The linking of solid mechanics with thermodynamics through the mathematics of differentials

K. Davey
Department of Mechanical, Aerospace and Civil Engineering, Manchester The University of Manchester, UK E-mail: keith.davey@manchester.ac.uk

Abstract Engineers are familiar with the concept of differentials but tend to think of these in an informal sense, that is, as infinitesimally small quantities. It is argued in this paper that differentials should be presented to engineering students in a more formal way. Differentials can be formalised by defining then as functions on tangent spaces. This enables their numerical evaluation and makes clear the nature of their behaviour when restricted to tangent subspaces. It is shown that through a proper understanding of differentials the physics of many processes can be more fully appreciated. Inexact and exact differentials and their link to path dependency are examined in the context of state spaces. The work differential is considered and it is shown how knowledge of the state space combined with a simple test for path dependency provides real insight. Differentials allow the unification of solution methodologies for thermodynamics and solid mechanics. Work differentials and state spaces associated with elasticity and plasticity theory are examined in the paper. Reversibility in a one-dimensional state space provides for the explicit formulation of work and is applicable to elasticity. The use of monotonicity in replace of reversibility is examined for plasticity theory. The idea of conservative and non-conservative forces in mechanics is also considered in the context of state spaces and differentials. Examples are presented in the paper to illustrate the wide applicability of differentials across subject boundaries.

Key words differentials; mechanics; thermodynamics

Introduction

Engineering students typically follow the teaching of differentials through standard texts such as Kreyszig [1] and Stroud [2]. The formula $\mathrm{d}f = f'(x)\mathrm{d}x$ is often presented to define $\mathrm{d}f$, but how satisfactory is this when $\mathrm{d}x$ is undefined? Typically a graph of the type shown in Fig. 1 is presented to give a pictorial understanding of $\mathrm{d}x$. However, it is clear from Fig. 1 that $\mathrm{d}x$ is not properly defined. In fact, the common engineers' view that differentials are infinitesimal is not supported by this figure, although clearly $\mathrm{d}x$ could be made arbitrarily small. The idea that differentials are infinitesimal quantities possibly stems from the definition of f' as a limiting process.

$$f'(x) = \frac{\mathrm{d}f}{\mathrm{d}x}(x) = \lim_{\varepsilon \to 0}\left(\frac{f(x+\varepsilon) - f(x)}{(x+\varepsilon) - x}\right) = \lim_{\varepsilon \to 0}\left(\frac{f(x+\varepsilon) - f(x)}{\varepsilon}\right) \tag{1}$$

It is natural to associate $\mathrm{d}f(x)$ with $f(x + \varepsilon) - f(x)$ and $\mathrm{d}x$ with ε, where ε is small. In addition, the use of $\mathrm{d}x$ as a measure quantity for integration reinforces the idea that $\mathrm{d}x$ is infinitesimal. It is evident that these are loose definitions and do not properly define $\mathrm{d}f$ and $\mathrm{d}x$ in a rigorous mathematical sense. The engineers' viewpoint that $\mathrm{d}f$ is an infinitesimal quantity identified with the change in f has merit. It

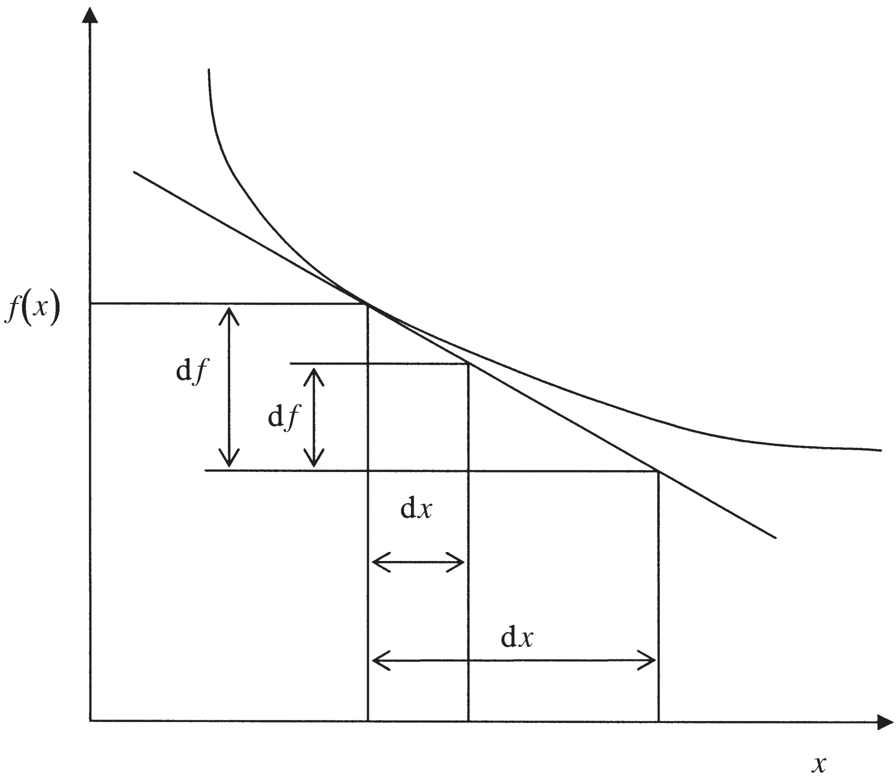

Fig. 1 *Definition of a differential.*

facilitates a broad grasp of the underpinning physics involved. For example, the first law of thermodynamics, written in differential form, is $\mathrm{d}U = \tilde{\mathrm{d}}W + \tilde{\mathrm{d}}Q$, which is immediately understood in this context. Here $\tilde{\mathrm{d}}$ refers to an inexact differential and note the sign convention for heat and work, which are positive if acting on a system. The symbol d is reserved for exact differentials and it will be made clear in the paper that d can be considered as an operator. The quantities $\mathrm{d}U$, $\tilde{\mathrm{d}}W$ and $\tilde{\mathrm{d}}Q$ indicate infinitesimal changes in internal energy, work and heat, respectively. The idea here is that if you do work and/or apply heat to a system, then its internal energy will increase. However, this relationship embodies many concepts that are hidden by the idea of infinitesimal changes. The differential $\mathrm{d}U$ is not referring to spatial changes but to change in U with respect to thermodynamic (state) variables such as temperature, T, and pressure, P. This raises the point that the space on which $\mathrm{d}U$ is defined needs to be specified. It may not be possible to relate W and Q to state variables but it is always possible for reversible (and some non-reversible) changes, for $\tilde{\mathrm{d}}W$ and $\tilde{\mathrm{d}}Q$ to be related to differentials of these variables. This raises a couple of concepts: first, $\tilde{\mathrm{d}}W$ and $\tilde{\mathrm{d}}Q$ are possibly inexact differentials; and second, their formal specification appears to require some form of relationship between exact differentials. The differential df specified above is exact by the very virtue that $f(x)$ is defined. It is not difficult to see why engineering students have difficulty understanding what $\tilde{\mathrm{d}}W$ and $\tilde{\mathrm{d}}Q$ signify beyond the hazy concept of infinitesimal change.

