{\text{rel}}(t_{\leftrightarrow}) \cdot \mathbf{n}|}, \end{aligned} \quad (5.3)$$

where τ indicates the microscopic time scale, $\sum_{\alpha\beta \cap P}$ indicates the summation over all bonds $\alpha - \beta$ crossing the plane P , $\sum_{\alpha \leftrightarrow P}$ indicates summation over all particles that cross P in the time interval $[t, t + \tau]$,³⁷ $\mathbf{v}_\alpha^{\text{rel}}$ denotes the local relative velocity of particle α , and t_{\leftrightarrow} indicates the time at which the particle crosses the plane. The correct form for $\mathbf{v}_\alpha^{\text{rel}}$ is not immediately obvious. Below, we derive equation (5.3) and obtain an explicit expression for $\mathbf{v}_\alpha^{\text{rel}}$.

We start with the Hardy stress in (5.2). Recall from Section 4.3 that if the averaging domain is taken to be a generalized cylinder \mathcal{C}_h of height h , the spatially-averaged traction field, $\mathbf{t}_w(\mathbf{x}, \mathbf{n}; t)$, on a surface passing through \mathbf{x} with normal \mathbf{n} is

$$\begin{aligned} \mathbf{t}_w(\mathbf{x}, \mathbf{n}; t) = & \lim_{h \rightarrow 0} \sigma_{w_h} \mathbf{n} = \lim_{h \rightarrow 0} (\sigma_{w_h, \text{v}} \mathbf{n} + \sigma_{w_h, \text{k}} \mathbf{n}) \\ =: & \mathbf{t}_{w, \text{v}} + \mathbf{t}_{w, \text{k}}. \end{aligned} \quad (5.4)$$

Using (5.2), we rewrite the potential part and kinetic part of (5.4) as

$$\mathbf{t}_{w, \text{v}}(\mathbf{x}, \mathbf{n}; t) = \frac{1}{2\tau} \lim_{h \rightarrow 0} \int_t^{t+\tau} \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} [-\mathbf{f}_{\alpha\beta} \otimes (\mathbf{x}_\alpha - \mathbf{x}_\beta) b_h(\mathbf{x}; \mathbf{x}_\alpha, \mathbf{x}_\beta)] dt, \quad (5.5a)$$

$$\mathbf{t}_{w, \text{k}}(\mathbf{x}, \mathbf{n}; t) = -\frac{1}{\tau} \lim_{h \rightarrow 0} \int_t^{t+\tau} \sum_{\alpha} m_\alpha w(\mathbf{x}_\alpha - \mathbf{x}; h) \mathbf{v}_\alpha^{\text{rel}}(t; h) (\mathbf{v}_\alpha^{\text{rel}}(t; h) \cdot \mathbf{n}) dt, \quad (5.5b)$$

where b_h denotes the bond function for a generalized cylinder of height h . Also note the dependence of $\mathbf{v}_\alpha^{\text{rel}}$ on h in (5.5b). Let us first consider the potential part of the traction in (5.5a). As h approaches zero, the generalized cylinder will no longer contain complete bonds. Assuming a constant weighting function, the bond function b_h equals the fraction of the length of the bond lying within the generalized cylinder per unit volume:

$$b_h(\mathbf{x}; \mathbf{x}_\alpha, \mathbf{x}_\beta) = \frac{1}{hA} \frac{h}{|(\mathbf{x}_\alpha - \mathbf{x}_\beta) \cdot \mathbf{n}|} = \frac{1}{A |(\mathbf{x}_\alpha - \mathbf{x}_\beta) \cdot \mathbf{n}|}, \quad (5.6)$$

³⁷A particle is counted multiple times if it crosses the plane multiple times.

for any bond $\alpha - \beta$ crossing the cylinder. Therefore (5.5a) takes the form

$$t_{w,v}(\mathbf{x}, \mathbf{n}; t) = \frac{1}{A\tau} \int_t^{t+\tau} \sum_{\alpha\beta \cap P} \left[-f_{\alpha\beta} \frac{(\mathbf{x}_\alpha - \mathbf{x}_\beta) \cdot \mathbf{n}}{|(\mathbf{x}_\alpha - \mathbf{x}_\beta) \cdot \mathbf{n}|} \right] dt. \quad (5.7)$$

Note that the $1/2$ factor is dropped because of the definition of the summation in the above equation. This is the first term in (5.3). Turning to the kinetic part of the traction in (5.5b), we interchange the summation and integral to obtain

$$t_{w,k}(\mathbf{x}, \mathbf{n}; t) = -\frac{1}{\tau} \lim_{h \rightarrow 0} \sum_{\alpha \in \mathcal{C}_h} \int_{t_1(\alpha; h)}^{t_2(\alpha; h)} m_\alpha w(\mathbf{x}_\alpha - \mathbf{x}; h) \mathbf{v}_\alpha^{\text{rel}}(t; h) (\mathbf{v}_\alpha^{\text{rel}}(t; h) \cdot \mathbf{n}) dt, \quad (5.8)$$

where $t_1(\alpha; h)$ and $t_2(\alpha; h)$ are the times of entry and exit of particle α , respectively, from a cylinder of height h . The summation in the above equation is over all particles that are in the generalized cylinder during the time interval $[t, t + \tau]$, with a particle counted k times if it enters and exits the cylinder k times. Multiplying and dividing the above equation by $t_2(\alpha; h) - t_1(\alpha; h)$ and substituting in w , we have

$$\begin{aligned} t_{w,k}(\mathbf{n}) &= -\frac{1}{A\tau} \lim_{h \rightarrow 0} \sum_{\alpha \in \mathcal{C}_h} \frac{t_2(\alpha; h) - t_1(\alpha; h)}{h} \frac{\int_{t_1(\alpha; h)}^{t_2(\alpha; h)} m_\alpha \mathbf{v}_\alpha^{\text{rel}}(t; h) (\mathbf{v}_\alpha^{\text{rel}}(t; h) \cdot \mathbf{n}) dt}{t_2(\alpha; h) - t_1(\alpha; h)} \\ &= -\frac{1}{A\tau} \sum_{\alpha \leftrightarrow P} \lim_{h \rightarrow 0} \frac{t_2(\alpha; h) - t_1(\alpha; h)}{h} m_\alpha \mathbf{v}_\alpha^{\text{rel}}(t_{\leftrightarrow}) (\mathbf{v}_\alpha^{\text{rel}}(t_{\leftrightarrow}) \cdot \mathbf{n}), \end{aligned} \quad (5.9)$$

where we have used the Lebesgue differentiation theorem [19] in the last equality. Note that the interchange of limit and summation in the above step is valid since we can assume that the summation for any \mathcal{C}_h is a finite summation which is physically meaningful. Since the averaging domain moves with a continuum velocity we note that

$$\lim_{h \rightarrow 0} \frac{t_2(\alpha; h) - t_1(\alpha; h)}{h} = \frac{1}{|\mathbf{v}_\alpha^{\text{rel}}(t_{\leftrightarrow}) \cdot \mathbf{n}|}. \quad (5.10)$$

In words, this equality states that the net time spent by particle α in the cylinder, divided by its height, is equal to the inverse of the velocity of particle α along the axis of the cylinder. This is correct in the limit, $h \rightarrow 0$, where particles only enter and exit the cylinder at its ends. Substituting (5.10) into (5.9), we have

$$\begin{aligned} t_{w,k}(\mathbf{n}) &= -\frac{1}{A\tau} \sum_{\alpha \leftrightarrow P} \frac{m_\alpha \mathbf{v}_\alpha^{\text{rel}}(t_{\leftrightarrow}) (\mathbf{v}_\alpha^{\text{rel}}(t_{\leftrightarrow}) \cdot \mathbf{n})}{|\mathbf{v}_\alpha^{\text{rel}}(t_{\leftrightarrow}) \cdot \mathbf{n}|} \\ &= -\frac{1}{A\tau} \sum_{\alpha \leftrightarrow P} m_\alpha \mathbf{v}_\alpha^{\text{rel}}(t_{\leftrightarrow}) \text{sign}(\mathbf{v}_\alpha^{\text{rel}}(t_{\leftrightarrow}) \cdot \mathbf{n}). \end{aligned} \quad (5.11)$$

This is the second term in (5.3). Note that

$$\mathbf{v}_\alpha^{\text{rel}}(t) = \lim_{h \rightarrow 0} \mathbf{v}_\alpha^{\text{rel}}(t; h) = \mathbf{v}_\alpha(t) - \lim_{h \rightarrow 0} \mathbf{v}(\mathbf{x}; h). \quad (5.12)$$

Hence, we have implicitly assumed that $\lim_{h \rightarrow 0} \mathbf{v}(\mathbf{x}; h)$ is well-defined for our averaging domain (plane P) which is a limit of the generalized cylinder \mathcal{C}_h . In the

following calculation, we show that $\mathbf{v}(\mathbf{x}; h)$ is well-defined and its exact form is derived.

We know that for a generalized cylinder

$$\begin{aligned} \mathbf{v}(\mathbf{x}; h) &:= \frac{\frac{1}{\tau} \int_t^{t+\tau} \sum_{\alpha} m_{\alpha} w(\mathbf{x}_{\alpha}(t) - \mathbf{x}; h) \mathbf{v}_{\alpha}(t) dt}{\frac{1}{\tau} \int_t^{t+\tau} \sum_{\beta} m_{\beta} w(\mathbf{x}_{\beta}(t) - \mathbf{x}) dt} \\ &= \frac{\sum'_{\alpha} \int_{t_1(\alpha; h)}^{t_2(\alpha; h)} m_{\alpha} \mathbf{v}_{\alpha}(t) dt}{\sum'_{\beta} \int_{t_1(\beta; h)}^{t_2(\beta; h)} m_{\beta} dt} \\ &= \frac{\sum'_{\alpha} m_{\alpha} [\mathbf{x}_{\alpha}(t_2(\alpha; h)) - \mathbf{x}_{\alpha}(t_1(\alpha; h))]}{\sum'_{\beta} m_{\beta} [t_2(\beta; h) - t_1(\beta; h)]}, \end{aligned} \quad (5.13)$$

where \sum' indicates summation over those particles that cross P in the time interval $[t, t + \tau]$, including multiple entries and exits.

Considering the limit $h \rightarrow 0$ of the first partial fraction of the last equation we have

$$\lim_{h \rightarrow 0} \frac{m_{\alpha} \frac{\mathbf{x}_{\alpha}(t_2(\alpha; h)) - \mathbf{x}_{\alpha}(t_1(\alpha; h))}{t_2(\alpha; h) - t_1(\alpha; h)}}{\sum'_{\beta} m_{\beta} \frac{t_2(\beta; h) - t_1(\beta; h)}{t_2(\alpha; h) - t_1(\alpha; h)}} = \frac{m_{\alpha} \mathbf{v}_{\alpha}(t_{\leftrightarrow})}{\sum'_{\beta} m_{\beta} |\mathbf{v}_{\alpha}(t_{\leftrightarrow}) \cdot \mathbf{n} / \mathbf{v}_{\beta}(t_{\leftrightarrow}) \cdot \mathbf{n}|}, \quad (5.14)$$

using the fact that in the limit $h \rightarrow 0$, $(t_2(\beta; h) - t_1(\beta; h)) / (t_2(\alpha; h) - t_1(\alpha; h))$, which is the ratio of the times spent by particles β and α in one of their sojourns into the cylinder, is equal to the inverse ratio of their normal velocities. Using (5.14) and taking the limit $h \rightarrow 0$ of (5.13), we obtain

$$\mathbf{v}(\mathbf{x}) = \lim_{h \rightarrow 0} \mathbf{v}(\mathbf{x}; h) = \sum_{\alpha \leftrightarrow P} \frac{m_{\alpha} \mathbf{v}_{\alpha}(t_{\leftrightarrow})}{\sum'_{\beta} m_{\beta} |\mathbf{v}_{\alpha}(t_{\leftrightarrow}) \cdot \mathbf{n} / \mathbf{v}_{\beta}(t_{\leftrightarrow}) \cdot \mathbf{n}|}. \quad (5.15)$$

Note that the above expression for the continuum velocity is far from intuitive. One might expect the continuum velocity to be the average velocity of particles crossing the surface, but this is not true. It is clear from the above equation that the averaging is not trivial.

From the relationship between the Tsai traction in (5.3) and the Hardy stress tensor in (5.2), it is apparent that the Tsai traction is a more local quantity than the Hardy stress tensor. The Tsai traction performs better than the Hardy stress in systems with free surfaces. This was studied by Cheung and Yip [7] for a one-dimensional case, in which virial stress and Tsai stress are compared (the virial stress is a special case of Hardy stress as shown in the next section).

The Tsai traction definition can be used to evaluate the stress tensor at a point by evaluating the traction on three perpendicular planes.³⁸ However, it is not clear from the perspective put forward by Tsai [63] whether the resulting stress tensor would

³⁸For example, if the normals to the planes are aligned with the axes of a Cartesian coordinate system with basis vectors \mathbf{e}_i , then $\mathbf{t}(\mathbf{e}_1)$ would give the components $\sigma_{11}, \sigma_{21}, \sigma_{31}$, $\mathbf{t}(\mathbf{e}_2)$ would give the components $\sigma_{12}, \sigma_{22}, \sigma_{32}$, and $\mathbf{t}(\mathbf{e}_3)$ would give the components $\sigma_{13}, \sigma_{23}, \sigma_{33}$.

be symmetric or even well-defined, i.e., it is not clear if another choice of planes will give suitably transformed components of the same stress tensor. Our derivation suggests that a stress tensor constructed from the Tsai traction should be well-defined and symmetric, at least in a weak sense, since it is a limit of the Hardy stress, which has these properties. The numerical experiments presented in Section 6, suggest that the Tsai traction is invariant with respect to the position of the Tsai plane P and the resulting stress tensor is symmetric.

5.3 Virial stress tensor

In this section, we show that the virial stress tensor derived in Section 2 and in Appendix A can be re-derived from the time-averaged version of the Hardy stress given in (5.2). The expression for the virial stress tensor is obtained from (5.2) as a special case for a weighting function which is constant on its support. The bond function, b , in (5.2) is evaluated approximately using its definition (4.16) by only counting those bonds $\alpha - \beta$ that lie entirely within the averaging domain and neglecting the bonds that cross the averaging domain. Hence, $b(\mathbf{x}; \mathbf{x}_\alpha, \mathbf{x}_\beta)$ is given by

$$b(\mathbf{x}; \mathbf{x}_\alpha, \mathbf{x}_\beta) = \begin{cases} 1/\text{vol}(\Omega_{\mathbf{x}}) & \text{if bond } \alpha - \beta \in \Omega_{\mathbf{x}}, \\ 0 & \text{otherwise,} \end{cases} \quad (5.16)$$

where $\Omega_{\mathbf{x}}$ denotes the averaging domain centered at \mathbf{x} . Substituting (5.16) into (5.2), we have

$$\boldsymbol{\sigma}(\mathbf{x}, t) = \frac{1}{\tau \text{vol}(\Omega_{\mathbf{x}})} \int_t^{t+\tau} \left[- \sum_{\alpha \in \Omega_{\mathbf{x}}} m_\alpha \mathbf{v}_\alpha^{\text{rel}} \otimes \mathbf{v}_\alpha^{\text{rel}} + \frac{1}{2} \sum_{\substack{\alpha, \beta \in \Omega_{\mathbf{x}} \\ \alpha \neq \beta}} [-\mathbf{f}_{\alpha\beta} \otimes (\mathbf{x}_\alpha - \mathbf{x}_\beta)] \right] dt, \quad (5.17)$$

which is identical to (A.27) in Appendix A. It is clear from this that the virial stress tensor is only an approximation and tends to the Hardy stress as the volume of the averaging domain is increased. This is because the ratio of the measure of bonds that cross the surface to those which are inside the averaging domain decreases as the size of the domain increases. The difference between the virial stress tensor and the Tsai traction was analytically calculated for a one-dimensional chain by Tsai (see [63]). Since, taking the averaging domain size to infinity is equivalent to taking the thermodynamic limit in this context, the Hardy and virial stress expressions become identical in this limit. Since the virial theorem was also derived in Section 2 for the case of equilibrium statistical mechanics, it follows that the Irving–Kirkwood–Noll procedure is consistent with the results of equilibrium statistical mechanics in the thermodynamic limit.

5.4 DA stress tensor

It was seen in Section 4.2, that the DA stress tensor, defined in (4.33), is derived using an appropriate generator in the systematized Murdoch–Hardy procedure. However,

unlike the Hardy stress, the DA stress cannot be derived from the Irving–Kirkwood–Noll procedure. It is also worth noting that the stress tensor given by (4.33) is in general non-symmetric, and only under very special conditions yields a symmetric tensor [45].

