Contents of issue 6 vol. XLVI

Polish Academy of Sciences

Institute of Fundamental Technological Research

Archives of Mechanics

Archiwum Mechaniki Stosowanej

volume 46 issue 6

Polish Scientific Publishers PWN

Warszawa 1994

ARCHIVES OF MECHANICS IS DEVOTED TO
Theory of elasticity and plasticity • Theory of nonclassical continua • Physics of continuous media • Mechanics of discrete media • Nonlinear mechanics • Rheology • Fluid gas-mechanics • Rarefied gas • Thermodynamics

---

Printed in Poland, Editorial Office: Świętokrzyska 21,
00-049 Warszawa (Poland)

Arkuszy wydawniczych 16. Arkuszy drukarskich 13,25
Papier offset. kl. III 70 g. B1. Oddano do składania w listopadzie 1994 r.
Druk ukończono w lutym 1995 r.
Skład i łamanie: IPPT PAN
Druk i oprawa: Drukarnia Braci Grodzickich, Żabieniec ul. Przelotowa 7

Arch. Mech., 46, 6, pp. 819–828, Warszawa 1994

Thermosolutal hydromagnetic instability of a compressible and partially ionized plasma in porous medium

SUNIL (HAMIRPUR)

THE THERMOSOLUTAL instability of a compressible and partially ionized plasma in porous medium is considered in the presence of a uniform vertical magnetic field to include the effects of collisions. The stable solute gradient and magnetic field are found to introduce oscillatory modes in the system for $(C_p/g)\beta > 1$. For stationary convection, the magnetic field and stable solute gradient are found to have stabilizing effects, whereas the medium permeability has a destabilizing effect on the system for $(C_p/g)\beta > 1$. The effect of compressibility is found to postpone the onset of convection. The collisional effects disappear for stationary convection. The sufficient conditions for the non-existence of overstability are obtained.

1. Introduction

A COMPREHENSIVE account of thermal instability, under varying assumptions of hydromagnetics, has been given by CHANDRASEKHAR [1]. VERONIS [2] has studied the problem of thermohaline convection in a layer of fluid heated and salted from below. The physics is quite similar in the stellar case in that helium acts like salt in raising the density and in diffusing more slowly than heat. The conditions under which convective motions are important in stellar atmospheres are usually far removed from consideration of single component fluid and rigid boundaries and therefore, it is desirable to consider a fluid acted on by a solute gradient and free boundaries. The problem of the onset of thermal instability in the presence of a solute gradient is of great importance because of its application to atmospheric physics and astrophysics, especially in the case of the ionosphere and the outer layers of the solar atmosphere.

A partially ionized plasma represents a state which often exists in the Universe and there are several situations where the interaction between the ionized and neutral gas components becomes important in cosmic physics. The existence of such situations follows from ALFVÉN's [3] theory on the origin of the planetary system, in which a high-ionization rate is suggested to appear from collisions between a plasma and a neutral gas cloud, and by the absorption of plasma waves due to ion-neutral collisions such as occur in the solar photosphere and chromosphere and in cool interstellar clouds. SHARMA [4] has studied the thermal hydromagnetic instability of a partially ionized plasma for incompressible case. SHARMA and SUNIL [5] have studied the thermosolutal instability of a partially ionized Hall plasma in a porous medium in presence of uniform vertical magnetic field. In another study, SHARMA and SUNIL [6] have considered the Rayleigh–Taylor

instability of a partially ionized plasma in a porous medium, in presence of a variable magnetic field.

When the fluids are compressible, the equations governing the system become quite complicated. To simplify the set of equations governing the flow of compressible fluids, SPIEGEL and VERONIS [7] have made the following assumptions:

i) the depth of the fluid layer is much smaller than the scale height as defined by them, and

ii) the fluctuations in density, pressure and temperature, introduced due to motion, do not exceed their total static variations.

Under the above assumptions, SPIEGEL and VERONIS [7] have found that the flow equations are the same as those for incompressible fluids, except that static temperature gradient is replaced by its excess over the adiabatic one. SHARMA and SUNIL [8] have studied the thermal instability of compressible Hall plasma in the presence of suspended particles.

Generally, it is accepted that comets consist of a dusty "snowball" of a mixture of frozen gases which, in the process of their journey, changes from solid to gas and vice versa. The physical properties of comets, meteorites and interplanetary dust strongly suggest the importance of porosity in astrophysical context (MC DONNEL [9]). The compressibility, collisions between ionized and neutral particles and medium porosity effects are likely to be important in astrophysical situations like stellar interiors and atmospheres and in geophysical situations like Earth's molten core. The present paper, therefore, deals with the thermosolutal instability of a compressible and partially ionized plasma in porous medium in the presence of uniform vertical magnetic field to include the effects of collisions.

2. Perturbation equations

Here we consider an infinite, horizontal, compressible and composite plasma layer, consisting of a finitely (electrical) conducting, ionized component of density ρ and, neutral component of density ρ_d, of thickness d, and acted on by a uniform vertical magnetic field $\mathbf{H}(0, 0, H)$ and gravity field $\mathbf{g}(0, 0, -g)$. This layer is heated and soluted from below so that the temperatures, densities and solute concentrations at the bottom surface $z = 0$ are T_0, ρ_0, and C_0 and at the upper surface $z = d$ are T_d, ρ_d and C_d, respectively, and that a uniform temperature gradient $\beta(= |dT/dz|)$ and uniform solute gradient $\beta'(= |dC/dz|)$ are maintained. This plasma layer is assumed to be flowing through an isotropic and homogeneous porous medium of porosity ε and medium permeability k_1. Also we assume that both the ionized gas and the neutral gas behave like continuum fluids and that the effects on the neutral component resulting from the presence of magnetic field, porosity and the fields of gravity and pressure are neglected.

SPIEGEL and VERONIS [8] defined f as any of the state variables pressure (p), density (ρ) or temperature (T), and expressed these in the form

$$f(x, y, z, t) = f_m + f_0(z) + f'(x, y, z, t), \tag{2.1}$$

where f_m is the constant space average of f, f_0 is the variation in the absence of motion, and f' is the fluctuation resulting from motion.

The initial state is, therefore, a state in which the density, pressure, temperature, solute concentration and velocity at any point in the plasma are given by

$$\rho = \rho(z), \qquad p = p(z), \qquad T = T(z), \qquad C = C(z), \qquad \mathbf{q} = (0, 0, 0), \tag{2.2}$$

respectively, where

$$\begin{aligned} T(z) &= T_0 - \beta z, \qquad C(z) = C_0 - \beta' z, \\ p(z) &= p_m - g \int_0^z (\rho_m + \rho_0)\, dz, \\ \rho(z) &= \rho_m[1 - \alpha_m(T - T_m) + \alpha'_m(C - C_m) + K_m(p - p_m)], \\ \alpha_m &= -\left(\frac{1}{\rho}\frac{\partial \rho}{\partial T}\right)_m, \qquad \alpha'_m = -\left(\frac{1}{\rho}\frac{\partial \rho}{\partial C}\right)_m, \\ K_m &= \left(\frac{1}{\rho}\frac{\partial \rho}{\partial p}\right)_m. \end{aligned} \tag{2.3}$$

Here we restrict our study to the infinitesimal peturbations, so that the motion-induced perturbations in density and pressure do not exceed, in the order of magnitude, their total static variations. Also, the depth of a layer of fluid is assumed to be much less than any scale height as defined by SPIEGEL and VERONIS [8]. Let δp, $\delta\rho$, θ, γ, $\mathbf{h}(h_x, h_y, h_z)$, $\mathbf{q}(u, v, w)$ and $\mathbf{q}_d(l, r, s)$ denote, respectively, the perturbations in presssure p, density ρ, temperature T, solute concentration C, magnetic field $\mathbf{H}$, ionized component velocity and neutral component velocity; μ, $\nu(= \mu/\rho_m)$, μ_e, κ, κ', g/C_p, η and ν_c stand for viscosity, kinematic viscosity, magnetic permeability, thermal diffusivity, solute diffusivity, adiabatic gradient, resistivity and collisional frequency between two components of composite medium. Then the linearized hydromagnetic perturbation equations relevant to the problem are

$$\begin{aligned} \frac{1}{\varepsilon}\frac{\partial \mathbf{q}}{\partial t} = &-\left(\frac{1}{\rho_m}\right)\nabla\delta p + \mathbf{q}\left(\frac{\delta\rho}{\rho_m}\right) + \nu\nabla^2\mathbf{q} + \frac{\mu_e}{4\pi\rho_m}(\nabla \times \mathbf{h}) \times \mathbf{H} \\ &+ \frac{\rho_d \nu_c}{\rho_m \varepsilon}(\mathbf{q}_d - \mathbf{q}) - \frac{\nu}{k_1}\mathbf{q}, \end{aligned} \tag{2.4}$$

$$\nabla \cdot \mathbf{q} = 0, \tag{2.5}$$

$$\frac{\partial \mathbf{q}_d}{\partial t} = -\nu_c(\mathbf{q}_d - \mathbf{q}), \tag{2.6}$$

$$E\frac{\partial \theta}{\partial t} = \left(\beta - \frac{g}{C_p}\right) w + \kappa \nabla^2 \theta, \tag{2.7}$$

$$E'\frac{\partial \gamma}{\partial t} = \beta' w + \kappa' \nabla^2 \gamma, \tag{2.8}$$

$$\varepsilon\frac{\partial \mathbf{h}}{\partial t} = (\mathbf{H} \cdot \nabla)\,\mathbf{q} + \varepsilon\eta\nabla^2 \mathbf{h}, \tag{2.9}$$

$$\nabla \cdot \mathbf{h} = 0, \tag{2.10}$$

where $E = \varepsilon + (1 - \varepsilon)(\rho_s C_s/\rho C)$ is a constant and E' is a constant analogous to E but corresponding to the solute rather than heat. ρ, C and ρ_s, C_s stand for density and specific heat of ionized component and solid (porous metrix) material, respectively. Here in the equation of motion (2.6) for the neutral component, there will be an equal and opposite term to that in the equation of motion (2.4) for ionized component, and the effects on the neutral component resulting from the presence of magnetic field, porosity and the fields of gravity and pressure are neglected.

The equation of state

$$\rho = \rho_m \left[1 - \alpha(T - T_m) + \alpha'(C - C_m)\right], \tag{2.11}$$

contains the thermal coefficient of expansion α and an analogous solute coefficient α'. The change in density is caused mainly by the temperature and solute concentration, and the suffix m refers to values at the reference level $z = 0$. The change in density $\delta\rho$, caused by the perturbation θ and γ, is given by

$$\delta\rho = -\rho_m(\alpha\theta - \alpha'\gamma). \tag{2.12}$$

3. Dispersion relation

Here we analyze the disturbances in normal modes; we assume that the perturbation quantities are of the form

$$[w, \theta, \gamma, h_z] = [W(z), \Theta(z), \Gamma(z), K(z)] \exp(ik_x x + ik_y y + nt), \tag{3.1}$$

where k_x, k_y are wave numbers along the x and y-directions, respectively, $k = (k_x^2 + k_y^2)^{1/2}$ is the resultant wave number, and n is the growth rate which is, in general, a complex constant.

Assuming that x, y, z stand for the coordinates in the new unit of length d and letting $a = kd$, $\sigma = nd^2/\nu$, $p_1 = \nu/\kappa$, $p_2 = \nu/\eta$, $q = \nu/\kappa'$, $\alpha_0 = \rho_d/\rho$,

$G = (C_p/g)\beta$, $P_l = k_1/\varepsilon d^2$ and $D = d/dz$, Eqs. (2.4) – (2.10) with the help of (2.11) and expression (3.1), written in nondimensional form become

$$\left[D^2 - a^2 - \sigma\left(1 + \frac{\alpha_0 \nu_c d^2/\nu}{\sigma + \nu_c d^2/\nu}\right) - \frac{1}{P_l}\right](D^2 - a^2)W - \left(\frac{g d^2 \varepsilon}{\nu}\right) a^2(\alpha\Theta - \alpha' \Gamma) + \left(\frac{\mu_e H d\varepsilon}{4\pi \rho_m \nu}\right)(D^2 - a^2)DK = 0, \tag{3.2}$$

$$(D^2 - a^2 - Ep_1\sigma)\Theta = -\frac{d^2}{\kappa}\left(\beta - \frac{g}{C_p}\right) W, \tag{3.3}$$

$$(D^2 - a^2 - E'q\sigma)\Gamma = -\frac{\beta' d^2}{\kappa'} W, \tag{3.4}$$

$$(D^2 - a^2 - p_2\sigma)K = -\left(\frac{Hd}{\eta\varepsilon}\right) DW. \tag{3.5}$$

Here we assume that the temperature and concentrations at the boundaries are kept fixed, the plasma layer is confined between two free boundaries and the adjoining medium is electrically nonconducting. The boundary conditions appropriate to the problem are

$$\left.\begin{array}{l} W = D^2 W = \Theta = \Gamma = DK = 0 \\ \text{and } h_x, h_y, h_z \text{ are continuous} \end{array}\right\} \text{at } z = 0 \text{ and } 1. \tag{3.6}$$

The case of two free boundaries, though a little artificial, is most appropriate for the stellar atmosphere. Using the above boundary conditions, it can be shown that all the even order derivatives of W must vanish for $z = 0$ and 1, and hence the proper solution of W, characterizing the lowest mode, is

$$W = W_0 \sin \pi z, \tag{3.7}$$

where W_0 is a constant.

Eliminating Θ, Γ and K between Eqs. (3.2) – (3.5) and substituting the proper solution $W = W_0 \sin \pi z$ in the resultant equation, we obtain the dispersion relation

$$R_1 x = \varepsilon^{-1}\left(\frac{G}{G-1}\right)\left\{(1+x)(1+x+iEp_1\sigma)\left[1 + x + i\sigma_1\left(1 + \frac{\alpha_0 \nu_c d^2/\nu}{i\sigma_1 \pi^2 + \nu_c d^2/\nu}\right) + \frac{1}{P}\right] + S_1 x\varepsilon \frac{(1+x+iEp_1\sigma_1)}{(1+x+iE'q\sigma_1)} + Q_1 \frac{(1+x)(1+x+iEp_1\sigma_1)}{(1+x+ip_2\sigma_1)}\right\}, \tag{3.8}$$

where

$$R_1 = \frac{g\alpha\beta d^4}{\nu\kappa\pi^4}, \qquad S_1 = \frac{g\alpha'\beta' d^4}{\nu\kappa'\pi^4}, \qquad Q_1 = \frac{\mu_e H^2 d^2}{4\pi\rho_m \nu\eta\pi^2},$$

$$\rho = \pi^2 P_l, \qquad i\sigma_1 = \frac{\sigma}{\pi^2} \qquad \text{and} \qquad x = \frac{a^2}{\pi^2}.$$

4. Stability of the system and oscillatory modes

Here we examine the possibility of the effect of oscillatory modes, if any, on the stability problem due to the presence of magnetic field and solute gradient. Multiplying Eq. (3.2) by W^*, the complex conjugate of W, integrating over the range of z, and making use of Eqs. (3.3)–(3.5) together with the boundary conditions, we obtain

$$I_1 + \left[\sigma\left(1 + \frac{\alpha_0\nu_c d^2/\nu}{\sigma + \nu_c d^2/\nu}\right) + \frac{1}{P_l}\right] I_2 + \frac{g\alpha'\kappa' a^2\varepsilon}{\nu\beta'}(I_5 + E'q\sigma^* I_6) + \frac{\varepsilon^2\mu_e\eta}{4\pi\rho_m\nu}(I_7 + p_2\sigma^* I_8) = \frac{C_p\alpha\kappa a^2\varepsilon}{\nu(G-1)}(I_3 + Ep_1\sigma^* I_4), \tag{4.1}$$

where

$$I_1 = \int_0^1 (|D^2W|^2 + 2a^2|DW|^2 + a^4|W|^2)\,dz,$$

$$I_2 = \int_0^1 (|DW|^2 + a^2|W|^2)\,dz,$$

$$I_3 = \int_0^1 (|D\Theta|^2 + a^2|\Theta|^2)\,dz, \qquad I_4 = \int_0^1 (|\Theta|^2)\,dz,$$

$$I_5 = \int_0^1 (|D\Gamma|^2 + a^2|\Gamma|^2)\,dz, \qquad I_6 = \int_0^1 (|\Gamma|^2)\,dz, \tag{4.2}$$

$$I_7 = \int_0^1 (|D^2K|^2 + 2a^2|DK|^2 + a^4|K|^2)\,dz,$$

$$I_8 = \int_0^1 (|DK|^2 + a^2|K|^2)\,dz.$$

The integrals $I_1 - I_8$ are all positive definite. Putting $\sigma = \sigma_r + i\sigma_1$ and equating the real and imaginary parts of Eq. (4.1), we obtain

$$(4.3)\qquad I_1 + \left[\frac{(\sigma_r + \nu_c d^2/\nu)(\sigma_r^2 + \sigma_r(1+\alpha_0)\nu_c d^2/\nu) + \sigma_1^2(\sigma_r + \alpha_0\nu_c d^2/\nu)}{[(\sigma_r + \nu_c d^2/\nu)^2 + \sigma_1^2]} + \frac{1}{P_l}\right] I_2 + \frac{g\alpha'\kappa' a^2\varepsilon}{\nu\beta'}(I_5 + E'q\sigma_r I_6) + \frac{\varepsilon^2\mu_e\eta}{4\pi\rho_m\nu}(I_7 + p_2\sigma_r I_8) = \frac{C_p\alpha\kappa a^2\varepsilon}{\nu(G-1)}(I_3 + Ep_1\sigma_r I_4),$$

and

$$(4.4)\qquad i\sigma_1\left[\frac{\sigma_r^2 + \sigma_1^2 + 2\sigma_r\nu_c d^2/\nu + (1+\alpha_0)(\nu_c d^2/\nu)^2}{[(\sigma_r + \nu_c d^2/\nu)^2 + \sigma_1^2]} I_2 + \frac{C_p\alpha\kappa a^2\varepsilon}{\nu(G-1)} Ep_1 I_4 - \frac{g\alpha'\kappa' a^2\varepsilon}{\nu\beta'} E'qI_6 - \frac{\varepsilon^2\mu_e\eta}{4\pi\rho_m\nu} p_2 I_8\right] = 0.$$

It follows from Eq. (4.4) that if $G > 1$ and if the magnetic field and solute gradient are absent, $\sigma_1 = 0$, which means that the oscillatory modes are not allowed and the principle of exchange of stabilities is satisfied for a porous medium in the absence of magnetic field and solute gradient. The oscillatory modes are introduced due to the presence of a magnetic field and a solute gradient, which were non-existent in their absence.

5. The stationary convection

When the instability sets in as stationary convection, the marginal state will be characterized by $\sigma = 0$. Putting $\sigma = 0$, the dispersion relation (3.8) reduces to

$$(5.1)\qquad R_1 = \varepsilon^{-1}\left(\frac{G}{G-1}\right)\left[\left(\frac{1+x}{x}\right)\left\{(1+x)^2 + \frac{1+x}{P} + Q_1\right\} + S_1\varepsilon\right],$$

which expresses the modified Rayleigh number R_1 as a function of the dimensionless wave number x and the parameters S_1, Q_1, P and G; let the nondimensional number G accounting for the compressibility effects be also kept fixed. Then we find that

$$(5.2)\qquad \overline{R}_c = \varepsilon^{-1}\left(\frac{G}{G-1}\right) R_c,$$

where $\overline{R}_c$ and R_c denote respectively, the critical Rayleigh number in the presence and in the absence of compressibility. The effect of compressibility consists, thus, in postponing the onset of thermal instability. Hence, compressibility has a

stabilizing effect. $G > 1$ is relevant here. The cases $G < 1$ and $G = 0$ correspond to negative and infinite values of critical Rayleigh numbers in the presence of compressibility, which are not relevant in the present study.

To investigate the effects of magnetic field, medium permeablity and stable solute gradient, we examine the behaviour of dR_1/dQ_1, dR_1/dP and dR_1/dS_1 analytically. It follows from Eq. (5.1) that

$$\frac{dR_1}{dQ_1} = \varepsilon^{-1}\left(\frac{G}{G-1}\right)\left(\frac{1+x}{x}\right), \tag{5.3}$$

$$\frac{dR_1}{dP} = -\varepsilon^{-1}\left(\frac{G}{G-1}\right)\left[\frac{(1+x)^2}{xP^2}\right], \tag{5.4}$$

and

$$\frac{dR_1}{dS_1} = \left(\frac{G}{G-1}\right). \tag{5.5}$$

Thus for stationary convection, the magnetic field and stable solute gradient are found to have stabilizing effects, whereas the medium permeability has a destabilizing effect on the system for $G > 1$.

6. The overstable case

Here we discuss the possibility of whether instability may occur as overstability. Since for overstability we wish to determine the critical Rayleigh numbers for the onset of instability via the state of pure oscillations, it suffices to find the conditions for which (3.8) will admit the solutions with real values of σ_1.

If we equate the real and imaginary parts of (3.8) and eliminate R_1 between them, we obtain

$$A_3C_1^3 + A_2C_1^2 + A_1C_1 + A_0 = 0, \tag{6.1}$$

where we have put $C_1 = \sigma_1^2$, $b = 1 + x$, and

$$A_3 = bP^2E'^2q^2p_2^2\pi^4\left[b\left(\frac{Ep_1}{p_2}+1\right) + Ep_1\left(\frac{\alpha_0\nu_c d^2}{\pi^2\nu}+\frac{1}{P}\right)\right], \tag{6.2}$$

$$A_0 = P^2\pi^4\left[\left\{Ep_1 + (1+\alpha_0)\left(\frac{\nu_c d^2}{\pi^2\nu}\right)^2\right\}b^6 + \left\{\frac{Ep_1}{P}\left(\frac{\nu_c d^2}{\pi^2\nu}\right)^2\right\}b^5 \tag{6.3}$$

$$+ \left\{Q_1(Ep_1 - p_2)\left(\frac{\nu_c d^2}{\pi^2\nu}\right)^2\right\}b^4 + \left\{\varepsilon S_1(b-1)(Ep_1 - E'q)\left(\frac{\nu_c d^2}{\pi^2\nu}\right)^2\right\}b^3\Bigg].$$

Since σ_1 is real for overstability, the three values of $C_1(=\sigma_1^2)$ are positive. Equation (6.1) is cubic in C_1 and the product of the roots is $(-A_0/A_3)$, and if this is to be positive, then $A_0 < 0$ since from (6.2), $A_3 > 0$.

Equation (6.3) shows that this is clearly impossible, i.e. A_0 is always positive if

$$Ep_1 > p_2 \quad \text{and} \quad Ep_1 > E'q, \tag{6.4}$$

which implies that

$$\kappa < \eta E \quad \text{and} \quad E'\kappa < E\kappa'. \tag{6.5}$$

Then if $\kappa < \eta E$ and $E'\kappa < E\kappa'$ are satisfied, overstability is impossible and the principle of exchange of stabilities holds good. $\kappa < \eta E$ and $E'\kappa < E\kappa'$ are, therefore, the sufficient conditions for the non-existence of overstability, the violation of which does not necessarily imply occurrence of overstability.

7. Conclusion

A compressible and partially ionized plasma layer heated and soluted from below in a porous medium occurs pretty often and is of considerable importance in cosmic physics, e.g. in the solar photosphere and chromosphere and in cool interstellar clouds, in geophysics, in atmospheric physics and astrophysics, especially in the case of the ionosphere and the outer layers of the solar atmosphere. For stationary convection, the magnetic field and stable solute gradient are found to have stabilizing effects, whereas the medium permeability has a destabilizing effect on the system for $(C_p/g)\beta > 1$. The effect of compressibility is found to postpone the onset of convection. The oscillatory modes are introduced due to the presence of a magnetic field and a solute gradient, which were non-existent in their absence. The sufficient conditions for the non-existence of overstability has been obtained.

References

1. S. CHANDRASEKHAR, *Hydrodynamic and hydromagnetic stability*, Dover Publications, New York 1981.
2. G. VERONIS, *On finite amplitude instability in thermohaline convection*, J. Marine Res., **23**, 1–17, 1965.
3. H. ALFVÉN, *The origin of the Solar System*, Clarendon Press, Oxford 1954.
4. R.C. SHARMA, *Thermal hydromagnetic instability of a partially ionized medium*, Physica, **81C**, 199–204, 1976.
5. R.C. SHARMA and SUNIL, *Thermosolutal instability of a partially-ionized Hall plasma in porous medium*, Astrophysics and Space Science, **194**, 303–311, 1992.
6. R.C. SHARMA and SUNIL, *Rayleigh–Taylor instability of a partially ionized plasma in a porous medium in presence of a variable magnetic field*, Z. Naturforschung, **47a**, 1227–1231, 1992.

7. E.A. SPIEGEL and G. VERONIS, *On the Boussinesq approximation for a compressible fluid*, Astrophys. J., **131**, 442, 1960.
8. R.C. SHARMA and SUNIL, *Thermal instability of compressible Hall plasma in the presence of suspended particles*, Ganita, **44**, 1, 1–11, 1993.
9. J.A.M. MC DONNEL, *Cosmic dust*, John Wiley and Sons, Toronto 1978, p.330.

DEPARTMENT OF APPLIED SCIENCES,
REGIONAL ENGINEERING COLLEGE, HAMIRPUR, INDIA.

Received May 23, 1994.

Arch. Mech., 46, 6, pp. 829–838, Warszawa 1994

Stability of flow of a viscoelastic fluid over a stretching sheet

B.S. DANDAPAT (CALCUTTA),
L.E. HOLMEDAL and H.I. ANDERSSON (TRONDHEIM)

THE EFFECT of fluid elasticity on the stability of laminar boundary layer flow past a linearly stretching surface is investigated. The non-Newtonian fluid considered is the viscoelastic Walters liquid B''. An exact analytical solution describes the basic flow, which is perturbed by three-dimensional disturbances that are periodic in the lateral direction. A linear stability analysis is performed by means of the Method of Weighted Residuals. The analysis shows that the flow is stable with respect to the Taylor–Görtler-type perturbations considered, and the decay rate of the disturbances increases with increasing wave number. The effect of viscoelasticity is explored, and it is concluded that fluid elasticity has a stabilizing influence on the flow as long as the wavelength of the disturbances does not exceed the viscoelastic length scale.

1. Introduction

THE INVESTIGATION of boundary layer flow of an incompressible viscous fluid caused by the motion of a rigid plane sheet in its own plane was initiated by SAKIADIS [13], who observed that the entrainment of the ambient fluid makes this boundary layer different from that associated with Blasius flow over a fixed flat plate. This study was modelled for applications in the polymer industry when a polymer sheet is extruded continuously from a die, with a tacit assumption that the sheet is *inextensible*. However, in real situations one has to encounter also the boundary layer flow over a *stretching* sheet, since the polymer sheet is sometimes being stretched while being drawn from a thin slit.