It is well appreciated by engineers that contour integration and differentials are closely related. The concepts of state variables and functions are closely related to the concept of path independence, which in turn depends on exact differentials. The fact that U can be explicitly related to thermodynamic variables immediately relays the fact that dU is an exact differential.

Differentials also appear in mechanics and appear closely related to the first law of thermodynamics. The force, F, acting on a mass, m, moving in a one-dimensional space is related to acceleration via the identity $F = m\mathrm{d}v/\mathrm{d}t$ or $F\mathrm{d}x = m(\mathrm{d}v/\mathrm{d}x)(\mathrm{d}x/\mathrm{d}t)\mathrm{d}x = (\mathrm{d}[mv^2/2]/\mathrm{d}x)\mathrm{d}x$, or equivalently $\tilde{\mathrm{d}}W = \mathrm{d}K$, where $\tilde{\mathrm{d}}W = F\mathrm{d}x$ and kinetic energy $K = mv^2/2$. Note the implicit assumption that velocity is a function of x here and consequently dv is an exact differential. The relationship $\tilde{\mathrm{d}}W = \mathrm{d}K$ appears to be nothing more than the first law of thermodynamics with $\mathrm{d}Q = 0$ and $U = K$. However, unlike the first law, these differentials do not refer to changes in state variables. Moreover, heat loss could well be involved here if friction forces are present. This observation reinforces the point that the space on which the differential is formed needs to be carefully stipulated. These concepts are explored in greater depth below.

In the next section, the idea is introduced of Cartesian product spaces, on which scalar and vector fields are defined. Differential forms are presented as functions whose domain is a tangent space. This concept is familiar to mathematicians working in the area of differential geometry. However, the concept is not common to engineering mathematics texts and consequently not considered on many engineering courses in universities.

In the section that follows, the concepts of exact and inexact differentials and contour integrals are discussed. The ideas of differentials are then discussed in the context of thermodynamic state spaces. This is followed by a look at elasticity and plasticity theory, and in particular on the definition of state spaces. Differentials for mechanics are then discussed, and simple problems are presented to illustate energy transport. Also explored in the section on mechanics is the idea of reversibility in a Cartesian space. In each section, simple problems are considered to reinforce the understanding of the concepts discussed. In the Discussion section, the educational issues surrounding the introduction of the ideas presented here are considered. It is argued that only through a dedicated course could the ideas be presented in the context of fluid mechanics, thermodynamics, solid mechanics and machines. Although the mathematics of differentials can be presented on mathematics courses, it is only through their application in spaces relevant to engineering that better understanding will be achieved.

Tangent spaces and differential forms

Consider the scalar function ϕ being specified at a point $\mathbf{p}$ on a Cartesian product space $\Re^n = \Re \times \Re \times \ldots \times \Re$, where $\mathbf{p}$ is an n-tuple of the form $(p_1, p_2, \ldots, p_n)$. Non-dimensional coordinate functions $s_1 \cdots s_n$ are specified on $\Re^n$ such that $s_i(\mathbf{p}) =$

p_i. The scalar function ϕ is defined as a composite function $\phi(s_1, s_2, \ldots, s_n)$. Thus $\phi:\Re^n \rightarrow \Re$, $s_i:\Re^n \rightarrow \Re$ (or $\boldsymbol{s}:\Re^n \rightarrow \Re^n$) and $\phi \circ \boldsymbol{s}:\Re^n \rightarrow \Re$, where the symbol $\circ$ denotes function composition.

The coordinate function is labelled s_i here rather than x_i to allow for generalisation where s_i may well be thermodynamic variables but alternatively spatial variables. These variables are assumed to be non-dimensional, which can always be arranged, but this restriction is revisited in the next section. It is important to consider whether $\phi(s_1, s_2, \ldots, s_n)$ is differentiable with respect to each variable s_i. In general, differentiability will be assumed, although it is possible to relax this requirement on certain contours in $\Re^n$, possibly associated with phase change.

For brevity, the notation $\phi_{\boldsymbol{p}} = \phi(p_1, p_2, \ldots, p_n) = \phi(s_1(\boldsymbol{p}), s_2(\boldsymbol{p}), \ldots, s_n(\boldsymbol{p})) = \phi \circ \boldsymbol{s}(\boldsymbol{p})$ is to be adopted and, where convenient, $\boldsymbol{p}$ will be dropped also. Consider also the definition of a coordinate frame in $\Re^n$, that is, let $\underline{\mathrm{e}}_1 = (1, 0, \ldots, 0)$ $\underline{\mathrm{e}}_2 = (0, 1, 0, \ldots, 0), \ldots, \underline{\mathrm{e}}_n = (0, 0, \ldots, 0, 1)$, where it is evident that any point $\boldsymbol{p}$ can be represented by $\boldsymbol{p} = \Sigma_{i=1}^n p_i \underline{\mathrm{e}}_i$. There are, of course, other possible coordinate frames for $\Re^n$. Also, it is useful to define a coordinate frame at each point $\boldsymbol{p}$ in $\Re^n$ denoted $\{\underline{\mathrm{E}}_1^{\boldsymbol{p}}, \underline{\mathrm{E}}_2^{\boldsymbol{p}}, \ldots, \underline{\mathrm{E}}_n^{\boldsymbol{p}}\}$ where $\underline{\mathrm{E}}_i^{\boldsymbol{p}}$ are independent orthogonal unit vectors. Each $\underline{\mathrm{E}}_i^{\boldsymbol{p}}$ can be related to the canonical frame $\{\underline{\mathrm{e}}_1, \underline{\mathrm{e}}_2, \ldots, \underline{\mathrm{e}}_n\}$ via expression of the form $\underline{\mathrm{E}}_i^{\boldsymbol{p}} = \Sigma_{j=1}^n \mathrm{q}_{\boldsymbol{p}}^{ij} \underline{\mathrm{e}}_j$ or, more formally, $\underline{\mathrm{E}}_i \circ \boldsymbol{s} = \Sigma_{j=1}^n \underline{\mathrm{e}}_j q^{ij} \circ \boldsymbol{s}$, where $q^{ij} \circ \boldsymbol{s}$ are assumed to be differentiable functions of s_i, apart possibly at certain contours in $\Re^n$. The square matrix $\boldsymbol{Q}^{\boldsymbol{p}}$, whose coefficients are $q_{\boldsymbol{p}}^{ij}$, is orthogonal and hence invertible. Each $\underline{\mathrm{E}}_i \circ \boldsymbol{s}$ is a vector field and two equivalent methods of defining an arbitrary regular vector field are available at this point:

$$\underline{\mathrm{v}} = \sum_{i=1}^n v_i \underline{\mathrm{e}}_i \quad \text{and} \quad \underline{\mathrm{v}} = \sum_{i=1}^n V_i \underline{\mathrm{E}}_i$$

where v_i and V_i are differentiable functions of s_j and are related through the coefficients q^{ij}. The space $T^{\boldsymbol{p}}$ defined by $T^{\boldsymbol{p}} = \mathrm{span}\,\{\underline{\mathrm{E}}_1^{\boldsymbol{p}}, \underline{\mathrm{E}}_2^{\boldsymbol{p}}, \ldots, \underline{\mathrm{E}}_{\mathrm{n}}^{\boldsymbol{p}}\}$ is called a tangent space at point $\boldsymbol{p}$ along with any subspace of this space.

A differential form (1-form) on $T^{\boldsymbol{p}}$ is defined to be a linear function φ: $T^{\boldsymbol{p}} \rightarrow \Re$, that is, $\varphi(\alpha\,\underline{\mathrm{v}}_{\boldsymbol{p}} + \beta\,\underline{\mathrm{w}}_{\boldsymbol{p}}) = \alpha\varphi(\underline{\mathrm{v}}_{\boldsymbol{p}}) + \beta\varphi(\underline{\mathrm{w}}_{\boldsymbol{p}})$, where $\alpha, \beta \in \Re$. One possible approach is to associate each φ with some $\underline{\mathrm{w}}_{\boldsymbol{p}}$ belonging to $T^{\boldsymbol{p}}$ via

$$\varphi\left(\underline{\mathrm{v}}_{\boldsymbol{p}}\right) = \left\langle \underline{\mathrm{w}}_{\boldsymbol{p}}, \underline{\mathrm{v}}_{\boldsymbol{p}} \right\rangle \tag{2}$$

where $\langle \bullet, \bullet \rangle$ denotes an inner product. This expression allows for the definition $\mathrm{d}s_i(\bullet) = \langle \underline{\mathrm{e}}_i, \bullet \rangle$, where $\mathrm{d}s_i(\underline{\mathrm{v}}) = \langle \underline{\mathrm{e}}_i, \Sigma_{j=1}^n \underline{\mathrm{e}}_j v_j \rangle = \Sigma_{j=1}^n v_j \langle \underline{\mathrm{e}}_i, \underline{\mathrm{e}}_j \rangle = \Sigma_{j=1}^n v_j \delta_{ij} = v_i$. Each differential form is associated with some $\underline{\mathrm{w}}_{\boldsymbol{p}}$ on $T^{\boldsymbol{p}}$ and in particular $\mathrm{d}s_i$ is associated with $\underline{\mathrm{e}}_i$. Note also that $\underline{\mathrm{w}} = \Sigma_{i=1}^n \underline{\mathrm{e}}_i w_i$, so $\varphi(\underline{\mathrm{v}}) = \langle \underline{\mathrm{w}}, \underline{\mathrm{v}} \rangle = \langle \Sigma_{i=1}^n \underline{\mathrm{e}}_i w_i, \underline{\mathrm{v}} \rangle = \Sigma_{i=1}^n w_i \langle \underline{\mathrm{e}}_i, \underline{\mathrm{v}} \rangle = \Sigma_{i=1}^n w_i \mathrm{d}s_i(\underline{\mathrm{v}})$. Thus, any differential form on $T^{\boldsymbol{p}}$ can be represented by a linear combination of $\mathrm{d}s_i$, that is, $\{\mathrm{d}s_i\}$ is a basis for forms in much the same way as $\{\underline{\mathrm{e}}_i\}$ is for vectors. Note also that given a vector $\Sigma_{i=1}^n w_i \underline{\mathrm{e}}_i$, the associated differential form is simply $\Sigma_{i=1}^n w_i \mathrm{d}s_i$. For example, the expression

$$\mathrm{d}\phi = \sum_{i=1}^n \frac{\partial \phi}{\partial s_i} \mathrm{d}s_i \tag{3}$$

is immediately associated with the gradient vector $\nabla\phi = \Sigma_{i=1}^n (\partial\phi/\partial s_i)\underline{\mathrm{e}}_i$.

The symbol d in equation (3) can be considered as an operator (exterior derivative) of ϕ. Higher-order differential forms can be obtained with the exterior derivative but these are not considered further here [3, 4].

Curves and exact and inexact differentials

A curve $\boldsymbol{\beta}$ in $\Re^n$ is defined by the mapping $\boldsymbol{\beta}$: $[0, \ell] \rightarrow \Re^n$, where $[0, \ell]$ is a closed interval in $\Re$. Attention is restricted here to smooth curves or possibly piecewise smooth curves. Consider then $s \in [0, \ell]$ and the differentiation of $\boldsymbol{\beta}$ with respect to s. It is convenient for the variable s to be the distance measured along the curve because in this case the derivative $\boldsymbol{\beta}' = \mathrm{d}\boldsymbol{\beta}/\mathrm{d}s = \underline{\mathrm{T}}$, where $\underline{\mathrm{T}}$ is a unit tangent vector. The concept of distance along a curve is obvious if s_i are simply spatial distances of the same dimension. However, in the case of thermodynamic variables of different dimension, the concept of length in a dimensional form of the spaces defined in the section above on tangent spaces is meaningless. The restriction imposed in that section on insisting that each s_i is non-dimensional has the merit that the concepts apply to both state and spatial spaces. The problem with dimension can often be glossed over on engineering courses, as insufficient regard is often given to the space on which calculations are to be performed. Certain concepts do not require length to be defined and in this case the minor inconvenience of non-dimensionalisation is not required.