5.5 Uniqueness of the macroscopic stress tensor

Three possible sources of non-uniqueness for the stress tensor have been identified in our discussion:

1. Given that there are multiple potential extensions (see Page 24), different force decompositions are possible and hence different pointwise stress tensors can be obtained.
2. For a given pointwise stress tensor, a new pointwise stress, which also satisfies the balance of linear momentum, can be obtained by adding on an arbitrary tensor field with zero divergence.
3. The generalization of the Irving–Kirkwood–Noll procedure in Section 3.6 to arbitrary “paths of interaction” leads to the possibility of non-symmetric expressions for the pointwise stress tensor.

We address the first two issues in this section. The third source of non-uniqueness is only possible in systems where the discrete particles making up the system possess internal structure, such as internal polarization or spin. For systems of discrete particles without internal structure only straight bonds are possible due to symmetry arguments. We leave the discussion of particles with internal structure to future work.

Uniqueness and potential energy extensions

The first source of non-uniqueness of the stress tensor is related to the potential energy extension discussed in Section 3.4. We show below that the macroscopic stress tensor, calculated as a spatial average of the pointwise stress tensor with constant weighting function, is always unique in the thermodynamic limit (see footnote 1 on page 5), i.e., the difference between the spatially-averaged pointwise stress tensors resulting from two different extensions tends to zero, as the volume of the averaging domain is increased.

The discussion below is limited to 5-body potentials since it can be easily extended to any interatomic potential. We first show that the contribution due to any cluster of 5 particles within the averaging domain is zero. Without loss of generality, we may assume that our system consists of 5 particles interacting with an interatomic potential energy given by

$$\mathcal{V}_{\text{int}} = \widehat{\mathcal{V}}(\mathbf{x}_1, \dots, \mathbf{x}_5). \quad (5.18)$$

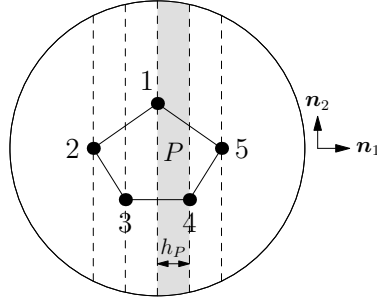


Fig. 5.1 A cluster of 5 particles that lie completely inside the averaging domain, does not contribute to the ambiguity in the stress tensor.

Let $\mathcal{V}_{\text{int}}(\zeta_{12}, \dots, \zeta_{45})$ and $\mathcal{V}_{\text{int}}^*(\zeta_{12}, \dots, \zeta_{45})$ be two different extensions of \mathcal{V}_{int} from the shape space \mathcal{S} to \mathbb{R}^{10} (see Section 3.4), and for any $\mathbf{s} = (r_{12}, \dots, r_{45}) \in \mathcal{S}$, let

$$\mathbf{f}_{\alpha\beta}(\mathbf{x}_1, \dots, \mathbf{x}_5) := \frac{\partial \mathcal{V}_{\text{int}}}{\partial \zeta_{\alpha\beta}}(\mathbf{s}) \frac{\mathbf{x}_\beta - \mathbf{x}_\alpha}{r_{\alpha\beta}}, \quad (5.19)$$

$$\mathbf{f}_{\alpha\beta}^*(\mathbf{x}_1, \dots, \mathbf{x}_5) := \frac{\partial \mathcal{V}_{\text{int}}^*}{\partial \zeta_{\alpha\beta}}(\mathbf{s}) \frac{\mathbf{x}_\beta - \mathbf{x}_\alpha}{r_{\alpha\beta}}, \quad (5.20)$$

be their corresponding force decompositions. Let $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}^*$ denote the resulting pointwise stress tensors in the Irving–Kirkwood–Noll procedure from \mathcal{V}_{int} and $\mathcal{V}_{\text{int}}^*$, respectively. Let $\Omega_{\mathbf{x}}$ denote the averaging domain³⁹ centered at \mathbf{x} that is used to calculate the Hardy stress tensor. Using (5.2b) and noting that all the bonds lie within Ω , the difference between the Hardy stress tensors resulting from these two representations, for the special case of a constant weighting function, is given by

$$\Delta \boldsymbol{\sigma}(\mathbf{x}, t) := \boldsymbol{\sigma}_w - \boldsymbol{\sigma}_w^* = \frac{1}{2\tau \text{vol}(\Omega_{\mathbf{x}})} \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \int_t^{t+\tau} [-\Delta \mathbf{f}_{\alpha\beta} \otimes (\mathbf{x}_\alpha - \mathbf{x}_\beta)] dt, \quad (5.21)$$

where $\Delta \mathbf{f}_{\alpha\beta} := \mathbf{f}_{\alpha\beta} - \mathbf{f}_{\alpha\beta}^*$.

We would like to show that $\Delta \boldsymbol{\sigma} \mathbf{n}_1 = \mathbf{0}$, where \mathbf{n}_1 is the normal vector as shown in Fig. 5.1. The essential idea is to interchange the integration and summation in (5.21) and split the terms appearing in the summation into fractions, such that each fraction yields a zero contribution to $\Delta \boldsymbol{\sigma} \mathbf{n}_1$. In order to show this, we partition the averaging domain into regions such that no region contains a particle in its interior and the partition surfaces are perpendicular to the normal (see Fig. 5.1). Let h_P denote the width of the partition P . Using (5.21), we can now write $\Delta \boldsymbol{\sigma} \mathbf{n}_1$ as

$$\begin{aligned} \Delta \boldsymbol{\sigma} \mathbf{n}_1 &= \frac{1}{2\tau \text{vol}(\Omega_{\mathbf{x}})} \int_t^{t+\tau} \sum_P \sum_{\alpha\beta \cap P} \left[-\Delta \mathbf{f}_{\alpha\beta}(\mathbf{x}_\alpha - \mathbf{x}_\beta) \cdot \mathbf{n}_1 \frac{h_P}{|(\mathbf{x}_\alpha - \mathbf{x}_\beta) \cdot \mathbf{n}_1|} \right], \\ &= \frac{1}{2\tau \text{vol}(\Omega_{\mathbf{x}})} \int_t^{t+\tau} \sum_P h_P \Delta \mathbf{F}_P, \end{aligned} \quad (5.22)$$

³⁹For simplicity assume that the averaging domain is convex.

where $\sum_{\alpha\beta\cap P}$ denotes the summation over the bonds crossing partition P , and

$$\Delta \mathbf{F}_P = \sum_{\alpha\beta\cap P} \left[-\Delta \mathbf{f}_{\alpha\beta} \frac{(\mathbf{x}_\alpha - \mathbf{x}_\beta) \cdot \mathbf{n}_1}{|(\mathbf{x}_\alpha - \mathbf{x}_\beta) \cdot \mathbf{n}_1|} \right] \quad (5.23)$$

is the *net* force on particles on one side of the partition due to particles on the other side. Since both representations give the same total force on each particle, the force difference, or net force, on each particle is zero and therefore, $\Delta \mathbf{F}_P = \mathbf{0}$. For example, for the partition shown in the figure,

$$\Delta \mathbf{F}_P = -2(\Delta \mathbf{f}_{51} + \Delta \mathbf{f}_{43}). \quad (5.24)$$

Since $\Delta \mathbf{f}_{45} = -\Delta \mathbf{f}_{54}$, we have

$$\begin{aligned} \Delta \mathbf{F}_P &= -2(\Delta \mathbf{f}_{51} + \Delta \mathbf{f}_{43} + \Delta \mathbf{f}_{45} + \Delta \mathbf{f}_{54}) \\ &= -2(\Delta \mathbf{f}_{43} + \Delta \mathbf{f}_{45}) - 2(\Delta \mathbf{f}_{51} + \Delta \mathbf{f}_{54}) \\ &= -2\Delta \mathbf{f}_4^{\text{int}} - 2\Delta \mathbf{f}_5^{\text{int}} = \mathbf{0} + \mathbf{0} = \mathbf{0}. \end{aligned} \quad (5.25)$$

Hence, $\Delta \boldsymbol{\sigma} \mathbf{n}_1 = \mathbf{0}$. Undertaking a similar argument in the other directions, we see that $\Delta \boldsymbol{\sigma} \mathbf{n}_i = \mathbf{0}$. These results together imply that $\Delta \boldsymbol{\sigma} = \mathbf{0}$. Given this, we can conclude that any cluster of particles that lies entirely within the averaging domain does not contribute to the spatial average of the difference between two stress definitions. Consequently, the only non-zero contribution comes from those clusters for which the bonds connecting its particles cross the averaging domain. Since this contribution scales as surface area, it tends to zero as volume tends to infinity.

Uniqueness and the addition of a divergence-free field to the stress

The second source of non-uniqueness of the stress tensor involves the addition to it of a divergence-free field. This issue is partly addressed by the result (shown in Section 5.3) that the spatially-averaged pointwise stress converges to the virial stress in the thermodynamic limit (see footnote 1 on page 5). Consider the pointwise stress, $\boldsymbol{\sigma}$, obtained through the Irving–Kirkwood–Noll procedure, which satisfies the balance of linear momentum, and a new pointwise stress, $\hat{\boldsymbol{\sigma}} = \boldsymbol{\sigma} + \tilde{\boldsymbol{\sigma}}$, where $\text{div}_{\mathbf{x}} \tilde{\boldsymbol{\sigma}} = \mathbf{0}$. Clearly, $\hat{\boldsymbol{\sigma}}$ also satisfies the balance of linear momentum and is therefore also a valid solution. The spatially-averaged stress obtained from the new definition is

$$\hat{\boldsymbol{\sigma}}_w(\mathbf{x}, t) = \int_{\mathbb{R}^3} w(\mathbf{y} - \mathbf{x}) \hat{\boldsymbol{\sigma}}(\mathbf{y}, t) d\mathbf{y} = \int_{\mathbb{R}^3} w(\mathbf{y} - \mathbf{x}) (\boldsymbol{\sigma}(\mathbf{y}, t) + \tilde{\boldsymbol{\sigma}}(\mathbf{y}, t)) d\mathbf{y}. \quad (5.26)$$

We showed in Section 5.3 that in the thermodynamic limit, the spatially-averaged pointwise stress, $\boldsymbol{\sigma}$, converges to the virial stress. We also expect $\hat{\boldsymbol{\sigma}}_w$ to equal the virial stress in this limit (since any macroscopic stress must converge to this value under equilibrium conditions). Therefore, (5.26) reduces to

$$\lim_{\text{TD}} \int_{\mathbb{R}^3} w(\mathbf{y} - \mathbf{x}) \tilde{\boldsymbol{\sigma}}(\mathbf{y}, t) d\mathbf{y} = \mathbf{0}, \quad (5.27)$$

where \lim_{TD} refers to the thermodynamic limit. Equation (5.27) places a strong constraint on allowable forms for $\tilde{\boldsymbol{\sigma}}$, the implications of which are left for future work.

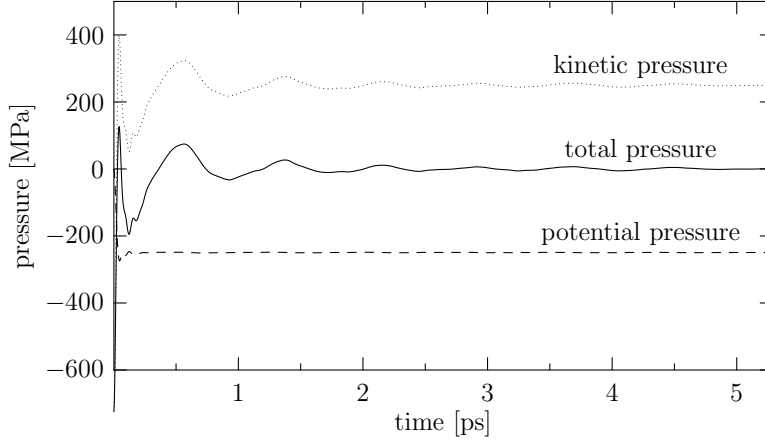


Fig. 6.1 The virial pressure as a function of time is plotted for an isolated cube of aluminum at 300K. The total pressure/virial pressure is the sum of the kinetic and potential pressures.

6 Numerical Experiments

In this section, we describe several numerical experiments, involving molecular dynamics and lattice statics simulations, conducted to capture differences in the spatially-averaged stress measures derived in Section 5. We consider the Hardy stress defined in (5.2), the Tsai traction defined in (5.3), the virial stress defined in (5.17) and the DA stress defined in (4.33). We will sometimes refer to these as the “microscopic definitions” or the “microscopically-based stress tensors”.

6.1 Experiment 1

We begin with the study of the kinetic part of the stress tensor. From the discussion in Section 5, it is clear that unlike the definition for the potential part of the stress tensor, there is no ambiguity in the definition for the kinetic part of stress. However, the kinetic part of the stress may appear to be at odds with the continuum definition of stress that is stated solely in terms of the forces acting between different parts of the body. The need for the kinetic part of stress becomes apparent when considering an ideal gas, where the potential interaction term is zero by definition and therefore it is the kinetic term that is wholly responsible for the transmission of pressure.

To demonstrate that the kinetic term in the stress tensor does indeed exist, we perform the following constant energy molecular dynamics simulation of an isolated cube. The cube, consisting of 4000 aluminum atoms in a face-centered cubic (fcc) arrangement ($10 \times 10 \times 10$ unit cells), is floating freely in a vacuum. The atoms interact according to an EAM potential for aluminum due to Ercolessi and Adams [16]:

$$\mathcal{V}_{\text{int}}^{\text{EAM}} = \frac{1}{2} \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \mathcal{V}_{\alpha\beta}(r_{\alpha\beta}) + \sum_{\alpha} \mathcal{U}_{\alpha}(\rho_{\alpha}), \quad \rho_{\alpha} = \sum_{\substack{\beta \\ \beta \neq \alpha}} f_{\beta}(r_{\alpha\beta}). \quad (6.1)$$

Here \mathcal{U}_α , called the *embedding function*, is the energy required to embed particle α in the electron density, ρ_α , due to the surrounding particles, and $f_\beta(r_{\alpha\beta})$ is the electron density of particle β at \mathbf{x}_α . The initial positions of the atoms are randomly perturbed by a small amount relative to their zero temperature equilibrium positions and the system is evolved by integrating the equations of motion. The initial perturbation is adjusted so that the temperature of the cube is about 300K (small fluctuations in temperature are expected since temperature is not controlled in the simulation). Since the block is unconstrained, we expect the stress, $\boldsymbol{\sigma}$, in the box and consequently the pressure, defined by $p = -\frac{1}{3} \text{tr } \boldsymbol{\sigma}$, to be zero. The virial expression for calculating the pressure follows from (A.23) as

$$p = \frac{1}{3V} \left[\sum_{\alpha} m_{\alpha} \overline{\|\mathbf{v}_{\alpha}\|^2} - \frac{1}{2} \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \overline{\|\mathbf{f}_{\alpha\beta}\| r_{\alpha\beta}} \right], \quad (6.2)$$

where

$$\mathbf{f}_{\alpha\beta} = \frac{\partial \mathcal{V}_{\text{int}}^{\text{EAM}}}{\partial \mathbf{r}_{\alpha\beta}} \frac{\mathbf{x}_{\beta} - \mathbf{x}_{\alpha}}{r_{\alpha\beta}}. \quad (6.3)$$

The three curves shown in Fig. 6.1 are the potential and kinetic parts of the pressure and the total pressure as a function of time, calculated using (6.2). As expected the total pressure tends to zero as the system equilibrates. However, the potential and kinetic parts are *non-zero*, converging to values that are equal and opposite such that their sum is zero. More interestingly, the kinetic part is not insignificant for our system. This clearly shows that kinetic part cannot be neglected even when considering solid systems. This can be quantified by noting that the kinetic part in (6.2) is simply the temperature per unit volume given by the equipartition theorem [26], $k_B T = 2\mathcal{T}/3N$, where \mathcal{T} is the kinetic energy. Therefore (6.2) reduces to

$$p = \frac{1}{V} \left[N k_B T - \frac{1}{6} \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \overline{\mathcal{V}'_{\alpha\beta} r_{\alpha\beta}} \right]. \quad (6.4)$$

For example at 300 K, $k_B T = 0.02585$ eV. The lattice spacing for the system considered is equal to 4.032 \AA . Hence, the volume per atom is $V/N = 4.023^3/4 = 16.387 \text{ \AA}^3$ and the kinetic pressure is 1.577 meV/\AA^3 . This translates to 252.394 MPa , which is a considerable stress.

6.2 Experiment 2

It is clear from Section 6.1, that the kinetic stress is a sizable quantity and cannot be neglected. In this experiment, we further explore the interplay between the potential and kinetic parts of the stress.