To this end CRANE [6] and MC CORMACK and CRANE [11] studied the boundary layer flow of a Newtonian fluid caused by the stretching of an elastic flat sheet, which moves in its own plane with a velocity varying linearly with the distance from a fixed point due to the application of a uniform force. This problem has been extended to a special class of non-Newtonian fluids known as second-order fluids by RAJAGOPAL, NA and GUPTA [12] who obtained similarity solutions of the boundary layer equations numerically. DANDAPAT and GUPTA [7] examined the same problem with heat transfer and found an exact analytical solution of the nonlinear equation governing the self-similar flow and heat transfer. Recently, ANDERSSON and DANDAPAT [2] extended the Newtonian boundary layer flow problem considered by CRANE [6] to an important class of non-Newtonian fluids obeying the inelastic power-law model.

In spite of the growing literature on flow over stretching sheets and its obvious importance in polymer and electrochemical industry, it is surprising to note that the corresponding stability analysis does not seem to have received any adequate

attention so far. Although it is a known fact that unless the above flows are stable, products of the desired physical characteristics can not be achieved. To the best of our knowledge only BHATTACHARYYA and GUPTA [4] and TAKHAR, ALI and GUPTA [14] have investigated the linear stability of viscous flow of a Newtonian fluid over a stretching sheet (whithout and with magnetic field) with respect to three-dimensional disturbances of Taylor-Görtler type.

The aim of the present study is to examine the stability of the flow of a viscoelastic liquid due to the stretching of a sheet while issuing from a thin slit. More specifically, the non-Newtonian fluid is a Walters liquid B'', for which an exact analytical similarity solution exists for the basic two-dimensional boundary layer problem. The linear stability of this basic flow with respect to three-dimensional disturbances is considered by means of the Method of Weighted Residuals.

2. Mathematical formulation and stability analysis

2.1. Rheological model and governing equations

We consider the flow of a viscoelastic fluid which arises due to the stretching of an impermeable flat sheet. The elastic sheet issuing from a thin slit at the origin ($x = 0, y = 0$) is being stretched with a speed proportional to the distance x from the slit, where the x-axis is directed along the sheet. Further we assume the fluid, which is confined to the half-space $y > 0$ above the sheet, to be a Walters liquid B'', and obey the rheological equations of state (WALTERS [16, 17]):

$$p_{ik} = -pg_{ik} + p'_{ik}, \tag{2.1}$$

where p_{ik} is the stress tensor, p is an arbitrary isotropic pressure, g_{ik} is the metric tensor of a fixed coordinate system x_i, and

$$p'_{ik} = 2\eta_0 e^{(1)}_{ik} - 2\kappa_0 \frac{\delta}{\delta t} e^{(1)}_{ik} \tag{2.2}$$

with η_0 being the viscosity of the fluid, κ_0 is the viscoelastic parameter, $e^{(1)}_{ik}$ is the rate-of-strain tensor, and the corotational derivative for a covariant tensor b_{ik} is defined as

$$\frac{\delta b_{ik}}{\delta t} = \frac{\partial b_{ik}}{\partial t} + u_m \frac{\partial b_{ik}}{\partial x_m} - \frac{\partial u_k}{\partial x_m} b_{im} - \frac{\partial u_i}{\partial x_m} b_{mk}, \tag{2.3}$$

where u_i is the velocity vector. It should be pointed out here that the constitutive model (2.1), (2.2) is valid only for short-memory liquids, i.e. for short relaxation times. Using the equations of state (2.1), (2.2), the governing equations of fluid motion in a Cartesian coordinate system can be written in the general form,

following BEARD and WALTERS [3]:

$$\rho\left[\frac{\partial u_i}{\partial t}+u_k\frac{\partial u_i}{\partial x_k}\right]=-\frac{\partial p}{\partial x_i}+\eta_0\frac{\partial^2 u_i}{\partial x_k\partial x_k}-\kappa_0\left[\frac{\partial}{\partial t}\left(\frac{\partial^2 u_i}{\partial x_k\partial x_k}\right)+u_m\frac{\partial^3 u_i}{\partial x_m\partial x_k\partial x_k}-\frac{\partial u_i}{\partial x_m}\frac{\partial^2 u_m}{\partial x_k\partial x_k}-2\frac{\partial^2 u_i}{\partial x_m\partial x_k}\frac{\partial u_m}{\partial u_k}\right], \tag{2.4}$$

where ρ is the constant fluid density.

2.2. Unperturbed state

The problem in hand has basic velocity components $[u_0(x,y), v_0(x,y), 0]$. Assuming the boundary layer approximations, the equations of continuity and momentum can be written as

$$\frac{\partial u_0}{\partial x}+\frac{\partial v_0}{\partial y}=0, \tag{2.5}$$

$$\left(u_0\frac{\partial u_0}{\partial x}+v_0\frac{\partial u_0}{\partial y}\right)=\nu\frac{\partial^2 u_0}{\partial y^2}-\kappa_0^*\left[\frac{\partial}{\partial x}\left(u_0\frac{\partial^2 u_0}{\partial y^2}\right)+\frac{\partial u_0}{\partial y}\frac{\partial^2 v_0}{\partial y^2}+v_0\frac{\partial^3 u_0}{\partial y^3}\right], \tag{2.5'}$$

where $\nu=\eta_0/\rho$ and $\kappa_0^*=\kappa_0/\rho$. The boundary conditions are:

$$u_0=cx, \qquad v_0=0 \quad \text{at} \quad y=0, \tag{2.6}$$

$$u_0\to 0 \quad \text{as} \quad y\to\infty, \tag{2.7}$$

where c is a positive constant. In deriving Eqs. (2.5), (2.5') it was assumed that, in addition to the usual boundary layer approximations, the contribution due to the normal stresses is of the same order of magnitude as that due to the shear stresses. Thus, both ν/c and κ_0^* are $0(\delta^2)$, δ being the thickness of the boundary layer which forms over the stretching sheet.

An exact analytical similarity solution of the above system was given by DANDAPAT and GUPTA [7] as

$$u_0=cxf'(\eta), \qquad v_0=-(c\nu)^{1/2}f(\eta), \qquad \eta=\left(\frac{c}{\nu}\right)^{1/2}y, \tag{2.8}$$

where the dimensionless stream function f becomes

$$f(\eta)=\frac{1}{Q}\left(1-e^{-Q\eta}\right), \qquad Q=(1-\kappa_0^*c/\nu)^{-1/2}. \tag{2.9}$$

The basic flow is therefore crucially dependent upon the dimensionless parameter Q, for which the deviation from unity represents a measure of the degree of viscoelasticity.

2.3. Stability analysis

We shall now examine the stability of the above system of solutions (2.8) – (2.9) for three-dimensional disturbances. The similarity of this form of solution with that of two-dimensional stagnation-point flow suggests, following GÖRTLER [9] and HÄMMERLIN [10], that instabilities may occur in the form of Taylor – Görtler vortices. Let us therefore consider the perturbed state as

$$\begin{aligned} \overline{u} &= u_0(x, y) + u(x, y, z, t), \\ \overline{v} &= v_0(x, y) + v(x, y, z, t), \\ \overline{w} &= w(x, y, z, t), \\ \overline{p} &= p_0 + p(x, y, z, t), \end{aligned} \tag{2.10}$$

where u_0 and v_0 are given by (2.8) and (2.9), p_0 denotes the constant basic pressure and w is the perturbation velocity component along the z-direction. Following GÖRTLER [9], we assume that all the perturbed quantities have periodicity in the direction normal to the basic flow, with usual exponential time-dependence as follows:

$$\begin{aligned} u &= cx u_1(\eta) \cos(\alpha z) e^{\beta t}, \\ v &= -(c\nu)^{1/2} v_1(\eta) \cos(\alpha z) e^{\beta t}, \\ w &= \nu \alpha w_1(\eta) \sin(\alpha z) e^{\beta t}, \\ p &= \rho \nu c p_1(\eta) \cos(\alpha z) e^{\beta t}. \end{aligned} \tag{2.11}$$

Let us now introduce the perturbed state expressed by Eqs. (2.10) into the three-dimensional continuity equation and the time-dependent momentum equation (2.4). After linearization and subtraction of the basic solution (2.8), a set of partial differential equations for the perturbations is obtained. Provided the disturbances are of the Görtler-type (2.11), the perturbation equations reduce to the following set of ordinary differential equations:

$$u_1 - v_1' + \overline{\alpha}^2 w_1 = 0, \tag{2.12}$$

$$\begin{aligned} u_1'' + f u_1' - \left(\overline{\alpha}^2 + 2f' + \overline{\beta}\right) u_1 + f'' v_1 + \gamma \Big[f u_1''' - \left(\overline{\beta} + 2f'\right) u_1'' \\ - \left(\overline{\alpha}^2 f - 3f''\right) u_1' + \overline{\beta}\overline{\alpha}^2 u_1 - f'' v_1'' - 2f''' v_1' + \left(\overline{\alpha}^2 f'' + f''''\right) v_1 \Big] = 0, \end{aligned} \tag{2.13}$$

$$\begin{aligned} v_1'' + f v_1' + \left(f' - \overline{\beta} - \overline{\alpha}^2\right) v_1 + \gamma \Big[f v_1''' - \left(\overline{\beta} + 3f'\right) v_1'' \\ - \left(\overline{\alpha}^2 f + 3f''\right) v_1' + \left(\overline{\beta}\overline{\alpha}^2 + f''' + \overline{\alpha}^2 f'\right) v_1 \Big] = -p_1', \end{aligned} \tag{2.14}$$

$$\begin{aligned} w_1'' + f w_1' - \left(\overline{\alpha}^2 + \overline{\beta}\right) w_1 + \gamma \Big[f w_1''' - \left(\overline{\beta} + 2f'\right) w_1'' \\ - \left(\overline{\alpha}^2 f + f''\right) w_1' + \overline{\beta}\overline{\alpha}^2 w_1 \Big] = -p_1, \end{aligned} \tag{2.15}$$

where a prime denotes the derivative with respect to η, and

$$\gamma = \kappa_0^* c/\nu, \qquad \overline{\alpha}^2 = \alpha^2 \nu/c, \qquad \overline{\beta} = \beta/c$$

are dimensionless viscoelasticity, wave number, and growth rate, respectively. The respective boundary conditions are

$$u_1 = v_1 = w_1 = 0 \quad \text{at} \quad \eta = 0 \quad \text{and} \quad \eta \to \infty \tag{2.16}$$

which assure that the disturbances vanish at infinity and conform to no-slip at the sheet. Using (2.16) in (2.12) gives $v_1' = 0$ at $\eta = 0$ and $\eta = \infty$. Thus, the boundary conditions become

$$u_1 = v_1 = v_1' = 0 \quad \text{at} \quad \eta = 0 \quad \text{and} \quad \eta \to \infty. \tag{2.17}$$

Now, w_1 is eliminated from (2.12) and (2.15) to obtain an equation for p_1. Then p_1' can be eliminated from (2.14), and the resulting equation combines with the differentiated equation (2.13) to yield:

$$\begin{aligned} &v_1'''' + f v_1''' + \left(f' - 2\overline{\alpha}^2 - \overline{\beta}\right) v_1'' + \left(f'' - \overline{\alpha}^2 f\right) v_1' \\ &\quad + \left\{f''' + \overline{\alpha}^2\left(\overline{\beta} + \overline{\alpha}^2 - f'\right)\right\} v_1 - 2f' u_1' - 2f'' u_1 + \gamma\left[4f'' u_1'' + 4f''' u_1'\right. \\ &+ f v_1'''''' - \left(\overline{\beta} + f'\right) v_1'''' - \left(4f'' + 2\overline{\alpha}^2 f\right) v_1''' + \left(2\overline{\beta}\overline{\alpha}^2 + 2\overline{\alpha}^2 f' - 4f'''\right) v_1'' \\ &\quad + \left\{\overline{\alpha}^2 f'' - f'''' + \overline{\alpha}^2\left(\overline{\alpha}^2 f + 3f''\right)\right\} v_1' + \left\{f'''''' + \overline{\alpha}^2 f'''\right. \\ &\quad \left.\left. - \overline{\alpha}^2\left(\overline{\beta}\overline{\alpha}^2 + f''' + \overline{\alpha}^2 f'\right)\right\} v_1\right] = 0. \end{aligned} \tag{2.18}$$

Following BHATTACHARYYA and GUPTA [4] and introducing the transformation:

$$T = e^{-Q\eta}, \qquad L = -T\frac{d}{dT} \tag{2.19}$$

in the final set (2.13) and (2.18), we obtain

$$\begin{aligned} &Q^2 L^2 u_1 + (1-T) L u_1 - \left(\overline{\alpha}^2 + 2T + \overline{\beta}\right) u_1 - QT v_1 \\ &\quad + \gamma\left[Q^2(1-T) L^3 u_1 - Q^2\left(\overline{\beta} + 2T\right) L^2 u_1 - \left\{\overline{\alpha}^2(1-T) + 3Q^2 T\right\} L u_1\right. \\ &\quad \left. + \overline{\beta}\overline{\alpha}^2 u_1 + Q^3 T L^2 v_1 - 2Q^3 T L v_1 - QT\left(Q^2 + \overline{\alpha}^2\right) v_1\right] = 0 \end{aligned} \tag{2.20}$$

and

$$\begin{aligned} &Q^4 L^4 v_1 + Q^2(1-T) L^3 v_1 + Q^2\left(T - 2\overline{\alpha}^2 - \overline{\beta}\right) L^2 v_1 \\ &\quad - \left\{Q^2 T + \overline{\alpha}^2(1-T)\right\} L v_1 + \left\{Q^2 T + \overline{\alpha}^2\left(\overline{\beta} + \overline{\alpha}^2 - T\right)\right\} v_1 \end{aligned} \tag{2.21}$$

(2.21) [cont.]
$$-2QTLu_1 + 2QTu_1 + \gamma\left[Q^4(1-T)L^5v_1 - Q^4\left(\overline{\beta}+T\right)L^4v_1 + Q^2\left\{4Q^2T - 2\overline{\alpha}^2(1-T)\right\}L^3v_1 + Q^2\left\{2\overline{\alpha}^2\overline{\beta} + T\left(2\overline{\alpha}^2 - 4Q^2\right)\right\}L^2v_1 - \left\{\overline{\alpha}^2Q^2T - Q^4T - \overline{\alpha}^2\left[\overline{\alpha}^2(1-T) - 3Q^2T\right]\right\}Lv_1 + \left\{Q^5T + \overline{\alpha}^2Q^2T - \overline{\alpha}^2\left[\overline{\beta}\overline{\alpha}^2 + T\left(Q^2+\overline{\alpha}^2\right)\right]\right\}v_1 - 4Q^3TL^2u_1 + 4Q^3TLu_1\right] = 0,$$

while the transformed boundary conditions become

(2.22) $$u_1 = v_1 = Lv_1 = 0 \quad \text{at} \quad T = 0 \quad \text{and} \quad T = 1.$$

At this stage one may recall that BHATTACHARYYA and GUPTA [4] considered the same problem for a Newtonian fluid. Thus, the above eigenvalue problem reduces to that of BHATTACHARYYA and GUPTA [4] if $\gamma = 0$ and $Q = 1$.

2.4. Solution technique

In order to solve the eigenvalue problem defined by the set (2.20) – (2.22), we shall use the Method of Weighted Residuals described, for example, by FINLAYSON [8]. Both u_1 and v_1 are expanded in terms of trial functions which should satisfy the boundary conditions (2.22) in the form

(2.23) $$u_1 = \sum_{i=1}^{\infty} A_i u_i(T) \quad \text{and} \quad v_1 = \sum_{i=1}^{\infty} B_i v_i(T).$$

By substituting (2.23) in (2.20) and (2.21) and using the respective orthogonality conditions, i.e. the residues are orthogonal to the trial functions, we finally obtain the set of equations

(2.24) $$A_iC_{ij} + B_iD_{ij} = 0,$$

(2.25) $$A_iE_{ij} + B_iF_{ij} = 0,$$

where the coefficients C_{ij}, D_{ij}, E_{ij} and F_{ij} are defined in Appendix I.

The Eqs. (2.24) and (2.25) have a non-trivial solution only if the determinant of the coefficients of A_i and B_i vanishes, i.e.

(2.26) $$\begin{vmatrix} C_{ij} & D_{ij} \\ E_{ij} & F_{ij} \end{vmatrix} = 0.$$

Now, the trial functions

(2.27) $$u_i = T^i(1-T) \quad \text{and} \quad v_i = T^i(1-T)^2$$

used by BHATTACHARYYA and GUPTA [4] were adopted, and it is readily observed that $u_i(T)$ and $v_i(T)$ satisfy (2.22). For a given combination of γ and $\overline{\alpha}$, the growth rate $\overline{\beta}$ can be determined as the eigenvalue of the matrix in Eq. (2.26). Although qualitatively correct results can be obtained even by using 1-term or 2-term expansions, see for example BHATTACHARYYA and GUPTA [4], the present results have been obtained by using 10-term trial functions.

3. Results and discussions

The dimensionless decay rate of the disturbances $-\overline{\beta}$, has been plotted versus the dimensionless wave number $\overline{\alpha}$ for different values of the viscoelastic parameter γ in Fig. 1. Since the decay rate $-\overline{\beta}$ is positive throughout the wave-number

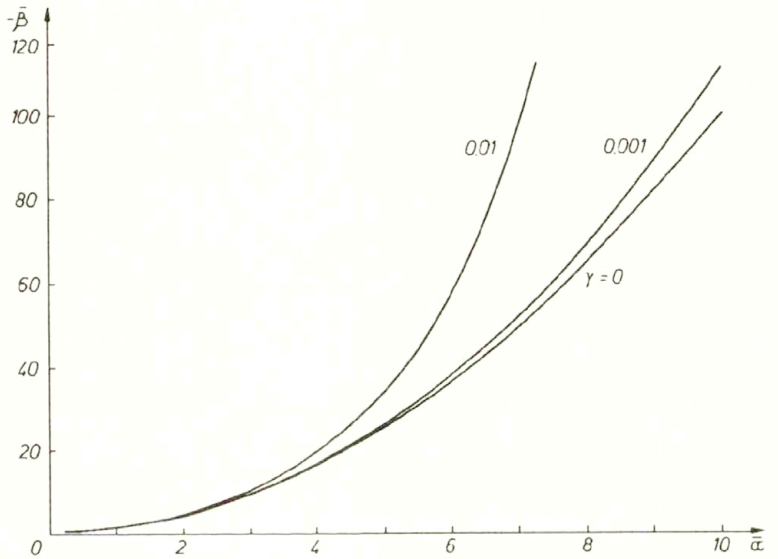


FIG. 1. Predicted variation of decay rate versus wave number for different values of the viscoelastic parameter.

range considered, i.e. the growth rate $+\beta$ is negative, there are no unstable modes and the flow is stable to infinitesimal disturbances of the Taylor–Görtler type. This is in keeping with the classical stability arguments for flow along curved streamlines given by von KÁRMÁN [15]. Based on simple "displaced particle" arguments, he considered a fluid element being displaced outwards from the centre of curvature with its angular momentum (product of local velocity and local radius of streamline curvature) being conserved. Then, if the angular momentum of the flow increases outwards, the velocity of the perturbed element is lower than its surroundings and the radial pressure gradient will force the displaced element inwards to its original position, i.e. the situation is stable. Thus, although geometrically similar, the centrifugal instability tends to destabilize the stagnation point flow considered by GÖRTLER [9] and HÄMMERLIN [10], while the flow along the stretching sheet is being stabilized by the streamline curvature. Although these considerations are based on purely frictionless analysis, it may nevertheless be conjectured that this inviscid mechanism is the most influential even in the actual viscous flow. The variation of the decay rate $-\overline{\beta}$ with $\overline{\alpha}$ in Fig. 1 shows moreover that high wave number disturbances are more effectively damped than low wave-number modes. In fact, $-\overline{\beta}$ tends asymptotically to a constant value of 0.735 as $\overline{\alpha}$ approaches zero.

The stability curves for $\gamma > 0$ show that, for a given wave number, the decay rate $-\overline{\beta}$ increases with increasing parameter values γ. Moreover, it can be seen that the increased damping is most prominent at high $\overline{\alpha}$, i.e. for disturbances whose wavelength is shorter than the viscous length-scale $L = (\nu/c)^{1/2}$. Since a certain amount of the energy associated with the disturbances is used to overcome the inherent force due to fluid elasticity, less energy becomes available to induce instabilities. Thus, with the provision that γ is small, it can be concluded that viscoelasticity has a stabilizing influence on the disturbances.

The dimensionless parameter γ can be identified as a Weissenberg number or, alternatively, as the squared ratio between the viscoelastic length-scale $(\kappa_0^*)^{1/2}$ and the viscous length L. Stabilization due to viscoelasticity does obviously not occur when the wavelength of the disturbances is large compared with the viscoelastic length-scale, i.e. for $\overline{\alpha}^2 \ll 1$.

The present findings are in contrast to the results of CHAN MAN FONG and WALTERS [5]. They considered the stability of plane Poiseuille flow with respect to wave-like disturbances and concluded that, according to linear theory, the presence of viscoelasticity has a destabilizing effect on the parallel flow. However, the difference in stability behaviour between the flow past a stretching sheet and plane Poiseuille flow is due to the absence of streamline curvature in the latter case.

It may be interesting to recall that TAKHAR *et al.* [14] investigated the effect of a uniform magnetic field on the stability of the flow over a stretching sheet with respect to Taylor–Görtler-like disturbances. They found that the decay rate of the imposed disturbances increased with increasing magnetic field, which therefore turned out to have a stabilizing influence on the flow. This is not surprising since ANDERSSON [1] observed that an external magnetic field has the same effect on the unperturbed basic flow as fluid viscoelasticity, namely to reduce the boundary layer thickness as well as the velocity within the boundary layer.

4. Closing remarks

Finally, we would like to emphasize that the rheological model considered, i.e. the Walters liquid B'', was developed for viscoelastic substances with short memories. Moreover, the non-Newtonian behaviour due to viscoelasticity was regarded as a perturbation of the Newtonian viscous flow, so that the viscoelastic parameter γ should always be appreciably below unity.

It has been observed (ANDERSSON [1]) that the same boundary layer equations for the steady and two-dimensional basic flow u_0, v_0 can be derived for a second-order fluid with gradually fading memory. However, in spite of the formal equivalence between flows of a second-order fluid and Walters liquid B'' over a stretching sheet, the rheological models are completely different and so are the general equations of motion. The results of the present analysis can therefore not be carried over to second-order fluids.

Appendix I

Definitions of C_{ij}, D_{ij}, E_{ij} and F_{ij} in Eqs. (2.23) and (2.24) are provided in the following. Here, the simplifying notation $< f >= \int_0^1 f\, dT$ and $D = d/dT$ has been introduced for convenience.

$$
\begin{aligned}
C_{ij} &= a_1 \langle u_j T D u_i \rangle + a_2 \left\langle u_j T^2 D u_i \right\rangle + a_3 \left\langle u_j T^2 D^2 u_i \right\rangle + a_4 \left\langle u_j T^3 D^2 u_i \right\rangle \\
&\quad - a_5 \left\langle u_j T^3 D^3 u_i \right\rangle + a_6 \left\langle u_j T^4 D^3 u_i \right\rangle - a_7 \langle u_j T u_i \rangle - a_8 \langle u_j u_i \rangle , \\
D_{ij} &= b_1 \left\langle u_j T^3 D^2 v_i \right\rangle + b_2 \left\langle u_j T^2 D v_i \right\rangle - b_3 \langle u_j T v_i \rangle , \\
E_{ij} &= a_9 \langle v_j T u_i \rangle + a_{10} \left\langle v_j T^2 D u_i \right\rangle - a_{11} \left\langle v_j T^3 D^2 u_i \right\rangle , \\
F_{ij} &= b_4 \langle v_j T D v_i \rangle + b_5 \left\langle v_j T^2 D v_i \right\rangle + b_6 \left\langle v_j T^2 D^2 v_i \right\rangle + b_7 \left\langle v_j T^3 D^2 v_i \right\rangle \\
&\quad + b_8 \left\langle v_j T^3 D^3 v_i \right\rangle + b_9 \left\langle v_j T^4 D^3 v_i \right\rangle + b_{10} \left\langle v_j T^4 D^4 v_i \right\rangle + b_{11} \left\langle v_j T^5 D^4 v_i \right\rangle \\
&\quad - b_{12} \left\langle v_j T^5 D^5 v_i \right\rangle + b_{13} \left\langle v_j T^6 D^5 v_i \right\rangle + b_{14} \langle v_j T v_i \rangle + b_{15} \langle v_j v_i \rangle ,
\end{aligned}
$$

where the coefficients a_i and b_i are given by the formulae

$$
\begin{aligned}
a_1 &= Q^2 - 1 + \gamma \left(\overline{\alpha}^2 - Q^2 \overline{\beta} - Q^2 \right) , \\
a_2 &= 1 + \gamma \left(2Q^2 - \overline{\alpha}^2 \right) , \qquad a_3 = Q^2 - \gamma \left(3Q^2 + Q^2 \overline{\beta} \right) , \\
a_4 &= \gamma Q^2 , \qquad a_5 = \gamma Q^2 , \qquad a_6 = \gamma Q^2 , \\
a_7 &= 2 , \qquad a_8 = \overline{\alpha}^2 + \overline{\beta} - \gamma \overline{\beta} \overline{\alpha}^2 , \\
a_9 &= 2Q , \qquad a_{10} = 2Q - 8\gamma Q^3 , \qquad a_{11} = 4\gamma Q^3 , \\
b_1 &= \gamma Q^3 , \qquad b_2 = 3\gamma Q^3 , \qquad b_3 = Q + \gamma \left[Q \left(Q^2 + \overline{\alpha}^2 \right) \right] , \\
b_4 &= Q^2 \left(Q^2 - 1 - 2\overline{\alpha}^2 - \overline{\beta} \right) + \overline{\alpha}^2 + \gamma \left[Q^2 \left(1 + \overline{\beta} \right) \left(2\overline{\alpha}^2 - Q^2 \right) - \overline{\alpha}^4 \right] , \\
b_5 &= 3Q^2 - \overline{\alpha}^2 + \gamma \left[\overline{\alpha}^2 \left(4Q^2 + \overline{\alpha}^2 \right) - 9Q^4 \right] , \\
b_6 &= Q^2 \left(7Q^2 - 3 - 2\overline{\alpha}^2 - \overline{\beta} \right) + \gamma \left[Q^2 \overline{\alpha}^2 \left(6 + 2\overline{\beta} \right) - Q^4 \left(15 + 7\overline{\beta} \right) \right] , \\
b_7 &= 4Q^2 - \gamma \left(8Q^4 + 4Q^2 \overline{\alpha}^2 \right) , \\
b_8 &= Q^2 (6Q^2 - 1) - \gamma \left[Q^2 \left(25Q^2 + 6\overline{\beta} Q^2 - 2\overline{\alpha}^2 \right) \right] , \\
b_9 &= Q^2 + \gamma \left[Q^2 \left(15Q^2 - 2\overline{\alpha}^2 \right) \right] , \\
b_{10} &= Q^4 - \gamma \left[Q^4 \left(10 + \overline{\beta} \right) \right] , \qquad b_{11} = 9\gamma Q^4 , \\
b_{12} &= \gamma Q^4 , \qquad b_{13} = \gamma Q^4 , \qquad b_{14} = Q^2 - \overline{\alpha}^2 + \gamma \left(Q^5 - \overline{\alpha}^4 \right) , \\
b_{15} &= \overline{\alpha}^2 \left(\overline{\beta} + \overline{\alpha}^2 \right) - \gamma \overline{\alpha}^4 \overline{\beta} .
\end{aligned}
$$

Acknowledgement

The authors are indebted to Prof. A.S. GUPTA and Dr. F. IRGENS for valuable comments and stimulating discussions.