It is clear that at a point $\boldsymbol{p}$ in $\Re^n$ the tangent vector $\underline{\mathrm{T}}$ belongs to $T^{\boldsymbol{p}}$. Observe that

$$\mathrm{d}\phi(T) = \sum_{i=1}^{n} \frac{\partial \phi}{\partial s_i} \mathrm{d}s_i(T) = \sum_{i=1}^{n} \frac{\partial \phi}{\partial s_i} \frac{\mathrm{d}s_i}{\mathrm{d}s} = \frac{\mathrm{d}\phi}{\mathrm{d}s} \tag{4}$$

and consequently $\mathrm{d}\phi = (\mathrm{d}\phi/\mathrm{d}s)\mathrm{d}s$.

It is clear that $\mathrm{d}\phi$ is an exact differential and integration along an arbitrary curve gives:

$$\int_C \mathrm{d}\phi = \int_0^{\ell} \frac{\mathrm{d}\phi}{\mathrm{d}s} \mathrm{d}s = \phi \circ \boldsymbol{\beta}(\ell) - \phi \circ \boldsymbol{\beta}(0) \tag{5}$$

demonstrating path independence.

Note that, in general, a differential form $\tilde{\mathrm{d}}\varphi = \Sigma_{i=1}^{n} \gamma_i \mathrm{d}s_i$ is an exact differential if and only if $\partial\gamma_i/\partial s_j = \partial\gamma_j/\partial s_i$.

Thermodynamic state spaces

Thermodynamics is principally concerned with the concept of state and movement of a system from one state to another. For a *PVT*-closed system the state variables are pressure, P, temperature, T, and volume, V. Useful functions dependent on any two of these variables are internal energy, U, enthalpy, $H = U + PV$, entropy, S, Gibbs free energy, $G = H - TS$, and Helmholtz free energy, $F = U - TS$. The first law of thermodynamics introduced in the introduction is $\mathrm{d}U = \tilde{\mathrm{d}}W + \tilde{\mathrm{d}}Q$, where it is recognised that $\mathrm{d}U$ is an exact differential and in general $\tilde{\mathrm{d}}W$ and $\tilde{\mathrm{d}}Q$ are inexact differentials. However, in the case of reversible changes $\tilde{\mathrm{d}}W = -P\mathrm{d}V$ and $\tilde{\mathrm{d}}Q = T\mathrm{d}S$, where this last expression arises as a consequence of the second law of thermodynamics. Substitution of these terms into $\mathrm{d}U = \tilde{\mathrm{d}}W + \tilde{\mathrm{d}}Q$ gives $T\mathrm{d}S = \mathrm{d}U + P\mathrm{d}V$, which is the central equation of thermodynamics. The importance of this equation should not be underestimated, as it consists of state functions only and thus universally applicable on a stable equilibrium surface.

It is useful to examine the thermodynamic concepts in the light of the theory presented in the two previous sections. An important consideration not discussed in those sections is the extent of the space on which the thermodynamic variables are defined. For example, the thermodynamic temperature, or its non-dimensional equivalent, for most engineering applications, is limited to non-negative values only. It is clear that only a subspace of $\Re^n$ is involved. It is common practice in thermodynamics to consider the existence of inverse relationships, where thermodynamic functions take on the role of thermodynamic variables, that is, no distinction is made between functions and variables. Note, however, that specific quantities such as u, h, f and s should be used in place of U, H, F and S to remove the dependency on mass. In general, inverse functions may exist only on a subspace of the space under consideration. The issue of non-dimensionality can be conveniently dealt with by dividing each state variable by unity. This is convenient because the numerical values of the state variables remain unchanged.

Consider now the simple system depicted in Fig. 2. The convention adopted here for work is that work on the system is positive. This convention is typically utilised by physicists but is also utilised by engineers interested in solid mechanics, because it is the system (i.e. the solid) which is of principal interest. The work differential of the applied force, F, is $\tilde{\mathrm{d}}W = F\mathrm{d}x$. If equilibrium with the pressure is assumed,

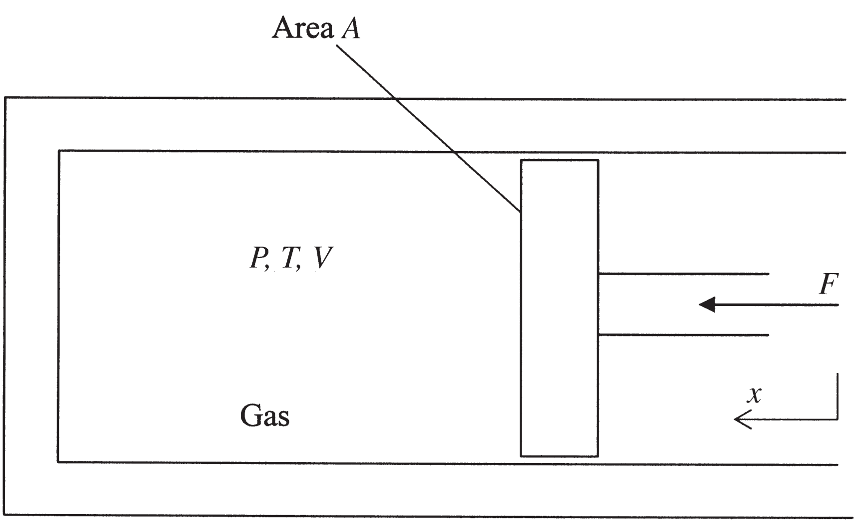

Fig. 2 *Pressurisation of a gas by a frictionless piston.*

International Journal of Mechanical Engineering Education 32/4

then $PA = F$, and note that $\mathrm{d}V = -A\mathrm{d}x$, giving $\tilde{\mathrm{d}}W = -P\mathrm{d}V$. The work differential is related to the exact differential $\mathrm{d}V$, but is $-P\mathrm{d}V$ exact? Consider the possibility that W can be explicitly represented in terms of the state variables V and P, that is, $W(V, P)$. Then equation 3 gives $\tilde{\mathrm{d}}W = \mathrm{d}W = \partial W/\partial P\mathrm{d}P + \partial W/\partial V\mathrm{d}V$, which on comparison with $\tilde{\mathrm{d}}W = -P\mathrm{d}V$ gives $\partial W/\partial P = 0$ and $\partial W/\partial V = -P$. Hence, by the condition stipulated at the end of the previous section, an exact differential requires $\partial(-P)/\partial P = 0$, which is clearly contradictory and consequently $\tilde{\mathrm{d}}W = -P\mathrm{d}V$ is an inexact differential and W cannot be explicitly represented in terms of the state variables V and P. Consider, however, the compression/expansion process to be adiabatic, that is, $\mathrm{d}Q = 0$, and the first law of thermodynamics gives $\mathrm{d}U = \tilde{\mathrm{d}}W = -P\mathrm{d}V$. The word 'adiabatic' is defined in the engineering sense, that is, no heat, with no implication of reversibility. It is clear that $\tilde{\mathrm{d}}W$ is an exact differential here (i.e. $\tilde{\mathrm{d}}W = \mathrm{d}W$), being equal to the exact differential $\mathrm{d}U$. This appears to contradict the analysis above, which ruled out the possibility that W can be explicitly represented in terms of the state variables V and P. However, the above analysis did not exclude the possibility of W being the function of one variable such as V. It is clear that reversibility on a one-dimensional space implies path independence. Thus the differential of $W(V)$ is $\mathrm{d}W = (\mathrm{d}W/\mathrm{d}V)\mathrm{d}V = -P\mathrm{d}V$, giving $\mathrm{d}W/\mathrm{d}V = -P$ or $W(V) = -\int_{V_0}^{V} P(V)\mathrm{d}V$. This section illustrates the importance of differentials in thermodynamics but also that the space on which the differential is formed is required to be well defined.