Consider a crystalline solid at a relatively low temperature under uniform stress. The atoms will vibrate about their mean positions with an amplitude that is small relative to the nearest-neighbor spacing. Now imagine placing a Tsai plane P between

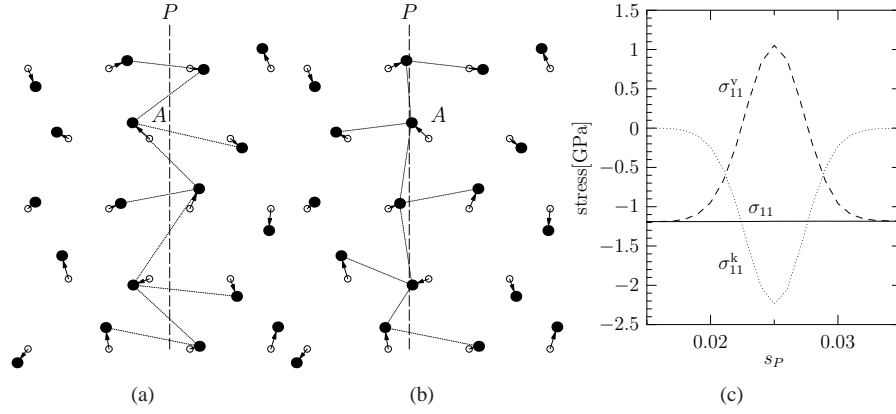


Fig. 6.2 The effect of the position of the Tsai plane on the potential and kinetic parts of stress. Frames (a) and (b) show schematic diagrams of a two-dimensional triangular lattice with (a) the Tsai plane positioned midway between the lattice planes and (b) the Tsai plane positioned almost on top of a lattice plane. The open circles correspond to the ideal lattice positions. The black circles are atoms that are shown in mid-vibration relative to the lattice site as indicated by the arrow. The Tsai plane is indicated by a vertical dashed line. The bonds crossing it appear as dotted lines. Frame (c) shows the plot of the kinetic part of stress σ_{11}^k , potential part of stress σ_{11}^v and the total stress σ_{11} , as a function of the normalized position $s_P = (x_P - x_L)/\Delta x$ of the Tsai plane P , where x_P is the position of P , x_L is the position of the lattice plane, and Δx is the spacing between the lattice planes.

two crystal lattice planes and measuring the traction across it. If P is midway between the lattice planes (see Fig. 6.2(a)), we expect that relatively few atoms will cross P and that consequently the kinetic stress will be small or even zero. In contrast, if P is close to the lattice plane there will be many such crossings and the kinetic stress will be large in magnitude. This seems to suggest that the traction will change as a function of the position of P , which would be incorrect since the system is under uniform stress. The reason that this does not occur is that every time an atom crosses P , the bonds connected with it reverse directions with respect to P , changing a positive contribution to the contact stress to a negative one and vice versa (see the bonds connected with atom A in Fig. 6.2(a) and Fig. 6.2(b)). This effect on the potential part of the stress exactly compensates for the change in magnitude of the kinetic stress leaving the total stress constant. This is demonstrated numerically in Fig. 6.2(c). This graph shows the results obtained from a molecular dynamics simulation of the system described in Section 6.1, with periodic boundary conditions. The periodic length of the box is set based on the zero temperature lattice spacing. Consequently upon heating by a temperature change of ΔT , a compressive stress is built up in the box according to

$$\epsilon = s : \sigma + I \alpha_T \Delta T = 0, \quad (6.5)$$

where s is the elastic compliance tensor and α_T is the coefficient of thermal expansion. Inverting this relation for an fcc crystal with cubic symmetry oriented along the crystallographic axes, we have

$$\sigma_{11} = \sigma_{22} = \sigma_{33} = -(c_{11} + 2c_{12})\Delta T = \sigma, \quad (6.6)$$

with the rest of the stress components zero. In (6.6), c_{ij} are the elastic constants of the material. Substituting in the appropriate values for Ercolessi-Adams EAM aluminum [16] ($c_{11} = 118.1$ GPa, $c_{12} = 62.3$ GPa, $\alpha_T = 1.6 \times 10^{-5} \text{K}^{-1}$) and $\Delta T = 310 \text{K}$, gives $\sigma = -1.2$ GPa. We see that the total stress in Fig. 6.2(c) is constant regardless of the position of the Tsai plane and equal to the expected value of -1.2 GPa computed above. However, the kinetic and potential parts change dramatically. When the Tsai plane is away from the lattice planes ($s_P = \pm 0.1$), the kinetic stress is zero and the entire stress is given by the potential part of the stress. As the Tsai plane moves closer to a lattice plane ($|s_P| \rightarrow 0$), the kinetic stress becomes more negative (increasing in magnitude) and the potential part of stress increases in exactly the right amount to compensate for this effect. When the Tsai plane is right on top of a lattice plane ($s_P = 0$), both the kinetic stress and potential stress are maximum in magnitude, but their sum remains equal to the constant total stress. This is a striking demonstration of the interplay between the kinetic and potential parts of the stress tensor.

6.3 Experiment 3 and 4

In this section, the predictions of the microscopically-based stress tensors are compared with analytical solutions from elasticity theory for two simple boundary-value problems. This is a revealing test, since stress is a continuum concept and therefore the microscopic definitions should reproduce the results of a continuum theory under the same conditions. We perform two numerical experiments. In each experiment, an atomistic boundary-value problem is set up, and the values computed from the discrete system are compared with the “exact” result computed from elasticity theory for the same problem using material properties predicted by the interatomic potential used in the atomistic calculations. The numerical experiments are conducted at zero-temperature since there is no controversy regarding the form of the kinetic stress which is the same for all stress definitions. Therefore, a comparison at zero temperature is sufficient to probe the differences between the stress measures, at least under equilibrium conditions. The properties we are interested in studying are:

1. Symmetry of the stress tensor.
2. Convergence of the stress tensor to the continuum value with the size of the averaging domain (a three-dimensional volume in the case of virial, Hardy and DA stresses and a plane in the case of the Tsai traction).

Inter-atomic Model

The numerical experiments in this section are carried out using a Lennard-Jones potential. The exact choice of material parameters is unimportant, since the objective of the experiment is to compare the values obtained from the microscopically-based stress for the discrete system with the “exact” values obtained from the continuum elasticity theory for the same material. The Lennard-Jones parameters, ϵ and σ , are therefore arbitrarily set to 1. The potential has the following form:

$$\phi(r) = 4 \left[\frac{1}{r^{12}} - \frac{1}{r^6} \right] - 0.0078r^2 + 0.0651. \quad (6.7)$$

Note that the above equation has been rendered dimensionless by expressing lengths in units of σ and energy in units of ϵ . As seen in the above equation, the Lennard-Jones potential is modified by the addition of a quadratic term. This is done to ensure that $\phi(r_{\text{cut}}) = 0$ and $\phi'(r_{\text{cut}}) = 0$, where $r_{\text{cut}} = 2.5$, denotes the cutoff radius for the potential. We refer to this as the “modified Lennard-Jones potential”. The ground state of this potential is an fcc crystal with a lattice constant of $a = 1.556$ and elastic constants, $c_{11} = 87.652$, $c_{12} = c_{44} = 50.379$. The conventional elastic moduli associated with the cubic elastic constants are [34]:

$$E = (c_{11}^2 + c_{11}c_{12} - 2c_{12}^2)/(c_{11} + c_{12}) = 50.877, \quad (6.8)$$

$$\mu = c_{44} = 50.379, \quad (6.9)$$

$$\nu = c_{12}/(c_{11} + c_{12}) = 0.365, \quad (6.10)$$

where E is Young’s modulus, μ is the shear modulus, and ν is Poisson’s ratio. (In the above, elastic constants are given in units of ϵ/σ^3 . Poisson’s ratio is dimensionless.)

Experiment 3: Dependence of the microscopically-based stress on the averaging domain size

The main aim of this experiment is to study the dependence of the stress given by various definitions (Hardy, Tsai, virial and DA) on the size of the averaging domain. We consider the special case of uniform uniaxial loading with $\sigma_{11} = 1$ (all other stress components zero). Our system is a cube of $10 \times 10 \times 10$ unit cells (4000 atoms) with periodic boundary conditions applied in all directions. To impose the uniaxial loading, the periodic lengths l_i ($i = 1, 2, 3$) in the three directions are modified according to the linear elastic solution for uniform straining:

$$l_1 = 10a(1 + \sigma_{11}/E) = 10.197, \quad (6.11)$$

$$l_2 = l_3 = 10a(1 - \nu\sigma_{11}/E) = 9.928. \quad (6.12)$$

We then compute the stress at the center of the periodic cell, while increasing the size of the averaging domain. In comparing the different stress definitions, the domain size is set by the Tsai plane which is taken to be a square normal to the 1-direction with the same dimension w in the 2 and 3-directions. The averaging domain for the Hardy, virial and DA stresses is a sphere of diameter $d = w$. The weighting function, $w(r)$, for the Hardy stress is taken to be constant with a suitable mollifying function,

$$w(r) = \begin{cases} c & \text{if } r < R - \epsilon \\ \frac{1}{2}c \left[1 - \cos\left(\frac{R-r}{\epsilon}\pi\right) \right] & \text{if } R - \epsilon < r < R \\ 0 & \text{otherwise} \end{cases}, \quad (6.13)$$

where c is chosen appropriately to normalize w . The results are presented in Fig. 6.3, where the stress σ_{11} is plotted as a function of the normalized domain size, $s = w/10a$. (Recall that the applied value is $\sigma_{11} = 1$.) We make the following qualitative observations based on these results:

1. The Hardy stress converges to the exact value most quickly of all stress definitions and has the least noise.

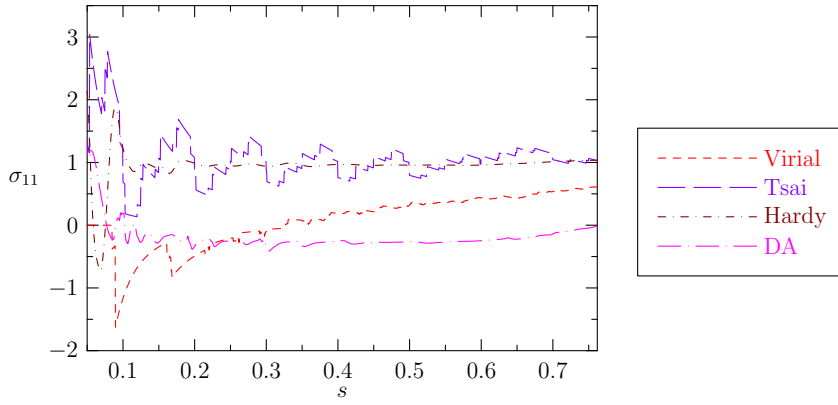


Fig. 6.3 Plot showing the dependence of σ_{11} , calculated using different definitions on the averaging domain size. The variable s represents the ratio of the domain size to the length of the system (10 unit cells).

2. The normal stress computed from the Tsai traction oscillates about the exact value with a fluctuation amplitude that decays rather slowly with domain size. The oscillations reflect the symmetry of the crystal as new bonds enter the calculation with increasing plane size.
3. The virial stress is always smaller than the Hardy and Tsai stresses since it does not take into account the bonds that cross out of the averaging domain. It appears to be converging towards the exact value, but convergence is slow and even at the maximum domain size studied, the virial stress still has a significant error.
4. The DA stress is much smaller than all other stresses due to greater averaging.

Experiment 4: A plate with a hole under tension

We now consider one of the classical elasticity boundary-value problems: an infinite plate with a hole subjected to uniaxial tension σ_∞ at infinity. This is traditionally named the *Kirsch* problem for an isotropic material model. Our objective is to compare the microscopically-based stresses computed for a discrete system set up for the Kirsch problem with the exact solution. A complication in making this comparison is that the fcc Lennard-Jones material we are considering is crystalline with cubic symmetry and is not isotropic. We must therefore compare the discrete solution with the more general solution for the Kirsch problem from the theory of elasticity for anisotropic media [34]. For anisotropic materials, the stress concentration⁴⁰ at the hole is no longer 3 (as it is for an isotropic material), but depends on the elastic constants of the material. For the elastic constants of the Lennard-Jones model in (6.7), we obtain a stress concentration of 2.408. In addition to the overall stress concentration, the analytical solution provides the complete stress field about the hole. We can therefore compare the microscopically-based stress fields with the continuum result.

⁴⁰The stress concentration is defined as the ratio of the maximum stress to the applied stress σ_∞ . The maximum stress for the Kirsch problem occurs at the circumference of the hole.

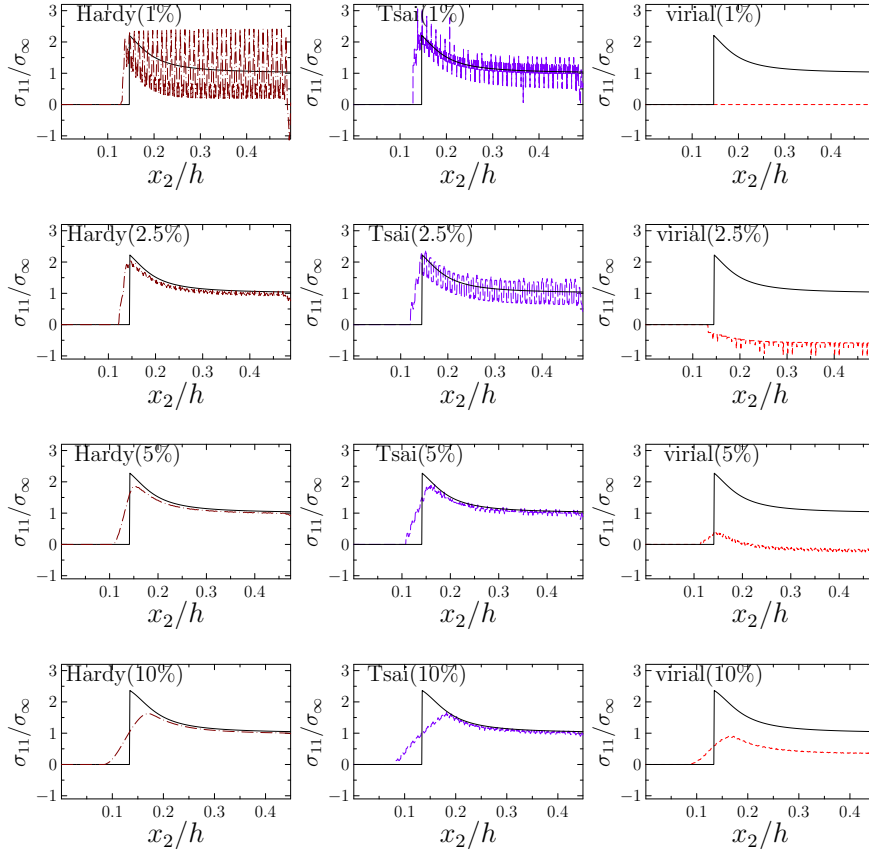


Fig. 6.4 Normalized σ_{11} component of the stress along the $x_1 = 0$ line for an anisotropic plate with a hole subjected to uniaxial tension in the 1-direction. The x_2 coordinate is normalized by the height h of the atomistic model. The exact solution for an infinite plate obtained from anisotropic linear elasticity (black solid line) is compared with the results obtained from the three microscopic definitions in the three columns: Hardy, Tsai and virial. The four rows correspond to four different averaging domains constituting 1%, 2.5%, 5% and 10% of h .

In order to model an infinite elastic space, we consider a large square plate oriented along the crystallographic axes consisting of 367,590 atoms, with a hole of radius $25a$, where a is the lattice constant. The plate is constructed by stacking $100 \times 100 \times 10$ unit cells and excluding the atoms that lie within the radius of the hole. The relatively large system size helps to ensure that the variation of the continuum stress is small on the lengthscale of the lattice spacing and minimizes boundary effects near the hole. The atoms interact according to the modified Lennard-Jones potential given in (6.7).

As before, the averaging domain size is set by the length and width of the Tsai plane, with the Hardy, virial and DA stress using a sphere of diameter equal to the length of the Tsai plane. The system is loaded by $\sigma_{11} = \sigma_{\infty}$, by displacing the atoms according to the exact solution from continuum mechanics for linear elastic

anisotropic media [34]. The applied stress is sufficiently small, so that the assumption of material linearity and the small strain approximation inherent in the elasticity theory provide a good approximation for the behavior of the system. After the atoms are displaced, the stress tensor is evaluated on a uniformly-distributed grid of points on the mid-plane of the plate located at $x_3 = 0$. A grid of 100×100 points is chosen to evaluate the virial, Tsai and Hardy expressions and a grid of 30×30 is chosen to evaluate the DA stress tensor. A coarser grid is used for the DA stress due to the higher computational cost of this calculation.