References

1. H.I. ANDERSSON, *MHD flow of a viscoelastic fluid past a stretching surface*, Acta Mechanica, **95**, 227–230, 1992.
2. H.I. ANDERSSON and B.S. DANDAPAT, *Flow of a power-law fluid over a stretching sheet*, Stability Appl. Anal. Continuous Media, **1**, 339–347, 1991.
3. D.W. BEARD and K. WALTERS, *Elastico-viscous boundary-layer flows I. Two-dimensional flow near a stagnation point*, Proc. Camb. Phil. Soc., **60**, 667–674, 1964.
4. S.N. BHATTACHARYYA and A.S. GUPTA, *On the stability of viscous flow over a stretching sheet*, Quart. Appl. Math., **43**, 359–367, 1985.
5. C.F. CHAN MAN FONG and K. WALTERS, *The solution of flow problems in the case of materials with memory. Part II*, J. Mécanique, **4**, 439–453, 1965.
6. L.J. CRANE, *Flow past a stretching plate*, J. Appl. Math. Phys. (ZAMP), **21**, 645–647, 1970.
7. B.S. DANDAPAT and A.S. GUPTA, *Flow and heat transfer in a viscoelastic fluid over a stretching sheet*. Int. J. Non-linear Mech., **24**, 215–219, 1989.
8. B.A. FINLAYSON, *The method of weighted residuals and variational principles*, Academic Press, New York 1972.
9. H. GÖRTLER, *Dreidimensionale Instabilität der ebenen Staupunktströmung gegenüber wirbelartigen Störungen*, 50 Jahre Grenzschichtforschung. Friedr. Vieweg und Sohn, Braunschweig, pp. 304–314, 1955.
10. A. HÄMMERLIN, *Zur Instabilitätstheorie der ebenen Staupunktströmung*, 50 Jahre Grenzschichtforschung. Friedr. Vieweg und Sohn, Braunschweig, pp. 315–327, 1955.
11. P.D. MCCORMACK and L. CRANE, *Physical fluid dynamics*, Academic Press, New York 1970.
12. K.R. RAJAGOPAL, T.Y. NA and A.S. GUPTA, *Flow of a viscoelastic fluid over a stretching sheet*, Rheol. Acta, **23**, 213–215, 1984.
13. B.C. SAKIADIS, *Boundary-layer behaviour on continuous solid surfaces. I. Boundary-layer equations for two-dimensional and axisymmetric flow*, AIChE J., **7**, 26–28, 1961.
14. H.S. TAKHAR, M.A. ALI and A.S. GUPTA, *Stability of magnetohydrodynamic flow over a stretching sheet*, Liquid Metal Hydrodynamics, J. LIELPETERIS and R. MOREAU [Eds.], Kluwer, Dordrecht, pp. 465–471, 1989.
15. T. VON KÁRMÁN, *Some aspects of the turbulence problem*, Proc. 4th Int. Congr. Appl. Mech. Cambridge, pp. 34, 1934.
16. K. WALTERS, *The motion of an elastico-viscous liquid contained between coaxial cylinders (II)*, Quart. J. Mech. Appl. Math., **113**, 4, 444–461, 1960.
17. K. WALTERS, *Non-Newtonian effects in some elastico-viscous liquids whose behaviour at small rates of shear is characterized by a general linear equation of state*, Quart. J. Mech. Appl. Math., **15**, 63–76, 1962.

PHYSICS AND APPLIED MATHEMATICS UNIT
INDIAN STATISTICAL INSTITUTE, CALCUTTA, INDIA
and
FACULTY OF MECHANICAL ENGINEERING
THE NORWEGIAN INSTITUTE OF TECHNOLOGY, TRONDHEIM, NORWAY.

Received June 21, 1994.

Arch. Mech., 46, 6, pp. 839–854, Warszawa 1994

On the interface modelling of crystal growth processes Part I. Thermostatic considerations

W. KOSIŃSKI (WARSZAWA) and S. TABACOVA(*) (NAPLES)

THE BASIC SET of balance laws governing the phase change processes have been presented using both the singular surface and layer approaches. As a particular case, the solidification of a sphere from melt in a quasi-static formulation has been investigated. In the layer approach the thickness of the finite slab has been related to a small parameter, which depends only on the material constants. The interfacial layer has been regarded as a boundary layer between the liquid and the solid phase. It has been found that the solutions corresponding to the singular surface approach are the zero-order terms of the asymptotic expansions in the small parameter of the layer approach solutions.

Basic notations in the order introduced

G	domain containing material,
G_t^-	solid bulk phase in the singular surface approach,
G_t^+	liquid bulk phase in the singular surface approach,
S_t	phase-change interface surface,
Σ_t	reference surface for layer approach,
$\mathbf{v}$	particle (bulk) velocity,
$\mathbf{V}$	particle (surface) velocity,
H	mean curvature of the surface S_t or Σ_t,
c_n	normal speed of the surface S_t or Σ_t,
ψ	general density functions in $G\backslash S_t$,
ψ_s	general density function on S_t for singular surface approach,
$\mathbf{w}$	general density efflux in $G\backslash S_t$,
$\mathbf{w}_s$	general density efflux on S_t for singular surface approach,
t	time [sec],
$\widehat{G}_t^{\pm}$	liquid (+) and solid (−) bulks in the layer approach,
$\mathbf{n}$	unit normal of S_t or Σ_t oriented towards G_t^+ or $\widehat{G}_t^+$,
$\mathbf{V}_{\|}$	tangential part of $\mathbf{V}$,
div_s	surface divergence operator defined on S_t or Σ_t,
$\frac{\delta}{\delta t}$	displacement derivative,
(r, ϕ, φ)	spherical coordinates in G,
$R(t)$	position of S_t or Σ_t in time [cm],
$\mathbf{x}(\phi, \varphi, t)$	convected parametric representation of S_t or Σ_t,
ρ	mass density [g/cm^3],
$\mathbf{T}$	Cauchy stress tensor,
e	specific internal energy [J/g],
$\mathbf{q}$	heat flux [W/cm^2],
θ	absolute temperature [deg],
κ	thermal conduction coefficient [W/deg cm],

(*) On leave from Technical University – Plovdiv, Sankt Petersburg 61, 4000 Plovdiv, Bulgaria.

$\kappa^{\pm}$ thermal conduction coefficients in $G_t^{\pm}$ or $\widehat{G}_t^{\pm}$ [W/deg cm],
$\mathbf{S}$ surface stress tensor on S_t,
σ surface tension coefficient of S_t,
grad_s surface gradient operator on S_t or Σ_t,
$\mathbf{1}_s$ metric tensor of S_t or Σ_t,
Z_t layer dividing the two bulks,
$S_t^{\pm}$ boundaries of Z_t,
$z^{\pm}$ distance of $S_t^{\pm}$ from Σ_t,
l parameter measuring the distance of points in Z_t from Σ_t-layer coordinate in Z_t,
ψ^s general density function on Σ_t in layer approach,
$j(l, \mathbf{x})$ ratio between the surface elements of Z_t and of Σ_t,
K Gaussian curvature of S_t or Σ_t,
$\mathbf{b}$ curvature tensor of S_t or Σ_t,
R^* spherical container radius [cm],
θ^{∞} spherical container wall temperature [deg],
θ_m solidification temperature [deg],
ε crystal seed radius [cm],
Q_0 negative heat flux from the crystal seed at the center of the solid sphere G_t^- or $\widehat{G}_t^-$ [W],
$\rho^{\pm}$ mass density in $G_t^{\pm}$ or $\widehat{G}_t^{\pm}$ [g/cm^3],
$\theta_{\min}$ minimum absolute temperature on the crystal seed at the center of the solid sphere G_t^- or $\widehat{G}_t^-$ [cm],
$e^{\pm}$ specific internal energy in $G_t^{\pm}$ or $\widehat{G}_t^{\pm}$ [J/g],
L latent heat [J/g],
R_0 initial radius of the solidified sphere G_t^- [cm],
R_f final radius of the solidified sphere G_t^- [cm],
$\Delta\theta$ temperature change in the layer Z_t [deg],
τ $= R_0/V_S$, solidification time [sec],
V_S solidification velocity [cm/sec],
PeS $= R_0^2\rho^+ L/\Delta\theta\kappa^+\tau \gg 1$, dimensionless parameter,
δ $= (\mathrm{PeS})^{-1/2} \ll 1$, dimensionless interfacial layer thickness – small parameter,
$c_v^{\pm}$ heat capacities in $G_t^{\pm}$ or $\widehat{G}_t^{\pm}$ [J/g deg].

1. Introduction

THE PROBLEM of crystal growth mechanism for most of the cases is still an open question, due to the indirect character of the experimental data. Generally, two types of interfaces could be observed (BRICE [4]): the diffusive ones, where the phase-change occupies several atomic layers, and the thermodynamic properties of the atomic layers gradually change from the properties corresponding to the melt (liquid phase) to the properties of the crystal (solid phase); and the second one – sharp (singular) interfaces regarded at an abrupt transition when the phase-change takes place only in one atomic layer and the crystal faces are then almost atomically flat. The transition from one type to another, or the existence of one of these types depends mainly on the entropy of melting.

Several methods modelling the interfaces at phase-change processes could be found in the literature. Some of those methods are rather general and are used

to model different physical or physico-chemical processes taking place between two immiscible fluids, or between solids and fluids – accounting for adsorption, solidification or melting. We shall try to adapt those models for the two observed types of interfaces at crystal growth from melt.

In the case of macroscopic equilibrium, the Gibbs' modelling (GIBBS [13]) of the interface as a surface with a uniform tension is quite satisfactory and is still applied by many authors. Apart from Gibbs' classical models, in the modern literature three classes of dynamic models exist:

i) A singular surface approach (SCRIVEN [23], SLATTERY [24], MOECKEL [20], NAPOLITANO [21], ROMANO [22], KOSIŃSKI [15], KOSIŃSKI [16]), in which the interface is modelled as a $2D$ region, the dynamics of which is analogous to that of the ordinary $3D$ world of bulk phases, but with new physical fields on the interface. Moreover, the mechanical and thermodynamic balance laws with surface singularities are formulated on it.

ii) A layer approach (BUFF [5], ALBANO *et al.* [2], ADAMSON [1], ALTS and HUTTER [3]) – on the basis of the extended Gibbs' ideas, the fields on singular surfaces are correlated with mean values of excess fields defined over the $3D$ interface region of finite thickness. In this approach the interface region is a kind of boundary layer, in which the excess quantities represent the differences between the actual interfacial fields and averages of extrapolations of bulk fields into the layer. A dividing surface is located somewhere in the transition (interface) zone and then the bulk quantities are extrapolated up to this surface by stipulating. The main problem of this model consists in introducing surface excess densities (quantities) to compensate the error introduced by replacing the exact (true) quantities by the extrapolated quantities in the transition zone.

iii) A finite slab approach (GATIGNOL [12], GATIGNOL and SEPPECHER [11], KOSIŃSKI and ROMANO [17], DELL'ISOLA and KOSIŃSKI [9], DELL'ISOLA and KOSIŃSKI [10], KOSIŃSKI [18]), in which two dividing surfaces are introduced, making the boundary between the single phase bulk media and the interface zone; in the latter case, a multi-phase behaviour is observed, in which the confined matter possesses constitutive properties different from the surrounding bulk phases and the field quantities of the layer are the true ones. Consequently, in dealing with this model the formulation of an initial-boundary value problem will be different from that in which excess quantities appear.

In second and third models an averaging procedure is applied in which integration along the thickness is performed to get mean quantities defined as surface fields. In the second model one relates those quantities to deviations between exact and extrapolated quantities in the layer, and in the third, the mean quantities are defined as line integrals of the exact fields on some reference (e.g. mean) surface located between the previous two.

All these models are phenomenological and treat only the macroscopic approach without entering into the microscopic level of the processes.

The aim of the present paper is to compare the singular surface approach and

the finite slab one in modelling interfacial phenomena and to show the applicability of the second one. Moreover, by solving a rather simple solidification problem we want to show that a solution corresponding to the singular surface approach is the zero order term of the asymptotic expansion in the small parameter of the solution in the finite slab approach.

As a particular case we shall observe the solidification of an elastic sphere and its interface evolution in time. A comparison is made between the results obtained by both approaches. It is shown that under some physical hypotheses the interfacial layer could be treated as a "phase-change" boundary layer and its thickness depends only on the solidification mechanism specific for a given material. The evolution time in the singular interface approach could be regarded as the zero order term in the asymptotic expansion (in a small parameter related to the layer thickness) of the evolution time corresponding to the finite slab interface approach.

2. Singular surface approach

In mechanics of continua the concept of a moving surface carrying disturbances in fields is a well-known and often used model of a wave front. A moving surface, however, may carry not only disturbances, but also physical properties different from those of the surrounding media. As an example, one can consider the adsorption phenomena or the direct interaction of two different phases of a material which can be modelled by the motion of a surface separating two well-behaved material (bulk) media, while attributing to the surface the physical properties of a phase change. Such a model will be first presented here with some major simplification by neglecting motion of the media.

Consider a material that consists of two different pure phases: solid (G_t^-) and liquid (G_t^+), occupying a region $G := G_t^- \cup G_t^+ \cup S_t$ in the $3D$ space E^3 as shown in the Fig. 1. The interface S_t is the melting isotherm surface and G is a closed material body. Let $\mathbf{v}$ (see notations) denote the particle (bulk) velocity in $G \setminus S_t$. If the interface possesses a non-vanishing surface mass distribution (like in adsorption problems), then the interfacial momentum appears in the analysis and an extra velocity vector $\mathbf{V}$ of "surface" particles needs to be introduced. Let H denote the distribution of the mean curvature of the surface S_t, and c_n – its normal speed of the displacement, which measures the rate of change of the position of S_t at different times in the normal direction of the surface. Note that c_n is the only intrinsic kinematical quantity – a scalar field – for the moving surface $\{S_t\}$.

For a general density function ψ with efflux $\mathbf{w}$ defined in $G \setminus S_t$ and its concentration ψ_s with efflux $\mathbf{w}_s$ on S_t, in the absence of source fields we have the local balance law split into two parts (KOSIŃSKI [15])

$$\frac{\partial \psi}{\partial t} + \operatorname{div}(\psi \mathbf{v} + \mathbf{w}) = 0 \qquad \text{in} \quad G \setminus S_t, \tag{2.1}$$

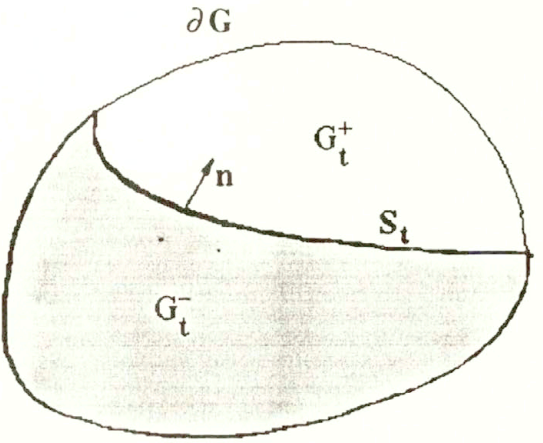


FIG. 1. Singular surface interface.

$$\frac{\delta \psi_s}{\delta t} + \operatorname{div}_s(\psi_s \mathbf{V}_{||} + \mathbf{w}_s) - 2c_n H \psi_s = [\![\psi(\mathbf{v}\cdot\mathbf{n} - c_n)]\!] + [\![\mathbf{w}\cdot\mathbf{n}]\!] \qquad \text{on} \quad S_t; \tag{2.2}$$

here $\mathbf{V}_{||}$ is the tangential part of $\mathbf{V}$, div_s denotes the surface divergence operator defined on S_t, $\frac{\delta}{\delta t}$ is the displacement derivative (KOSIŃSKI [16]) and the square brackets $[\![\cdot]\!]$ denote the jump of the function inside of them. Moreover, the flux $\mathbf{w}_s$ has only tangential components. In the further part the velocity vector $\mathbf{V}$ will be neglected and, due to the particular choice of the parametric representation of the moving surface S_t (i.e. as the family of parallel spheres in the convected parametrization $r = R(t)$), the displacement derivative will be identical with the partial time derivative.

In the case when the balance of mass ψ equals ρ – the bulk mass density function, $\mathbf{w}$ vanishes and then we have:

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{v}) = 0. \tag{2.3}$$

For the linear momentum $\mathbf{w}$ equals the negative Cauchy stress tensor $\mathbf{T}$, while ψ is equal to the product $\rho\mathbf{v}$. In the case of balance of energy $\psi = \rho e$ plus kinetic energy, where e is the specific internal energy and $\mathbf{w} = -\mathbf{v}\mathbf{T} + \mathbf{q}$, where $\mathbf{q}$ is the bulk heat flux, assumed to be governed by the Fourier law

$$\mathbf{q} = -\kappa \operatorname{grad}\theta, \tag{2.4}$$

with θ as the absolute temperature and κ – the thermal conduction coefficient, supposed to be different for the different phases: κ^+ for G_t^+ and κ^- for G_t^-.

Now, with $\mathbf{v} = \mathbf{V} = 0$ the balance equations for linear and angular momentum and energy will be

$$\operatorname{div}\mathbf{T} = 0 \qquad \text{and} \qquad \mathbf{T} = \mathbf{T}^T, \tag{2.5}$$

$$\rho\left(\frac{\partial e}{\partial t}\right) = -\operatorname{div}\mathbf{q}. \tag{2.6}$$

On the interface, if ψ_s and $\mathbf{w}_s$ are zeros, we get for the mass

$$[\![\rho c_n]\!] = 0, \tag{2.7}$$

for the linear and angular momentum

$$\begin{aligned} [\![\mathbf{T}\cdot\mathbf{n}]\!] &= \operatorname{div}_s\mathbf{S}, \\ \mathbf{n}\cdot\mathbf{S} &= 0, \qquad \mathbf{S}^{12} = \mathbf{S}^{21} \end{aligned} \tag{2.8}$$

and for the energy

$$[\![-\rho e c_n + \mathbf{q}\cdot\mathbf{n}]\!] = 0. \tag{2.9}$$

Here $\mathbf{S}$ is the surface (interfacial) stress (tension) tensor; if it is taken to be isotropic with a surface tension coefficient σ, then $\mathbf{S} = \frac{\sigma}{2}\mathbf{1}_s$ and $\operatorname{div}_s\mathbf{S} = 0.5\operatorname{grad}_s\sigma + \sigma H\mathbf{n}$, where $\mathbf{1}_s$ is the metric tensor of S_t.

We are not going to discuss the second law of thermodynamics, because as a special case we shall restrict ourselves only to equilibrium crystal growth processes. However, this does not exclude the future possibility of extending our research to non-equilibrium situations.

3. Layer-finite slab approach

The model of an interface as a finite slab presented here uses the concept of a moving surface in a continuum, as a reference surface, similarly to that in the shell theory.

In this approach the region G is divided into two subregions $\hat{G}_t^{\pm}$ by a narrow layer Z_t with regular surface boundaries $S_t^{\pm}$ and a reference surface Σ_t located between them (cf. Fig. 2), to which the mean interfacial fields will be referred (KOSIŃSKI [18]).

The balance equations valid for the bulk media are the same as in the singular surface approach i.e., Eqs. (2.3), (2.5) and (2.6), however, they hold in different regions. Before writing down the balance laws for the finite slab (layer), we are going to define the surface fields as integrals of the true $3D$ quantities over the layer coordinates (l is the parameter along the normal line to the reference surface Σ_t):

$$\psi^s := \int_{Z^-}^{Z^+} \psi(\mathbf{x} + l\mathbf{n}, t)\, j(l, \mathbf{x})\, dl =: <\psi j> \tag{3.1}$$

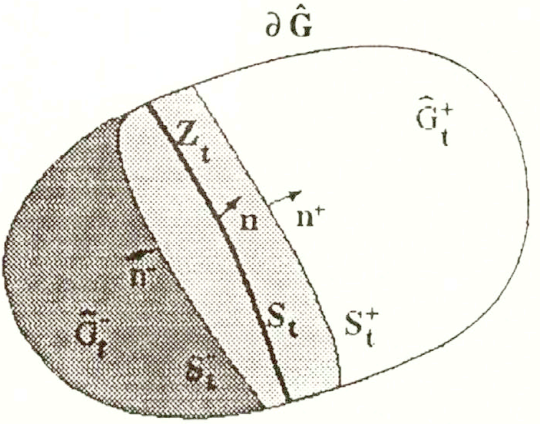


FIG. 2. Layer interface.

if a general density function ψ is considered and $j := 1 - 2Hl + Kl^2$, with H and K – the mean and Gaussian curvatures. The choice of forms of the functions in the layer depends on constitutive assumptions made for every particular case discussed. The reference surface could be chosen starting from different physical hypotheses, every time obtaining different reference surfaces. The variety of methods known in the literature is listed briefly in (KOSIŃSKI [18]).

Then the general balance law for the layer Z_t will be:

$$\frac{\delta \psi^s}{\delta t} + \operatorname{div}_s \mathbf{W}^s\{\psi\} - 2Hc_n\psi^s = \left[\!\left[j(\psi(\mathbf{v}\cdot\mathbf{n} - c_n) + \mathbf{w}\cdot\mathbf{n}) \right]\!\right], \tag{3.2}$$

where

$$\begin{aligned} \mathbf{W}^s\{\psi\} &:\equiv \langle \psi \otimes (\mathbf{v} + l \operatorname{grad}_s c_n)\mathbf{A}_s(l)\rangle + \langle \mathbf{w}\cdot\mathbf{A}_s(l)\rangle, \\ A_s(l) &:= \mathbf{1}_s + l(\mathbf{b} - 2H\mathbf{1}_s), \end{aligned} \tag{3.3}$$

$\mathbf{b}$ is the curvature tensor of Σ_t and $[\![g]\!] := g(\mathbf{x} + z^-\mathbf{n}(\mathbf{x}), t) - g(\mathbf{x} + z^+\mathbf{n}(\mathbf{x}), t)$ is valid for an arbitrary field g defined in G.

The particular balance laws are defined as in the singular surface case (KOSIŃSKI [18]).

4. Crystal growth of a solid sphere in a melt

Let us consider the solidification of a sphere $\hat{G}_t^-$ with radius R in a spherical container with radius R^* filled with its melt $\hat{G}_t^+$, as shown in Fig. 3. The phase change process is governed by the continuous negative heat source (heat sink or cooling) from a crystal seed initialy situated at the center of the sphere, while the container wall is isothermal with temperature $\theta^\infty \gg \theta_m$, where θ_m is the

solidification temperature. Assuming that at every moment of time the solid phase is at thermostatic equilibrium with the liquid phase, we shall try to model the solidification process (in the absence of body forces, e.g. gravity, the spherical shape of the boundary is the equilibrium one) in time, applying the two types of approach presented in the previous two sections.

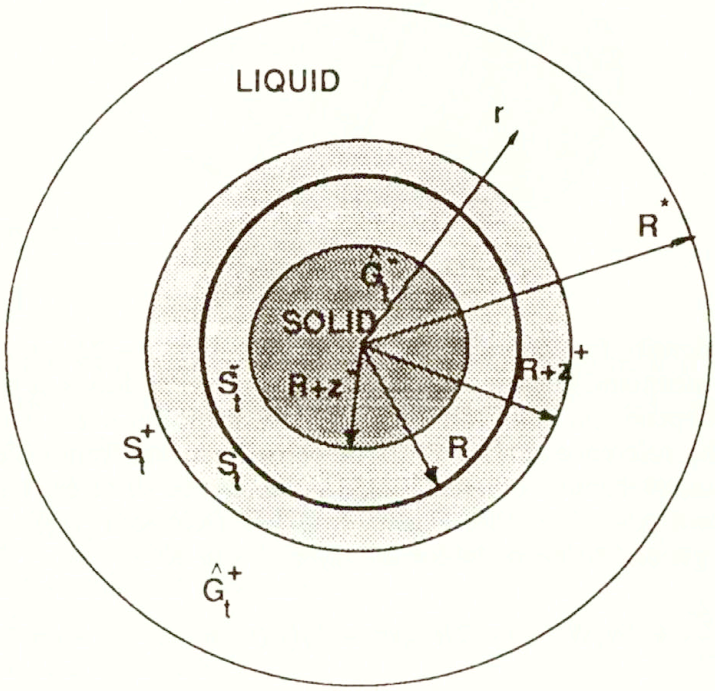


FIG. 3. Crystal growth of a sphere.

The solid and liquid phases are assumed to be isotropic and incompressible and the sphere is assumed to crystallize with a spherical symmetry, i.e. spherical coordinate system (r, ϕ, φ) is proposed, where $r = R(t)$ will describe the solidification front propagation in time. The mean and Gauss curvature are $H = -1/R$ and $K = 1/R^2$, and $\mathrm{div}_s \cdot = 0$.

In the present (first) part of our work we shall confine ourselves only to the energy and mass balance, while in the second part we shall consider the details of the deformation in the layer, the constitutive equation for the stress tensor and the relation with the energy.

The crystal growth processes (e.g. for semiconductor materials) are usually slow processes and the solidification velocity is much smaller than the conduction one (cf. TABACOVA [25]). In order to take into account the solidification front (interface) evolution in time, it is appopriate to treat the problem in the solidification time scale, which means that the temperature distribution in the bulk phases for the both approaches is stationary. In this way, Eq. (2.6) is reduced to

(the densities of the solid and liquid phase are assumed constant):

$$\operatorname{div} \mathbf{q} = 0. \tag{4.1}$$

In this formulation of our problem all the field quantities are functions of the space variable r only.

The boundary condition for the temperature function θ at $r = R^*$ is $\theta = \theta^\infty$ and at $r = \varepsilon$ is $-\kappa^- \frac{\partial \theta}{\partial r} = \frac{Q_0}{4\pi\varepsilon^2}$, where $Q_0 < 0$ and $c_1 \geq\geq |Q_0| \geq c_2 > 0$, c_1, c_2 are constants to be defined later in the text, and ε is greater than the critical radius of a crystal nucleus (CHALMERS [7]).