Elasticity

It is of interest to examine elasticity in the context of state spaces and examine the important role of differentials. Assume the existence of a uniform state of stress and a state of free stress. It is well appreciated that elasticity theory is founded on relationships between nominal stress and strain. Consider the uniaxial deformation process depicted in Fig. 3. The work differential of the applied force, F, is $\tilde{\mathrm{d}}W = F\mathrm{d}[u(\ell_0)]$. Note that the material displacement, u, is a linear function of x_1, that is, $u(x_1) = (x_1/\ell_0)u(\ell_0)$. Let nominal stress and strain be defined by $\sigma = F/A_0$ and $\varepsilon = \mathrm{d}u/\mathrm{d}x_1 = u(\ell_0)/\ell_0$, respectively. It follows that $\tilde{\mathrm{d}}W = F\mathrm{d}[u(\ell_0)] = V_0\sigma\mathrm{d}\varepsilon$. It is important to appreciate what $\mathrm{d}[u(\ell_0)]$ signifies. The differential is not with respect to x_1, since $u(\ell_0)$ is essentially a constant in this respect and $\mathrm{d}[u(\ell_0)]$ would simply be zero. This is an illustration of the importance of defining the space on which the differential is defined. Essentially, $u(\ell_0)$ is taking the role of a state variable and the equality $\mathrm{d}[u(\ell_0)] = \ell_0^{-1}\mathrm{d}\varepsilon$ indicates that ε is an alternative if not better choice for the state variable. More generally, $\mathrm{d}W = V_0\sigma_{ij}\mathrm{d}\varepsilon_{ij}$, where summation is over repeating suffices and $\varepsilon_{ij} = (\partial u_i/\partial x_j + \partial u_j/\partial x_i)/2$. The central equation of thermodynamics, applicable to this case, is $\mathrm{d}U = T\mathrm{d}S + V_0\sigma_{ij}\mathrm{d}\varepsilon_{ij}$. A principal assumption of elastic deformation is that the work performed by the stress field is stored as potential energy with no significant temperature change. With the assumption that the process is isothermal, the central equation gives $\mathrm{d}F = \mathrm{d}(U - TS) = V_0\sigma_{ij}\mathrm{d}\varepsilon_{ij} = \tilde{\mathrm{d}}W$ and it immediately follows that $\tilde{\mathrm{d}}W$ is an exact differential (i.e. $\tilde{\mathrm{d}}W = \mathrm{d}W$) and thus any contour integral of $\mathrm{d}W$ is path independent. This is of fundamental importance in elasticity theory, as it leads to the concept of superposition. The point here is that any particular deformation

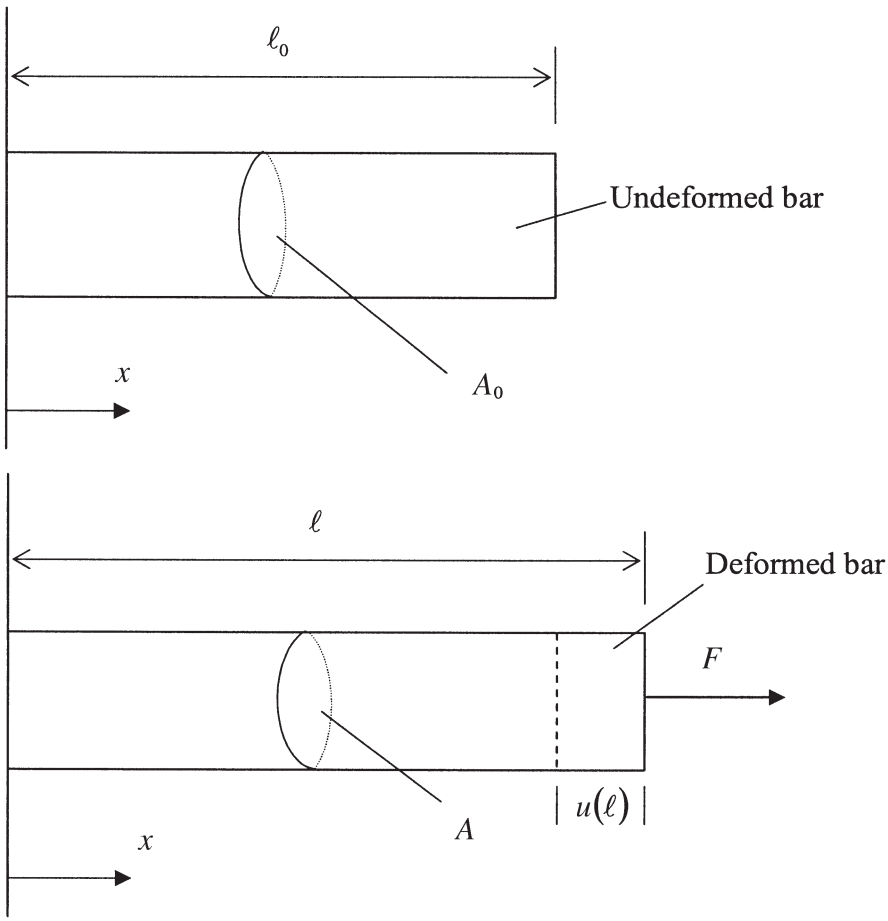

Fig. 3 *Uniaxial deformation of a prismatic bar.*

can be achieved in an infinite number of ways associated with different paths in the state space. Note that the state variables are the six independent terms of the strain tensor:

$$\underline{\underline{\varepsilon}} = \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \varepsilon_{13} \\ \varepsilon_{21} & \varepsilon_{22} & \varepsilon_{23} \\ \varepsilon_{31} & \varepsilon_{32} & \varepsilon_{33} \end{bmatrix} \qquad (7)$$