First, we consider the σ_{11} component of the stress along the $x_1 = 0$ line, where we expect the maximum value at the hole surface. The results are plotted in Fig. 6.4, which shows a comparison between the exact value and the three microscopic stress definitions (Hardy, Tsai and virial) for four different averaging domains ranging from 1% to 10% of the height h of the atomistic model. We see that the Hardy and Tsai stresses faithfully follow the exact solution, but then drop-off as their averaging domain overlaps with the hole. This drop-off reflects the fact that the microscopically-based stress measures are bulk expressions. The smaller the averaging domain, the closer the microscopic measures can approach the exact stress concentration at the hole surface, however, this increased fidelity comes at the cost of significantly large fluctuation about the exact value. The virial stress is identically zero for the smallest averaging domain because it is too small to contain complete bonds. For the same reason, the Hardy stress experiences very large fluctuations and a nearly constant average value. For larger averaging domains, the Hardy stress has smaller fluctuations than the other stress definitions.

The reason that the drop-off effect described above is so pronounced in this simulation, is that the system is very small by continuum standards. If instead of a hole with a radius of $25a$, we studied a plate with a hole 100 or 1000 times larger, using the same-sized averaging domain, the spatially-averaged expressions would get much closer to the correct value before dropping off over the same lengthscale as seen in Fig. 6.4. However, microscopic stress measures are often computed for small system sizes and therefore the difficulties presented in the figure are typical of realistic atomistic simulation.

Next, we explore the stress field over the entire plane. The color density plots given in Fig. 6.5, show variation of σ_{11} in the mid-plane of the plate. It can be seen that the stress within the hole is zero. Comparing the plots for σ_{11} of the virial stress (Fig. 6.5(b)), Tsai stress (Fig. 6.5(c)), Hardy stress (Fig. 6.5(d)) and the DA stress (Fig. 6.5(e)) with the exact solution (Fig. 6.5(a)), we see that the first three definitions capture the overall variation in σ_{11} , whereas the DA stress does not. However, it is clear that the microscopically-based stress in all of the cases is smeared relative to the stress given by the exact solution and none reach the exact stress concentration of 2.408. This is a result of the averaging procedure involved in all the definitions as explained above. Although the DA stress tensor plotted in Fig. 6.5(e)⁴¹ captures the variations in the field, it is much smaller in magnitude compared to the exact solution. This is because of the greater degree of averaging involved in the DA stress

⁴¹Fig. 6.5(e) and Fig. 6.8 are generated from a much coarser grid compared to the other plots due to the computational expense of the DA stress definition.

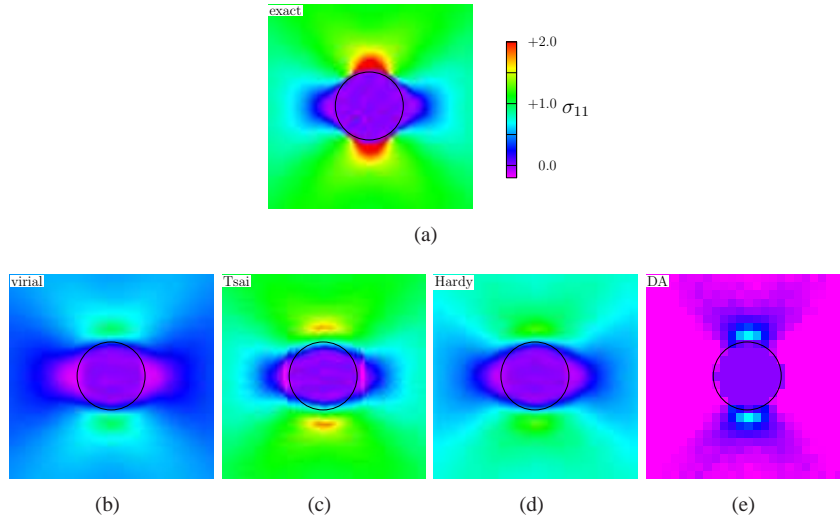


Fig. 6.5 Color density plots of σ_{11} are plotted on a common scale: (a) exact (b) virial (c) Tsai (d) Hardy (e) DA. Results plotted for an averaging domain size of 10% of the height of the model.

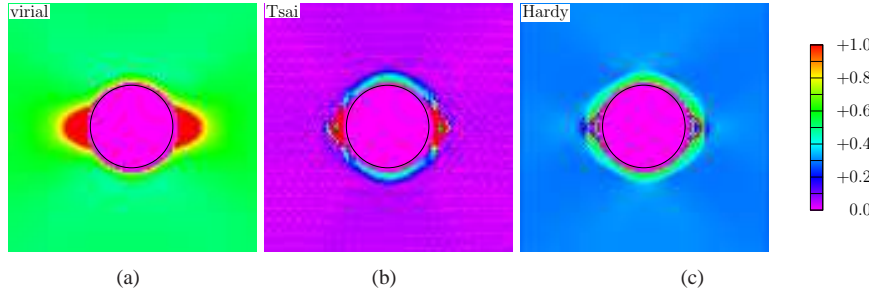


Fig. 6.6 Color density plots of error in σ_{11} , defined as the absolute value of $(\sigma_{11} - \sigma_{11}^{\text{exact}})/\sigma_{11}^{\text{exact}}$, where σ_{11} is the stress calculated using (a) virial (b) Tsai and (c) Hardy stress definitions.

tensor. Overall, the Hardy stress is less noisy than the virial or Tsai definitions due to the smoothing afforded by the weighting function. (This is hard to see in the figure.) The stress computed from the Tsai traction, in particular, is more noisy since the averaging is limited to a plane compared with the volume averaging of the Hardy and virial definitions. However, this more localized definition enables the Tsai stress to approach the exact stress concentration most closely of all of the microscopic definitions. Similar results were observed by Cheung and Yip [7] for the stress near a free surface.

The relative error in σ_{11} for the three microscopic definitions is shown in Fig. 6.6. Of the three definitions, the stress computed from the Tsai traction is generally more accurate, followed by the Hardy stress and then the virial stress. As noted above, the Tsai stress does particularly well in capturing the variations in the stress field close to the hole where the fact that it is localized in one direction is particularly helpful.

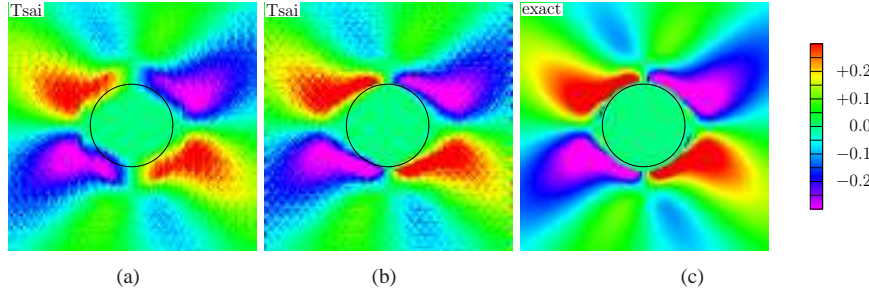


Fig. 6.7 Color density plots of the shear stress components computed from the Tsai traction, (a) σ_{12} and (b) σ_{21} , and (c) the exact shear stress, plotted on a common scale.

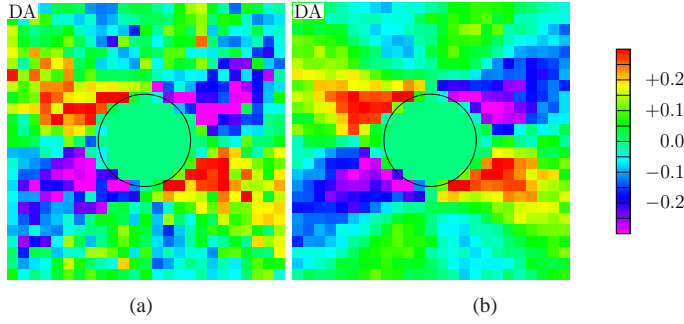


Fig. 6.8 Color density plots of the DA shear stress components, (a) σ_{12} and (b) σ_{21} , plotted on a common scale.

It is also interesting to examine the shear stress components. Fig. 6.7 shows the exact result from the continuum solution and the σ_{12} and σ_{21} components of the stress tensor computed from the Tsai traction from two different planes, one normal to the 1 direction and the other normal to the 2 direction. We see that the Tsai stress reproduces the exact distribution and appears generally symmetric. This suggests that the symmetry of the Hardy stress is preserved while taking the limit to arrive at the Tsai traction. The DA stress tensor is in general non-symmetric [45], but from Fig. 6.8(a) and Fig. 6.8(b) we observe that in this case, σ_{12} and σ_{21} appear similar. Interestingly, in contrast to the normal stress, the magnitude of shear stress is captured by the DA stress definition, at least for the case studied here. The reason for this is not obvious.

Overall, we can summarize our results as follows. Of the three definitions studied, the Hardy stress is generally preferred. It tends to be the smoothest and provides good accuracy away from surfaces as long as the lengthscale over which the continuum fields vary is large relative to the atomic spacing. In situations where either of those conditions break down, the Tsai traction provides a better localized measure of stress. The virial stress is less accurate than both. From a computational standpoint, the virial stress has the advantage of being easiest to compute. The evaluation of the bond function in the Hardy stress makes it slightly more expensive to compute, but comparable to the virial stress. The Tsai traction is most difficult and time consuming

to compute, since it requires the detection of bonds and atoms that cross a given plane during the averaging process. Furthermore, this evaluation must be performed for three separate planes in order to obtain the full stress tensor in three dimensions.

7 Summary and future work

In this paper, we provide a unified interpretation and possible generalization of all commonly used definitions for the stress tensor for a discrete system of particles. The macroscopic stress in a system under conditions of thermodynamic equilibrium are derived using the ideas of canonical transformations within the framework of classical statistical mechanics. The stress in non-equilibrium systems is obtained in a two-step procedure:

1. The Irving–Kirkwood–Noll procedure [27, 48] is applied to obtain a *pointwise* (microscopic) stress tensor. The stress consists of a kinetic part σ_k and a potential part σ_v . The potential part of the stress is obtained for multi-body potentials which have a continuously differentiable extension from the shape space of the system to a higher-dimensional Euclidean space.⁴² This generalizes the original Irving–Kirkwood–Noll approach that was limited to pair potentials. This generalization is obtained based on the important result that for any multi-body potential with a continuously differentiable extension, the force on a particle in a discrete system can always be expressed as a sum of *central* forces. In other words, the *strong law of action and reaction* is always satisfied.
2. The pointwise stress obtained in the first step is spatially averaged to obtain the macroscopic stress.

This two-step procedure provides a general unified framework from which various stress definitions can be derived including *all* of the main definitions commonly used in practice. In particular, it is shown that the two-step procedure leads directly to the stress tensor derived by Hardy in [23]. The traction of Cauchy and Tsai [5, 6, 63] is obtained from the Hardy stress in the limit that the averaging domain is collapsed to a plane. The virial stress of Clausius and Maxwell [8, 40, 41] is an approximation of the Hardy stress tensor for a uniform weighting function where bonds that cross the averaging domain are neglected. The Hardy stress and virial stress become identical in the thermodynamic limit. In this manner, clear connections are established between all of the major stress definitions in use today.

The unified framework described above yields a *symmetric* stress tensor for *all* interatomic potentials which have an extension, when used with the standard Irving–Kirkwood–Noll procedure. However, there are materials in nature, such as liquid crystals, which can have non-symmetric stress tensors. In order to explore the possibility of non-symmetric stress, the Irving–Kirkwood–Noll procedure is generalized to *curved paths of interaction* as suggested in [52]. This involves the generalization of Noll’s lemmas in [48], originally derived for straight bonds, to arbitrary curved paths as defined in this paper. These generalized lemmas, lead to a non-symmetric

⁴²Most practical interatomic potentials satisfy this conditions.

stress tensor when applied within the Irving–Kirkwood–Noll procedure. It is postulated that curved paths of interaction may be important in systems with internal degrees of freedom, such as liquid crystals and objective structures [28]. This is left for future work.

One of the key points addressed in this paper is the uniqueness of the stress tensor. Three possible sources of non-uniqueness of the stress tensor are identified and addressed:

1. Different pointwise stress definitions can be obtained for different potential energy extensions. This is demonstrated through a simple one-dimensional example. We also show that regardless of the uniqueness of the pointwise stress tensor, the *macroscopic* stress tensor obtained through a procedure of spatial averaging is unique since the difference resulting from alternative pointwise stress tensors tends to zero as the volume of the averaging domain is increased.
2. The pointwise stress tensor is obtained by solving the balance of linear momentum, $\text{div}_x \sigma_v = h(x)$, where σ_v is the potential part of the stress tensor, and $h(x)$ is a known function. The Irving–Kirkwood–Noll procedure leads to a closed-form solution to this problem. However, an arbitrary tensor field $\tilde{\sigma}$ with zero divergence can be added to σ_v without violating the balance of linear momentum. We argue that in the thermodynamic limit, the non-equilibrium stress obtained through our unified two-step process must converge to the virial stress of equilibrium statistical mechanics. This is similar to the argument made by Wajnryb et al. [64]. This condition is satisfied by the general stress expression that we obtain. Any divergence-free stress $\tilde{\sigma}$ added to this stress must therefore also disappear under equilibrium conditions. This greatly restricts the allowable forms of $\tilde{\sigma}$.
3. The generalization of the Irving–Kirkwood–Noll procedure from straight bonds to arbitrary curved paths of interaction implies the existence of multiple stress tensors for a given system. However, the existence of curved bonds implies the existence of internal structure for the discrete particles, a possibility already discussed by Kirkwood in [29]. For a system of point masses without internal structure, only straight bonds are possible due to symmetry arguments, and therefore this source of non-uniqueness is removed. The general case of non-symmetric stress must be addressed within the context of an appropriate multipolar theory as discussed by Pitteri [49]. We leave this to future work.

In addition to the unified framework described above which is based on the Irving–Kirkwood–Noll procedure, we also investigated the Murdoch–Hardy procedure [23, 43] of defining continuum fields as direct spatial averages of microscopic quantities. We demonstrate that this approach can be systematized by adopting a non-local continuum perspective and introducing suitable generator functions. The various stress definitions resulting from the Murdoch–Hardy procedure, such as the Hardy, virial and the “double-averaged” stress (suggested by Murdoch in [47]) can be derived from this unified framework. Although we share the concern regarding the ambiguity of the probability density functions used in the Irving–Kirkwood–Noll procedure that led Murdoch to develop the direct spatial averaging approach [46], we feel that since these probability density functions exist *in principle*, the Irving–

Kirkwood–Noll formalism is the correct framework to phrase the problem in with approximations introduced later to derive practical expressions.

Finally, numerical experiments involving molecular dynamics and lattice statics simulations are conducted to study the various stress definitions derived in this paper. It is generally observed that the Hardy stress definition appears to be most accurate and converges most quickly with the averaging domain size. In situations where a more localized measure of stress is needed, such as near surfaces or defects, the Tsai traction can be used instead. The virial stress is less accurate than the other two definitions and converges most slowly with averaging domain size. Its main advantage is its simple form and low computational cost. One of the most interesting results, which requires further study, comes from Experiment 2 of a crystalline system under uniform hydrostatic stress. Fig. 6.2(c) shows that although the potential and kinetic parts of the Tsai traction largely depend on the position of the Tsai plane between two adjacent lattice planes, the total stress remains constant. This calculation provides a striking demonstration of the interplay between the kinetic and potential parts of the stress tensor.

A Derivation of the virial stress tensor

The original virial theorem was a scalar equation credited to Clausius [8], which gives a definition for the pressure in a gas. Maxwell [40] extended this result to a tensor version which gives a definition for stress. We present here a more modern version of the virial theorem partly based on the derivation by Marc and McMillan [38].