4.1. Singular surface interface

The interface S_t is described by $r = R(t)$ and it corresponds to the melting (solidification) temperature isotherm $\theta = \theta_m$. The displacement velocity has only the normal component $c_n = \dot{R}$ (where the upper dot is for the differentiation with respect to time). Then from Eqs. (2.7) and (4.1) we get:

$$\rho^+ = \rho^- = \text{const}, \tag{4.2}$$

$$\theta(r,t) = \theta_m - \frac{Q_0}{4\pi\kappa^-}\left(\frac{1}{R} - \frac{1}{r}\right) \qquad \text{in} \quad G_t^-, \tag{4.3}$$

$$\theta(r,t) = \theta_m + \frac{R^*(\theta^\infty - \theta_m)}{R^* - R}\left(1 - \frac{R}{r}\right) \qquad \text{in} \quad G_t^+. \tag{4.4}$$

If $\lim_{r\to\varepsilon} \theta(r) \geq \theta_{\min}$, then from Eq. (4.3) the upper limit c_1 for $|Q_0|$ is obtained.

The energy on S_t has a jump ($e^+ = e^- + L$, where L is the latent heat), while the temperature remains constant. The Eq. (2.9) represents the classical Stefan phase change condition:

$$[\![\kappa \operatorname{grad} \theta]\!] = \rho^+ L \dot{R}, \tag{4.5}$$

and the interface evolution is described by the first order ordinary differential equation:

$$\dot{R}\, \Phi_s(R) = d - cR, \tag{4.6}$$

where

$$\Phi_s(R) = \rho^+ L R^2 (R^* - R), \qquad c = \kappa^+(\theta^\infty - \theta_m)R^* + \frac{|Q_0|}{4\pi}, \qquad d = \frac{|Q_0|}{4\pi}R^*$$

with the initial condition $R(0) = R_0$ and $R_0 > \varepsilon$. The solution of Eq. (4.6) is $R = R(t)$, where $R \in [R_0, R_f]$ and R_f is the final radius of the solidified sphere.

The necessary condition for the crystal growth, i.e. the solid sphere enlarging with time, is $d - cR_f \geq 0$ or:

$$|Q_0| \geq c_2, \qquad \text{where} \qquad c_2 = 4\pi\kappa^+ R_f(\theta^\infty - \theta_m)/(1 - R_f/R^*). \tag{4.7}$$

The solidification time $t_s(R_f)$ is the value of the inverse function of $R(t)$ calculated at $R = R_f$ and can be given by the analytical expression:

$$t_s(R_f) = I(R_f) - I(R_0), \tag{4.8}$$

where

$$I(R) = \frac{\rho^+ L}{c^3}\left\{-\frac{(d-cR)^3}{3c} + \frac{(d-cR)^2}{2c}(3d - cR^*) + d(3d - 2cR^*)R + \frac{d^2}{c}(d - cR^*)\ln(d - cR)\right\} + \text{const}.$$

4.2. Constant thickness layer interface

Before postulating the initial value problem, we shall make some physical assumptions necessary for writing the constitutive laws in the layer:

i) The narrow interfacial layer with thickness $\delta R_0 \ll R_0$ is a $3D$ region and for it the bulk balance laws (2.3), (2.6) should have to be fulfilled. However, in the layer the field functions suffer great changes, e.g., in our case the internal energy changes with the quantity of the latent heat when the deviation of the temperature function $\Delta\theta$ is very small ($\Delta\theta \ll \theta_m$) or almost zero (for the singular surface case $\Delta\theta = 0$). Making a dimensionless analysis of (2.6) by means of the characteristic parameters τ – solidification time, δR_0 – interfacial thickness, L – latent heat, $\Delta\theta$ – temperature change, we get:

$$\delta^2 \rho' \left(\frac{\partial e'}{\partial t'}\right) = (\text{PeS})^{-1}\frac{\partial q'}{\partial r'}, \qquad \text{where} \quad q' \sim \frac{\partial \theta'}{\partial r'}$$

and the primes are for the dimensionless quantities. If the parameter $(\text{PeS})^{-1}$ in front of the highest order derivative is much smaller than one, then the thickness of the "boundary layer" (we shall call it a Stefan boundary layer or a phase change boundary layer) between the liquid and the solid phase is

$$z^+ - z^- = \delta R_0, \qquad \text{where} \quad \delta = (\text{PeS})^{-1/2}. \tag{4.9}$$

From now on, it is convenient to use the theory of perturbations and asymptotic expansions (COLE [8]) for our "boundary layer" problem. If $\Delta\theta = 0$, then $\delta \to 0$ and we get the singular surface approach. It is appropriate to seek all the unknown functions F in an asymptotic expansion of the small parameter δ:

$$F = F_0 + \delta F_1 + \delta^2 F_2 + \ldots, \tag{4.10}$$

where F_0 corresponds to the singular surface approach solution. If the constitutive law for the heat flux $\mathbf{q}$ in the layer is known, then the problem could be treated with the conventional technique of the matched asymptotic expansions in the boundary layer and in the bulk regions. In the present work we propose the layer approach given in Sec. 3 instead of calculating the $3D$ equations in the layer region, but a future comparison between the two methods will be substantial.

To our best knowledge the interfacial thickness was not determined in the other models treating phase-change processes. In HUTTER [14] it is a free phenomenological constant on which the transport coefficients in the layer depend.

ii) The internal energy e is a linear function of the layer coordinate l:

$$e(l) = e^- + L(l - z^-)/(z^+ - z^-), \tag{4.11}$$

where $e(z^\pm) = e^\pm$.

iii) There exists a non-monotone relation of the van der Waals type (LANDAU and LIFSHITZ [19]) between the temperature and the internal energy in the layer region (cf. Fig. 4), where the Maxwell approximation with $\Delta\theta = 0$ corresponds to the classical Stefan condition (i.e. singular surface approach).

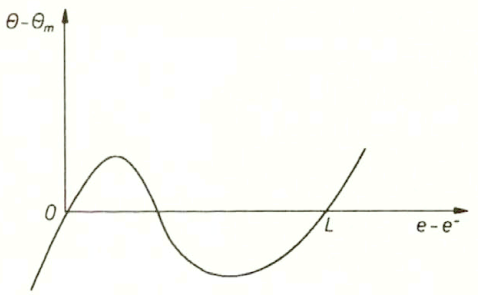


FIG. 4. Energy – temperature approximate graph.

iv) The density ρ is a regular function of the layer coordinate l.

v) The temperature function θ is approximated as cubic polynomial of $e - e^-$:

$$\theta(e) = \theta_m + \frac{e - e^-}{c_v^-} - \frac{(2c_v^+ + c_v^-)(e - e^-)^2}{Lc_v^+c_v^-} + \frac{(c_v^+ + c_v^-)(e - e^-)^3}{L^2c_v^+c_v^-}, \tag{4.12}$$

where $c_v^\pm$ are the heat capacities in bulk phases and $\theta(e^\pm) = \theta_m$ and $\left.\frac{d\theta}{de}\right|_{e=e^\pm} = \frac{1}{c_v^\pm}$.

vi) The reference surface Σ_t is supposed to be the isotherm at which the temperature is the melting one. The condition $\theta(e(0)) = \theta_m$ leads to the relation between the thickness z^+ and z^-:

$$z^+ = -z^-(c_v^-/c_v^+) \tag{4.13}$$

which together with (4.9) gives the explicit expressions of both the quantities,

$$z^+ = \delta R_0 \cdot \frac{c_v^-}{(c_v^+ + c_v^-)}, \qquad z^- = -\delta R_0 \cdot \frac{c_v^+}{(c^+ + c_v^-)}. \tag{4.14}$$

Using the expression $j(l) = (1 + l/R)^2$ and the proposed assumptions, the following forms of (3.2) in the case of the mass ($\psi = \rho$, $\psi^s = \rho^s$, $\mathbf{w} = 0$, $\mathbf{v} = 0$) and energy ($\psi = \rho e$, $\psi^s = \rho^s e^s$, $\mathbf{w} = \mathbf{q} = (-\kappa \frac{\partial \theta}{\partial r}, 0, 0)$, $\mathbf{v} = 0$) balances are obtained:

$$\frac{\partial \rho^s}{\partial t} + \frac{2\rho^s}{R}\dot{R} = \dot{R}\left\{\left(1 + \frac{z^+}{R}\right)^2 \rho^+ - \left(1 + \frac{z^-}{R}\right)^2 \rho^-\right\}, \tag{4.15}$$

$$\begin{aligned} \rho^s \frac{\partial e^s}{\partial t} = \dot{R}\left\{\left(1 + \frac{z^+}{R}\right)^2 \rho^+ (e^+ - e^s) - \left(1 + \frac{z^-}{R}\right)^2 \rho^- (e^- - e^s)\right\} \\ + \left(1 + \frac{z^+}{R}\right)^2 \kappa^+ \left.\frac{\partial \theta}{\partial r}\right|_{r=R+z^+} - \left(1 + \frac{z^-}{R}\right)^2 \kappa^- \left.\frac{\partial \theta}{\partial r}\right|_{r=R+z^-}. \end{aligned} \tag{4.16}$$

Due to condition iv) stating that ρ is a regular function of the thickness (polynomial of l) and to the relation (3.1) applied to ρ^s, we can deduce that ρ^s does not possess zero order term in δ. On the other hand, analyzing both sides of (4.15), the zero order term appears only on its right-hand side ([1]). This implies that if the bulk densities are expressed by (4.10), then the zero order densities of the solid and the liquid phase should have to be equal ($\rho^+ = \rho^- = \text{const}$), confirming the result of the singular surface approach (4.2). Consequently, the surface density obtained after applying (3.1) is:

$$\rho^s = \delta R_0 \rho^+ \left\{1 - \frac{\delta R_0}{R} \cdot \frac{(c_v^+ - c_v^-)}{(c_v^+ + c_v^-)} + \left(\frac{\delta R_0}{R}\right)^2 \cdot \frac{\left(c_v^{+2} - c_v^+ c_v^- + c_v^{-2}\right)}{3(c_v^+ + c_v^-)^2}\right\}. \tag{4.17}$$

The surface energy density $\rho^s e^s$ is obtained in a similar manner from (4.11) and (3.1):

$$e^s = \frac{\rho^+}{\rho^s}\left\{e_0^s + \frac{2e_1^s}{R} + \frac{e_2^s}{R^2}\right\}, \qquad e_0^s = \delta R_0 (e^- + L/2), \tag{4.18}$$

where

$$e_1^s = \frac{(\delta R_0)^2}{2(c_v^+ + c_v^-)}\left[e^-(c_v^- - c_v^+) + L(2c_v^- - c_v^+)/3\right],$$

$$e_2^s = \frac{(\delta R_0)^3}{3(c_v^+ + c_v^-)^2}\left[e^-\left(c_v^{+2} - c_v^+ c_v^- + c_v^{-2}\right) + L\left(c_v^{+2} - 2c_v^+ c_v^- + 3c_v^{-2}\right)/4\right].$$

([1]) We have assumed that the time derivatve of ρ^s is of the same order as the function itself.

It is evident that zero order terms of (4.17) and (4.18) are zeros, as it was assumed in the singular surface approach, ie. $\rho^s = 0$, $\rho^s e^s = 0$.

For the temperature in the bulks, the solutions are

$$\theta(r,t) = \theta_m - \frac{Q_0}{4\pi\kappa^-}\left(\frac{1}{R+z^-} - \frac{1}{r}\right) \qquad \text{in} \quad \hat{G}_t^-, \tag{4.19}$$

$$\theta(r,t) = \theta_m + \frac{R^*(\theta^\infty - \theta_m)}{R^* - (R+z^+)}\left(1 - \frac{(R+z^+)}{r}\right) \qquad \text{in} \quad \hat{G}_t^+. \tag{4.20}$$

If θ as given above is expanded in the form (4.10), then the zero order terms are the singular surface approach solutions (4.3) and (4.4), respectively.

Now, having the explicit forms for ρ^s and e^s in terms of R and material constants, Eqs. (4.15)–(4.16) lead to the following initial value problem:

$$\dot{R}\,\Phi_L(R) = d - cR\left(1 + \frac{z^+}{R}\right), \tag{4.21}$$

where

$$\Phi_L(R) = -\rho^+\left\{\frac{1}{(c_v^+ + c_v^-)}\cdot\left(\frac{\delta R_0}{R}\right)^2\left[c_v^- - c_v^+ + \left(\frac{\delta R_0}{R}\right)\right.\right.$$
$$\left.\cdot\frac{2(c_v^{+2} - c_v^+ c_v^- + c_v^{-2})}{3(c_v^+ + c_v^-)}\right]e^s - \frac{2}{R^2}\left(e_1^s + \frac{e_2^s}{R}\right) - \left(1 + \frac{z^+}{R}\right)^2(e^+ - e^s)$$
$$\left. + \left(1 + \frac{z^-}{R}\right)^2(e^- - e^s)\right\}\cdot R^2\left[R^* - (R+z^+)\right],$$

with $R(0) = R_0 - z^-$.

As in the singular surface case, we get some restriction for the heat source at the center of the solid sphere:

$$|Q_0| \geq c_2, \tag{4.22}$$

where

$$c_2 = 4\pi\kappa^+ R_f\,(\theta^\infty - \theta_m)/(1 - R_f)/R^*)$$
$$\times\left[1 + \delta R_0\cdot\frac{c_v^-}{(c_v^+ + c_v^-)}\cdot\frac{R^*}{R_f(R^* - R_f)} + O(\delta^2)\right].$$

The solidification time for a sphere of radius R_f (i.e. with mean surface radius $R_f - z^-$) given in its asymptotic expansion in δ is

$$t_l(R_f) = t_s(R_f) + \delta(J(R_f) - J(R_0)) + O(\delta^2), \tag{4.23}$$

where

$$J(R) = \rho^+ L R_0 \left\{ \frac{R^2}{2c} - \frac{d(d - cR^*)}{c^3} \sum_{k=2}^{\infty} \left(\frac{cR}{d}\right)^k \left(1 - \frac{1}{k}\right) \right\} + \text{const}.$$

From (4.8) and (4.22) it is clear that the time necessary for the solidification of a sphere of radius R_f in the case of the layer approach is greater ($J(R_f) - J(R_0) > 0$) than the solidification time in the singular surface case, and it does not explicitly depend on the material constants, but only on the process parameters and on the layer thickness.

If at the initial time, the sphere of radius R_0 is assumed to be composed of a solid sphere of radius $R_0 - \delta R_0$, and an interfacial layer with thickness δR_0, i.e. $R(0) = R_0 - z^+$, then the solidification time for a sphere of radius R_f represented as a solid sphere ($R_f - \delta R_0$) surrounded by a interfacial layer (δR_0) could be expressed in a similar manner as (4.22). Here the disturbance due to the interface is

$$J(R) = \rho^+ L R_0 \left\{ -\frac{R^2}{2c} + \frac{d(d - cR^*)}{c^3} \sum_{k=2}^{\infty} \left(\frac{cR}{d}\right)^k \frac{1}{k} \right\} + \text{const} \tag{4.24}$$

which implies that the solidification time t_l is less than the corresponding time t_s ($J(R_f) - J(R_0) < 0$). In the first case the layer occurs as an obstacle for the solidification of one and the same volume, while in the second – the solidification of the smaller volume is not disturbed enough by the layer and its solidification time is still less than that of the larger solid volume. This means that we could expect an intermediate case, where a part of the interfacial layer is inside the sphere and the other part outside it, for which the solidification times t_l and t_s would be equal. This result will be of importance when applying the present model for approximation of the classical Stefan problem (singular surface approach) in its enthalpy formulation (cf. TABACOVA [25]) and its numerical realization for geometrically more complex cases. A similar smoothing, but resulting from a differernt phase field model, is used succesfully (CAGINALP and SOCOLOVSKY [6]) for computing the interface propagation when the interface thickness is a free parameter.

5. Conclusions

The basic balance laws for the phase change (solidification) processes have been presented using the singular surface approach and the layer (finite slab) approach. The investigation of a sphere solidification from melt in a quasi-static formulation has been observed as a special case. Under the assumption of great dimensionless number PeS, the conventional boundary layer theory has been applied for the interfacial layer as a "phase-change boundary layer" between

the liquid and the solid phases. Its thickness is related to a small parameter and the asymptotic expansions in it have been sought for all the field functions. For the internal energy and the temperature distribution in the layer, a van der Waals relation has been considered. The internal energy has been assumed to be a linear function of the layer coordinate, and the reference (mean) surface Σ_t has been accepted to be an isotherm of the melting temperature. Finally, it occurs that the layer thickness depends only on the substance thermal parameters and not on the process parameters (external heat distribution or heat sources). The comparison between the two approaches leads to the conclusion that the solution corresponding to the singular surface approach are the zero order terms in the assymptotic expansion of the solutions obtained by the layer approach. For the observed case, it is found that the time necessary for the solidification of a sphere with a prescribed radius in the layer approach is greater than the time in the singular surface approach by an additional part of the small parameter order. Moreover, if the sphere in the layer approach is regarded as composed by an inner solid sphere and an interfacial annular layer, it solidifies faster than the sphere in the singular surface approach and the time difference is of the small parameter order.

Acknowledgement

A part of this work was done when one of the authors (W. KOSIŃSKI) was visiting MARS Center, Naples and LMM of Université Paris VI. The possibility of having fruitful discussions at those places are greatly appreciated.

References

1. A.W. ADAMSON, *Physical chemistry of surfaces*, Interscience, New York-London 1982.
2. A.M. ALBANO, D. BEDEAUX and J. VLIEGER, *On the description of interfacial properties using singular densities and currents at a dividing suface*, Physica, **99A**, 293–304, 1979.
3. T. ALTS and K. HUTTER, *Continuum description of the dynamics and thermodynamics of phase boundaries between ice and water. Parts III and IV*, J. Non-Equilib. Thermodyn., **13**, 301–329, **14**, 1–22, 1989.
4. C. BRICE, *The growth of crystals from the melt*, North-Holland, Amsterdam 1963.
5. F.P. BUFF, *The theory of capillarity*, [in:] Handbuch der Physik, vol. 10, Springer, Berlin-Heildelberg-New York, 1960, pp. 281–304.
6. G. CAGINALP and E.A. SOCOLOVSKY, *Computation of sharp phase boundaries by spreading: The planar and spherically symmetric cases*, J. Comput. Physics, **95**, 85–100, 1991.
7. B. CHALMERS, *Principle of solidification*, Robert E. Krieger Publ. Co., Malabar, Florida 1982.
8. J.COLE, *Perturbation methods in applied mathematics*, Ginn and Company, Boston 1968.
9. F. DELL'ISOLA and W. KOSIŃSKI, *Deduction of thermodynamic balance laws for bidimensional nonmaterial directed continua modelling interphase layers*, Arch. Mech., **45**, 3, 333–359, 1993.
10. F. DELL'ISOLA and W. KOSIŃSKI, *The interface between phases as a layer. Part Ia*, [in:] Proc. Vth Meeting: Waves and Stability in Continuous Media, Sorrento, October 1989, S. RIONERO [Ed.], World Scientific, Singapore 1991, pp. 108–113.
11. R. GATIGNOL and P. SEPPECHER, *Modelization of fluid-fluid interfaces with material properties*, J. Mech. Theor. et Appl., **Numero special**, 225–247, 1986.

12. R. GATIGNOL, *Liquid-vapour interfacial conditions*, Rev. Romaine des Sci. Tech. Mech. Appl., **3**, 255–271, 1987.
13. J.W. GIBBS, *The scientific papers of J. Willard Gibbs*, vol. I, Dover Publications, New York 1961.
14. K. HUTTER, *The physics of ice-water phase change surface*, [in:] Modelling Macroscopic Phenomena at Liquid Boundaries, CISM, W. KOSIŃSKI and A.I. MURDOCH [Eds.], Springer Verlag, 1991, pp. 159–216.
15. W. KOSIŃSKI, *Thermodynamics of singular surfaces and phase transitions*, [in:] Free Moving Boundary Problems: Applications and Theory, vol. III, A. BOSSAVIT, A. DAMLAMIAN and A.M. FREMOND [Eds.], Pitman, Boston-London-Melbourne, 1985, pp. 140–151.
16. W. KOSIŃSKI, *Field singularities and wave analysis in continuum mechanics*, Ellis Horwood and PWN-Polish Scientific Publishes, Chichester, Warsaw, 1986.
17. W. KOSIŃSKI and A. ROMANO, *Evolution balance laws in two-phase problems*, XVIIth IUTAM Congress in Grenoble, August 1988, also Arch. Mech., **41**, 2-3, 255–266, 1989.
18. W. KOSIŃSKI, *A phenomenological approach to fluid-phase interfaces*, [in:] Modelling Macroscopic Phenomena at Liquid Boundaries, CISM, W. KOSIŃSKI and A.J. MURDOCH [Eds.], Springer Verlag, 1991, pp. 65–107.
19. L.D. LANDAU and E.M. LIFSHITZ, *Statistical physics*, Pergamon Press, London 1958.
20. G.P. MOECKEL, *Thermodynamics of an interface*, Arch. Rational Mech. Analysis, **57**, 3, 255–280, 1974.
21. L.G. NAPOLITANO, *Thermodynamics and dynamics of pure interfaces*, Acta Astronautica, **5**, 655–670, 1978.
22. A. ROMANO, *Thermodynamic of a continuum with an interface and Gibbs' rule*, Ricerche di Matematica, **31**, 2, 277–294, 1983.
23. L.E. SCRIVEN, *Dynamics of a fluid interface*, Chem. Engng. Sci., **12**, 98–108, 1960.
24. J.C. SLATTERY, *General balance equation for a phase interface*, Ind. Engng. Chem. Fundamentals, **6**, 1, 108–115, 1967.
25. S. TABACOVA, *On the two-dimensional numerical modelling of thermal processes in Czochralski crystal growth from melt*, Control and Cybernetics, **18**, 3–4, 7–30, 1989.

POLISH ACADEMY OF SCIENCES
INSTITUTE OF FUNDAMENTAL TECHNOLOGICAL RESEARCH
and
MICROGRAVITY ADVANCED RESEARCH AND SUPPORT CENTER (MARS)
VIA DIOCLEZIANO 328, 80125 NAPLES, ITALY.

Received July 8, 1994.

Arch. Mech., 46, 6, pp. 855–879, Warszawa 1994

A rate-independent thermoplastic theory at finite deformation(*)

Z. P. HUANG (BEIJING)

IN THIS PAPER, a thermoplastic constitutive theory of a rate-independent material at finite deformation is systematically formulated in both the "entropy-strain" and its conjugate "temperature-stress" spaces. Based on the thermomechanical postulate of dissipative work proposed by the present author, the normality condition and the corresponding constitutive equations are derived. The strain measure in the constitutive equations is in the spirit of Lagrangian mechanics introduced by Hill, with any fixed configuration as the reference configuration. The stress is a work-conjugate variable of the strain. The transformation rules and invariance properties for different strain measures and reference configurations are investigated in detail. It is shown that the Eulerian formulation based on the current configuration or on the intermediate stress-free configuraton (if it exists) may be considered as a special case of the present result.

1. Introduction

A NUMBER OF competing theories on the elastic-plastic constitutive relations for finite deformation [1] are known. The diversity of these theories may be attributed to the differences in the following factors [2]: 1) Foundations of the irreversible thermodynamics; 2) The kinematics of large elastic-plastic deformation; 3) The choice of thermodynamic state variables, especially the internal variables and their evolution equations; 4) The existence of the yield surface (or the loading surface); 5) The formulation of constitutive equations.

Among the various controversial non-equilibrium thermodynamic theories, probably the local accompanying state model [3, 4] is considered to be the most acceptable one due to its simplicity, flexibility, clear physical contents and its mathematical consistency. Adopting this model, we may choose the (thermal) state variables to be specific internal energy ε, strain tensor $\mathbf{E}$ and a finite set of (scalar, vectorial, or tensorial) internal variables ξ_α ($\alpha = 1, 2, \ldots$) to describe a properly selected constrained equilibrium state, and to write down the Gibbs equation. Thus the specific entropy η and the (absolute) temperature θ of the accompanying state may also be defined. The fundamental inequality for the rate of production of the accompanying entropy $\dot{\Theta}$ can be expressed as

$$\dot{\Theta} = \rho_0 \dot{\eta} - \rho_0 \left(\frac{r}{\theta}\right) + \nabla_0 \cdot \left(\frac{\mathbf{q}_0}{\theta}\right) = \frac{1}{\theta}(\mathbf{T} - \mathbf{T}^a) : \dot{\mathbf{E}} + \frac{1}{\theta} A^\alpha \dot{\xi}_\alpha + \mathbf{q}_0 \cdot \nabla_0 \left(\frac{1}{\theta}\right) \geq 0, \tag{1.1}$$

(*) Supported by National Natural Science Foundation of China.

where ρ_0 is the density in the reference configuration, r is the specific rate of external heat supply, $\mathbf{q}_0$ is the heat flux defined in the reference configuration, $\mathbf{T}$ and $\mathbf{T}^a$ are the stress tensors conjugate to $\mathbf{E}$, measured relative to the real nonequilibrium state and to the accompanying equilibrium state, respectively. A^α is the affinity conjugate to ξ_α while $A^\alpha \dot{\xi}_\alpha$ (summation over α) constitutes the rate of dissipation that cannot be recovered during the real process. From inequality (1.1), we obtain

$$\frac{1}{\theta} = \frac{\partial \eta}{\partial \epsilon}, \qquad \mathbf{T}^a = -\rho_0 \theta \frac{\partial \eta}{\partial \mathbf{E}}, \qquad A^\alpha = \rho_0 \theta \frac{\partial \eta}{\partial \xi_\alpha}. \tag{1.2}$$

If $(\theta, \mathbf{E}, \xi_\alpha)$ are chosen to be "independent" state variables, then Eqs. (1.2) may equivalently be written as

$$-\eta = \frac{\partial \psi}{\partial \theta}, \qquad \mathbf{T}^a = \rho_0 \frac{\partial \psi}{\partial \mathbf{E}}, \qquad A^\alpha = -\rho_0 \frac{\partial \psi}{\partial \xi_\alpha}, \tag{1.3}$$

where $\psi = \varepsilon - \theta\eta$ is the specific (Helmholtz) free energy. Similarly, when choosing $(\rho_0\eta, \mathbf{E}, \xi_\alpha)$ to be state variables, we obtain

$$\theta = \frac{\partial \varepsilon}{\partial \eta}, \qquad \mathbf{T}^a = \rho_0 \frac{\partial \varepsilon}{\partial \mathbf{E}}, \qquad A^\alpha = -\rho_0 \frac{\partial \varepsilon}{\partial \xi_\alpha}. \tag{1.4}$$

Suppose that the strain $\mathbf{E}$ and η may be expressed uniquely by $\mathbf{T}^a$ and θ by inverting Eqs. (1.3) for fixed ξ_α, then by choosing $(\theta, \mathbf{T}^a, \xi_\alpha)$ to be the state variables and defining the specific Gibbs free energy: $G = \psi - \mathbf{T}^a : \mathbf{E}$, we have

$$-\eta = \frac{\partial G}{\partial \theta}, \qquad \mathbf{E} = -\rho_0 \frac{\partial G}{\partial \mathbf{T}^a}, \qquad A^\alpha = -\rho_0 \frac{\partial G}{\partial \xi_\alpha}. \tag{1.5}$$

It should be pointed out that the strain $\mathbf{E}$ in Eqs. (1.1)–(1.5) is in the spirit of Lagrangian mechanics introduced by HILL [5], with any fixed configuration as the reference configuration. A family of strain tensors can be expressed as

$$\mathbf{E} = \sum_{k=1}^{3} e(\lambda_k) \mathbf{N}_k \otimes \mathbf{N}_k, \tag{1.6}$$

where e is a smooth and monotone scale function satisfying $e(1) = 0$, $e'(1) = 1$, λ_k and $\mathbf{N}_k$ $(k = 1, 2, 3)$ are eigenvalues and unit eigenvectors of $(\mathbf{F}^T \cdot \mathbf{F})^{1/2}$ respectively, $\mathbf{F}$ is the deformation gradient from the reference configuration to the current configuration.