It is evident that the requirement that $\mathrm{d}W = V_0\sigma_{ij}\mathrm{d}\varepsilon_{ij}$ is an exact differential places a restriction on the relationship between σ_{ij} and ε_{ij}, namely $\partial\sigma_{ij}/\partial\varepsilon_{\ell m} = \partial\sigma_{\ell m}/\partial\varepsilon_{ij}$. Stress can be conveniently defined in this case via the Helmholtz free energy, since $\mathrm{d}F = (\partial F/\partial\varepsilon_{ij})\mathrm{d}\varepsilon_{ij} = V_0\sigma_{ij}\mathrm{d}\varepsilon_{ij}$ provides $\sigma_{ij} = V_0^{-1}(\partial F/\partial\varepsilon_{ij})$. Defining strain energy density to be $\omega = F/V_0$ gives the familiar result $\sigma_{ij} = \partial\omega/\partial\varepsilon_{ij}$. Note that

$$\partial\sigma_{ij}/\partial\varepsilon_{\ell m}=\partial^2\omega/\partial\varepsilon_{ij}\partial\varepsilon_{\ell m}=\partial^2\omega/\partial\varepsilon_{\ell m}\partial\varepsilon_{ij}=\partial\sigma_{\ell m}/\partial\varepsilon_{ij} \tag{8}$$

Consequently, path independency is solely reliant on the differentiability of ω. An example of a quadratic strain energy density function is:

$$\omega=\frac{1}{2}\mathrm{A}\left(\varepsilon_{11}^2+\varepsilon_{22}^2+\varepsilon_{33}^2\right)+\mathrm{B}\left(\varepsilon_{11}\varepsilon_{22}+\varepsilon_{22}\varepsilon_{33}+\varepsilon_{33}\varepsilon_{11}\right)+\frac{1}{2}\mathrm{C}\left(\varepsilon_{12}^2+\varepsilon_{23}^2+\varepsilon_{31}^2\right) \tag{9}$$

where A, B and C are material constants, which can be related to the Hookean constants: $\mathrm{A}=\mathrm{E}(1-\nu)/(1+\nu)(1-2\nu)$, $\mathrm{B}=\mathrm{E}\nu/2(1+\nu)(1-2\nu)$ and $\mathrm{C}=2\mathrm{G}=\mathrm{E}/(1+\nu)$, where E is Young's modulus, G is the shear modulus and ν Poisson's ratio.

The stresses are:

$$\sigma_{11}=\frac{\partial\omega}{\partial\varepsilon_{11}}=\mathrm{A}\varepsilon_{11}+\mathrm{B}\left(\varepsilon_{22}+\varepsilon_{33}\right)=\frac{\mathrm{E}}{(1+\nu)(1-2\nu)}\left[(1-\nu)\varepsilon_{11}+\nu\left(\varepsilon_{22}+\varepsilon_{33}\right)\right]$$

$$\sigma_{12}=\frac{\partial\omega}{\partial\varepsilon_{12}}=2\mathrm{G}\varepsilon_{12}\ \ldots \tag{10}$$

Plasticity

It is of interest to examine the use of differentials in the context of an irreversible process. Consider again the uniaxial deformation of the rod depicted in Fig. 3. Since plastic deformation is associated with large geometric changes, stress and strain are defined with reference to the current configuration rather than the original. True stress and true strain increments are defined as $\sigma=F/A$ and $\tilde{\mathrm{d}}\varepsilon=\ell^{-1}\mathrm{d}\ell$, respectively. It is useful to reflect on the precise meaning of the expression $\tilde{\mathrm{d}}\varepsilon=\ell^{-1}\mathrm{d}\ell$. A possible misconception is to consider the differential $\mathrm{d}\ell$ to be defined as exact in the spatial domain, and to view the length as a function of the spatial variable x_1. If this is true, then $\mathrm{d}\ell=(\mathrm{d}\ell/\mathrm{d}x_1)\mathrm{d}x_1$, $\tilde{\mathrm{d}}\varepsilon=\ell^{-1}(\mathrm{d}\ell/\mathrm{d}x_1)\mathrm{d}x_1$ and $\tilde{\mathrm{d}}\varepsilon(e_1)=\ell^{-1}(\mathrm{d}\ell/\mathrm{d}x_1)\mathrm{d}x_1(e_1)=\ell^{-1}(\mathrm{d}\ell/\mathrm{d}x_1)$. Few students would object if this theory was placed before them. The problem here is that $\ell(x_1)$ is meaningless, and $\mathrm{d}\ell$ needs to be viewed in the light of a one-dimensional tangent space, where ℓ is the coordinate function. It is common practice in plasticity theory to specify values for $\tilde{\mathrm{d}}\varepsilon$, which is meaningful and understandable in this context since $\tilde{\mathrm{d}}\varepsilon(e_1)=\ell^{-1}\mathrm{d}\ell(e_1)=\ell^{-1}$.

Another issue is the question of the plasticity differential $\tilde{\mathrm{d}}\varepsilon$ being exact. Since the limit of our attention is on a one-dimensional space spanned by e_1, it might be expected that this is the case. In the one-dimensional state space considered in the section 'Thermodynamic State Spaces', reversibility was required. However, this has more to do with definition of the state variables, establishing functionality. Here it is clear that ε is properly defined as a function of ℓ, that is, $\varepsilon(\ell)=\ln(\ell/\ell_0)$, hence $\tilde{\mathrm{d}}\varepsilon=\mathrm{d}\varepsilon$. It is recognised that ε does not identify the material state, despite the use of empirical relationships of the form $\sigma=\mathrm{A}\varepsilon^{\mathrm{n}}$ and $\sigma=\mathrm{A}(\mathrm{B}+\varepsilon)^{\mathrm{n}}$, where A, B and n are experimentally determined material constants. Equations of this type are obtained under monotonic loading conditions, typically from a uniaxial tensile test. The point is that if a bar is extended then ε will increase but if it is subsequently compressed then these empirical relationships do not apply. The question that arises is how a state variable can be defined for an irreversible process. It is clear that ε takes on this role if it is monotonically increasing, which provides a clue. Consider the work differential $\tilde{d}W = V\sigma_{ij}d\varepsilon_{ij}$, where summation is over repeating suffices and it is assumed that elastic effects are negligible. The differential can be written as $\tilde{d}W = V\underline{\underline{\sigma}} : \underline{\underline{d\varepsilon}} = V\underline{\underline{\sigma}}' : \underline{\underline{d\varepsilon}}$, where

$$\underline{\underline{d\varepsilon}} = \begin{bmatrix} d\varepsilon_{11} & d\varepsilon_{12} & d\varepsilon_{13} \\ d\varepsilon_{21} & d\varepsilon_{22} & d\varepsilon_{23} \\ d\varepsilon_{31} & d\varepsilon_{32} & d\varepsilon_{33} \end{bmatrix}, \underline{\underline{\sigma}} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix} \tag{11}$$

and V is volume, $\underline{\underline{\sigma}}' = \underline{\underline{\sigma}} - \sigma_m\underline{\underline{I}}$, and where $\sigma_m = (\sigma_{11} + \sigma_{22} + \sigma_{33})/3$ and volume constancy is assumed, that is, $d\varepsilon_{11} + d\varepsilon_{22} + d\varepsilon_{33} = 0$.