Consider a system of N interacting point masses. The position of each point mass is given by

$$\mathbf{x}_\alpha = \mathbf{x} + \mathbf{r}_\alpha, \quad (\text{A.1})$$

where \mathbf{x} is the position of the center of mass of the system of particles and \mathbf{r}_α is the position of each point mass relative to the center of mass. From Newton's second law, we have

$$\mathbf{f}_\alpha = \dot{\mathbf{p}}_\alpha, \quad (\text{A.2})$$

where \mathbf{f}_α is the force acting on particle α , and \mathbf{p}_α is its linear momentum given by

$$\mathbf{p}_\alpha = m_\alpha(\dot{\mathbf{x}} + \dot{\mathbf{r}}_\alpha) = m_\alpha(\dot{\mathbf{x}} + \mathbf{v}_\alpha^{\text{rel}}). \quad (\text{A.3})$$

In (A.3), m_α is the mass of particle α and $\mathbf{v}_\alpha^{\text{rel}} = \dot{\mathbf{r}}_\alpha$ is its relative velocity with respect to the center of mass. Since the position of the center of mass \mathbf{x} is given by

$$\mathbf{x} = \frac{\sum_\alpha m_\alpha \mathbf{x}_\alpha}{\sum_\alpha m_\alpha}, \quad (\text{A.4})$$

we have the following identities which will be used later:

$$\sum_\alpha m_\alpha \mathbf{r}_\alpha = \mathbf{0}, \quad \sum_\alpha m_\alpha \mathbf{v}_\alpha^{\text{rel}} = \mathbf{0}. \quad (\text{A.5})$$

Next, we apply a tensor product with \mathbf{r}_α to both sides of (A.2) to give

$$\mathbf{r}_\alpha \otimes \mathbf{f}_\alpha = \mathbf{r}_\alpha \otimes \dot{\mathbf{p}}_\alpha. \quad (\text{A.6})$$

On using the identity

$$\frac{d}{dt}(\mathbf{r}_\alpha \otimes \mathbf{p}_\alpha) = \mathbf{v}_\alpha^{\text{rel}} \otimes \mathbf{p}_\alpha + \mathbf{r}_\alpha \otimes \dot{\mathbf{p}}_\alpha, \quad (\text{A.7})$$

equation (A.6) becomes

$$\frac{d}{dt}(\mathbf{r}_\alpha \otimes \mathbf{p}_\alpha) = \mathcal{W}_\alpha + 2\mathcal{T}_\alpha, \quad (\text{A.8})$$

where $\mathcal{W}_\alpha = \mathbf{r}_\alpha \otimes \mathbf{f}_\alpha$ is the *virial tensor* and $\mathcal{T}_\alpha = \frac{1}{2}(\mathbf{v}_\alpha^{\text{rel}} \otimes \mathbf{p}_\alpha)$ is the *kinetic tensor* of particle α . Equation (A.8) is called the *dynamical tensor virial theorem*. This “theorem”, which is simply an alternative form for the balance of linear momentum, becomes useful in a modified form after making the assumption that the atoms in the system are in *stationary motion*. This means that there exists a time scale τ , which is short relative to macroscopic processes but long relative to the characteristic time of the atomic system, over which the atoms remain close to their original positions with bounded positions and velocities. This condition is satisfied for a system of atoms undergoing thermal vibrations about their mean positions as expected in a solid at moderate temperature. To exploit this property of the system, we define the time average of any quantity f over the time τ as

$$\bar{f} = \frac{1}{\tau} \int_0^\tau f(t) dt \quad (\text{A.9})$$

and apply this averaging to (A.8):

$$\frac{1}{\tau}(\mathbf{r}_\alpha \otimes \mathbf{p}_\alpha) \Big|_0^\tau = \overline{\mathcal{W}_\alpha} + 2\overline{\mathcal{T}_\alpha}. \quad (\text{A.10})$$

Assuming that $\mathbf{r}_\alpha \otimes \mathbf{p}_\alpha$ is bounded, and a separation of time scales between microscopic and continuum processes exists, the term on the left-hand side can be made as small as desired by taking τ sufficiently

large. Therefore the above equation is reduced to $\overline{\mathcal{W}}_\alpha = -2\overline{\mathcal{T}}_\alpha$. Summing over all particles gives the *tensor virial theorem*:

$$\overline{\mathcal{W}} = -2\overline{\mathcal{T}}, \quad (\text{A.11})$$

where $\overline{\mathcal{W}}$ is the time-averaged virial tensor of the system,

$$\overline{\mathcal{W}} = \sum_{\alpha} \overline{\mathbf{r}_{\alpha} \otimes \mathbf{f}_{\alpha}}, \quad (\text{A.12})$$

and $\overline{\mathcal{T}}$ is the time-averaged kinetic tensor of the system,

$$\overline{\mathcal{T}} = \frac{1}{2} \sum_{\alpha} m_{\alpha} \overline{\mathbf{v}_{\alpha}^{\text{rel}} \otimes \dot{\mathbf{x}}}. \quad (\text{A.13})$$

The expression for $\overline{\mathcal{T}}$ can be simplified by substituting in (A.1) and noting that due to separation of time scales $\dot{\mathbf{x}}$ is constant on the atomistic time scale τ , so that

$$\overline{\mathcal{T}} = \frac{1}{2} \sum_{\alpha} m_{\alpha} \overline{\mathbf{v}_{\alpha}^{\text{rel}} \otimes \mathbf{v}_{\alpha}^{\text{rel}}} + \left[\sum_{\alpha} m_{\alpha} \overline{\mathbf{v}_{\alpha}^{\text{rel}}} \right] \otimes \dot{\mathbf{x}}. \quad (\text{A.14})$$

The term in the square brackets is zero due to (A.5), and thus

$$\overline{\mathcal{T}} = \frac{1}{2} \sum_{\alpha} m_{\alpha} \overline{\mathbf{v}_{\alpha}^{\text{rel}} \otimes \mathbf{v}_{\alpha}^{\text{rel}}}. \quad (\text{A.15})$$

It is important to note that the virial theorem in (A.11) applies equally to continuum systems at rest as well as those that are not in macroscopic equilibrium and are therefore in a state of motion. This statement hinges on the separation of scales assumption according to which continuum motion occurs so slowly relative to atomistic processes so as to be essentially constant on that time scale.

The virial theorem leads to a definition for stress by considering the idea that the forces on the particles in the system can be divided into two parts: an internal part, $\mathbf{f}_{\alpha}^{\text{int}}$, resulting from the interaction of particle α with other particles in the system and an external part, $\mathbf{f}_{\alpha}^{\text{ext}}$, due to its interaction with atoms outside the system and due to external fields,

$$\mathbf{f}_{\alpha} = \mathbf{f}_{\alpha}^{\text{int}} + \mathbf{f}_{\alpha}^{\text{ext}}. \quad (\text{A.16})$$

In terms of continuum variables, the external part of the force can be identified with the traction, \mathbf{t} , that the surrounding medium applies to the system of particles and the body force, $\rho \mathbf{b}$, acting on it, where ρ is the mass density and \mathbf{b} is body force per unit mass.⁴³ Substituting (A.16) into (A.12), divides the virial tensor for the system into internal and external parts,

$$\overline{\mathcal{W}} = \overline{\mathcal{W}}_{\text{int}} + \overline{\mathcal{W}}_{\text{ext}} = \sum_{\alpha} \overline{\mathbf{r}_{\alpha} \otimes \mathbf{f}_{\alpha}^{\text{int}}} + \sum_{\alpha} \overline{\mathbf{r}_{\alpha} \otimes \mathbf{f}_{\alpha}^{\text{ext}}}. \quad (\text{A.17})$$

Rewriting the external virial as⁴⁴

$$\overline{\mathcal{W}}_{\text{ext}} = \sum_{\alpha} \overline{\mathbf{r}_{\alpha} \otimes \mathbf{f}_{\alpha}^{\text{ext}}} := \int_{\Omega} \mathbf{x} \otimes \rho \mathbf{b} dV + \int_{\partial\Omega} \mathbf{x} \otimes \mathbf{t} dA, \quad (\text{A.18})$$

where Ω is the domain occupied by the system of particles and $\partial\Omega$ is a continuous closed surface bounding the particles and separating them from surrounding particles. The variable \mathbf{x} is a position vector within

⁴³These density, traction and body forces are *pointwise* continuum fields, i.e., they are defined at all points but do not include the spatial averaging implicit in the macroscopic continuum description. See Section 3.1 for more details on pointwise fields.

⁴⁴We accept this step as an ansatz due to the many assumptions involved, one of which being that Ω is large enough to express the external forces acting on Ω in the form of the continuum traction \mathbf{t} .

Ω and on the surface $\partial\Omega$. Substituting the Cauchy relation, $\mathbf{t} = \boldsymbol{\sigma}\mathbf{n}$, where $\boldsymbol{\sigma}$ is the pointwise Cauchy stress, into the above equation and applying the divergence theorem, we have

$$\overline{\mathbf{W}}_{\text{ext}} = \int_{\Omega} [\mathbf{x} \otimes \rho \mathbf{b} + \text{div}_{\mathbf{x}}(\mathbf{x} \otimes \boldsymbol{\sigma})] dV = \int_{\Omega} [\boldsymbol{\sigma}^T + \mathbf{x} \otimes (\text{div}_{\mathbf{x}} \boldsymbol{\sigma} + \rho \mathbf{b})] dV. \quad (\text{A.19})$$

We assume that the pointwise fields satisfy the same balance laws as the macroscopic continuum fields. Therefore, the term $(\text{div}_{\mathbf{x}} \boldsymbol{\sigma} + \rho \mathbf{b})$ is zero under equilibrium conditions (see (3.23)). We therefore have that

$$\overline{\mathbf{W}}_{\text{ext}} = V \boldsymbol{\sigma}^T, \quad (\text{A.20})$$

where we have defined the continuum stress field as the average value of $\boldsymbol{\sigma}$ over the domain Ω :

$$\boldsymbol{\sigma}_{\text{av}} := \frac{1}{V} \int_{\Omega} \boldsymbol{\sigma} dV, \quad (\text{A.21})$$

Here V is the volume of Ω .⁴⁵ Substituting (A.20) into (A.17) and then into the virial theorem in (A.11) gives

$$\boldsymbol{\sigma}_{\text{av}} = -\frac{1}{V} [\overline{\mathbf{W}}_{\text{int}} + 2\overline{\mathcal{T}}]^T. \quad (\text{A.22})$$

Substituting in (A.15) and the definition of $\overline{\mathbf{W}}_{\text{int}}$ from (A.17), we have

$$\boldsymbol{\sigma}_{\text{av}} = -\frac{1}{V} \left[\sum_{\alpha} \overline{\mathbf{f}_{\alpha}^{\text{int}} \otimes \mathbf{r}_{\alpha}} + \sum_{\alpha} m_{\alpha} \overline{\mathbf{v}_{\alpha}^{\text{rel}} \otimes \mathbf{v}_{\alpha}^{\text{rel}}} \right]. \quad (\text{A.23})$$

This is the virial stress tensor. It is an expression for the Cauchy stress tensor given entirely in terms of atomistic quantities. Finally to demonstrate the symmetry of the virial stress, we rewrite $\mathbf{f}_{\alpha}^{\text{int}}$ as the sum over its decomposition:

$$\mathbf{f}_{\alpha}^{\text{int}} = \sum_{\substack{\beta \\ \beta \neq \alpha}} \mathbf{f}_{\alpha\beta}, \quad (\text{A.24})$$

where $\mathbf{f}_{\alpha\beta}$ are the terms in the central-force decomposition corresponding to a potential energy extension as explained in Section 3.4. Substituting (A.24) into (A.23), we obtain

$$\boldsymbol{\sigma}_{\text{av}} = -\frac{1}{V} \left[\sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \overline{\mathbf{f}_{\alpha\beta} \otimes \mathbf{r}_{\alpha}} + \sum_{\alpha} m_{\alpha} \overline{\mathbf{v}_{\alpha}^{\text{rel}} \otimes \mathbf{v}_{\alpha}^{\text{rel}}} \right]. \quad (\text{A.25})$$

Now recalling that $\mathbf{f}_{\alpha\beta} = -\mathbf{f}_{\beta\alpha}$, we have the following identity

$$\sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \mathbf{f}_{\alpha\beta} \otimes \mathbf{r}_{\alpha} = \frac{1}{2} \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} (\mathbf{f}_{\alpha\beta} \otimes \mathbf{r}_{\alpha} + \mathbf{f}_{\beta\alpha} \otimes \mathbf{r}_{\beta}) = \frac{1}{2} \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \mathbf{f}_{\alpha\beta} \otimes (\mathbf{r}_{\alpha} - \mathbf{r}_{\beta}). \quad (\text{A.26})$$

Substituting (A.26) into (A.25), we have

$$\boldsymbol{\sigma}_{\text{av}} = -\frac{1}{V} \left[\frac{1}{2} \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \overline{\mathbf{f}_{\alpha\beta} \otimes (\mathbf{r}_{\alpha} - \mathbf{r}_{\beta})} + \sum_{\alpha} m_{\alpha} \overline{\mathbf{v}_{\alpha}^{\text{rel}} \otimes \mathbf{v}_{\alpha}^{\text{rel}}} \right]. \quad (\text{A.27})$$

This expression shows that the virial stress is symmetric, since $\mathbf{f}_{\alpha\beta}$ is parallel to $\mathbf{r}_{\alpha} - \mathbf{r}_{\beta}$.

⁴⁵The definition of this volume is somewhat arbitrary. One can possibly define V as the total volume of the Voronoi cells of the atoms lying within Ω .

B Distance geometry

In Section 3.4, we saw that the interatomic potential energy, \mathcal{V}_{int} , of a set of N particles can be defined as a function on a $3N - 6$ dimensional manifold embedded in $\mathbb{R}^{N(N-1)/2}$, where each point on this manifold represents an $N(N-1)/2$ -tuple of real numbers which correspond to the distances between the N particles in \mathbb{R}^3 . We also noted that these $N(N-1)/2$ distances must satisfy certain geometric constraints in order to be physically meaningful (see footnote 19). In this appendix, we discuss the nature of these geometric constraints.

An entire field of geometry, referred to as *distance geometry*, has emerged to describe the geometry of sets of points in terms of the distances between them. The subject was first treated systematically by Karl Menger in the early twentieth century and summarized in the book by Blumenthal [3]. More recent references include [4, 10, 11, 25, 56]. Our discussion below is partly based on [50], which provided a particularly clear explanation of the subject.

The key function in distance geometry is the *Cayley-Menger determinant*, $\chi : \mathbb{R}^{N(N-1)/2} \rightarrow \mathbb{R}$, which is defined as

$$\chi(\zeta_{12}, \dots, \zeta_{1N}, \zeta_{23}, \dots, \zeta_{(N-1)N}) = \det \begin{bmatrix} 0 & s_{12} & s_{13} & \cdots & s_{1N} & 1 \\ s_{12} & 0 & s_{23} & \cdots & s_{2N} & 1 \\ s_{13} & s_{23} & 0 & \cdots & s_{3N} & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ s_{1N} & s_{2N} & s_{3N} & \cdots & 0 & 1 \\ 1 & 1 & 1 & \cdots & 1 & 0 \end{bmatrix}, \quad (\text{B.1})$$

where

$$s_{\alpha\beta} = \zeta_{\alpha\beta}^2. \quad (\text{B.2})$$

For a system of N points, $\{\mathbf{x}_1, \dots, \mathbf{x}_N\}$, embedded in \mathbb{R}^3 , the Cayley-Menger determinant evaluated at the point, $(r_{12}, \dots, r_{(N-1)N})$, where $r_{\alpha\beta} = \|\mathbf{x}_\alpha - \mathbf{x}_\beta\|$, is related to the volume V_{N-1} of a simplex of N points in an $(N-1)$ -dimensional space through the relation:

$$\chi(r_{12}, \dots, r_{(N-1)N}) = \frac{2^{N-1}((N-1)!)^2}{(-1)^N} V_{N-1}^2(\mathbf{x}_1, \dots, \mathbf{x}_N). \quad (\text{B.3})$$

For $N = 2$,

$$\chi(r_{12}) = 2L^2, \quad (\text{B.4})$$

where $L = \sqrt{s_{12}}$ is the length of the segment defined by the two points. For $N = 3$,

$$\chi(r_{12}, r_{13}, r_{23}) = -16A^2, \quad (\text{B.5})$$

where A is the area of the triangle defined by the three points. For $N = 4$,

$$\chi(r_{12}, \dots, r_{34}) = 288V^2, \quad (\text{B.6})$$

where V is the volume of the tetrahedron defined by the four points. For $N \geq 5$, we must have

$$\chi(r_{12}, \dots, r_{(N-1)N}) = 0, \quad (\text{B.7})$$

for points in \mathbb{R}^3 since any simplex with five or more points has zero volume in three-dimensional space.⁴⁶

We now seek to go in the opposite direction. Rather than computing the squared distances $\{s_{\alpha\beta}\}$ from a set of points and using the Cayley-Menger determinants to compute volumes, we seek to verify that a set of distances actually corresponds to a set of points in three-dimensional space. In technical terms, we want the points associated with the distances to be *embeddable* in \mathbb{R}^3 . In order for this to be the case the following conditions must be satisfied:⁴⁷

⁴⁶This is easier to visualize in two-dimensional space. In that case, a simplex with four vertices (a tetrahedron) would have zero volume since its vertices would be confined to a plane. The same applies to higher-order simplexes and the corresponding higher-order volumes.

⁴⁷Actually, a somewhat stronger theorem can be proved. If the points are numbered in such a way that the first four points satisfy conditions 1 and 2, then conditions 3 and 4 need only be applied to groups of points that include these four points. See [3] for details.