It is worth noting that the reference configuration does not necessarily have to be the natural configuration at the initial instant, i.e., any real or fictitious configuration (e.g., the intermediate stress-free configuration – if it exists) may be taken as the reference configuration. So, by the transformation of generalized

coordinate, the kinematics of the large elastic-plastic deformation used in Mandel-Lee theory may result from that used in Hill-Rice theory or Green-Naghdi theory as a transition from a fixed coordinate descripton to a moving coordinate description [6, 7]. One of the essential differences among the existing theories is in the choice of the internal variables and their evolution laws. For example, if $\mathbf{F}$ can be multiplicatively decomposed into elastic part $\mathbf{F}_e$ and plastic part $\mathbf{F}_p$ [8]

$$\mathbf{F} = \mathbf{F}_e \cdot \mathbf{F}_p, \tag{1.7}$$

then $(\mathbf{F}_p^T \cdot \mathbf{F}_p)^{1/2}$ may be taken as an internal variable in Green, Naghdi and Casey's theory (e.g. [9, 10, 11]), whereas the director vectors attached to the material substructure in the intermediate stress-free configuration play an important role in the Mandel and Kratochvil theory (Ref.[12, 13, 14]), and especially in the Dafalias and Loret theory (Ref. [15, 16, 17]) that further provides an explicit constitutive relation for the plastic spin.

Nevertheless, it should be noted that if a certain kind of "quasi-thermodynamics" postulate such as Drucker's postulate or Ilyushin's postulate is utilized, the form of elastic-plastic constitutive equations will be greatly simplified. In this way, a detailed discussion on the choice of internal variables and their evolution equations may be bypassed to some extent, and only the investigations on the evolution laws of the loading surface is needed. The Ilyushin's postulate has been generalized to the finite deformation case by HILL and RICE [18] as well as by NAGHDI and TRAPP [19], by means of which isothermal elastic-plastic constitutive equations at large deformation were constructed. In this paper, the Ilyushin's postulate is further extended to the thermomechanical case. From this postulate, the normality condition and the corresponding thermoplastic constitutive equations are derived. The transformation rules and invariance properties for different strain measures and reference configurations are also discussed. It is seen that the Eulerian formulation based on the current configuration or on the intermediate stress-free configuration (if it exists) may be considered as a special case of the present result.

2. A rate-independent thermal elastic-plastic material

Now let us take $(\rho_0\eta, \mathbf{E}, \xi_\alpha)$ to be the state variables of a material element and assume the loading surface

$$g(\rho_0\eta, \mathbf{E}, \xi_\alpha) = 0 \tag{2.1}$$

to exist. Then the loading criterion of the thermoplastic material may be defined according to the sign of the scalar

$$\hat{g} = \frac{\partial g}{\rho_0 \partial \eta}(\rho_0 \dot{\eta}) + \frac{\partial g}{\partial \mathbf{E}} : \dot{\mathbf{E}}\,. \tag{2.2}$$

With the use of the symbol

$$(2.3)\qquad <\hat{g}>=\begin{cases}0, & \text{when } g<0,\\ 0, & \text{when } g=0 \text{ and } \hat{g}<0,\\ 0, & \text{when } g=0 \text{ and } \hat{g}=0,\\ \hat{g} & \text{when } g=0 \text{ and } \hat{g}>0,\end{cases}$$

the evolution equations of the internal variables may be formally written as

$$(2.4)\qquad \dot{\xi}_\alpha = \lambda \Xi_\alpha(\rho_0\eta, \mathbf{E}, \xi_\beta) <\hat{g}> \qquad (\alpha,\beta = 1,2,\ldots),$$

where λ is a parameter, which can be determined from the consistency relation

$$(2.5)\qquad 1 + \lambda \frac{\partial g}{\partial \xi_\alpha}\Xi_\alpha = 0 \quad \text{(summation over } \alpha)$$

and Ξ_α is the α-th functon of state variables, the exact expression of which is left unspecified. Clearly, Eq. (2.3) may be considered as generalization of the strain-space formulation proposed by NAGHDI and TRAPP [20].

If the material element is in a thermoelastic state with $<\hat{g}>= 0$ (or $\dot{\xi}_\alpha = 0$), then inequality (1.1) will be reduced to

$$(2.6)\qquad (\mathbf{T} - \mathbf{T}^a) : \dot{\mathbf{E}} - \frac{1}{\theta}\mathbf{q}_0 \cdot \nabla_0\theta \geq 0.$$

Noting that for a rate-independent material the stress $\mathbf{T}$ does not depend on $\dot{\mathbf{E}}$ and that inequality (2.6) is satisfied for an arbitrary strain rate $\dot{\mathbf{E}}$ with an arbitrarily distributed uniform temperature field, we have

$$(2.7)\qquad \mathbf{T} = \mathbf{T}^a = \rho_0 \frac{\partial \varepsilon}{\partial \mathbf{E}}\bigg|_{(\rho_0\eta,\xi_\alpha)}.$$

Since the thermoelastic state inside the loading surface may arbitrarily approach the thermoelastic-plastic state on the loading surface, we may conclude that Eq. (2.7) should also remain valid during elastic-plastic loading (i.e., for $g = 0$, $\hat{g} > 0$). In the following $\mathbf{T}^a$ will always be replaced by $\mathbf{T}$ which represents the actual stress conjugate to $\mathbf{E}$. Then inequality (1.1) may be rewritten as

$$(2.8)\qquad A^\alpha \dot{\xi}_\alpha - \frac{1}{\theta}\mathbf{q}_0 \cdot \nabla_0\theta \geq 0 \qquad \text{(summation over } \alpha).$$

It can be seen from (1.4) and (2.4) that A^α and $\dot{\xi}_\alpha$ are independent of $\nabla_0\theta$. If we further assume that $\mathbf{q}_0$ is independent of $\dot{\eta}$ and $\dot{\mathbf{E}}$, inequality (2.8) can then be replaced by stronger ones

$$\mathbf{q}_0 \cdot \nabla_0\theta \leq 0$$

and

$$A^{\alpha}\,\dot{\xi}_{\alpha} \geq 0, \tag{2.9}$$

where (2.9) is usually called the Kelvin inequality.

From the above discussion, relations betwen $(\rho_0\eta, \mathbf{E})$ and their conjugate variables $(\theta, \mathbf{T})$ for given ξ_α may be expressed as

$$\theta = \left.\frac{\partial \varepsilon}{\partial \eta}\right|_{(\mathbf{E},\xi_\alpha)}, \qquad \mathbf{T} = \rho_0 \left.\frac{\partial \varepsilon}{\partial \mathbf{E}}\right|_{(\rho_0\eta,\xi_\alpha)}, \tag{2.10}$$

and

$$\rho_0\eta = -\rho_0 \left.\frac{\partial G}{\partial \theta}\right|_{(\mathbf{T},\xi_\alpha)}, \qquad \mathbf{E} = -\rho_0 \left.\frac{\partial G}{\partial \mathbf{T}}\right|_{(\theta,\xi_\alpha)}, \tag{2.11}$$

where $\varepsilon = \varepsilon(\rho_0\eta, \mathbf{E}, \xi_\alpha)$ and $-G = -G(\theta, \mathbf{T}, \xi_\alpha)$ may be regarded as constitutive potentials with the internal variables ξ_α $(\alpha = 1, 2, \ldots)$ being parameters.

Along an entropy-strain trajectory, the rates of temperature and stress can be written as

$$\begin{aligned} \dot{\theta} &= \mathcal{L}_0(\rho_0\,\dot{\eta}) + \boldsymbol{\mathcal{L}}_1 : \dot{\mathbf{E}} + \hat{\theta} < \hat{g} >, \\ \dot{\mathbf{T}} &= \boldsymbol{\mathcal{L}}_1(\rho_0\,\dot{\eta}) + \boldsymbol{\mathcal{L}} : \dot{\mathbf{E}} + \hat{\mathbf{T}} < \hat{g} >, \end{aligned} \tag{2.12}$$

where

$$\begin{aligned} \mathcal{L}_0 &= \frac{1}{\rho_0}\frac{\partial \theta}{\partial \eta} = \frac{1}{\rho_0}\frac{\partial^2 \varepsilon}{\partial \eta^2}, \\ \boldsymbol{\mathcal{L}}_1 &= \frac{\partial \theta}{\partial \mathbf{E}} = \frac{\partial^2 \varepsilon}{\partial \eta \partial \mathbf{E}} = \frac{1}{\rho_0}\frac{\partial \mathbf{T}}{\partial \eta}, \\ \boldsymbol{\mathcal{L}} &= \frac{\partial \mathbf{T}}{\partial \mathbf{E}} = \rho_0 \frac{\partial^2 \varepsilon}{\partial \mathbf{E} \partial \mathbf{E}} \qquad \text{(the adiabatic modulus)}, \end{aligned} \tag{2.13}$$

and

$$\begin{aligned} \hat{\theta} &= \lambda \sum_{\alpha} \frac{\partial \theta}{\partial \xi_\alpha} \Xi_\alpha, \\ \hat{\mathbf{T}} &= (\hat{\mathbf{T}})^T = \lambda \sum_{\alpha} \frac{\partial \mathbf{T}}{\partial \xi_\alpha} \Xi_\alpha. \end{aligned} \tag{2.14}$$

Similarly, along a temperature-stress trajectory, the rates of entropy and strain can be written as

$$\begin{aligned} \rho_0\,\dot{\eta} &= \mathcal{M}_0\,\dot{\theta} + \boldsymbol{\mathcal{M}}_1 : \dot{\mathbf{T}} + \rho_0 \hat{\eta} < \hat{g} >, \\ \dot{\mathbf{E}} &= \boldsymbol{\mathcal{M}}_1\,\dot{\theta} + \boldsymbol{\mathcal{M}} : \dot{\mathbf{T}} + \hat{\mathbf{E}} < \hat{g} >, \end{aligned} \tag{2.15}$$

where

$$
\begin{aligned}
&\mathcal{M}_0 = \rho_0 \frac{\partial \eta}{\partial \theta} = -\rho_0 \frac{\partial^2 G}{\partial \theta^2}, \\
&\boldsymbol{\mathcal{M}}_1 = \rho_0 \frac{\partial \eta}{\partial \mathbf{T}} = -\rho_0 \frac{\partial^2 G}{\partial \theta \partial \mathbf{T}} = \frac{\partial \mathbf{E}}{\partial \theta} \\
&\qquad \text{(the thermal expansion coefficient at fixed } \mathbf{T} \text{ and } \xi_\alpha), \\
&\boldsymbol{\mathcal{M}} = \frac{\partial \mathbf{E}}{\partial \mathbf{T}} = -\rho_0 \frac{\partial^2 G}{\partial \mathbf{T} \partial \mathbf{T}} \\
&\qquad \text{(the isothermal compliance),}
\end{aligned} \tag{2.16}
$$

and

$$
\begin{aligned}
&\hat{\eta} = \lambda \sum_\alpha \frac{\partial \eta}{\partial \xi_\alpha} \Xi_\alpha, \\
&\hat{\mathbf{E}} = (\hat{\mathbf{E}})^T = \lambda \sum_\alpha \frac{\partial \mathbf{E}}{\partial \xi_\alpha} \Xi_\alpha.
\end{aligned} \tag{2.17}
$$

In the above equations $\mathcal{L}_0$ and $\mathcal{M}_0$ are scalars, $\boldsymbol{\mathcal{L}}_1$ and $\boldsymbol{\mathcal{M}}_1$ are second order symmetric tensors, and $\boldsymbol{\mathcal{L}}$ and $\boldsymbol{\mathcal{M}}$ are fourth order symmetric tensors, the components of which satisfy the relations

$$
\begin{aligned}
&\mathcal{L}^{ABKL} = \mathcal{L}^{BAKL} = \mathcal{L}^{ABLK} = \mathcal{L}^{KLAB}, \\
&\mathcal{M}_{ABKL} = \mathcal{M}_{BAKL} = \mathcal{M}_{ABLK} = \mathcal{M}_{KLAB}.
\end{aligned}
$$

The relation between (2.13) and (2.16) can be easily obtained during a thermoelastic unloading (with $\hat{g} = 0$),

$$
\begin{aligned}
&\mathcal{L}_0 \mathcal{M}_0 + \boldsymbol{\mathcal{L}}_1 : \boldsymbol{\mathcal{M}}_1 = 1, \\
&\mathcal{L}_0 \boldsymbol{\mathcal{M}}_1 + \boldsymbol{\mathcal{L}}_1 : \boldsymbol{\mathcal{M}} = \mathcal{M}_0 \boldsymbol{\mathcal{L}}_1 + \boldsymbol{\mathcal{L}} : \boldsymbol{\mathcal{M}}_1 = 0, \\
&\boldsymbol{\mathcal{L}}_1 \otimes \boldsymbol{\mathcal{M}}_1 + \boldsymbol{\mathcal{L}} : \boldsymbol{\mathcal{M}} = \overset{(1)}{\mathbf{I}},
\end{aligned} \tag{2.18}
$$

where $\overset{(1)}{\mathbf{I}}$ is the fourth order identity tensor, and its components may be written as

$$
\overset{(1)}{\mathbf{I}}{}^{PQ}_{..KL} = \frac{1}{2} \left(\delta^P_K \delta^Q_L + \delta^P_L \delta^Q_K \right). \tag{2.19}
$$

Since the quantities which appear in Eqs. (2.18) are independent of $(\rho_0 \dot{\eta}, \dot{\mathbf{E}})$ and $(\dot{\theta}, \dot{\mathbf{T}})$, so Eqs. (2.18) is also valid during the thermoelastic-plastic loading.

Hence the rates of the temperature and the stress may be additively decomposed into two parts:

$$\begin{aligned} \dot{\theta}_e &= \mathcal{L}_0(\rho_0\,\dot{\eta}) + \mathcal{L}_1 \,:\dot{\mathbf{E}}, \\ \dot{\mathbf{T}}_e &= \mathcal{L}_1(\rho_0\,\dot{\eta}) + \mathcal{L} \,:\dot{\mathbf{E}}, \end{aligned} \tag{2.20}$$

and

$$\dot{\theta}_p = \hat{\theta} <\hat{g}>, \qquad \dot{\mathbf{T}}_p = \hat{\mathbf{T}} <\hat{g}>, \tag{2.21}$$

which are the elastic and the inelastic parts of the rates of the temperature and the stress, respectively. Similarly for the rates of the entropy and the strain, we have

$$\begin{aligned} \rho_0\,\dot{\eta}_e &= \mathcal{M}_0\,\dot{\theta} + \mathcal{M}_1 \,:\dot{\mathbf{T}}, \\ \dot{\mathbf{E}}_e &= \mathcal{M}_1\,\dot{\theta} + \mathcal{M} \,:\dot{\mathbf{T}}, \end{aligned} \tag{2.22}$$

and

$$\rho_0\,\dot{\eta}_p = \rho_0\hat{\eta} <\hat{g}>, \qquad \dot{\mathbf{E}}_p = \hat{\mathbf{E}} <\hat{g}>, \tag{2.23}$$

which are the elastic and the inelastic parts of the rates of the entropy and the strain, respectively.

From Eqs. (2.12), (2.15) and (2.18), relations between $(\hat{\theta}, \hat{\mathbf{T}})$ and $(\rho_0\hat{\eta}, \hat{\mathbf{E}})$ may also be obtained.

$$\hat{\theta} = -(\mathcal{L}_0\rho_0\hat{\eta} + \mathcal{L}_1 \,:\hat{\mathbf{E}}), \qquad \hat{\mathbf{T}} = -(\mathcal{L}_1\rho_0\hat{\eta} + \mathcal{L} \,:\hat{\mathbf{E}}), \tag{2.24}$$

$$\rho_0\hat{\eta} = -(\mathcal{M}_0\hat{\theta} + \mathcal{M}_1 \,:\hat{\mathbf{T}}), \qquad \hat{\mathbf{E}} = -(\mathcal{M}_1\hat{\theta} + \mathcal{M} \,:\hat{\mathbf{T}}). \tag{2.25}$$

For a given loading surface (Eq. (2.1)), we can write, with the aid of (2.11), a corresponding loading surface in the temperature-stress space using the formula

$$f(\theta, \mathbf{T}, \xi_\alpha) = g(\rho_0\eta(\theta, \mathbf{T}, \xi_\beta), \mathbf{E}(\theta, \mathbf{T}, \xi_\beta), \xi_\alpha) = 0. \tag{2.26}$$

The outward normal to $f = 0$ in a seven-dimensional temperature-stress space $\left(\frac{\partial f}{\partial\theta}, \frac{\partial f}{\partial\mathbf{T}}\right)$ may be expressed by an outward normal to $g = 0$ in a seven-dimensional entropy-strain space $\left(\frac{\partial g}{\rho_0\partial\eta}, \frac{\partial g}{\partial\mathbf{E}}\right)$ in the following matrix form

$$\begin{pmatrix} \dfrac{\partial f}{\partial\theta} \\ \dfrac{\partial f}{\partial\mathbf{T}} \end{pmatrix} = \begin{pmatrix} \mathcal{M}_0 & \mathcal{M}_1 \\ \mathcal{M}_1 & \mathcal{M} \end{pmatrix} \begin{pmatrix} \dfrac{\partial g}{\rho_0\partial\eta} \\ \dfrac{\partial g}{\partial\mathbf{E}} \end{pmatrix} \tag{2.27}$$

or

$$\begin{pmatrix} \dfrac{\partial g}{\rho_0 \partial \eta} \\ \dfrac{\partial g}{\partial \mathbf{E}} \end{pmatrix} = \begin{pmatrix} \mathcal{L}_0 & \boldsymbol{\mathcal{L}}_1 \\ \boldsymbol{\mathcal{L}}_1 & \boldsymbol{\mathcal{L}} \end{pmatrix} \begin{pmatrix} \dfrac{\partial f}{\partial \theta} \\ \dfrac{\partial f}{\partial \mathbf{T}} \end{pmatrix}. \tag{2.27 [cont.]}$$

Hence from Eqs. (2.24), (2.25) and (2.27), we have

$$\rho_0 \hat{\eta} \frac{\partial g}{\rho_0 \partial \eta} + \hat{\mathbf{E}} : \frac{\partial g}{\partial \mathbf{E}} = -\left(\hat{\theta} \frac{\partial f}{\partial \theta} + \hat{\mathbf{T}} : \frac{\partial f}{\partial \mathbf{T}} \right), \tag{2.28}$$

and during the thermoelastic-plastic loading (i.e., $\hat{g} > 0$), Eq. (2.2) can be rewritten in the form

$$\hat{g} = \left(\frac{\partial f}{\partial \theta} \mathcal{L}_0 + \frac{\partial f}{\partial \mathbf{T}} : \boldsymbol{\mathcal{L}}_1 \right) (\rho_0 \dot{\eta}) + \left(\frac{\partial f}{\partial \theta} \boldsymbol{\mathcal{L}}_1 + \frac{\partial f}{\partial \mathbf{T}} : \boldsymbol{\mathcal{L}} \right) : \dot{\mathbf{E}}$$
$$= \hat{f} - \left(\frac{\partial f}{\partial \theta} \hat{\theta} + \frac{\partial f}{\partial \mathbf{T}} : \hat{\mathbf{T}} \right) \hat{g},$$

where

$$\hat{f} = \frac{\partial f}{\partial \theta} \dot{\theta} + \frac{\partial f}{\partial \mathbf{T}} : \dot{\mathbf{T}}. \tag{2.29}$$

It follows that during the thermoelastic-plastic loading, we have

$$\Phi = \hat{f}/\hat{g} \qquad (g = 0, \quad \hat{g} > 0), \tag{2.30}$$

where Φ is defined by

$$\Phi = 1 + \frac{\partial f}{\partial \theta} \hat{\theta} + \frac{\partial f}{\partial \mathbf{T}} : \hat{\mathbf{T}} = 1 - \left[\frac{\partial g}{\rho_0 \partial \eta} (\rho_0 \hat{\eta}) + \frac{\partial g}{\partial \mathbf{E}} : \hat{\mathbf{E}} \right]. \tag{2.31}$$

Clearly, the quotient Φ distinguishes three types of material response, namely hardening, softening and perfectly plastic behaviour, by its positive, negative and zero values. This may also be considered as a generalization of the related discussion in Ref. [10] and [21].

3. The thermomechanical postulate and the normality rule

It was emphasized in Ref. [4] that the description of material behaviour makes no sense other than with respect to the given mechanical and thermal actions. Two materials which behave identically under usual circumstances may exhibit totally different behaviour in other circumstances. Hence the classification of materials by "pairs of material-processes" is preferred. In this paper, a phenomenological

rate-independent thermoplastic constitutive theory of polycrystalline metals is discussed. Suppose that the pair of material-processes under discussion satisfy the thermomechanical postulate proposed in Ref. [22], viz. for an arbitrary closed cycle in the entropy-strain space during time interval $[t_0, t_f]$, with η and $\mathbf{E}$ having, respectively, the same values at time t_0 and t_f, the following integral is non-negative:

$$\int_{t_0}^{t_f} \left(\rho_0 \theta \, \dot{\eta} + \mathbf{T} : \dot{\mathbf{E}}\right) dt \geq 0, \tag{3.1}$$

or equivalently,

$$\int_{t_0}^{t_f} \left(\rho_0 \, \dot{\varepsilon} + A^{\alpha} \, \dot{\xi}_{\alpha}\right) dt \geq 0. \tag{3.2}$$

Let us consider an interior point and a boundary point of the loading surface in the entropy-strain space, which are denoted by $(\rho\eta^{(1)}, \mathbf{E}^{(1)}, \xi_{\alpha})$ and $(\rho_0\eta^{(2)}, \mathbf{E}^{(2)}, \xi_{\alpha})$, respectively. If the material element undergoes a small thermoelastic-plastic loading from $(\rho_0\eta^{(2)}, \mathbf{E}^{(2)}, \xi_{\alpha})$ such that the internal variables have their increments $d\xi_{\alpha}$, then from inequality (3.1) we obtain [22]

$$\rho_0 \left\{ \left.\frac{\partial \varepsilon}{\partial \xi_{\alpha}}\right|_{(\rho_0\eta^{(1)}, \mathbf{E}^{(1)})} - \left.\frac{\partial \varepsilon}{\partial \xi_{\alpha}}\right|_{(\rho_0\eta^{(2)}, \mathbf{E}^{(2)})} \right\} d\xi_{\alpha} \geq 0 \qquad \text{(summation over } \alpha\text{)}. \tag{3.3}$$

With the use of the mean value theorem, inequality (3.3) may be rewritten as

$$\rho_0 \left\{ \frac{\partial^2 \varepsilon}{\partial \xi_{\alpha} \partial \eta} \Delta\eta + \frac{\partial^2 \varepsilon}{\partial \xi_{\alpha} \partial \mathbf{E}} : \Delta\mathbf{E} \right\} \Bigg|_{(\rho_0\eta^{(m)}, \mathbf{E}^{(m)})} d\xi_{\alpha} \geq 0 \tag{3.4}$$

(summation over α),

where $\Delta\eta = \eta^{(1)} - \eta^{(2)}$, $\Delta\mathbf{E} = \mathbf{E}^{(1)} - \mathbf{E}^{(2)}$, and the partial derivatives in the above expression are evaluated at the point $(\rho_0\eta^{(m)}, \mathbf{E}^{(m)})$ which is on the straight line connecting $(\rho_0\eta^{(1)}, \mathbf{E}^{(1)}, \xi_{\alpha})$ and $(\rho_0\eta^{(2)}, \mathbf{E}^{(2)}, \xi_{\alpha})$ in the entropy-strain space. When values of $\eta^{(1)}$ and $\mathbf{E}^{(1)}$ are very close to those of $\eta^{(2)}$ and $\mathbf{E}^{(2)}$, $(\rho_0\eta^{(m)}, \mathbf{E}^{(m)})$ may be replaced by $(\rho_0\eta^{(2)}, \mathbf{E}^{(2)})$ without changing the sign of (3.4). Thus from Eqs. (1.4), (2.7), (2.14) and (2.4), we have

$$\hat{\theta}(\rho_0 \Delta\eta) + \hat{\mathbf{T}} : \Delta\mathbf{E} \geq 0, \tag{3.5}$$

where $(\rho_0\Delta\eta, \Delta\mathbf{E})$ in (3.5) may be considered to be a vector in a seven-dimensional space starting from a boundary point $(\rho_0\eta^{(2)}, \mathbf{E}^{(2)})$ of the loading surface

to an arbitrary neighbouring interior point $(\rho_0\eta^{(1)}, \mathbf{E}^{(1)})$ of the loading surface. Hence for given ξ_α, $(\hat{\theta}, \hat{\mathbf{T}})$ must be directed along the inward normal to the loading surface $g = 0$ which is smooth at the boundary point $(\rho_0\eta^{(2)}, \mathbf{E}^{(2)})$ in the entropy-strain space. It follows that

$$\hat{\theta} = -\gamma^* \frac{\partial g}{\rho_0 \partial \eta}, \qquad \hat{\mathbf{T}} = -\gamma^* \frac{\partial g}{\partial \mathbf{E}}, \qquad (\gamma^* \geq 0) \tag{3.6}$$

and with the help of (2.24), (2.25), (2.27), we may deduce from (3.6) that

$$\rho_0 \hat{\eta} = \gamma^* \frac{\partial f}{\partial \theta}, \qquad \hat{\mathbf{E}} = \gamma^* \frac{\partial f}{\partial \mathbf{T}}, \qquad (\gamma^* \geq 0), \tag{3.7}$$

where the right-hand sides of (3.6) and (3.7) are evaluated on the loading surfaces $g = 0$ and $f = 0$, respectively.

It should be pointed out that from experimental observations the normality rule may not be valid in some cases. The discrepancy between the theory and experiments may be attributed to the following facts: 1) there are different ways to define the rate of plastic deformation; therefore, if the normality rule holds for one definition, it may not hold for other definitions; 2) the physical theory of plasticity and experimental results show that the shape and location of a yield (or loading) surface are sensitive to the magnitude of the additional plastic strain-increment used to define the yield stress. So, the normality rule may be violated when this plastic strain-increment is relatively large; 3) in this paper, the evolution equations of the internal variables have the form given by Eq. (2.4). This means that all dissipative mechanisms in the material are governed by the same loading criterion Eq. (2.3). If Eq. (2.4) is not valid, the loading surface loses its original meaning. In this case the normality rule does not exist anymore.