In classical plasticity, the flow of material is governed by the Levy Mises flow rule, $\underline{\underline{d\varepsilon}} = \tilde{d}\lambda\underline{\underline{\sigma}}'$. An effective stress, $\bar{\sigma}$, can be defined from the von Mises yield criterion, $\bar{\sigma}^2 = (3/2)\underline{\underline{\sigma}}' : \underline{\underline{\sigma}}'$. Similarly, the effective strain differential, $\tilde{d}\bar{\varepsilon}$, can be defined via the identity $\tilde{d}\bar{\varepsilon}^2 = (2/3)\underline{\underline{d\varepsilon}} : \underline{\underline{d\varepsilon}}$. These relationships can be expressed in terms of principal values:

$$\bar{\sigma}^2 = \frac{3}{2}\underline{\underline{\sigma}}' : \underline{\underline{\sigma}}' = \frac{1}{2}\left((\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2\right) \tag{12a}$$

$$\tilde{d}\bar{\varepsilon}^2 = \frac{2}{3}\underline{\underline{d\varepsilon}} : \underline{\underline{d\varepsilon}} = \frac{2}{9}\left((d\varepsilon_1 - d\varepsilon_2)^2 + (d\varepsilon_2 - d\varepsilon_3)^2 + (d\varepsilon_3 - d\varepsilon_1)^2\right) \tag{12b}$$

Substitution of $\underline{\underline{d\varepsilon}} = d\lambda\underline{\underline{\sigma}}'$ in $\tilde{d}\bar{\varepsilon}^2 = (2/3)\underline{\underline{d\varepsilon}} : \underline{\underline{d\varepsilon}}$ gives $\tilde{d}\bar{\varepsilon}^2 = (d\lambda)^2(2/3)\underline{\underline{\sigma}}' : \underline{\underline{\sigma}}' = (2/3)^2(\tilde{d}\lambda\bar{\sigma})^2$ or $\tilde{d}\lambda = (3/2)\bar{\sigma}^{-1}\tilde{d}\bar{\varepsilon}$. These definitions enable the work differential to be simplified as:

$$dW = V\underline{\underline{\sigma}}' : \underline{\underline{d\varepsilon}} = \tilde{d}\lambda V\underline{\underline{\sigma}}' : \underline{\underline{\sigma}}' = \left(\frac{3}{2}\frac{\tilde{d}\bar{\varepsilon}}{\bar{\sigma}}\right)V\left(\frac{2}{3}\bar{\sigma}^2\right) = V\bar{\sigma}\tilde{d}\bar{\varepsilon} \tag{13}$$

The variables $\bar{\sigma}$ and $\bar{\varepsilon}$ represent state variables for stress and strain, respectively. Both variables are positive and $\bar{\varepsilon}$ monotonically increases with deformation. Much algebraic manipulation of the differentials is involved in plasticity theory and the legitimacy of this manipulation is founded on the ability to evaluate differentials pointwise on tangent spaces. The choice of $\bar{\varepsilon}$ as a state variable means that $\tilde{d}\bar{\varepsilon} = d\bar{\varepsilon}$, by definition. Other state variables may be involved, but if $\bar{\sigma}$ is a function of $\bar{\varepsilon}$ only, then path dependence is not possible. Monotonicity essentially implies that there is only one possible path between two states, $\bar{\varepsilon}_1$ and $\bar{\varepsilon}_2$. Introducing another state variable such as temperature, T, raises the possibility of path dependency in the state space. The first law of thermodynamics gives $dU = \tilde{d}Q + V\bar{\sigma}d\bar{\varepsilon}$. It is apparent that with $\partial\bar{\sigma}/\partial T \neq 0$, no explicit expression is available for $W(\bar{\varepsilon}, T)$. Readers familiar with plasticity theory will appreciate that deformation, as described via the integration of $\underline{\underline{d\varepsilon}}$, is path dependent. However, it is important to define the space on which the integration is performed. In this case the space is $\Re^6$, where the coordinate functions are $\{\varepsilon_{11}, \varepsilon_{22}, \varepsilon_{33}, \varepsilon_{13}, \varepsilon_{23}, \varepsilon_{21}\}$.

Mechanics

Differentials also have a part to play in mechanics. Consider a point mass, M, moving along a curve under the action of a result force F in an Euclidean space E^3. The work differential is simply $\tilde{d}W = F \cdot dx$, where dx is defined to be $e_1 dx_1 + e_2 dx_2 + e_3 dx_3$ or (dx_1, dx_2, dx_3). If the mass is following some contour in E^3, then $dx = T ds$, where T is a unit tangent vector. This result follows because $dx(T) = e_1 dx_1(T) + e_2 dx_2(T) + e_3 dx_3(T) = T$ and $ds(T) = 1$. Newton's second law provides $F = M(d\boldsymbol{v}/dt)$, where in this case the velocity $\boldsymbol{v} = vT$. Thus, $\tilde{d}W = F \cdot dx = M(v(dT/dt) + (dv/dt)T) \cdot T ds = M(dv/dt)ds$ and with the assumption that v is a function of s, then $\tilde{d}W = F \cdot dx = M(dv/ds)(ds/dt)ds = dK$, where kinetic energy $K = Mv^2/2$ and $dK = (dK/ds)ds$. In examining the reversible adiabatic expansion/compression of a gas in the section 'Thermodynamic State Spaces', the equality $dU = \tilde{d}W$ gave $\tilde{d}W = dW$, that is, the work differential is exact. At first sight it appears that the equality $\tilde{d}W = dK$ implies a similar result. The difference here is that E^3 is not a state space and K is not a state variable. Thus, it is possible to move between two points in E^3, starting with the same K value but taking different paths to arrive with different values of K; that is, $\tilde{d}W = dK$ implies that K is not uniquely defined at points in E^3. If, however, $\tilde{d}W = dW$, then an energy potential Φ can be defined so that $dW = -d\Phi = -\nabla\Phi \cdot dx$. In this case $d(\Phi + K) = 0$ and integration between two points provides $\Phi_1 + K_1 = \Phi_2 + K_2$, that is, the total energy, $E = \Phi + K$, is invariant. Note that energy transfers between kinetic energy K and potential energy Φ take place without loss. For many important problems, losses are present and energy is converted into a form that, from a mechanics viewpoint, is non-usable (i.e. thermal energy). It is evident from the discussion above that losses will occur if $\tilde{d}W \neq dW$ and in this case thermal energy is involved. Clearly, the first and second laws of thermodynamics must be involved but this link is seldom explored in mechanics texts.