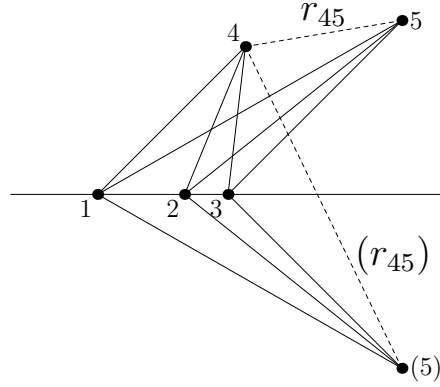


Fig. B.1 A cluster of 5 particles. The cluster is shown projected onto the plane normal to the plane defined by particles 1, 2 and 3 shown as a horizontal line. Atom 4 lies above this plane. There are two possible positions for atom 5 (at the same height above or below the 1-2-3 plane), where the distances $\{r_{12}, \dots, r_{35}\}$ are all the same and only distance r_{45} differs. The alternative position of atom 5 and corresponding distance r_{45} are shown in parentheses. Based on Fig. 2 in [39].

1. $\chi(r_{\alpha\beta}, r_{\alpha\gamma}, r_{\beta\gamma}) \leq 0, \quad \forall \alpha < \beta < \gamma,$
2. $\chi(r_{\alpha\beta}, r_{\alpha\gamma}, r_{\alpha\delta}, \dots, r_{\gamma\delta}) \geq 0, \quad \forall \alpha < \beta < \gamma < \delta,$
3. $\chi(r_{\alpha\beta}, r_{\alpha\gamma}, r_{\alpha\delta}, r_{\alpha\epsilon}, \dots, r_{\delta\epsilon}) = 0, \quad \forall \alpha < \beta < \gamma < \delta < \epsilon,$
4. $\chi(r_{\alpha\beta}, r_{\alpha\gamma}, r_{\alpha\delta}, r_{\alpha\epsilon}, r_{\alpha\zeta}, \dots, r_{\epsilon\zeta}) = 0, \quad \forall \alpha < \beta < \gamma < \delta < \epsilon < \zeta.$

For example, Condition 1 states that the Cayley-Menger determinants computed for all distinct triplets of points must be negative or zero. Conditions 2-4 apply similarly to sets of four, five and six points. The above conditions are easy to understand in terms of (B.4)-(B.7). They enforce the correct sign of areas and volumes in three-dimensional space and the degeneracy condition in (B.7).

As an example, let us see how this is applied for a cluster of $N = 5$ particles. There are $(5 \times 4)/2 = 10$ distances, and only $3 \times 5 - 6 = 9$ degrees of freedom. There is therefore one more interatomic distance than is needed to describe the configuration and indeed there is one Cayley-Menger determinant coming from condition 3:

$$\chi(r_{12}, \dots, r_{45}) = \det \begin{bmatrix} 0 & s_{12} & s_{13} & s_{14} & s_{15} & 1 \\ s_{12} & 0 & s_{23} & s_{24} & s_{25} & 1 \\ s_{13} & s_{23} & 0 & s_{34} & s_{35} & 1 \\ s_{14} & s_{24} & s_{34} & 0 & s_{45} & 1 \\ s_{15} & s_{25} & s_{35} & s_{45} & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix} = 0.$$

This expression can be expanded out leading to the following explicit expression [39]:⁴⁸

$$-\frac{1}{24} \sum_{\alpha=1}^5 \sum_{\beta=1}^5 \sum_{\gamma=1}^5 \sum_{\delta=1}^5 \sum_{\epsilon=1}^5 \left[24s_{\alpha\beta}s_{\beta\gamma}s_{\gamma\delta}s_{\delta\epsilon} - 6s_{\alpha\beta}s_{\beta\gamma}s_{\gamma\delta}s_{\delta\alpha} \right. \\ \left. - 8s_{\alpha\beta}s_{\gamma\delta}s_{\delta\epsilon}s_{\epsilon\gamma} - 12(s_{\alpha\beta})^2s_{\gamma\delta}s_{\delta\epsilon} + 3(s_{\alpha\beta})^2(s_{\gamma\delta})^2 \right] = 0, \quad (\text{B.8})$$

where in the above equation $s_{\alpha\beta} = s_{\beta\alpha}$ whenever $\alpha > \beta$. For a given set of nine squared distances, say $\{s_{11}, \dots, s_{35}\}$, (B.8) provides a quadratic equation for the tenth squared distance, s_{45} ,

$$A(s_{45})^2 + Bs_{45} + C = 0, \quad (\text{B.9})$$

⁴⁸Note that Martin [39] has a small typographical error in his relation.

where A , B and C are functions of the other squared distances. (See [30] for explicit expressions for A , B and C .) This means that there are two possible solutions for s_{45} (and hence also for r_{45}) when the other distances are set. This situation is demonstrated in Fig. B.1.

The above example shows that although only 9 degrees of freedom are necessary to characterize a cluster of 5 particles, selecting a subset of 9 distances is not sufficient. For this reason, any interatomic potential energy extension (see Section 3.4) must be expressed as a function of $N(N-1)/2$ arguments and not just an arbitrary subset of $3N-6$ of them.

C Useful lemmas

In his 1955 paper on the "Derivation of the fundamental equations of continuum thermodynamics from statistical mechanics" [48], Walter Noll proves two lemmas that play an important role in his derivation. In this section, for completeness, we derive Noll's first lemma and then extend it to arbitrary curved "paths of interaction". We then derive Noll's second lemma for this more generalized case.

Let $\mathbf{f}(\mathbf{v}, \mathbf{w})$ be a tensor-valued function of two vectors \mathbf{v} and \mathbf{w} , which satisfies the following three conditions:

1. $\mathbf{f}(\mathbf{v}, \mathbf{w})$ is defined for all \mathbf{v} and \mathbf{w} and is continuously differentiable .
2. There exists a $\delta > 0$, such that the auxiliary function $\mathbf{g}(\mathbf{v}, \mathbf{w})$, defined through

$$\mathbf{g}(\mathbf{v}, \mathbf{w}) := \mathbf{f}(\mathbf{v}, \mathbf{w}) \|\mathbf{v}\|^{3+\delta} \|\mathbf{w}\|^{3+\delta}, \quad (\text{C.1})$$

and its gradients $\nabla_{\mathbf{v}} \mathbf{g}$ and $\nabla_{\mathbf{w}} \mathbf{g}$ are bounded.

3. $\mathbf{f}(\mathbf{v}, \mathbf{w})$ is antisymmetric, i.e.,

$$\mathbf{f}(\mathbf{v}, \mathbf{w}) = -\mathbf{f}(\mathbf{w}, \mathbf{v}). \quad (\text{C.2})$$

Lemma C.1 *Under the conditions mentioned above, the following equation holds:*⁴⁹

$$\int_{\mathbf{y} \in \mathbb{R}^3} \mathbf{f}(\mathbf{x}, \mathbf{y}) d\mathbf{y} = -\frac{1}{2} \operatorname{div}_{\mathbf{x}} \int_{\mathbf{z} \in \mathbb{R}^3} \left[\int_{s=0}^1 \mathbf{f}(\mathbf{x} + s\mathbf{z}, \mathbf{x} - (1-s)\mathbf{z}) ds \right] \otimes \mathbf{z} d\mathbf{z}. \quad (\text{C.3})$$

Proof Conditions (1) and (2) guarantee absolute convergence and allow the order of integration to be swapped. From (C.2) we have

$$\int_{\mathbf{y} \in \mathbb{R}^3} \mathbf{f}(\mathbf{x}, \mathbf{y}) d\mathbf{y} = - \int_{\mathbf{y} \in \mathbb{R}^3} \mathbf{f}(\mathbf{y}, \mathbf{x}) d\mathbf{y}. \quad (\text{C.4})$$

Introduce new integration variables: on the left replace \mathbf{y} with $\mathbf{z} = \mathbf{x} - \mathbf{y}$ and on the right replace \mathbf{y} with $\mathbf{z} = \mathbf{y} - \mathbf{x}$. Thus,

$$\int_{\mathbf{z} \in \mathbb{R}^3} \mathbf{f}(\mathbf{x}, \mathbf{x} - \mathbf{z}) d\mathbf{z} = - \int_{\mathbf{z} \in \mathbb{R}^3} \mathbf{f}(\mathbf{x} + \mathbf{z}, \mathbf{x}) d\mathbf{z}. \quad (\text{C.5})$$

Note that a minus sign on the left-hand side is dropped by formally reversing the integration bounds. The two terms in (C.5) are equal to $\int_{\mathbf{y} \in \mathbb{R}^3} \mathbf{f}(\mathbf{x}, \mathbf{y}) d\mathbf{y}$, we can therefore write

$$\int_{\mathbf{y} \in \mathbb{R}^3} \mathbf{f}(\mathbf{x}, \mathbf{y}) d\mathbf{y} = \frac{1}{2} \int_{\mathbf{z} \in \mathbb{R}^3} [\mathbf{f}(\mathbf{x}, \mathbf{x} - \mathbf{z}) - \mathbf{f}(\mathbf{x} + \mathbf{z}, \mathbf{x})] d\mathbf{z}. \quad (\text{C.6})$$

Next, from the chain rule, we have

$$\nabla_{\mathbf{x}} \mathbf{f}(\mathbf{x} + s\mathbf{z}, \mathbf{x} - (1-s)\mathbf{z}) = \nabla_{\mathbf{v}} \mathbf{f} + \nabla_{\mathbf{w}} \mathbf{f}, \quad (\text{C.7})$$

where $s \in \mathbb{R}$. Similarly,

$$\frac{d}{ds} \mathbf{f}(\mathbf{x} + s\mathbf{z}, \mathbf{x} - (1-s)\mathbf{z}) = (\nabla_{\mathbf{v}} \mathbf{f} + \nabla_{\mathbf{w}} \mathbf{f}) \mathbf{z}. \quad (\text{C.8})$$

⁴⁹The expression in Noll's paper appears transposed relative to (C.3). This is because the gradient and divergence operations used by Noll are the transpose of our definitions. See end of Section 1.

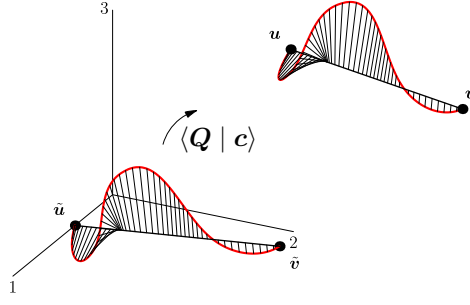


Fig. C.1 The property of a *path of interaction* as mentioned in the definition is illustrated in this figure. Two paths of interaction for the pairs (\mathbf{u}, \mathbf{v}) and $(\tilde{\mathbf{u}}, \tilde{\mathbf{v}})$ are shown in the figure.

Combining (C.7) and (C.8), we have

$$[\nabla_{\mathbf{x}} \mathbf{f}(\mathbf{x} + s\mathbf{z}, \mathbf{x} - (1-s)\mathbf{z})] \mathbf{z} = \frac{d}{ds} \mathbf{f}(\mathbf{x} + s\mathbf{z}, \mathbf{x} - (1-s)\mathbf{z}). \quad (\text{C.9})$$

Integrating the above equation with respect to s from 0 to 1 gives

$$\left[\nabla_{\mathbf{x}} \int_{s=0}^1 \mathbf{f}(\mathbf{x} + s\mathbf{z}, \mathbf{x} - (1-s)\mathbf{z}) ds \right] \mathbf{z} = \mathbf{f}(\mathbf{x} + \mathbf{z}, \mathbf{x}) - \mathbf{f}(\mathbf{x}, \mathbf{x} - \mathbf{z}). \quad (\text{C.10})$$

Substituting (C.10) into (C.6), and using the identity

$$(\nabla_{\mathbf{x}} \mathbf{T}) \mathbf{a} = \text{div}_{\mathbf{x}} (\mathbf{T} \otimes \mathbf{a}), \quad (\text{C.11})$$

where $\mathbf{T}(\mathbf{x})$ is a tensor of any order and \mathbf{a} is a vector which is not a function of \mathbf{x} , gives (C.3). \square

Lemma C.1 provides us a solution $\mathbf{F}(\mathbf{x})$, to the equation

$$\int_{\mathbf{y} \in \mathbb{R}^3} \mathbf{f}(\mathbf{x}, \mathbf{y}) d\mathbf{y} = \text{div}_{\mathbf{x}} \mathbf{F}(\mathbf{x}). \quad (\text{C.12})$$

It is clear that $\mathbf{F}(\mathbf{x})$ is only a single solution out of an infinite set of solutions that differ from it by a divergenceless second-order tensor field. The following lemma extends Lemma C.1 to accommodate other possible solutions of which the solution given in (C.3) is a special case. Before that we give the following definition:

Definition 3 For any \mathbf{u} and $\mathbf{v} \in \mathbb{R}^3$, the “the path of interaction” of \mathbf{u} and \mathbf{v} is a contour that joins \mathbf{u} and \mathbf{v} such that for any $\tilde{\mathbf{u}}$ and $\tilde{\mathbf{v}}$ in \mathbb{R}^3 , where $\|\mathbf{u} - \mathbf{v}\| = \|\tilde{\mathbf{u}} - \tilde{\mathbf{v}}\|$, the path of interaction of pairs (\mathbf{u}, \mathbf{v}) and $(\tilde{\mathbf{u}}, \tilde{\mathbf{v}})$ are related by a rigid body transformation $\langle \mathbf{Q} | \mathbf{c} \rangle$, for some $\mathbf{Q} \in \text{SO}(3)$ and $\mathbf{c} \in \mathbb{R}^3$, with $\mathbf{Q}(\mathbf{u} - \mathbf{v}) = \tilde{\mathbf{u}} - \tilde{\mathbf{v}}$ and $\mathbf{Q} = \mathbf{I}$, whenever $\mathbf{u} - \mathbf{v} = \tilde{\mathbf{u}} - \tilde{\mathbf{v}}$. See Fig. C.1.

Basically, this definition enforces the condition that the “path of interaction” is a contour whose shape is only a function of the distance between the points that it connects. The shape of the contour for a given distance is assumed to be dictated by the nature of bonding in the material. Here we assume this shape to be known. (See also footnote 28 on page 35).

Let $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ be a basis of \mathbb{R}^3 . For every $l > 0$, let $\Upsilon_l : [0, 1] \rightarrow \mathbb{R}^3$ be a continuously differentiable contour in \mathbb{R}^3 such that

$$\Upsilon_l(s) \cdot \mathbf{e}_1 = sl, \quad 0 \leq s \leq 1, \quad (\text{C.13a})$$

$$\Upsilon_l(0) = (0, 0, 0); \quad \Upsilon_l(1) = (l, 0, 0). \quad (\text{C.13b})$$

Fig. C.2 describes the properties of the contour Υ_l mentioned above.

By the definition of the *path of interaction*, it is clear that the shape of any path of interaction can be described by the contour Υ_l . Moreover, from its properties it is possible to explicitly define the path of

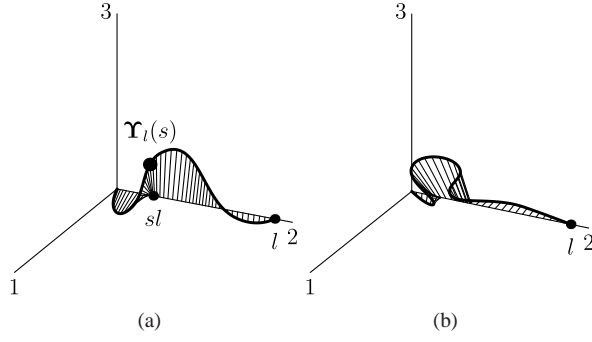


Fig. C.2 Frame (a) shows an admissible contour and frame (b) shows an inadmissible contour which violates property (C.13a) of Υ_l .

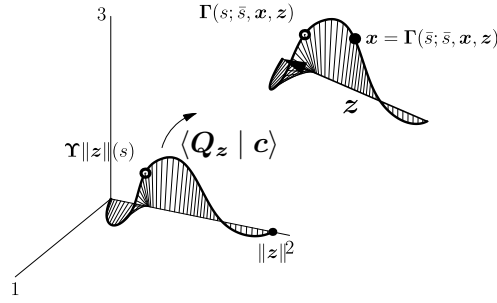


Fig. C.3 The contour $\Gamma(\cdot; \bar{s}, x, z)$ passes through x and the vector z is the difference between the two end points of the contour. It is related to the contour $\Upsilon_{||z||}$, such that any point $\Upsilon_{||z||}(s)$ is mapped to $\Gamma(\bar{s}; \bar{s}, x, z)$ (shown above as a hollow dot) by a rigid body motion $\langle Q_z | c \rangle$, where Q_z represents rotation and $c = x - Q_z \Upsilon_{||z||}(\bar{s})$, represents translation.

interaction as a function of Υ_l , a given point x through which it passes, and a vector z which connects its end points. More precisely, for any $0 \leq \bar{s} \leq 1$, $x \in \mathbb{R}^3$ and $z \in \mathbb{R}^3$, let $\Gamma(\cdot; \bar{s}, x, z) : [0, 1] \rightarrow \mathbb{R}^3$ denote a path of interaction in \mathbb{R}^3 , such that

$$\Gamma(\bar{s}; \bar{s}, x, z) = x + Q_z [\Upsilon_{||z||}(s) - \Upsilon_{||z||}(\bar{s})], \quad (\text{C.14})$$

for some $Q_z \in SO(3)$, satisfying

$$Q_z e_1 = -\frac{z}{||z||}, \quad (\text{C.15})$$

as shown in Fig. C.3. Let us now verify that Γ qualifies as a path of interaction. It is clear from (C.14) that the contours $\Gamma(\bar{s}; \bar{s}, x, z)$ and $\Upsilon_{||z||}(s)$ are related through a rigid body transformation, such that

$$\Gamma(\bar{s}; \bar{s}, x, z) = x, \quad (\text{C.16a})$$

$$\Gamma(1; \bar{s}, x, z) - \Gamma(0; \bar{s}, x, z) = -z. \quad (\text{C.16b})$$

Fig. C.3 describes the properties mentioned above.