4. Rate-type constitutive equations in thermoplasticity

By means of the normality condition (3.6) or (3.7), expression (2.31) may be rewritten in the form

$$\Phi = 1 - \gamma^* H, \tag{4.1}$$

where

$$H = \frac{\partial g}{\rho_0 \partial \eta} \frac{\partial f}{\partial \theta} + \frac{\partial g}{\partial \mathbf{E}} : \frac{\partial f}{\partial \mathbf{T}}. \tag{4.2}$$

The measure of strain-hardening behaviour of the material element may also be characterized by $h = \Phi/\gamma^*$ for positive γ^*, which gives

$$\gamma^* = \frac{1}{H + h}. \tag{4.3}$$

Hence during a thermoelastic-plastic loading, Eqs. (2.12) may be expressed in the form

$$
\begin{aligned}
\dot{\theta} &= \left[\mathcal{L}_0 - \left(\frac{1}{H+h}\right)\left(\frac{\partial g}{\rho_0 \partial \eta}\right)^2\right](\rho_0\,\dot{\eta}) + \left[\mathcal{L}_1 - \left(\frac{1}{H+h}\right)\frac{\partial g}{\rho_0 \partial \eta}\,\frac{\partial g}{\partial \mathbf{E}}\right] : \dot{\mathbf{E}}, \\
\dot{\mathbf{T}} &= \left[\mathcal{L}_1 - \left(\frac{1}{H+h}\right)\frac{\partial g}{\rho_0 \partial \eta}\,\frac{\partial g}{\partial \mathbf{E}}\right](\rho_0\,\dot{\eta}) + \left[\mathcal{L} - \left(\frac{1}{H+h}\right)\frac{\partial g}{\partial \mathbf{E}} \otimes \frac{\partial g}{\partial \mathbf{E}}\right] : \dot{\mathbf{E}},
\end{aligned} \tag{4.4}
$$

whereas during a thermoelastic unloading, Eqs. (2.12) becomes

$$
\begin{aligned}
\dot{\theta} &= \mathcal{L}_0(\rho_0\,\dot{\eta}) + \mathcal{L}_1 : \dot{\mathbf{E}}, \\
\dot{\mathbf{T}} &= \mathcal{L}_1(\rho_0\,\dot{\eta}) + \mathcal{L} : \dot{\mathbf{E}}\,.
\end{aligned} \tag{4.5}
$$

Correspondingly, along the temperature-stress trajectory, Eqs. (2.15), may be written as

$$
\begin{aligned}
\rho_0\,\dot{\eta} &= \mathcal{M}_0\,\dot{\theta} + \boldsymbol{\mathcal{M}}_1 : \dot{\mathbf{T}} + \gamma^* < \hat{g} > \frac{\partial f}{\partial \theta}\,, \\
\dot{\mathbf{E}} &= \boldsymbol{\mathcal{M}}_1\,\dot{\theta} + \boldsymbol{\mathcal{M}} : \dot{\mathbf{T}} + \gamma^* < \hat{g} > \frac{\partial f}{\partial \mathbf{T}}\,.
\end{aligned} \tag{4.6}
$$

During a thermoelastic-plastic loading ($\hat{g} > 0$), $\hat{g}$ may be expressed by $\hat{f}/(h\gamma^*)$ provided $\Phi \neq 0$. It then follows from (4.6) that for hardening and for softening behaviour ($h \neq 0$)

$$
\begin{aligned}
\rho_0\,\dot{\eta} &= \left[\mathcal{M}_0 + \left(\frac{1}{h}\right)\left(\frac{\partial f}{\partial \theta}\right)^2\right]\dot{\theta} + \left[\boldsymbol{\mathcal{M}}_1 + \left(\frac{1}{h}\right)\frac{\partial f}{\partial \theta}\,\frac{\partial f}{\partial \mathbf{T}}\right] : \dot{\mathbf{T}}, \\
\dot{\mathbf{E}} &= \left[\boldsymbol{\mathcal{M}}_1 + \left(\frac{1}{h}\right)\frac{\partial f}{\partial \theta}\,\frac{\partial f}{\partial \mathbf{T}}\right]\dot{\theta} + \left[\boldsymbol{\mathcal{M}} + \left(\frac{1}{h}\right)\frac{\partial f}{\partial \mathbf{T}} \otimes \frac{\partial f}{\partial \mathbf{T}}\right] : \dot{\mathbf{T}}\,.
\end{aligned} \tag{4.7}
$$

In the case of perfectly plastic behaviour ($h = 0$), the above expression is meaningless. But Eqs. (4.6) is still valid with γ^* replaced by $\left(\frac{1}{H}\right)$. Moreover, it is not dfficult to see that if in a thermoelastic-plastic state the material is exhibiting perfectly plastic behaviour ($h = 0$), then there is one and only one loading direction $(\rho_0\,\dot{\eta}, \dot{\mathbf{E}})$ which is the outward normal to the loading surface $f = 0$ in the temperature-stress space, such that for this looading direction, the rates of temperature and stress $(\dot{\theta}, \dot{\mathbf{T}})$ vanish.

In the following, we assume that the thermoelastic matrices

$$
(\mathbf{L}) = \begin{pmatrix} \mathcal{L}_0 & \mathcal{L}_1 \\ \mathcal{L}_1 & \mathcal{L} \end{pmatrix} \quad \text{and} \quad (\mathbf{M}) = \begin{pmatrix} \mathcal{M}_0 & \boldsymbol{\mathcal{M}}_1 \\ \boldsymbol{\mathcal{M}}_1 & \boldsymbol{\mathcal{M}} \end{pmatrix} \tag{4.8}
$$

are positive definite. This implies that $H > 0$. With the use of notations

$$
\begin{aligned}
\mathcal{L}_0^{ep} &= \mathcal{L}_0 - \left(\frac{1}{H+h}\right)\left(\frac{\partial g}{\rho_0 \partial \eta}\right)^2, \\
\mathcal{L}_1^{ep} &= \mathcal{L}_1 - \left(\frac{1}{H+h}\right)\frac{\partial g}{\rho_0 \partial \eta}\frac{\partial g}{\partial \mathbf{E}}, \\
\mathcal{L}^{ep} &= \mathcal{L} - \left(\frac{1}{H+h}\right)\frac{\partial g}{\partial \mathbf{E}} \otimes \frac{\partial g}{\partial \mathbf{E}},
\end{aligned} \tag{4.9}
$$

$$
(\mathbf{L}^{ep}) = \begin{pmatrix} \mathcal{L}_0^{ep} & \mathcal{L}_1^{ep} \\ \mathcal{L}_1^{ep} & \mathcal{L}^{ep} \end{pmatrix}, \tag{4.10}
$$

we obtain

$$
(\mathbf{L}^{ep}) \begin{pmatrix} \dfrac{\partial f}{\partial \theta} \\ \dfrac{\partial f}{\partial \mathbf{T}} \end{pmatrix} = \Phi(\mathbf{L}) \begin{pmatrix} \dfrac{\partial f}{\partial \theta} \\ \dfrac{\partial f}{\partial \mathbf{T}} \end{pmatrix}, \tag{4.11}
$$

or

$$
(\overline{\mathbf{M}}) \begin{pmatrix} \dfrac{\partial f}{\partial \theta} \\ \dfrac{\partial f}{\partial \mathbf{T}} \end{pmatrix} = \Phi \begin{pmatrix} \dfrac{\partial f}{\partial \theta} \\ \dfrac{\partial f}{\partial \mathbf{T}} \end{pmatrix}, \tag{4.12}
$$

where $(\overline{\mathbf{M}}) = (\mathbf{M})(\mathbf{L}^{ep})$ and $(\mathbf{M})$ is the inverse matrix of $(\mathbf{L})$. The above expression indicates that the seven-dimensional vector $\left(\dfrac{\partial f}{\partial \theta}, \dfrac{\partial f}{\partial \mathbf{T}}\right)^T$ is an eigenvector of the matrix $(\overline{\mathbf{M}})$ with eigenvalue Φ. Since in a seven-dimensional space there always exist six linearly independent vectors $(f_{\theta p}^*, \mathbf{f}_p^*)^T$, $(p = 1, 2, \ldots, 6)$ such that

$$
f_{\theta p}^* \frac{\partial g}{\rho_0 \partial \eta} + \mathbf{f}_p^* : \frac{\partial g}{\partial \mathbf{E}} = (f_{\theta p}^*, \mathbf{f}_p^*)(\mathbf{L}) \begin{pmatrix} \dfrac{\partial f}{\partial \theta} \\ \dfrac{\partial f}{\partial \mathbf{T}} \end{pmatrix} = 0 \qquad (p = 1, 2, \ldots, 6),
$$

and

$$
(f_{\theta r}^*, \mathbf{f}_r^*)(\mathbf{L}) \begin{pmatrix} f_{\theta p}^* \\ \mathbf{f}_p^* \end{pmatrix} = 0 \qquad \begin{pmatrix} r, p = 1, 2, \ldots, 6 \\ r \neq p \end{pmatrix},
$$

so, in view of the positive definiteness of the matrix $(\mathbf{L})$, we may conclude that $\left(\dfrac{\partial f}{\partial \theta}, \dfrac{\partial f}{\partial \mathbf{T}}\right)$ and $(f_{\theta p}^*, \mathbf{f}_p^*)$, $(p = 1, 2, \ldots, 6)$ are linearly independent and $(f_{\theta p}^*, \mathbf{f}_p^*)$,

($p = 1, 2, \ldots, 6$) are also eigenvectors of the matrix $(\overline{\mathbf{M}})$ with their eigenvalues being unities. Hence the determinant of the matrix $(\overline{\mathbf{M}})$ is equal in value to Φ, i.e.,

$$\Phi = \det(\mathbf{L}^{ep})/\det(\mathbf{L}). \tag{4.13}$$

Finally, from Eqs. (2.20), Eqs. (4.4) may be written in the matrix form as

$$\begin{pmatrix} \dot{\theta} \\ \dot{\mathbf{T}} \end{pmatrix} = (\mathbf{L}^{ep})(\mathbf{M}) \begin{pmatrix} \dot{\theta}_e \\ \dot{\mathbf{T}}_e \end{pmatrix}. \tag{4.14}$$

This means that Φ is the determinant of the matrix which transforms the seven-dimensional vector $(\dot{\theta}_e, \dot{\mathbf{T}}_e)$ into $(\dot{\theta}, \dot{\mathbf{T}})$.

5. The temperature-strain space formulation

In the previous sections we have introduced the thermoplastic constitutive equations in both "entropy-strain" and its conjugate "temperature-stress" spaces. But it may be convenient, in certain circumstances, to develop the "temperature-strain" space formulation of the thermoplastic theory as follows. Let us take $(\theta, \mathbf{E}, \xi_\alpha)$ as the state variables and rewrite Eqs. (1.3) in the form

$$\begin{aligned} \overline{\eta} &= \overline{\eta}(\theta, \mathbf{E}, \xi_\alpha) = -\frac{\partial \psi}{\partial \theta}, \\ \overline{\mathbf{T}} &= \overline{\mathbf{T}}(\theta, \mathbf{E}, \xi_\alpha) = \rho_0 \frac{\partial \psi}{\partial \mathbf{E}}, \end{aligned} \tag{5.1}$$

where $\psi = \psi(\theta, \mathbf{E}, \xi_\alpha)$ is the specific (Helmholtz) free energy, $\overline{\eta}$ and $\overline{\mathbf{T}}$ are entropy and stress, respectively, which are dependent on the variables $(\theta, \mathbf{E}, \xi_\alpha)$. The time rates of Eqs.(5.1) may be given by

$$\begin{aligned} \rho_0 \dot{\overline{\eta}} &= \rho_0 \left(\frac{C_E}{\theta}\right) \dot{\theta} + \overline{\mathcal{L}_1} : \dot{\mathbf{E}} + \rho_0 \hat{\overline{\eta}} < \hat{g} >, \\ \dot{\overline{\mathbf{T}}} &= -\overline{\mathcal{L}_1}\, \dot{\theta} + \overline{\mathcal{L}} : \dot{\mathbf{E}} + \hat{\overline{\mathbf{T}}} < \hat{g} >. \end{aligned} \tag{5.2}$$

With the help of identities

$$\begin{aligned} \overline{\mathbf{T}}(\theta, \mathbf{E}, \xi_\alpha) &= \mathbf{T}(\rho_0 \overline{\eta}(\theta, \mathbf{E}, \xi_\beta), \mathbf{E}, \xi_\alpha), \\ \eta &= \overline{\eta}(\theta(\rho_0 \eta, \mathbf{E}, \xi_\beta), \mathbf{E}, \xi_\alpha), \\ \mathbf{T} &= \overline{\mathbf{T}}(\theta, \mathbf{E}(\theta, \mathbf{T}, \xi_\beta), \xi_\alpha), \end{aligned} \tag{5.3}$$

coefficients in (5.2) may be written as

$$C_E = \theta \frac{\partial \overline{\eta}}{\partial \theta} = \frac{\theta}{\rho_0 \mathcal{L}_0}, \tag{5.4}$$

which is called the specific heat at fixed $\mathbf{E}$ and ξ_α, and

$$\begin{aligned} \overline{\mathcal{L}_1} &= -\frac{\partial \overline{\mathbf{T}}}{\partial \theta} = -\rho_0 \frac{\partial^2 \psi}{\partial \theta \partial \mathbf{E}} = \rho_0 \frac{\partial \overline{\eta}}{\partial \mathbf{E}} = -\frac{1}{\mathcal{L}_0} \mathcal{L}_1, \\ \overline{\mathcal{L}} &= \frac{\partial \overline{\mathbf{T}}}{\partial \mathbf{E}} = \rho_0 \frac{\partial^2 \psi}{\partial \mathbf{E} \partial \mathbf{E}} = \mathcal{L} - \frac{1}{\mathcal{L}_0} \mathcal{L}_1 \otimes \mathcal{L}_1, \end{aligned} \tag{5.5}$$

which represent a second order and a fourth order symmetric tensor, respectively.

The specific heat at fixed $\mathbf{T}$ and ξ_α is defined by

$$C_T = \theta \frac{\partial \eta}{\partial \theta} = \frac{\theta \mathcal{M}_0}{\rho_0}$$

and may be expressed through C_E by

$$C_T - C_E = \frac{\theta}{\rho_0} \overline{\mathcal{L}_1} : \mathcal{M}_1 = -C_E \mathcal{L}_1 : \mathcal{M}_1. \tag{5.6}$$

$\overline{\mathcal{L}}$ in Eqs. (5.5) may be regarded as the inverse of $\mathcal{M}$ in Eqs. (2.18), i.e.,

$$\overline{\mathcal{L}} : \mathcal{M} = \mathcal{M} : \overline{\mathcal{L}} = \overset{(1)}{\mathbf{I}}. \tag{5.7}$$

Hence from

$$\overline{\mathcal{L}_1} = \overline{\mathcal{L}} : \mathcal{M}_1 = \mathcal{M}_1 : \overline{\mathcal{L}}$$

or (5.8)

$$\mathcal{M}_1 = \mathcal{M} : \overline{\mathcal{L}_1} = \overline{\mathcal{L}}_1 : \mathcal{M}$$

and from Eqs. (2.18), (5.5)–(5.8), we obtain

$$\mathcal{L} : \mathcal{M}_1 = \left(\frac{C_T}{C_E}\right) \overline{\mathcal{L}} : \mathcal{M}_1$$

or

$$(\mathcal{M} : \mathcal{L}) : \mathcal{M}_1 = \left(\frac{C_T}{C_E}\right) \mathcal{M}_1. \tag{5.9}$$

This means that the six-dimensional vector $\mathcal{M}_1$ is an eigenvector of $\mathcal{M} : \mathcal{L}$ with eigenvalue $\dfrac{C_T}{C_E} = \mathcal{L}_0 \mathcal{M}_0$. Since in six-dimensional space there exist five vectors $\mathcal{L}_p$ $(p = 2, 3, \ldots, 6)$ such that

$$\mathcal{L}_1 : \mathcal{L}_p = 0, \qquad \mathcal{L}_p : \mathcal{L}_r = 0 \qquad \begin{pmatrix} r, p = 2, \ldots, 6 \\ r \neq p \end{pmatrix},$$

it then follows that every vector $\mathcal{L}_p$ $(p = 2, \ldots, 6)$ is an eigenvector of $\mathcal{M} : \mathcal{L} = {}^{(1)}\mathbf{I} - \mathcal{M}_1 \otimes \mathcal{L}_1$ with the eigenvalue unity. Hence we have

$$(\det \mathcal{M})(\det \mathcal{L}) = \frac{C_T}{C_E} = \mathcal{L}_0 \mathcal{M}_0 = 1 - \mathcal{L}_1 : \mathcal{M}_1. \tag{5.10}$$

With the use of Eq. (5.7), (5.8) and (2.15), $\rho_0 \hat{\bar{\eta}}$ and $\hat{\bar{\mathbf{T}}}$ in Eqs. (5.2) may also be expressed in terms of $\hat{\theta}$ and $\hat{E}$:

$$\rho_0 \hat{\bar{\eta}} = -\frac{1}{\mathcal{L}_0} \hat{\theta}, \qquad \hat{\bar{\mathbf{T}}} = -\bar{\mathcal{L}} : \hat{\mathbf{E}}. \tag{5.11}$$

So Eqs. (5.2) may be rewritten in the form

$$\begin{aligned} \rho_0 \dot{\bar{\eta}} &= \rho_0 \frac{C_E}{\theta} (\dot{\theta} - \hat{\theta} < \hat{g} >) + \bar{\mathcal{L}}_1 : \dot{\mathbf{E}}, \\ \dot{\bar{\mathbf{T}}} &= -\bar{\mathcal{L}}_1 \dot{\theta} + \bar{\mathcal{L}} : (\dot{\mathbf{E}} - \hat{\mathbf{E}} < \hat{g} >). \end{aligned} \tag{5.12}$$

The expression of the loading surface in the "temperature-strain" space can be found through the relation

$$\begin{aligned} \bar{g} &= \bar{g}(\theta, \mathbf{E}, \xi_\alpha) = \bar{g}(\theta(\rho_0 \eta, \mathbf{E}, \xi_\beta), \mathbf{E}, \xi_\alpha) = g(\rho_0 \eta, \mathbf{E}, \xi_\alpha) \\ \text{or} \quad \bar{g} &= \bar{g}(\theta, \mathbf{E}, \xi_\alpha) = \bar{g}(\theta, \mathbf{E}(\theta, \mathbf{T}, \xi_\beta), \xi_\alpha) = f(\theta, \mathbf{T}, \xi_\alpha), \end{aligned} \tag{5.13}$$

which leads to

$$\begin{aligned} &\frac{\partial g}{\rho_0 \partial \eta} = \frac{\partial \bar{g}}{\partial \theta} \mathcal{L}_0, \qquad \frac{\partial g}{\partial \mathbf{E}} = \frac{\partial \bar{g}}{\partial \theta} \mathcal{L}_1 + \frac{\partial \bar{g}}{\partial \mathbf{E}}, \\ &\frac{\partial f}{\partial \theta} = \frac{\partial \bar{g}}{\partial \theta} + \frac{\partial \bar{g}}{\partial \mathbf{E}} : \mathcal{M}_1, \qquad \frac{\partial f}{\partial \mathbf{T}} = \frac{\partial \bar{g}}{\partial \mathbf{E}} : \mathcal{M}, \\ \text{or} \quad &\frac{\partial g}{\partial \mathbf{E}} = \frac{\partial f}{\partial \mathbf{T}} : \bar{\mathcal{L}}. \end{aligned} \tag{5.14}$$

From the normality conditions (3.6) and (3.7), the inelastic part of $(-\rho_0 \dot{\bar{\eta}}, \dot{\bar{\mathbf{T}}})$ may be given by

$$\begin{aligned} -\rho_0 \hat{\bar{\eta}} \hat{g} &= \frac{1}{\mathcal{L}_0} \hat{\theta} \hat{g} = -\gamma^* \frac{\partial \bar{g}}{\partial \theta} \left(\frac{\hat{\bar{g}}}{1 + \hat{\theta} \dfrac{\partial \bar{g}}{\partial \theta}} \right), \\ \hat{\bar{\mathbf{T}}} \hat{g} &= -\bar{\mathcal{L}} : \hat{\mathbf{E}} \hat{g} = -\gamma^* \frac{\partial \bar{g}}{\partial \mathbf{E}} \left(\frac{\hat{\bar{g}}}{1 + \hat{\theta} \dfrac{\partial \bar{g}}{\partial \theta}} \right), \end{aligned}$$

where the symbol

$$\hat{\overline{g}} \triangleq \frac{\partial \overline{g}}{\partial \theta}\,\dot{\theta} + \frac{\partial \overline{g}}{\partial \mathbf{E}} : \dot{\mathbf{E}} = \left(1 + \hat{\theta}\frac{\partial \overline{g}}{\partial \theta}\right)\hat{g} \tag{5.15}$$

has been used. Therefore, the inelastic part of $(-\rho_0\,\dot{\overline{\eta}}, \dot{\overline{\mathbf{T}}})$ is directed along the inward normal to the loading surface $\overline{g} = 0$ in the "temperature-strain" space. Finally, it can be seen from (5.14) that the expression of Φ in Eq. (2.31) may also be written in the form

$$\Phi = 1 + \frac{\partial \overline{g}}{\partial \theta}\hat{\theta} - \frac{\partial \overline{g}}{\partial \mathbf{E}} : \hat{\mathbf{E}}. \tag{5.16}$$

Obviously, this is consistent with the isothermal plastic theory.

6. Invariance relations

In the above discussions the strain measure is based on a certain fixed configuration, which is not necessarily an unstressed one. Each change of the reference configuration and/or of the strain measure will lead to a corresponding change of the stress $\mathbf{T}$ and the parameters appearing in the constitutive equations. In this section the invariance relations developed by HILL for isothermal elasto-plasticity ([5, 23]) will be generalized to non-isothermal cases.

Suppose there are two ways of specifying the deformation by strains and densities in the reference configurations, $\mathbf{E}$, ρ_0 and $\mathbf{E}^*$, ρ_0^*, respectively. In the simplest case, it may be assumed that $\mathbf{E}$ is a known function of $\mathbf{E}^*$

$$\mathbf{E} = \mathbf{E}(\mathbf{E}^*). \tag{6.1}$$

Hence

$$\dot{\mathbf{E}} = \boldsymbol{\mathcal{A}} : \dot{\mathbf{E}}^*, \tag{6.2}$$

where

$$\boldsymbol{\mathcal{A}} = \frac{\partial \mathbf{E}}{\partial \mathbf{E}^*} \tag{6.3}$$

and its component may be expressed by

$$\mathcal{A}_{AB}^{\cdot\cdot\ MN} = \frac{\partial E_{AB}}{\partial E_{MN}^*}. \tag{6.4}$$

It may be further assumed that the specific internal energy ε, the specific entropy η, the absolute temperature θ and the rate of heat absorption per unit

reference mass $\frac{1}{\rho_0}\dot{Q}$ are invariant under the transformation of the reference configuration and the strain measure. Then the specific (Helmholtz) free energy ψ is also invariant. In view of the First Law of Thermodynamics,

$$\dot{Q} = \rho_0 r - \nabla_0 \cdot \mathbf{q}_0 = \rho_0 \dot{\varepsilon} - \mathbf{T} : \dot{\mathbf{E}} = \rho_0 \theta \dot{\eta} - A^\alpha \dot{\xi}_\alpha, \tag{6.5}$$

we see that

$$\frac{1}{\rho_0}\mathbf{T} : \dot{\mathbf{E}} \qquad \text{and} \qquad \frac{1}{\rho_0} A^\alpha \dot{\xi}_\alpha$$

are invariant. Therefore, from

$$\frac{1}{\rho_0}\mathbf{T} : \dot{\mathbf{E}} = \frac{1}{\rho_0^*}\mathbf{T}^* : \dot{\mathbf{E}}^*$$

we have

$$\frac{\rho_0}{\rho_0^*}\mathbf{T}^* = \mathbf{T} : \boldsymbol{\mathcal{A}} = \boldsymbol{\mathcal{A}}^T : \mathbf{T} \tag{6.6}$$

and

$$\frac{\rho_0}{\rho_0^*}(\dot{\mathbf{T}}^* - \boldsymbol{\mathcal{B}} : \dot{\mathbf{E}}^*) = \boldsymbol{\mathcal{A}}^T : \dot{\mathbf{T}}, \tag{6.7}$$

where

$$\boldsymbol{\mathcal{B}} = \boldsymbol{\mathcal{B}}^T = \frac{\rho_0^*}{\rho_0}\mathbf{T} : \frac{\partial \boldsymbol{\mathcal{A}}}{\partial \mathbf{E}^*} \tag{6.8}$$

is a fourth-order symmetric tensor with its components

$$\mathcal{B}^{KLMN} = \mathcal{B}^{MNKL} = \frac{\rho_0^*}{\rho_0} T^{AB} \frac{\partial^2 E_{AB}}{\partial E_{KL}^* \partial E_{MN}^*}.$$

Also, from Eqs. (5.1),

$$\overline{\mathbf{T}}^* = \rho_0^* \frac{\partial \psi}{\partial \mathbf{E}^*} = \rho_0^* \frac{\partial \psi}{\partial \mathbf{E}} : \boldsymbol{\mathcal{A}} = \frac{\rho_0^*}{\rho_0}\overline{\mathbf{T}} : \boldsymbol{\mathcal{A}}, \tag{6.9}$$

we may conclude that the transformation rules of $\overline{\mathbf{T}}$ and $\dot{\overline{\mathbf{T}}}$ are the same as those of $\mathbf{T}$ (Eq. (6.6)) and $\dot{\mathbf{T}}$ (Eq. (6.7)), respectively.

Now let us consider the transformation rules of the parameters appearing in Eqs. (2.12). Obviously, $\rho_0 \mathcal{L}_0 = \frac{\partial \theta}{\partial \eta}$ and the specific heat per unit reference mass at fixed strain and internal variables $C_E = \theta/(\rho_0 \mathcal{L}_0)$ are invariant, i.e.,

$$\frac{\rho_0^*}{\rho_0}\mathcal{L}_0^* = \mathcal{L}_0. \tag{6.10}$$

By combining

$$\mathcal{L}_1^* = \frac{\partial \theta^*}{\partial \mathbf{E}^*} = \frac{\partial \theta}{\partial \mathbf{E}} : \mathcal{A} = \mathcal{L}_1 : \mathcal{A} \tag{6.11}$$

with (5.5) and (6.10), it follows that

$$\frac{\rho_0}{\rho_0^*}\overline{\mathcal{L}}_1^* = \overline{\mathcal{L}}_1 : \mathcal{A}, \tag{6.12}$$

which means that $\overline{\mathcal{L}}_1$ has the transformation rule like $\mathbf{T}$ in Eq. (6.6). Since in seven-dimensional "entropy-strain" space there always exist six linearly independent unloading directions, so from (2,12), (6.7), and (6.11) we obtain

$$\frac{\rho_0}{\rho_0^*}(\mathcal{L}^* - \mathcal{B}) = \mathcal{A}^T : \mathcal{L} : \mathcal{A}. \tag{6.13}$$

Likewise, in view of (5.2) and (6.9),

$$\frac{\rho_0}{\rho_0^*}(\overline{\mathcal{L}}^* - \overline{\mathcal{B}}) = \mathcal{A}^T : \overline{\mathcal{L}} : \mathcal{A}, \tag{6.14}$$

where $\overline{\mathcal{B}}(= \mathcal{B})$ is a function of $(\theta, \mathbf{E}, \xi_\alpha)$. By eliminating $\mathcal{B}$ from (6.13) and (6.14),

$$\frac{\rho_0}{\rho_0^*}(\overline{\mathcal{L}} - \mathcal{L}^*) = \mathcal{A}^T : (\overline{\mathcal{L}} - \mathcal{L}) : \mathcal{A}$$

which shows that the bilinear form

$$\frac{1}{\rho_0} d\mathbf{E}^{(1)} : (\overline{\mathcal{L}} - \mathcal{L}) : d\mathbf{E}^{(2)}$$

is invariant, where $d\mathbf{E}^{(1)}$ and $d\mathbf{E}^{(2)}$ are any two differential increments of the strain.