Consider the problem depicted in Fig. 4 and the following equations:

$$\tilde{d}W_{res} = \tilde{d}W_{ncon} + dW_{con} = \tilde{d}W_{fric} + \tilde{d}W_{ext} + dW_{con} = dK \tag{14}$$

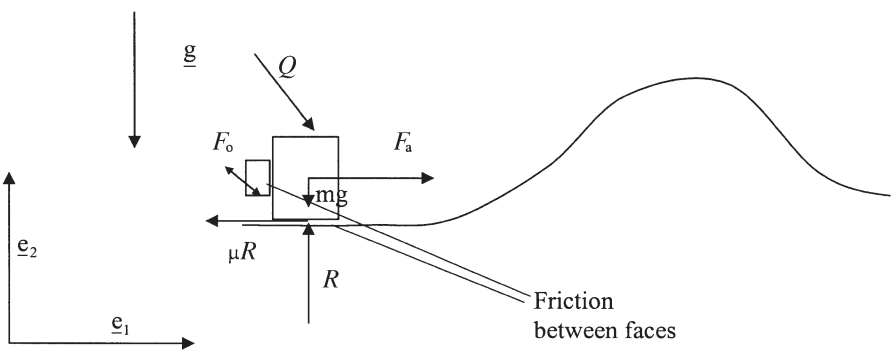

Fig. 4 *Mechanics of a point mass.*

and

$$\tilde{d}W_{fric} + \tilde{d}W_{ofric} + \tilde{d}Q_{ext} = dU \tag{15}$$

where the subscripts 'res', 'fric', 'con', 'ncon', 'ext' and 'ofric' denote 'resultant', 'friction', 'conservative', 'non-conservative', 'external' and 'other friction', respectively.

Equation 14 is a consequence of Newton's second law, where $\tilde{d}W_{res} = F_{res} \cdot dx$, $\tilde{d}W_{fric} = F_{fric} \cdot dx$, and so on. The term $\tilde{d}W_{ofric}$ refers to other frictional forces, which, as depicted in Fig. 4, do not affect the motion of the point mass in the direction under consideration. Note that equations 14 and 15 have a common term, namely $\tilde{d}W_{fric}$, yet the spaces on which the differentials dU and dK are defined are different. The work done by the conservative force is associated with energy potential, that is, $dW_{con} = -d\Phi = -\nabla\Phi \cdot dx$. Substitution of this and elimination of $\tilde{d}W_{fric}$ from equations 14 and 15 gives:

$$\tilde{d}W_{ext} - \tilde{d}W_{ofric} - \tilde{d}Q_{ext} = dK + d\Phi - dU \tag{16}$$

Equation 16 has differentials defined in different spaces and the question of its validity is raised. Consider the situation where $\tilde{d}W_{ext} = \tilde{d}W_{ofric} = \tilde{d}Q_{ext} = 0$ to give

$$dU = dK + d\Phi \Rightarrow U_2 - U_1 = K_2 - K_1 + \Phi_2 - \Phi_1 \tag{17}$$

Equation 17 is applied to a non-reversible process and care needs to taken, as it makes sense only when $U_2 - U_1 \geq 0$. These types of problem are useful for highlighting the types of energy transfer that are possible. Another interesting feature with problems where differentials are defined on two spaces is to do with path dependency. Questions of interest are: 'Are common differentials exact (or inexact) in both spaces?' and, similarly, 'Does path dependency in one space mean path dependency in the other?' Concepts of type have physical significance and are seldom explored on current engineering courses.

Discussion

It is clear that differentials play an important role in engineering science. This is not surprising because they are important to thermodynamics and this subject underpins many other engineering disciplines. The focus in this paper has been on solid mechanics but, equally, the physics of many fluid mechanics problems is well described with the use of differentials. Engineering students meet differentials in many subject areas but their definition is seldom rigorous. One of the critical points raised in this paper is the importance of correctly defining the space on which the differential is defined. This is not only mathematically sound practice but also provides for physical insight, as demonstrated for elasticity and plasticity theory above. Mechanics problems raised the possibility of equations with differentials defined on different spaces. These problems are seldom considered in university courses yet they are fundamental to understanding how energy is transferred between macro and micro scales. It is clear that placing differentials in a formal setting has significant advantages and provides for greater depth of understanding. The concept of tangent spaces provides a foundation for the use of differentials. However, it is recognised that many engineering students may not appreciate a subject that relies heavily on mathematics. Certainly such a subject could not be taught in the early part of an engineering degree course. However, a course founded on differentials in the final year would serve to integrate the apparently disparate themes that underpin themodynamics, fluid and solid mechanics, material theories, and so on. The bringing together of concepts makes for an appealing course, possibly of interest to some of the more able students.

Conclusions

This paper is concerned with the use of differentials for linking concepts across engineering subjects. The following conclusions can be drawn from this brief discussion:

(1) The space on which a differential is defined needs to be identified for the problem at hand and this provides for greater physical understanding.
(2) The notion of a differential as an infinitesimal quantity is useful to an extent but provides for limited understanding.
(3) The concepts of path dependency, exact and inexact differentials, reversibility, monotonicity, and their interrelationships provide for physical insight.
(4) The integration of thermodynamics in other subject areas is facilitated with the use of differentials.

References

[1] E. Kreyszig, *Advanced Engineering Mathematics*, 8th edn (John Wiley, Chichester, 1999).
[2] K. A. Stroud, *Engineering Mathematic*, 5th edn (Palgrave, Basingstoke, 2001).
[3] B. O'Neil, *Elementary Differential Geometry* (Academic Press, London, 1966).
[4] M. Crampin, F. A. E. Pirani, *Applicable Differential Geometry*, Mathematical Society Lecture Note Series 59 (Cambridge University Press, Cambridge, 1988).