From (C.16a), it follows that Γ passes through the point x and from (C.16b) it is clear that z is the vector joining its endpoints. Moreover Q_z in (C.14) is made independent of \bar{s} to ensure that the condition $Q_z = I$ whenever $x - y = u - v$, in the definition of the path of interaction, is satisfied.

For any $\mathbf{x} \in \mathbb{R}^3$, let

$$\mathcal{C}_{\mathbf{x}} = \{\Gamma(\cdot; \bar{s}, \mathbf{x}, \mathbf{z}) : 0 \leq \bar{s} \leq 1, \mathbf{z} \in \mathbb{R}^3\}, \quad (\text{C.17})$$

denote the set of all paths of interaction that pass through \mathbf{x} . Now for every \mathbf{y} on a contour $\Gamma \in \mathcal{C}_{\mathbf{x}}$, let \mathbf{y}_{\perp} denote the projection of \mathbf{y} on the line joining the end points of the path. It is easy to see that $\mathbf{y}_{\perp}(s; \bar{s}, \mathbf{x}, \mathbf{z})$ is given by

$$\mathbf{y}_{\perp}(s; \bar{s}, \mathbf{x}, \mathbf{z}) = \mathbf{x} + \mathbf{Q}_{\mathbf{z}} [s\|\mathbf{z}\|\mathbf{e}_1 - \mathbf{r}_{\|\mathbf{z}\|}(\bar{s})]. \quad (\text{C.18})$$

Therefore, from (C.16a) we have

$$\mathbf{x}_{\perp}(\bar{s}, \mathbf{x}, \mathbf{z}) = \mathbf{y}_{\perp}(\bar{s}; \bar{s}, \mathbf{x}, \mathbf{z}) = \mathbf{x} + \mathbf{Q}_{\mathbf{z}} [\bar{s}\|\mathbf{z}\|\mathbf{e}_1 - \mathbf{r}_{\|\mathbf{z}\|}(\bar{s})]. \quad (\text{C.19})$$

Using (C.15) this simplifies to

$$\mathbf{x}_{\perp}(\bar{s}, \mathbf{x}, \mathbf{z}) = \mathbf{x} - \bar{s}\mathbf{z} - \mathbf{Q}_{\mathbf{z}}\mathbf{r}_{\|\mathbf{z}\|}(\bar{s}). \quad (\text{C.20})$$

We now generalize Noll's Lemma C.1 to arbitrary paths of interaction.

Lemma C.2 *For given paths of interaction on $\mathbb{R}^3 \times \mathbb{R}^3$, and under conditions 1-3 on $\mathbf{f}(\mathbf{v}, \mathbf{w})$ given at the start of this appendix, the following equation holds:*

$$\int_{\mathbf{y} \in \mathbb{R}^3} \mathbf{f}(\mathbf{x}, \mathbf{y}) d\mathbf{y} = \frac{1}{2} \operatorname{div}_{\mathbf{x}} \int_{\mathbf{z} \in \mathbb{R}^3} \left[\int_{\bar{s}=0}^1 \mathbf{f}(\mathbf{x}_{\perp} + \bar{s}\mathbf{z}, \mathbf{x}_{\perp} - (1-\bar{s})\mathbf{z}) \right] \otimes \mathbf{Q}_{\mathbf{z}} \mathbf{r}'_{\|\mathbf{z}\|}(\bar{s}) d\bar{s} d\mathbf{z}, \quad (\text{C.21})$$

where $\mathbf{x}_{\perp}(\bar{s}, \mathbf{x}, \mathbf{z})$ is given by (C.20).

Proof From (C.6) we have

$$\int_{\mathbf{y} \in \mathbb{R}^3} \mathbf{f}(\mathbf{x}, \mathbf{y}) d\mathbf{y} = \frac{1}{2} \int_{\mathbf{z} \in \mathbb{R}^3} [\mathbf{f}(\mathbf{x}, \mathbf{x} - \mathbf{z}) - \mathbf{f}(\mathbf{x} + \mathbf{z}, \mathbf{x})] d\mathbf{z}. \quad (\text{C.22})$$

Next, from the chain rule, we have

$$\nabla_{\mathbf{x}} \mathbf{f}(\mathbf{x}_{\perp} + \bar{s}\mathbf{z}, \mathbf{x}_{\perp} - (1-\bar{s})\mathbf{z}) = (\nabla_{\mathbf{v}} \mathbf{f} + \nabla_{\mathbf{w}} \mathbf{f}) \nabla_{\mathbf{x}} \mathbf{x}_{\perp}, \quad (\text{C.23})$$

where $\bar{s} \in \mathbb{R}$. From (C.20) we have $\nabla_{\mathbf{x}} \mathbf{x}_{\perp} = \mathbf{I}$. Therefore

$$\nabla_{\mathbf{x}} \mathbf{f}(\mathbf{x}_{\perp} + \bar{s}\mathbf{z}, \mathbf{x}_{\perp} - (1-\bar{s})\mathbf{z}) = \nabla_{\mathbf{v}} \mathbf{f} + \nabla_{\mathbf{w}} \mathbf{f}. \quad (\text{C.24})$$

Similarly,

$$\frac{d}{d\bar{s}} \mathbf{f}(\mathbf{x}_{\perp} + \bar{s}\mathbf{z}, \mathbf{x}_{\perp} - (1-\bar{s})\mathbf{z}) = (\nabla_{\mathbf{v}} \mathbf{f} + \nabla_{\mathbf{w}} \mathbf{f}) \left(\frac{d\mathbf{x}_{\perp}}{d\bar{s}} + \mathbf{z} \right). \quad (\text{C.25})$$

From (C.20), we have

$$\frac{d\mathbf{x}_{\perp}}{d\bar{s}} = -\mathbf{z} - \mathbf{Q}_{\mathbf{z}} \mathbf{r}'_{\|\mathbf{z}\|}(\bar{s}). \quad (\text{C.26})$$

Substituting (C.26) into (C.25), we have

$$\frac{d}{d\bar{s}} \mathbf{f}(\mathbf{x}_{\perp} + \bar{s}\mathbf{z}, \mathbf{x}_{\perp} - (1-\bar{s})\mathbf{z}) = -(\nabla_{\mathbf{v}} \mathbf{f} + \nabla_{\mathbf{w}} \mathbf{f}) \mathbf{Q}_{\mathbf{z}} \mathbf{r}'_{\|\mathbf{z}\|}(\bar{s}). \quad (\text{C.27})$$

Combining (C.24) and (C.27), we have

$$\frac{d}{d\bar{s}} \mathbf{f}(\mathbf{x}_{\perp} + \bar{s}\mathbf{z}, \mathbf{x}_{\perp} - (1-\bar{s})\mathbf{z}) = -[\nabla_{\mathbf{x}} \mathbf{f}(\mathbf{x}_{\perp} + \bar{s}\mathbf{z}, \mathbf{x}_{\perp} - (1-\bar{s})\mathbf{z})] \mathbf{Q}_{\mathbf{z}} \mathbf{r}'_{\|\mathbf{z}\|}(\bar{s}). \quad (\text{C.28})$$

Integrating both sides over the interval $\bar{s} \in [0, 1]$, we have

$$\begin{aligned} \mathbf{f}(\mathbf{x}_{\perp}(1, \mathbf{x}, \mathbf{z}) + \mathbf{z}, \mathbf{x}_{\perp}(1, \mathbf{x}, \mathbf{z})) - \mathbf{f}(\mathbf{x}_{\perp}(0, \mathbf{x}, \mathbf{z}), \mathbf{x}_{\perp}(0, \mathbf{x}, \mathbf{z}) - \mathbf{z}) = \\ - \int_{\bar{s}=0}^1 [\nabla_{\mathbf{x}} \mathbf{f}(\mathbf{x}_{\perp} + \bar{s}\mathbf{z}, \mathbf{x}_{\perp} - (1-\bar{s})\mathbf{z})] \mathbf{Q}_{\mathbf{z}} \mathbf{r}'_{\|\mathbf{z}\|}(\bar{s}) d\bar{s}. \end{aligned} \quad (\text{C.29})$$

Using (C.20), (C.15), and property (C.13b) of \mathcal{Y}_l , we have

$$\mathbf{x}_\perp(0, \mathbf{x}, \mathbf{z}) = \mathbf{x} \quad ; \quad \mathbf{x}_\perp(1, \mathbf{x}, \mathbf{z}) = \mathbf{x}. \quad (\text{C.30})$$

Substituting, (C.30) into (C.29) and using the identity (C.11), we have

$$\mathbf{f}(\mathbf{x} + \mathbf{z}, \mathbf{x}) - \mathbf{f}(\mathbf{x}, \mathbf{x} - \mathbf{z}) = -\operatorname{div}_{\mathbf{x}} \int_{\bar{s}=0}^1 \mathbf{f}(\mathbf{x}_\perp + \bar{s}\mathbf{z}, \mathbf{x}_\perp - (1-\bar{s})\mathbf{z}) \otimes \mathbf{Q}_z \mathcal{Y}'_{\|\mathbf{z}\|}(\bar{s}) d\bar{s}. \quad (\text{C.31})$$

If we now substitute the above equation into (C.22), the lemma is proved. \square

Lemma C.3 *Let $\Omega \subset \mathbb{R}^3$, with a piecewise smooth surface S . For given paths of interaction on $\mathbb{R}^3 \times \mathbb{R}^3$, and under the conditions 1-3 on $\mathbf{f}(\mathbf{v}, \mathbf{w})$ given at the start of this appendix, the following equation holds:*

$$\begin{aligned} & \int_{\mathbf{x} \in \Omega} \int_{\mathbf{y} \in \Omega^c} \mathbf{f}(\mathbf{x}, \mathbf{y}) d\mathbf{y} d\mathbf{x} \\ &= \frac{1}{2} \int_S \int_{\mathbf{z} \in \mathbb{R}^3} \int_{\bar{s}=0}^1 \mathbf{f}(\mathbf{x}_\perp + \bar{s}\mathbf{z}, \mathbf{x}_\perp - (1-\bar{s})\mathbf{z}) (\mathbf{Q}_z \mathcal{Y}'_{\|\mathbf{z}\|}(\bar{s}) \cdot \mathbf{n}) d\bar{s} d\mathbf{z} dS(\mathbf{x}), \end{aligned} \quad (\text{C.32})$$

where $\mathbf{x}_\perp(\bar{s}, \mathbf{x}, \mathbf{z})$ is given by (C.20) and $\Omega^c := \mathbb{R}^3 \setminus \Omega$.

Proof We immediately see that due to antisymmetry of $\mathbf{f}(\mathbf{v}, \mathbf{w})$, we have

$$\int_{\mathbf{x} \in \Omega} \int_{\mathbf{y} \in \Omega} \mathbf{f}(\mathbf{x}, \mathbf{y}) d\mathbf{y} d\mathbf{x} = \mathbf{0}. \quad (\text{C.33})$$

Therefore, we have

$$\int_{\mathbf{x} \in \Omega} \int_{\mathbf{y} \in \Omega^c} \mathbf{f}(\mathbf{x}, \mathbf{y}) d\mathbf{y} d\mathbf{x} = \int_{\mathbf{x} \in \Omega} \int_{\mathbf{y} \in \mathbb{R}^3} \mathbf{f}(\mathbf{x}, \mathbf{y}) d\mathbf{y} d\mathbf{x}. \quad (\text{C.34})$$

From Lemma C.2, we have

$$\int_{\mathbf{y} \in \mathbb{R}^3} \mathbf{f}(\mathbf{x}, \mathbf{y}) d\mathbf{y} = \operatorname{div}_{\mathbf{x}} \mathbf{g}(\mathbf{x}), \quad (\text{C.35})$$

with

$$\mathbf{g}(\mathbf{x}) = \frac{1}{2} \int_{\mathbf{z} \in \mathbb{R}^3} \int_{\bar{s}=0}^1 \mathbf{f}(\mathbf{x}_\perp + \bar{s}\mathbf{z}, \mathbf{x}_\perp - (1-\bar{s})\mathbf{z}) \otimes \mathbf{Q}_z \mathcal{Y}'_{\|\mathbf{z}\|}(\bar{s}) d\bar{s} d\mathbf{z}, \quad (\text{C.36})$$

where $\mathbf{x}_\perp(\bar{s}, \mathbf{x}, \mathbf{z})$ is given by (C.20). From the divergence theorem we have

$$\int_{\mathbf{x} \in \Omega} \operatorname{div}_{\mathbf{x}} \mathbf{g}(\mathbf{x}) d\mathbf{x} = \int_S \mathbf{g}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) dS(\mathbf{x}). \quad (\text{C.37})$$

Using (C.34)-(C.37), we obtain (C.32). \square

D Derivation of the Hardy stress from the Murdoch–Hardy procedure

The Hardy stress, $\tilde{\sigma}_{w,v}^H$, was derived in Section 4.2 subject to the condition:

$$\operatorname{div}_{\mathbf{x}} \tilde{\sigma}_{w,v}^H(\mathbf{x}, t) = \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \mathbf{f}_{\alpha\beta} w(\mathbf{x}_\alpha - \mathbf{x}). \quad (\text{D.1})$$

Recall that the expression for $\tilde{\sigma}_{w,v}^H$ given in (4.32) was obtained by using the generator function \mathbf{g}^H (see (4.29)), which is actually a distribution. Due to the inherent obstacles present in this procedure (see footnote 34 on page 47), we now derive the expression for the Hardy stress given in (4.32), by avoiding the use of the Dirac delta distribution.

Definition D.1 Consider the following definitions for the function $\eta(\mathbf{x})$ and associated functions taken from [18]:

1. Define a mollifier $\eta \in C^\infty(\mathbb{R}^3)$ by

$$\eta(\mathbf{r}) := \begin{cases} C \exp\left(\frac{1}{\|\mathbf{r}\|^2 - 1}\right) & \text{if } \|\mathbf{r}\| < 1, \\ 0 & \text{if } \|\mathbf{r}\| \geq 1, \end{cases} \quad (\text{D.2})$$

where the constant $C > 0$ is selected so that $\int_{\mathbb{R}^3} \eta \, d\mathbf{r} = 1$.

2. For each $\epsilon > 0$, set

$$\eta_\epsilon(\mathbf{r}) := \frac{1}{\epsilon^3} \eta\left(\frac{\mathbf{r}}{\epsilon}\right). \quad (\text{D.3})$$

The family of functions η_ϵ are C^∞ and satisfy

$$\int_{\mathbb{R}^3} \eta_\epsilon \, d\mathbf{r} = 1. \quad (\text{D.4})$$

The support of η_ϵ is contained in a ball of radius ϵ centered at $\mathbf{0}$.