The transformation rules of the parameters in Eqs. (2.15) may be discussed as follows. From Eqs. (5.7), (5.8), (6.12) and (6.14), we have

$$\frac{\rho_0}{\rho_0^*}(\overline{\mathcal{L}}^* : \mathcal{M}_1^*) = \frac{\rho_0}{\rho_0^*}\overline{\mathcal{L}}_1^* = \mathcal{A}^T : \overline{\mathcal{L}}_1 = \mathcal{A}^T : \overline{\mathcal{L}} : \mathcal{M}_1 = \frac{\rho_0}{\rho_0^*}(\overline{\mathcal{L}}^* - \overline{\mathcal{B}}) : \mathcal{A}^{-1} : \mathcal{M}_1$$

and

$$\frac{\rho_0}{\rho_0^*}(\overset{(1)}{\mathbf{I}} - \mathcal{M}^* : \overline{\mathcal{B}}) = \mathcal{M}^* : \mathcal{A}^T : \mathcal{M}^{-1} : \mathcal{A}.$$

Therefore,

$$\mathcal{M}_1^* = (\overset{(1)}{\mathbf{I}} - \mathcal{M}^* : \mathcal{B}) : \mathcal{A}^{-1} : \mathcal{M}_1, \tag{6.15}$$

$$\mathcal{M} = \left(\frac{\rho_0^*}{\rho_0}\right)\mathcal{A} : (\overset{(1)}{\mathbf{I}} - \mathcal{M}^* : \mathcal{B})^{-1} : \mathcal{M}^* : \mathcal{A}^T. \tag{6.16}$$

The value of $\mathcal{M}_0^*$ may be given, from Eqs. (2.18), (5.7), (5.8), (6.12) and (6.15), by

$$\mathcal{M}_0^* = \frac{1}{\mathcal{L}_0^*} + \overline{\mathcal{L}}_1^* : \mathcal{M}_1^* = \frac{1}{\mathcal{L}_0^*} + \left(\frac{\rho_0^*}{\rho_0}\right)\overline{\mathcal{L}}_1 : \mathcal{A} : (\overset{(1)}{\mathbf{I}} - \mathcal{M}^* : \mathcal{B}) : \mathcal{A}^{-1} : \mathcal{M}_1.$$

This leads to the expression

$$\mathcal{M}_0^* = \left(\frac{\rho_0^*}{\rho_0}\right)\left[\mathcal{M}_0 - \mathcal{M}_1 : \overline{\mathcal{L}} : \mathcal{A} : \mathcal{M}^* : \mathcal{B} : \mathcal{A}^{-1} : \mathcal{M}_1\right]. \tag{6.17}$$

With the use of Eq. (6.16)

$$\mathcal{A} : \mathcal{M}^* = \left(\frac{\rho_0}{\rho_0^*}\right)\mathcal{M} : \mathcal{A}^{-T} : (\overset{(1)}{\mathbf{I}} - \mathcal{B} : \mathcal{M}^*),$$

Eq. (6.17) may be rewritten in the form

$$\frac{\mathcal{M}_0^*}{\rho_0^*} - \frac{\mathcal{M}_0}{\rho_0} = -\mathcal{M}_1 : \mathcal{A}^{-T} : (\mathcal{B} - \mathcal{B} : \mathcal{M}^* : \mathcal{B}) : \mathcal{A}^{-1} : \mathcal{M}_1 \tag{6.18}$$

or equivalently,

$$C_T^* - C_T = \theta \mathcal{M}_1 : \mathcal{A}^{-T} : (\mathcal{B} - \mathcal{B} : \mathcal{M}^* : \mathcal{B}) : \mathcal{A}^{-1} : \mathcal{M}_1. \tag{6.19}$$

Suppose that the loading surfaces in different strain measures satisfy

$$\frac{1}{\rho_0} g(\rho_0 \eta, \mathbf{E}, \xi_\alpha) = \frac{1}{\rho_0^*} g^*(\rho_0^* \eta^*, \mathbf{E}^*, \xi_\alpha). \tag{6.20}$$

It follows that

$$\frac{\partial g^*}{\rho_0^* \partial \eta^*} = \frac{\partial g}{\rho_0 \partial \eta} \quad \text{and} \quad \frac{\rho_0}{\rho_0^*}\frac{\partial g^*}{\partial \mathbf{E}^*} = \frac{\partial g}{\partial E} : \mathcal{A}. \tag{6.21}$$

Hence $\dfrac{\partial g}{\rho_0 \partial \eta}$ is invariant and the transformation rule of $\dfrac{\partial g}{\partial \mathbf{E}}$ is the same with that of $\mathbf{T}$. Also, from Eqs. (5.14), (6.11), (6.12) and (6.21), we see that

$$\frac{1}{\rho_0}\frac{\partial \overline{g}}{\partial \theta}$$

is invariant and the transformation rule of $\dfrac{\partial \overline{g}}{\partial \mathbf{E}}$ is the same with that of $\overline{\mathcal{L}}_1$ (or $\mathbf{T}$). It can be seen from (5.8) and (5.14) that the relation between $\dfrac{\partial f}{\partial \mathbf{T}}$ and $\dfrac{\partial \overline{g}}{\partial \mathbf{E}}$ is

the same as that between $\mathcal{M}_1$ and $\overline{\mathcal{L}}_1$. Therefore, the transformation rule of $\dfrac{\partial f}{\partial \mathbf{T}}$ should be the same as that of $\mathcal{M}_1$. Moreover, with the use of (5.14), (6.15) and (6.16), we obtain

$$(6.22)\qquad \frac{1}{\rho_0^*}\left(\frac{\partial f^*}{\partial \theta^*}\right) - \frac{1}{\rho_0}\left(\frac{\partial f}{\partial \theta}\right) = -\frac{1}{\rho_0}\frac{\partial \overline{g}}{\partial \mathbf{E}} : \mathcal{A} : \mathcal{M}^* : \mathcal{B} : \mathcal{A}^{-1} : \mathcal{M}_1$$

$$= -\frac{1}{\rho_0^*}\frac{\partial \overline{g}}{\partial \mathbf{E}} : \mathcal{M} : \mathcal{A}^{-T} : (\overset{(1)}{\mathbf{I}} - \mathcal{B} : \mathcal{M}^*) : \mathcal{B} : \mathcal{A}^{-1} : \mathcal{M}_1$$

$$= \frac{1}{\rho_0}\frac{\partial f}{\partial \mathbf{T}} : \mathcal{A}^{-T} : (\mathcal{B} : \mathcal{M}^* : \mathcal{B} - \mathcal{B}) : \mathcal{A}^{-1} : \mathcal{M}_1.$$

It remains to discuss the transformation rules of $\hat{\theta}$, $\hat{\mathbf{T}}$, $\rho_0\hat{\eta}$ and $\hat{\mathbf{E}}$. According to Eqs. (6.21) and (5.1), values of $\dfrac{1}{\rho_0}\hat{g}$ and $\rho_0\hat{\overline{\eta}}$ are invariant under transformation (6.1), whereas $\rho_0\hat{\mathbf{T}}$ and $\rho_0\hat{\overline{\mathbf{T}}}$ have the same transformation rule as that of $\mathbf{T}$. Hence, in accordance with Eqs. (5.11) and (6.10), it is seen that $\rho_0\hat{\theta}$ is invariant. Suppose that the normality condition holds. Then from (3.6) and (3.7) we see that $\rho_0\gamma^*$ (or $\rho_0/(H+h)$) is also invariant, and $\rho_0\hat{\eta}$ and $\rho_0\hat{\mathbf{E}}$ transform like $\dfrac{1}{\rho_0}\dfrac{\partial f}{\partial \theta}$ and $\dfrac{\partial f}{\partial \mathbf{T}}$, respectively.

According to the invariance theory developed by Hill, the choice of reference configuration can be arbitrary. Especially, the reference configuration may be chosen to be instantaneously coincident with the current configuration or the intermediate stress-free configuration (if it exists). Suppose that for the same deformation, the spectral decomposition of the strain $\mathbf{E}^*$ in (6.1) may be expressed by

$$(6.23)\qquad \mathbf{E}^* = \sum_{k=1}^{3} e^*(\lambda_k)\mathbf{N}_k \otimes \mathbf{N}_k,$$

where $e^*(\lambda_k)$ is another scale function different from that in (1.6). When the deformation is quite small, Eq. (6.1) may be written in a series form as

$$(6.24)\qquad \mathbf{E} = \mathbf{E}^* + (m - m^*)\mathbf{E}^* \cdot \mathbf{E}^* + \ldots,$$

where $m = \dfrac{1}{2}[1 + e''(1)]$, $m^* = \dfrac{1}{2}[1 + e^{*\prime\prime}(1)]$. Taking the current configuration as a reference configuration, we have

$$(6.25)\qquad \mathbf{E} = \mathbf{E}^* = 0, \qquad \rho_0 = \rho_0^*, \qquad \dot{\mathbf{E}} = \dot{\mathbf{E}}^* = \mathbf{D}, \qquad \mathbf{T} = \mathbf{T}^* = \boldsymbol{\tau},$$

where $\mathbf{D}$ is the rate of deformation, $\boldsymbol{\tau} = \tau^{AB}\mathbf{G}_A \otimes \mathbf{G}_B$ is the Kirchhoff stress and $\mathbf{G}_A$ $(A = 1, 2, 3)$ are covariant base vectors in the reference configuration. Hence

(6.3) and (6.8) reduce to

$$\mathcal{A} = \overset{(1)}{\mathbf{I}}, \qquad \mathcal{B} = (m - m^*)\mathcal{T}, \tag{6.26}$$

where

$$\mathcal{T} = \frac{1}{2}\left(\tau^{AM}G^{BN} + \tau^{BM}G^{AN} + \tau^{AN}G^{BM} + \tau^{BN}G^{AM}\right)\mathbf{G}_A \otimes \mathbf{G}_B \otimes \mathbf{G}_M \otimes \mathbf{G}_N$$

and G^{AB} are contravariant components of the metric tensor in the reference configuration. Specifically, when $\mathbf{E}^*$ is the logarithmic strain $\mathbf{E}^{(0)}$ ($m^* = 0$), the logarithmic stress rate will be the Jaumann rate of Kirchoff stress: $\dot{\mathbf{T}}^{(0)} = \overset{\nabla}{\boldsymbol{\tau}}$, and Eq. (6.7) reduces to

$$\dot{\mathbf{T}} = \overset{\nabla}{\boldsymbol{\tau}} - m\mathcal{T} : \mathbf{D} = \overset{\nabla}{\boldsymbol{\tau}} - m(\boldsymbol{\tau}\cdot\mathbf{D} + \mathbf{D}\cdot\boldsymbol{\tau}).$$

These results may also be found in a series of papers by Hill.

We now discuss the constitutive theory based on the decomposition (1.7). The internal variables ξ_α ($\alpha = 1, 2, \ldots, n$) may be taken as ζ_δ ($\delta = 1, 2, \ldots, n-1$) and $\mathbf{F}_p$. When the intermediate configuration is chosen to be the reference configuration, the state variables will become $(\theta, \mathbf{E}^e, \zeta_\delta)$, where $\mathbf{E}^e$ is the elastic strain expressed in terms of $\mathbf{E}$ and $\mathbf{F}_p$:

$$\mathbf{E}^e = \mathbf{E}^e(\mathbf{E}, \mathbf{F}_p). \tag{6.27}$$

The inverse relation of (6.27) is assumed to be

$$\mathbf{E} = \mathbf{E}(\mathbf{E}^e, \mathbf{F}_p).$$

Thus the expression of specific free energy may be given by

$$\psi_e(\theta, \mathbf{E}^e, \zeta_\delta) = \psi_e(\theta, \mathbf{E}^e(\mathbf{E}, \mathbf{F}_p), \zeta_\delta) = \psi(\theta, \mathbf{E}, \zeta_\delta, \mathbf{F}_p). \tag{6.28}$$

Since Eq. (6.27) is different from (6.1), the corresponding transformation rules should be revised accordingly. The stress $\overline{\mathbf{T}}^e$ conjugate to $\mathbf{E}^e$ may be defined by

$$\overline{\mathbf{T}}^e = \rho_e \frac{\partial \psi_e}{\partial \mathbf{E}^e} = \rho_e \frac{\partial \psi}{\partial \mathbf{E}} : \frac{\partial \mathbf{E}}{\partial \mathbf{E}^e} = \rho_e \frac{\partial \psi}{\partial \mathbf{E}} : \mathcal{A}_e, \tag{6.29}$$

where $\mathbf{F}_p$ is regarded as a parameter, ρ_e is the density in the intermediate configuration, and $\mathcal{A}_e = \dfrac{\partial \mathbf{E}}{\partial \mathbf{E}^e}$. It follows

$$\frac{\rho_0}{\rho_e}\overline{\mathbf{T}}^e = \overline{\mathbf{T}} : \mathcal{A}_e = \mathcal{A}_e^T : \overline{\mathbf{T}}. \tag{6.30}$$

Values of ζ_δ and $\mathbf{F}_p$ will remain unchanged during thermoelastic unloading. Hence the same transformation rules as above may be obtained (e.g. Eqs (6.7), (6.10), (6.11) ...), if $\mathcal{A}$ and $\mathcal{B}$ are replaced by $\mathcal{A}_e$ and

$$\mathcal{B}_e = \frac{\rho_e}{\rho_0}\overline{\mathbf{T}} : \frac{\partial \mathcal{A}_e}{\partial \mathbf{E}^e},$$

respectively.

In the following, the strain measure will specifically be taken as the Green strain, i.e.,

$$\mathbf{E}^e = \frac{1}{2}(\mathbf{F}_e^T \cdot \mathbf{F}_e - \mathbf{I})$$

and

$$\mathbf{E} = \mathbf{F}_p^T \cdot \mathbf{E}^e \cdot \mathbf{F}_p + \frac{1}{2}(\mathbf{F}_p^T \cdot \mathbf{F}_p - \mathbf{I}),$$

where $\mathbf{I}$ is the second order identity tensor. Hence,

$$(\mathcal{A}_e)_{AB}^{..MN} = \frac{1}{2}\left[(F_p)_{.A}^{M}(F_p)_{.B}^{N} + (F_p)_{.A}^{N}(F_p)_{.B}^{M}\right],$$
$$\mathcal{B}_e = 0$$

and (6.30) reduces to

$$\frac{\rho_0}{\rho_e}\overline{\mathbf{T}}^e = \mathbf{F}_p \cdot \overline{\mathbf{T}} \cdot \mathbf{F}_p^T. \tag{6.31}$$

For simplicity, it may be assumed that the value of density will not be changed due to purely plastic deformation, i.e., $\rho_e = \rho_0$. Then from (6.31), we have

$$\dot{\overline{\mathbf{T}}} = \mathbf{F}_p^{-1}\left[\dot{\overline{\mathbf{T}}}^{\,e} - \mathbf{L}_p \cdot \overline{\mathbf{T}}^e - \overline{\mathbf{T}}^e \cdot \mathbf{L}_p^T\right] \cdot \mathbf{F}_p^{-T}, \tag{6.32}$$

where $\mathbf{L}_p = \dot{\mathbf{F}}_p \cdot \mathbf{F}_p^{-1}$. This means that for fixed $\overline{\mathbf{T}}$,

$$\dot{\overline{\mathbf{T}}}^{\,e} = \mathbf{L}_p \cdot \overline{\mathbf{T}}^e + \overline{\mathbf{T}}^e \cdot \mathbf{L}_p^T. \tag{6.33}$$

If we further assume that from (6.29), the rate of elastic strain may be expressed by

$$\dot{\mathbf{E}}^{\,e} = \frac{\partial \mathbf{E}^e}{\partial \theta}\,\dot{\theta} + \frac{\partial \mathbf{E}^e}{\partial \overline{\mathbf{T}}^e} : \dot{\overline{\mathbf{T}}}^{\,e} + \frac{1}{2}(\dot{\mathbf{C}}_e)^p, \tag{6.34}$$

where

$$\mathbf{C}_e = \mathbf{F}_e^T \cdot \mathbf{F}_e,$$

and

$$(\dot{\mathbf{C}}_e)^p = \left.\frac{\partial \mathbf{C}_e}{\partial \zeta_\delta}\right|_{(\theta, \overline{\mathbf{T}}^e)} \dot{\zeta}_\delta \qquad \text{(summation over } \delta\text{)},$$

then the inelastic part of the strain rate may be written as

$$\dot{\mathbf{E}}_p = \mathbf{F}_p^T \cdot \left[(\mathbf{C}_e \cdot \mathbf{L}_p)_s + \frac{1}{2}(\dot{\mathbf{C}}_e)^p + \frac{\partial \mathbf{C}_e}{\partial \overline{\mathbf{T}}^e} : (\mathbf{L}_p \cdot \overline{\mathbf{T}}^e)_s\right] \cdot \mathbf{F}_p. \tag{6.35}$$

From $\mathbf{D} = \mathbf{F}^{-T} \cdot \dot{\mathbf{E}} \cdot \mathbf{F}^{-1}$, the rate of plastic deformation may be defined by

$$\mathbf{D}^p = (\mathbf{F}_e \cdot \mathbf{L}_p \cdot \mathbf{F}_e^{-1})_s + \mathbf{D}', \tag{6.36}$$

where

$$\mathbf{D}' = \mathbf{F}_e^{-T} \cdot \left[\frac{1}{2}(\dot{\mathbf{C}}_e)^p + \frac{\partial \mathbf{C}_e}{\partial \overline{\mathbf{T}}^e} : (\mathbf{L}_p \cdot \overline{\mathbf{T}}^e)_s\right] \cdot \mathbf{F}_e^{-1}.$$

From (6.36), an expression for the rate of plastic deformation is also suggested by many authors with $\mathbf{D}'$ being zero (cf. [16]).

The loading surface in the temperature-stress space is

$$f_e(\theta, \overline{\mathbf{T}}^e, \zeta_\delta, \mathbf{F}_p) = f(\theta, \overline{\mathbf{T}}, \zeta_\delta, \mathbf{F}_p) = 0.$$

It can be seen that for constant θ, if $\dot{\mathbf{E}}^p$ is directed along the outward normal to $f = 0$ $\left(\text{i.e., } \frac{\partial f}{\partial \overline{\mathbf{T}}}\right)$, then $\mathbf{F}_e^T \cdot \mathbf{D}^p \cdot \mathbf{F}_e$ must be directed along the outward normal to $f_e = 0$ $\left(\text{i.e., } \frac{\partial f_e}{\partial \overline{\mathbf{T}}^e}\right)$.

The rate of plastic work may be defined by

$$W_p = \overline{\mathbf{T}} : \dot{\mathbf{E}}_p = \frac{\rho_0}{\rho_e}\overline{\mathbf{T}}^e : (\mathbf{F}_e^T \cdot \mathbf{D}^p \cdot \mathbf{F}_e) = \frac{\rho_0}{\rho}\boldsymbol{\sigma} : \mathbf{D}^p, \tag{6.37}$$

where ρ is the density in the current configuration and $\boldsymbol{\sigma}$ is the Cauchy stress. In view of

$$\frac{\partial \mathbf{E}^e}{\partial \mathbf{F}_p} : \dot{\mathbf{F}}_p = -(\mathbf{C}_e \cdot \mathbf{L}_p)_s$$

for fixed θ and $\mathbf{E}$, the rate of dissipation of a material element (see Eq. (2.9)) may be written as

$$-\frac{\partial \psi}{\partial \xi_\alpha}\dot{\xi}_\alpha = -\left(\frac{\partial \psi}{\partial \mathbf{F}_p} : \dot{\mathbf{F}}_p + \frac{\partial \psi}{\partial \zeta_\delta}\dot{\zeta}_\delta\right) = \frac{1}{\rho_0}\overline{\mathbf{T}}^e : (\mathbf{C}_e \cdot \mathbf{L}_p)_s - \frac{\partial \psi_e}{\partial \zeta_\delta}\dot{\zeta}_\delta$$

$$= \frac{1}{\rho_0}W_p - \frac{1}{\rho}\boldsymbol{\sigma} : \mathbf{D}' - \frac{\partial \psi_e}{\partial \zeta_\delta}\dot{\zeta}_\delta,$$

which is always non-negative whereas the rate of plastic dissipative work W_p is not necessarily positive in the case of a strong Bauschinger effect.

7. Concluding remarks

In this paper, a rate-independent thermoplastic constitutive theory at finite deformation is discussed. The plastic deformation and the heat conduction is supposed to be uncoupled. So the constitutive equation for the heat flux is not provided. Within the constitutive framework based on internal state variables, one of the important and difficult problems is how to choose these internal variables and how to provide their evolution laws. Nevertheless, if for certain "pairs of material-processes" the thermomechanical postulate proposed here is valid, then the forms of thermoplastic constitutive equations will be greatly simplified. The problem will now reduce to seeking the matrix for the thermoelastic coefficients in the constitutive equations and providing the evolution laws of the loading surface.

Since parameters in the constitutive equations are dependent on the choice of the strain measure and the reference configuration, we have investigated the transformation rules and the invariance properties. It is shown that actually, the choice of the reference configuration is not substantial. Once the constitutive equations for some reference configuration are obtained, the constitutive equations based on current or intermediate (if it exists) configuration may also be determined.

Acknowledgement

The author would like to thank Professor REN WANG for his helpful discussion on this subject.

References

1. M.M. CARROLL, *Foundations of solid mechanics*, Appl. Mech. Rev., **38**, 10, 1301–1308, 1985.
2. ZHUPING HUANG, *On the constitutive framework in finite deformation plasticity*, Acta Mechanica Sinica, **21** (Sup), 31–41, 1989.
3. P. GERMAIN, Q.S. NGUYEN and P. SUQUET, *Continuum thermodynamics*, J. Appl. Mech., **50**, 1010–1020, 1983.
4. J. BATAILLE and J. KESTIN, *Irreversible processes and physical interpretation of rational thermodynamics*, J. Non-Equilib. Thermodyn., **4**, 229–258, 1979.
5. R. HILL, *Aspects of invariance in solid mechanics*, Advances in Appl. Mech., **18**, 1–75, 1978.
6. M. KLEIBER and B. RANIECKI, *Elastic-plastic materials at finite strains*, [in:] Plasticity Today – Modelling, Methods and Applications, A. SAWCZUK and G. BIANCHI [Eds.], pp.3–46, Elsevier Appl. Sci. Pub., 1983.
7. ZHUPING HUANG and ZHUPING DUAN, *On the decomposition of strain and strain rate in finite elastoplastic theory*, Advances in Mechanics, **20**, 1, 24–39, 1990.
8. E.H. LEE, *Elastic-plastic deformation at finite strains*, [in:] Topics in Plasticity (Anniversary Volume in Honor of Professor E.H. Lee), WEI H. YANG [Ed.], pp. 15–30, AM Press, 1991.
9. A.E. GREEN and P.M. NAGHDI, *A general theory of an elastic-plastic continuum*, Arch. Rational Mech. Anal., **18**, 251–281, 1965.
10. J. CASEY and P.M. NAGHDI, *Further constitutive results in finite plasticity*, Q. J. Mech. Appl. Math., **37**, 2, 231–259, 1984.
11. P.M. NAGHDI, *A critical review of the state of finite plasticity*, J. Appl. Math. Phys. (ZAMP), **41**, 315–394, 1990.

12. J. MANDEL, *Plasticité classique et viscoplasticité*, Courses and Lectures, No.97, International Center for Mechanical Sciences, Udine, Springer, New York 1971.
13. J. MANDEL, *Sur la définition de la vitesse de déformation élastique et sa relation avec la vitesse de contrainte*, Int. J. Solids Struct., **17**, 873–878, 1981.
14. J. KRATOCHVIL, *On a finite strain theory of elastic-inelastic materials*, Acta Mechanica, **16**, 127–142, 1973.
15. Y.F. DAFALIAS, *A missing link in the macroscopic constitutive formulation of large plastic deformations*, [in:] Plasticity Today – Modelling, Methods and Applications, A. SAWCZUK and G. BIANCHI [Eds.], pp. 135–151, Elsevier Appl. Sci. Pub., 1983.
16. Y.F. DAFALIAS, *Issues on the constitutive formulation at large elastoplastic deformations*, Acta Mechanica, **69**, 119–138, 1987; **73**, 121–146, 1988.
17. B. LORET, *On the effects of plastic rotation in the finite deformation of anisotropic elastoplastic materials*, Mechanics of Materials, **2**, 287–304, 1983.
18. R. HILL and J.R. RICE, *Elastic potentials and the structure of inelastic constitutive laws*, SIAM J. Appl. Math., **25**, 448–461, 1973.
19. P.M. NAGHDI and J.A. TRAPP, *Restrictions on constitutive equations of finitely deformed elastic-plastic materials*, Q. J. Mech. Appl. Math., **28**, 25–46, 1975.
20. P.M. NAGHDI and J.A. TRAPP, *The significance of formulating plasticity theory with reference to loading surfaces in strain space*, Int. J. Engng. Sci., **13**, 785–797, 1975.
21. M.M. CARROLL, *A rate-independent constitutive theory of finite inelastic deformation*, J. Appl. Mech., **54**, 15–21, 1987.
22. ZHUPING HUANG, *A thermomechanical postulate in finite plasticity*, Acta Sci. Nat. Univ. Pekinensis, **27**, 3, 317–322, 1991.
23. R. HILL, *On intrinsic eigenstates in plasticity with generalized variables*, Math. Proc. Camb. Phil. Soc., **93**, 177–189, 1983.

DEPARTMENT OF MECHANICS,
PEKING UNIVERSITY, BEIJING, CHINA.

Received July 8, 1994.

Arch. Mech., 46, 6, pp. 881–891, Warszawa 1994

Flow in a region with large density variation

Z. PERADZYŃSKI and E. ZAWISTOWSKA (WARSZAWA)

IN THIS PAPER we investigate the influence of viscosity (Reynolds number) on the flow through a "hot region" where the density changes abruptly. Our aim is to model the flow in the vicinity of the front of such a region by searching for approximate and simple formulae. The case of gradually changing density is also discussed. This kind of modelling can have various applications, e.g. in the description of flame propagation in the case of curved flame fronts.