3. If the function $h : \mathbb{R}^3 \rightarrow \mathbb{R}$ is locally integrable, define its mollification $h_\epsilon(\mathbf{r})$ as

$$h_\epsilon(\mathbf{r}) := \int_{\mathbb{R}^3} \eta_\epsilon(\mathbf{r} - \mathbf{y}) h(\mathbf{y}) \, d\mathbf{y}. \quad (\text{D.5})$$

We use the following property of mollifiers in our derivation:

$$h_\epsilon \rightarrow h \text{ almost everywhere as } \epsilon \rightarrow 0. \quad (\text{D.6})$$

For the proof of (D.6), refer to [18]. Equation (D.1) can be rewritten as

$$\operatorname{div}_{\mathbf{x}} \tilde{\boldsymbol{\sigma}}_{w,v}^H(\mathbf{x}, t) = \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \mathbf{f}_{\alpha\beta} \sqrt{w(\mathbf{x}_\alpha - \mathbf{x}) w(\mathbf{x}_\alpha - \mathbf{x})}, \quad (\text{D.7})$$

since $w(\mathbf{x}_\alpha - \mathbf{x}) > 0$. Now, since $\sqrt{w(\mathbf{x}_\beta - \mathbf{y})}$ is locally integrable, using property (D.6), we have

$$\sqrt{w(\mathbf{x}_\alpha - \mathbf{x})} = \lim_{\epsilon \rightarrow 0} \int_{\mathbb{R}^3} \sqrt{w(\mathbf{x}_\beta - \mathbf{y})} \eta_\epsilon(\mathbf{x}_\beta - \mathbf{x}_\alpha + \mathbf{x} - \mathbf{y}) \, d\mathbf{y}, \quad (\text{D.8})$$

where referring to (D.5), $h(\mathbf{y}) = \sqrt{w(\mathbf{x}_\beta - \mathbf{y})}$ and $\mathbf{r} = \mathbf{x}_\beta - \mathbf{x}_\alpha + \mathbf{x}$. Using (D.8), (D.7) can be rewritten as

$$\begin{aligned} \operatorname{div}_{\mathbf{x}} \tilde{\boldsymbol{\sigma}}_{w,v}^H(\mathbf{x}, t) &= \lim_{\epsilon \rightarrow 0} \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \int_{\mathbb{R}^3} \mathbf{f}_{\alpha\beta} \sqrt{w(\mathbf{x}_\alpha - \mathbf{x}) w(\mathbf{x}_\beta - \mathbf{y})} \eta_\epsilon(\mathbf{x}_\beta - \mathbf{x}_\alpha + \mathbf{x} - \mathbf{y}) \, d\mathbf{y} \\ &=: \lim_{\epsilon \rightarrow 0} \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \int_{\mathbb{R}^3} \mathbf{g}_{\alpha\beta}^\epsilon(\mathbf{x}, \mathbf{y}, t) \, d\mathbf{y}. \end{aligned} \quad (\text{D.9})$$

Now, note that the function $\mathbf{g}_{\alpha\beta}^\epsilon$ is anti-symmetric with respect to its arguments, \mathbf{x} and \mathbf{y} , for each $\epsilon > 0$ and satisfies all the necessary conditions for the application of Lemma C.1 in Appendix C. Therefore, from C.3 it follows that

$$\begin{aligned} \operatorname{div}_{\mathbf{x}} \tilde{\boldsymbol{\sigma}}_{w,v}^H(\mathbf{x}, t) &= \lim_{\epsilon \rightarrow 0} \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \left(-\frac{1}{2} \operatorname{div}_{\mathbf{x}} \int_{\mathbb{R}^3} \left[\int_{s=0}^1 \mathbf{f}_{\alpha\beta} \sqrt{w(\mathbf{x}_\alpha - \mathbf{x} - s\mathbf{z}) w(\mathbf{x}_\beta - \mathbf{x} + (1-s)\mathbf{z})} \right. \right. \\ &\quad \left. \left. \times \eta_\epsilon(\mathbf{x}_\beta - \mathbf{x}_\alpha + \mathbf{z}) \, ds \right] \otimes \mathbf{z} \, d\mathbf{z} \right) \\ &= \lim_{\epsilon \rightarrow 0} \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \left(-\frac{1}{2} \int_{\mathbb{R}^3} \int_{s=0}^1 \nabla_{\mathbf{x}} \left(\mathbf{f}_{\alpha\beta} \sqrt{w(\mathbf{x}_\alpha - \mathbf{x} - s\mathbf{z}) w(\mathbf{x}_\beta - \mathbf{x} + (1-s)\mathbf{z})} \right) \right. \\ &\quad \left. \times \eta_\epsilon(\mathbf{x}_\beta - \mathbf{x}_\alpha + \mathbf{z}) \mathbf{z} \, ds \, d\mathbf{z} \right), \end{aligned} \quad (\text{D.10})$$

where in the last equality we have used the identity given in (C.11) and the fact that η_ϵ is independent of \mathbf{x} in the above equation. Since w is positive, and if w is a continuously differentiable function, it follows that

$$\nabla_{\mathbf{x}} \left(\mathbf{f}_{\alpha\beta} \sqrt{w(\mathbf{x}_\alpha - \mathbf{x} - s\mathbf{z})w(\mathbf{x}_\beta - \mathbf{x} + (1-s)\mathbf{z})} \right) \mathbf{z} \quad (\text{D.11})$$

is a locally integrable function of \mathbf{z} . Therefore, by property D.6 and using the property, $\eta_\epsilon(\mathbf{r}) = \eta_\epsilon(-\mathbf{r})$, we have

$$\begin{aligned} \text{div}_{\mathbf{x}} \tilde{\boldsymbol{\sigma}}_{w,v}^H(\mathbf{x}, t) &= -\frac{1}{2} \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \int_{s=0}^1 \nabla_{\mathbf{x}} [\mathbf{f}_{\alpha\beta} w(\mathbf{x}_\alpha - \mathbf{x} - s(\mathbf{x}_\alpha - \mathbf{x}_\beta))] (\mathbf{x}_\alpha - \mathbf{x}_\beta) ds \\ &= \text{div}_{\mathbf{x}} \left[\frac{1}{2} \sum_{\substack{\alpha, \beta \\ \alpha \neq \beta}} \int_{s=0}^1 -[\mathbf{f}_{\alpha\beta} w((1-s)\mathbf{x}_\alpha + s\mathbf{x}_\beta - \mathbf{x}) \otimes (\mathbf{x}_\alpha - \mathbf{x}_\beta)] ds \right], \end{aligned} \quad (\text{D.12})$$

where referring to (D.5), $\mathbf{r} = \mathbf{x}_\alpha - \mathbf{x}_\beta$, $\mathbf{y} = \mathbf{z}$, and $h(\mathbf{z})$ is given in (D.11). Comparing both sides of (D.12), we arrive at the expression for the Hardy stress tensor given in (4.32).

Acknowledgements The authors would like to thank Dr. Marcel Arndt for assisting with the translation of Noll's paper [48]. In addition, the authors would like to thank Roger Fosdick, Richard D. James and Ryan S. Elliott for many stimulating conversations regarding various aspects of this paper and, in particular, the issue of the extension of potential energy functions. Additional helpful conversations with Ronald E. Miller, Eliot Fried, Richard B. Lehoucq and Andrew L. Ruina are gratefully acknowledged.

References

1. Arnold, V.I.: Mathematical methods of classical mechanics. Springer (1989)
2. Belytchko, T., Krongauz, Y., Organ, D., Fleming, M., Krysl, P.: Meshless methods: An overview and recent developments. *Computer Methods in Applied Mechanics and Engineering* **139**, 3–47 (1996)
3. Blumenthal, L.M.: Theory and applications of distance geometry. Second edn. Chelsea, New York (1970)
4. Braun, W.: Distance geometry and related methods for protein structure determination from nmr data. *Quarterly Reviews of Biophysics* **19**(3/4), 115–157 (1987)
5. Cauchy, A.: Exercices de mathématique, vol. 3, chap. Sur l'équilibre et le mouvement d'un système de points matériels sollicités par des forces d'attraction ou de répulsion mutuelle, pp. 227–252. Chez de Bure Frères, Paris (1928)
6. Cauchy, A.: Exercices de mathématique, vol. 3, chap. De la pression ou tension dans un système de points matériels, pp. 253–277. Chez de Bure Frères, Paris (1928)
7. Cheung, K.S., Yip, S.: Atomic level stress in an inhomogeneous system. *Journal of Applied Physics* **70**(10), 5688–5690 (1991)
8. Clausius, R.: On a mechanical theorem applicable to heat. *Philosophical Magazine* **40**, 122–127 (1870)
9. Cormier, J., Rickman, J.M., Delph, T.J.: Stress calculation in atomistic simulations of perfect and imperfect solids. *Journal of Applied Physics* **89**(1), 99–104 (2001)
10. Crippen, G.M.: A novel approach to calculation of conformation: distance geometry. *Journal of Computational Physics* **24**, 96–107 (1977)
11. Crippen, G.M., Havel, T.F.: Distance geometry and molecular conformation. Wiley, New York (1988)
12. Daw, M., Baskes, M.: Embedded-atom method: Derivation and application to impurities, surfaces, and other defects in metals. *Phys. Rev. B* **29**, 6443–6453 (1984)
13. Delph, T.J.: Conservation laws for multibody interatomic potentials. *Modeling and Simulation in Material Science and Engineering* **13**, 585–594 (2005)
14. Dummit, D.S., Foote, R.M.: Abstract algebra. John Wiley and Sons (2004)
15. E, W., Engquist, B., Li, X.T., Ren, W.Q., Vanden-Eijnden, E.: Heterogeneous multiscale methods: A review. *Comm. Comp. Phys.* **2**(3), 367–450 (2007)

16. Ercolessi, F., Adams, J.B.: Interatomic potentials from first-principles calculations - the force-matching method. *Europhysics Letters* **26**(8), 583–588 (1994)
17. Evans, D.J., Morriss, G.P.: *Statistical Mechanics of Nonequilibrium Liquids*. Academic Press, London (1990)
18. Evans, L.C.: *Partial differential equations*, vol. 19. American Mathematical Society (2002)
19. Folland, G.B.: *Real Analysis: Modern Techniques and Their Applications*. John Wiley and Sons (1999)
20. Fosdick, R.L.: (2009). Private communication
21. Fosdick, R.L., Virga, E.G.: A variational proof of the stress theorem of cauchy. *Archive for Rational Mechanics and Analysis* **105**, 95–103 (1989)
22. Goldstein, H.: *Classical mechanics*. Addison-Wesley (1965)
23. Hardy, R.J.: Formulas for determining local properties in molecular dynamics simulation: Shock waves. *Journal of Chemical Physics* **76**(01), 622–628 (1982)
24. Hardy, R.J., Root, S., Swanson, D.R.: Continuum properties from molecular simulations. *AIP Conference Proceedings* **620**, 363–366 (2002)
25. Havel, T.F., Kuntz, I.D., Crippen, G.M.: The theory and practice of distance geometry. *Bulletin of Mathematical Biology* **45**, 665–720 (1983)
26. Huang, K.: *Statistical Mechanics*. John Wiley and Sons (1963)
27. Irving, J.H., Kirkwood, G.: The statistical mechanics theory of transport processes. iv. the equations of hydrodynamics. *Journal of Chemical Physics* **18**(06), 817–829 (1950)
28. James, R.D.: Objective structures. *Journal of Mechanics and Physics of Solids* **54**, 2354–2390 (2006)
29. Kirkwood, G.: The statistical mechanical theory of transport processes. *Journal of Chemical Physics* **14**, 180–201 (1946)
30. Klapper, M.H., DeBrot, D.: Use of caley-menger determinants in the calculation of molecular structures. *Journal of Computational Physics* **37**, 56–69 (1980)
31. Knap, J., Ortiz, M.: An analysis of the quasicontinuum method. *J. Mech. Phys. Sol.* **49**(9), 1899–1923 (2001)
32. Lanczos, C.: *The variational principles of mechanics*. University of Toronto Press, Toronto (1986)
33. Lehoucq, R.B., Silling, S.A.: Force flux and the peridynamic stress tensor. *Journal of Mechanics and Physics of Solids* **56**, 1566–1577 (2008)
34. Lekhnitskii, S.G.: *Theory of elasticity of an anisotropic elastic body*. Holden-Day Series in Mathematical Physics (1963)
35. Love, A.E.H.: *A treatise on the mathematical theory of elasticity*. Cambridge University Press, Cambridge (1927)
36. Lutsko, J.F.: Stress and elastic constants in anisotropic solids: Molecular dynamics techniques. *Journal of Applied Physics* **64**(3), 1152–1154 (1988)
37. Malvern, L.E.: *Introduction to the mechanics of a continuous medium*. Prentice-Hall, Upper Saddle River (1969)
38. Marc, G., McMillan, W.G.: The virial theorem. *Advances in Chemical Physics* **58**, 209–361 (1985)
39. Martin, J.W.: Many body forces in metals and the Brugger elastic constants. *Journal of Physics C: Solid State Physics* **8**, 2837–2857 (1975)
40. Maxwell, J.C.: On reciprocal figures, frames and diagrams of forces. *Translations of the Royal Society of Edinburgh* **XXVI**, 1–43 (1870)
41. Maxwell, J.C.: Van der waals on the continuity of the gaseous and liquid states. *Nature* **10**, 477–80 (1874)
42. Morante, S., Rossi, G.C.: The stress tensor of a molecular system: An exercise in statistical mechanics. *The Journal of Chemical Physics* **125**, 034,101–1–034,101–11 (2006)
43. Murdoch, A.I.: The motivation of continuum concepts and relations from discrete considerations. *Journal of Applied Mathematics and Mechanics* **36**, 163–187 (1982)
44. Murdoch, A.I.: On the microscopic interpretation of stress and couple stress. *Journal of Elasticity* **71**, 105–131 (2003)
45. Murdoch, A.I.: A critique of atomistic definitions of the stress tensor. *Journal of Elasticity* **88**, 113–140 (2007)
46. Murdoch, A.I., Bedeaux, D.: On the physical interpretation of fields in continuum mechanics. *International Journal of Engineering Science* **31**(10), 1345–1373 (1993)
47. Murdoch, A.I., Bedeaux, D.: Continuum equations of balance via weighted averages of microscopic quantities. *Proceedings of the Royal Society of London A* **445**, 157–179 (1994)
48. Noll, W.: Die herleitung der grundgleichungen der thermomechanik der kontinua aus der statistischen mechanik. *Journal of Rational Mechanics and Analysis* **4**, 627–646 (1955)

49. Pitteri, M.: On a statistical-kinetic model for generalized continua. *Archive for Rational Mechanics and Analysis* **111**, 99–120 (1990)
50. Porta, J.M., Ros, L., Thomas, F., Torras, C.: A branch-and-prune solver for distance constraints. *IEEE Transactions on Robotics* **21**(2), 176–187 (2005)
51. Rudd, R.E., Broughton, J.Q.: Concurrent coupling of length scales in solid state systems. *Physical Status Solidi B* **217**, 251–291 (2000)
52. Schofield, P., Henderson, J.R.: Statistical mechanics of inhomogeneous fluids. *Proceedings of the Royal Society of London A* **379**, 232–246 (1982)
53. Shenoy, V.B., Miller, R., Tadmor, E., Rodney, D., Phillips, R., Ortiz, M.: An adaptive methodology for atomic scale mechanics: The quasicontinuum method. *J. Mech. Phys. Sol.* **47**, 611–642 (1999)
54. Shilkrot, L.E., Miller, R.E., Curtin, W.: Multiscale plasticity modeling: Coupled atomistic and discrete dislocation mechanics. *J. Mech. Phys. Sol.* **52**(4), 755–787 (2004)
55. Silling, S.A.: Reformulation of elasticity theory for discontinuities and long-range forces. *Journal of Mechanics and Physics of Solids* **48**, 175–209 (2000)
56. Sippl, M.J., Scheraga, H.A.: Caley-menger coordinates. *Proceedings of the National Academy of Sciences* **83**, 2283–2287 (1986)
57. Stillinger, H., Weber, A.: Computer simulation of local order in condensed phases of silicon. *Physical Review B* **31**, 5262–5271 (1985)
58. Tadmor, E.B., Miller, R.E.: A unified framework and performance benchmark of fourteen multiscale atomistic/continuum coupling methods. *Modeling and Simulation in Material Science and Engineering* **17**, 053,001 (2009)
59. Tadmor, E.B., Ortiz, M., Phillips, R.: Quasicontinuum analysis of defects in solids. *Philosophical Magazine A* **73**(6), 1529–1563 (1996)
60. Tersoff, J.: Empirical interatomic potential for silicon with improved elastic properties. *Phys. Rev. B* **38**(14), 9902–9905 (1988)
61. Truesdell, C.: *Essays in the history of mechanics*. Springer (1968)
62. Truesdell, C., Noll, W.: *The non-linear field theories of mechanics*. Springer (1965)
63. Tsai, D.H.: The virial theorem and stress calculation in molecular dynamics. *Journal of Chemical Physics* **70**(03), 1375–1382 (1979)
64. Wajnryb, E., Altenberger, A.R., Dahler, J.S.: Uniqueness of the microscopic stress tensor. *Journal of Chemical Physics* **103**(22), 9782–9787 (1995)
65. Weiner, J.H.: *Statistical Mechanics of Elasticity*. Second edn. Dover, Mineola (2002)
66. Xiao, S., Belytschko, T.: A bridging domain method for coupling continua with molecular dynamics. *Computer Methods in Applied Mechanics and Engineering* **193**, 1645–69 (2004)
67. Zhou, M.: A new look at the atomic level virial stress: On continuum-molecular system equivalence. *Proceedings of the Royal Society of London A* **459**, 2347–2392 (2003)
68. Zimmerman, J.A., III, E.B.W., Hoyt, J.J., Jones, R.E., Klein, P.A., Bammann, D.J.: Calculation of stress in atomistic simulation. *Modeling and Simulation in Material Science and Engineering* **12**, S319–S332 (2004)