1. Introduction

WHEN UNIFORM flow approaches the region of decreasing density, e.g. because of higher temperature, its behaviour resembles the flow past an obstacle. The pressure increases when approaching the hot region, velocity decreases and most of the streamlines avoid the region of low density; only a fraction of them enter the low density region. On the other hand, when entering the low density region the velocity of the gas increases due to a drop in the gas density, according to the continuity equation. The force which accelerates the fluid comes from the pressure gradient which is created due to the Bernoulli effect.

At first, we concentrate our attention on the flow in a region with two different values of fluid density, e.g. a "hot" ball ($\rho = \rho_*$ for $r < R$ and $\rho = \rho_0$ for $r > R$). In the case of flow through the "hot" ball the nonviscous solution was found analytically in [1] and can be compared with viscous solutions. In the spherical system of coordinates it is given by

$$
\begin{aligned}
v_r &= \begin{cases} v_0 \left(1 - 2A\dfrac{R^3}{r^3}\right)\cos\theta, & r > R, \\ v_0 \left(B + C\dfrac{r^2}{R^2}\right)\cos\theta, & r \leq R, \end{cases} \\
v_\theta &= \begin{cases} -v_0 \left(1 + A\dfrac{R^3}{r^3}\right)\sin\theta, & r > R, \\ -v_0 \left(B + 2C\dfrac{r^2}{R^2}\right)\sin\theta, & r \leq R. \end{cases}
\end{aligned} \tag{1.1}
$$

The constants A, B, C are determined from the jump condition for the momentum density $[\rho \mathbf{v}] = 0$ and momentum flux tensor $[\rho v_r \mathbf{v}] = 0$. They lead to

$$
A = \frac{2 - \varepsilon - \varepsilon B}{4 + \varepsilon}, \qquad C = \frac{3 - (2 + \varepsilon)B}{4 + \varepsilon},
$$

$$3A\frac{2-A}{2} + (1-A)^2 = -C\varepsilon\frac{3B+2C}{2} + \varepsilon(B+C)^2,$$

where $\varepsilon = \rho_*/\rho_0$ (note in [1] an error in the sign in the last equation). In [1] the authors derived also the asymptotic formulae for $\varepsilon \to 0$

$$A \approx \frac{1}{2} - \frac{1}{4}\sqrt{\frac{3\varepsilon}{2}}, \qquad B \approx \sqrt{\frac{3}{2\varepsilon}}, \qquad C \approx \frac{3}{4} - \sqrt{\frac{3}{8\varepsilon}}.$$

For example, the maximal velocity is $v_{\max} \approx \sqrt{\frac{3}{2\varepsilon}}$ and the axial velocity ($\theta = \pi$ and $r = R - 0$) of the fluid after entering the sphere is

$$v_B = (B+C)v_0 \approx \left(\frac{1}{2}\sqrt{\frac{3}{2\varepsilon}} + \frac{3}{4}\right) v_0.$$

This suggests that the velocity of the fluid penetrating a low density plume is proportional rather to $\sqrt{\rho_0/\rho}$, where ρ is the local fluid density, and not to ρ_0/ρ which would take place in a purely one-dimensional flow. It also appears that similar results can be obtained for shapes other than the sphere. In [2] such results were obtained for tube-like shapes of the hot region on the basis of what we call hydraulic approximation, which assumes the flow to be unidirectional and pressure to be constant on each perpendicular cross-section. In Sec. 2 we derive analytical approximate formula which gives the velocity at the front for streamline axisymmetrical shapes. Then in Sec. 3 we discuss the numerical evidence supporting the inverse proportional dependence of the flow velocity on the square root of the density, also for a continuous distribution of density.

2. Flow modelling

For the sake of simplicity, let us assume that the "hot region" has a streamline shape and is axisymmetrical. Let us also assume that the fluid is inviscid and its density changes abruptly from ρ_0 outside to ρ_* inside. We confine our attention to the vicinity of the symmetry axis which is also a streamline. If v_0 is the fluid velocity at infinity, and v_1 – the velocity just before the fluid enters the low density region, then from the Bernoulli equation (which still holds in the region of $\rho = \rho_0$) we have

$$p_1 = p_0 + \frac{1}{2}\rho_0\left(v_0^2 - v_1^2\right). \tag{2.1}$$

If p_* is the pressure inside, then $p_1 - p_*$ is the pressure drop which accelerates the fluid passing from the high density to the low density region. From the continuity equation at the boundary of the region we have

$$\rho_0 v_1 = \rho_* v_*, \tag{2.2}$$

where ρ_*, v_* are the parameters inside the low density region. The velocity v_1 is to be determined as a function of ρ_0, ρ_*, v_0.

From the momentum equation applied at the discontinuity we have the following jump condition

$$[\rho \mathbf{v} \otimes \mathbf{v} + pI] = 0, \tag{2.3}$$

which leads to

$$\rho_0 v_1^2 - \rho_* v_*^2 + p_1 - p_* = 0. \tag{2.4}$$

The velocity v_1 can be expressed by v_* from (2.2) and, similarly, p_1 from (2.1). The only quantity which remains undefined is $\frac{1}{\rho_0}(p_0 - p_*)$. Obviously $p_1 - p_0 = 0$ and $v_* = v_0$ if $\rho_* = \rho_0$. Therefore we can postulate

$$\frac{1}{\rho_0}(p_0 - p_*) = \alpha(v_0^2 - \varepsilon v_*^2) + \beta(v_0^2 - \varepsilon^2 v_*^2) + \cdots \tag{2.5}$$

where $\varepsilon = \rho_*/\rho_0$ and α, β are constants. Then we arrive at

$$v_*^2 = \frac{v_0^2}{\varepsilon} \frac{\alpha + \beta + 0.5}{1 + \alpha + (\beta - 0.5)\varepsilon}. \tag{2.6}$$

In the limit of $\varepsilon \to 0$ we should obtain the flow of inviscid and incompressible fluid around the "body" which has the shape of the "hot" region with $\rho_* \to 0$ (i.e. defined by $\rho(x) = \rho_* = 0$). Since we have assumed the streamline shape of the hot region, its front part can be treated in the first approximation as the part of the ball. Taking

$$\lim_{\varepsilon \to 0}(p_0 - p_*) = \frac{5}{8}\rho_0 v_0^2, \tag{2.7}$$

i.e. equal to the cavitation pressure for the ball [3], one obtains from (2.5) and (2.6)

$$\beta = \frac{1}{8}(5 + \alpha).$$

Applying it to (2.6), one finally arrives at a formula which surprisingly does not contain α

$$v_*^2 = \frac{9}{8}\frac{v_0^2}{\varepsilon}\frac{1}{1 + \frac{\varepsilon}{8}}. \tag{2.8}$$

Equation (2.8) gives for velocity in the front part of the low density region values slightly lower than the maximal velocity of the "hot" ball solution. For $\varepsilon \to 1$, v_* approaches v_0, whereas for $\varepsilon \to 0$, v_* behaves as $\sim v_0/\sqrt{\varepsilon}$ which is in agreement with the conclusion of [1].

3. Numerical results

Numerical computations were performed with the use of the numerical scheme described in [4] which is based on the integration of the conservative form of the Navier-Stokes equations along small contours determined by a grid. Since, in the discretization, central differences were used, the scheme was unstable for higher Reynolds numbers but was quite insensitive to large density gradients, even to jumps. The results for various density ratios (ρ_*/ρ from $1/4$ to $1/40$) and various Reynolds numbers are presented in Figs. 1–6. The Reynolds number is defined by $\mathrm{Re} = v_0 \frac{2R}{\eta}$. In Fig. 1 examples of axial velocity distribution on the symmetry axis of the "hot" ball as the function of nondimensional distance $z/2R$ are presented.

The flow at infinity is directed along z-axis in the positive direction. The numerical results for $\varepsilon = 0.1$ exhibit some overshoots of axial velocity at the front part of the ball as compared to nonviscous solution. This can be caused by the difference in the boundary conditions. We assume $\mathbf{v} = \mathbf{v}_0$ at a distance which is a little shorter than $2\times$ sphere radius from the center of the sphere, whereas in the analytic solution, $\mathbf{v} = \mathbf{v}_0$ for $z \to \infty$. An example of streamline pattern for $\varepsilon = 0.1$ and $\mathrm{Re} = 14$ is shown in Fig. 2.

The axial velocity as a function of the nondimensional distance $r/2R$ from the symmetry axis in the cross-section passing through the center of the sphere is shown in Fig. 3 ($\varepsilon = 0.1, \mathrm{Re} = 34$), and is compared with nonviscous solution ($\varepsilon = 0.1, \mathrm{Re} = \infty$). This distribution is flatter when compared with the inviscid solution. Additionally, velocity at the boundary of the ball is higher, $\approx 2.5 v_0$, whereas the inviscid solution gives values lower than $1.5 v_0$ (= limiting value for $\varepsilon \to 0$).

In practical applications the density is varying continuously. In Fig. 4 the axial velocity distributions at the symmetry axis are presented for various Reynolds numbers and a Gaussian-like density distribution

$$\rho = \rho_0 \left[1 - (1 - \varepsilon) \exp\left(- \frac{x^2 + y^2 + (z - z_0)^2}{R^2} \right) \right]. \tag{3.1}$$

In Fig. 5 the maximum of axial velocity at the symmetry axis is presented as a function of Reynolds number for different ε. The dashed lines are the asymptotic values, $\mathrm{Re} \to \infty$, calculated from the exact analytic solution which was discussed in the Introduction. The presented numerical values would be a little lower (a few percent) if the boundary condition were assumed far enough.

In Fig. 6 an example of radial a) and axial b) velocity components are presented for $\varepsilon = 0.09$, $\mathrm{Re} = 42$ and $R = 0.26$. The radial velocity is about ten times smaller than the axial velocity. Figure 6c presenting the ratio of axial velocity from the example in Fig. 6b divided by v_*, where v_* is computed from (2.8) at

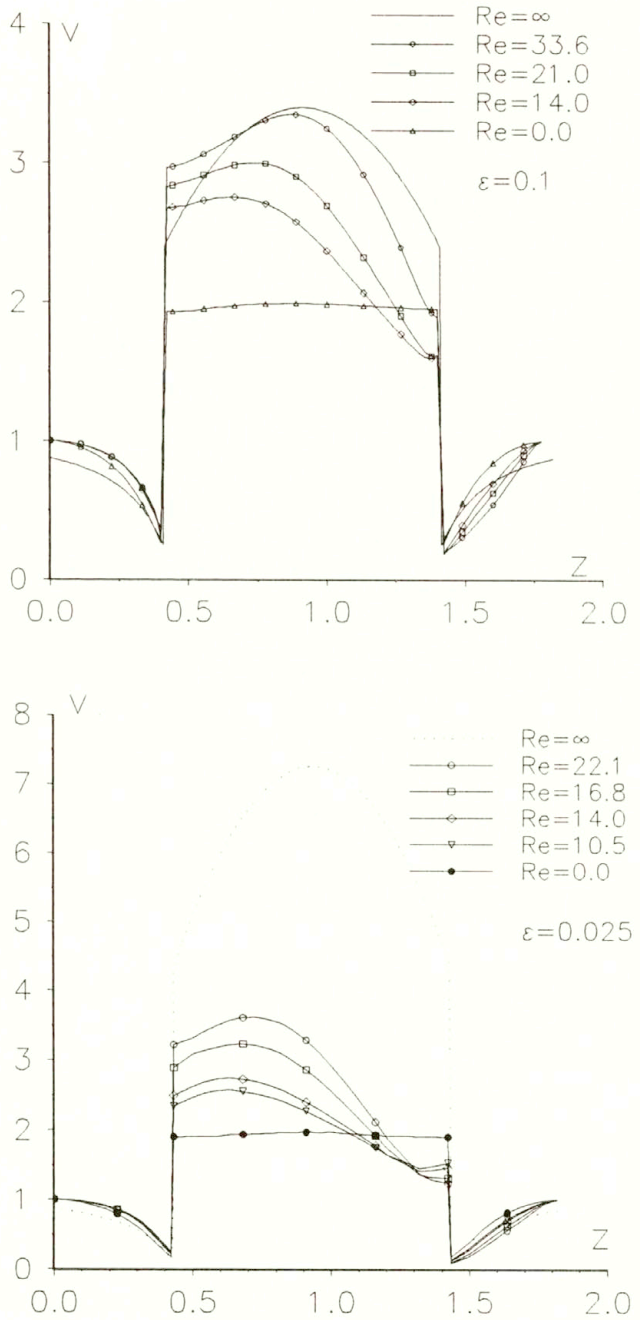


FIG. 1. Axial flow velocity distributions v/v_0 on the symmetry axis as a function of normalized distance $z/2R$ ("hot" ball) for various $\varepsilon = \rho_*/\rho_0$: a) $\varepsilon = 0.1$, b) $\varepsilon = 0.025$, and different Reynolds numbers.

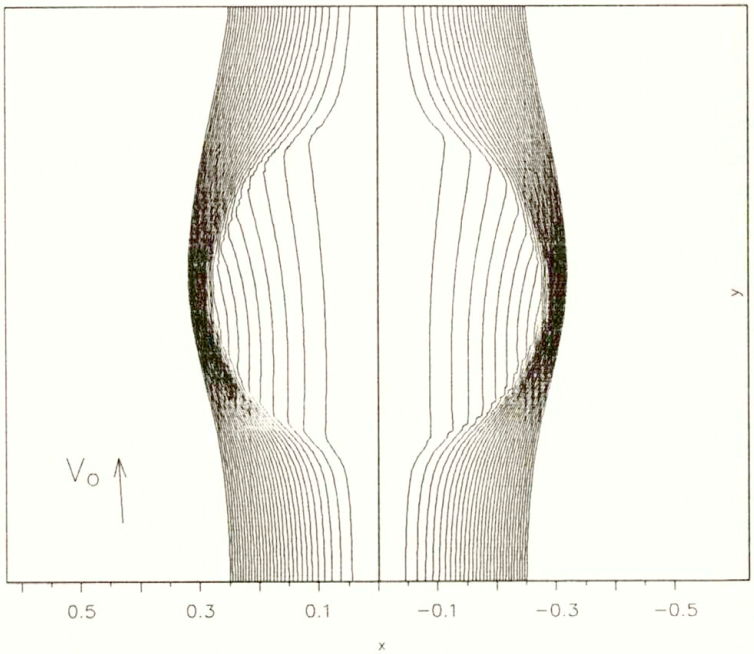


FIG. 2. Streamlines, $\varepsilon = 0.1$, Re = 14.

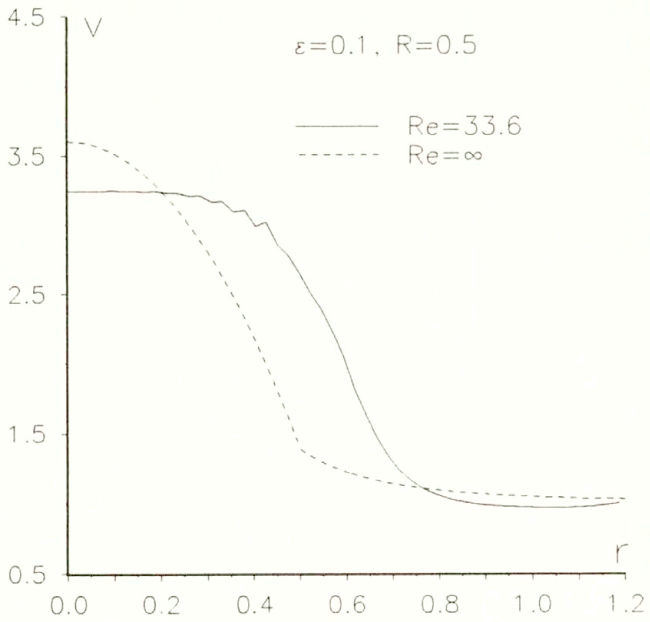


FIG. 3. Radial distribution of the axial velocity component in the cross-section of the sphere perpendicular to the symmetry axis.

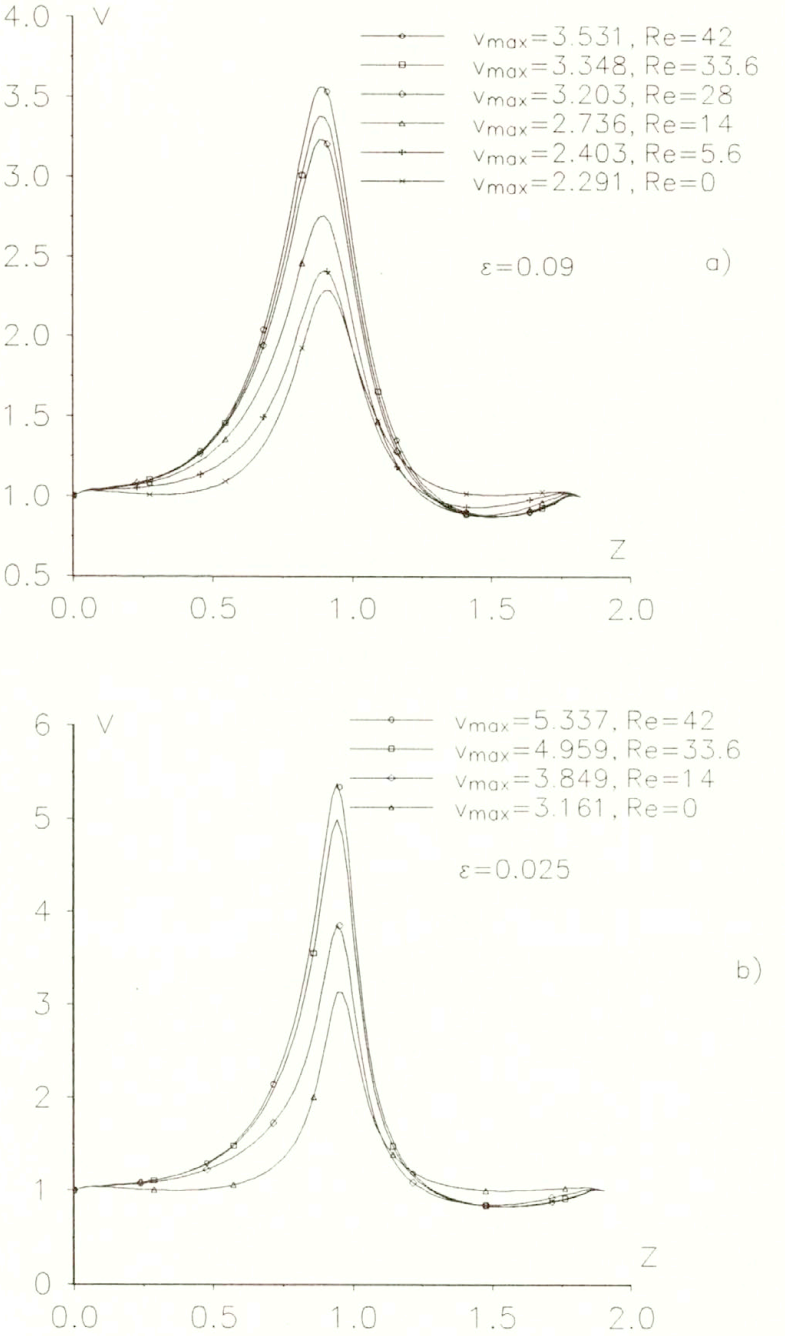


FIG. 4. Axial velocity distributions at $r = 0$, as a function of $z/2R$ for various Reynolds numbers and for density as in (3.1): a) $\varepsilon = 0.09$, b) $\varepsilon = 0.025$.

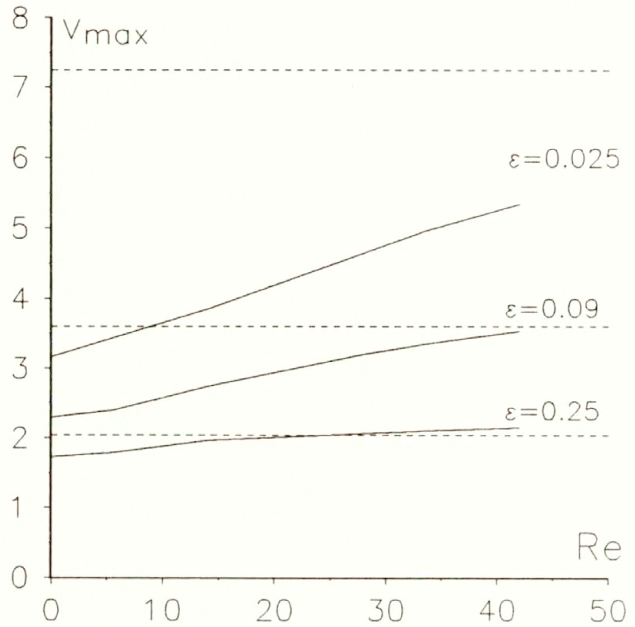
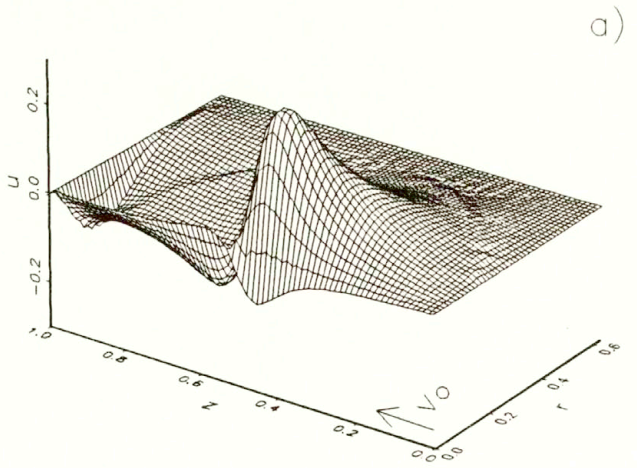


FIG. 5. Maximum of axial velocity at $r = 0$ as a function of Reynolds number and different ε (density as in (3.1)). Dashed lines are the maximal velocities of the analytic solution for "hot" ball, $\mathrm{Re} \to \infty$.



[FIG. 6a]

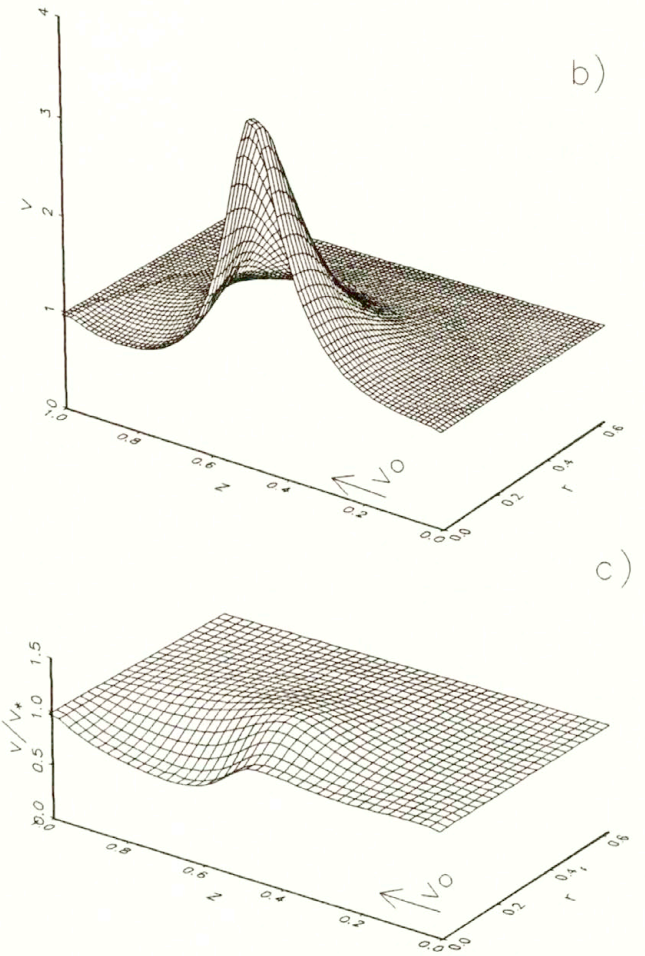


FIG. 6. Distribution of radial a) and axial b) velocity components in the case of continuous density distribution given by (3.1); c) the ratio of axial velocity from the example in b) divided by v_* computed from (2.8) at every point of the flow according to the local density; at the front part of the flow this ratio is almost exactly equal to 1, with $R = 0.26$ for $\varepsilon = 0.09$, Re = 42.

every point of the flow, according to the local value of density, confirms quite well the applicability of (2.8) to a large class of flows with variable fluid density.

4. Conclusion

For finite Reynolds numbers the following approximate formula fits well the numerical results

$$v(\mathrm{Re},\varepsilon) = \left(v(0,\varepsilon)\frac{\mathfrak{L}}{\varepsilon\mathrm{Re}} + v(\infty,\varepsilon)\right)\frac{\varepsilon\mathrm{Re}}{\varepsilon\mathrm{Re}+\mathfrak{L}}, \tag{4.1}$$

where the constant $\mathfrak{L}$ is approximately equal to 0.9, $v(0,\varepsilon)$ is the Stokes solution

and $v(\infty, \varepsilon)$ is the maximal velocity for the inviscid "hot" ball discussed in the Introduction. Also formula (2.8) can be used for $v(\infty, \varepsilon)$.

For very small Reynolds numbers (Re $\to$ 0) the numerical results can be approximated by the formula

$$v = \frac{\delta}{1 + (\delta - 1)\varepsilon} v_0, \tag{4.2}$$

where $\delta = 2$ for "hot" ball. Taking $\delta \approx 3$ one obtains rough estimates for "Gaussian-like" density distribution (3.1).

For large Reynolds numbers deviation from symmetry between the front and the rear of the flow are small; they become larger when Re gets smaller but then, gradually, one arrives at the symmetric solution of the Stokes flow, Re = 0. This appears to be true also for the continuous symmetric density distributions such as (3.1). As can be seen in Fig. 6c, for these intermediate values of Reynolds numbers formula (2.8) works better for the front part of the flow.

If we assume now, that (2.8) expresses the axial velocity component v_z, then the full velocity can be determined from the continuity equation (even in the time-dependent case)

$$\frac{1}{r}\frac{\partial}{\partial r}(\rho\, r\, v_r) + \frac{\partial}{\partial z}(\rho\, v_z) + \frac{\partial \rho}{\partial t} = 0$$

provided $\rho(x, t)$ is given. In the case of plane flows appropriate form of the continuity equation must be used.

REMARK. The approach similar to the one described in Sec. 2 can be used in the two-dimensional case. However, this time the results depends on α. The comparison with numerical results suggests $\alpha = 0$ which leads to

$$v_* = \sqrt{\frac{2}{\varepsilon}\frac{1}{1+\varepsilon}}\,.$$

This approach was used to determine the parameters of the laser sustained plasma in [5].

Acknowledgments

The first author was supported by KBN grant 336249203.

References

1. K.G. GUSKOV, YU.P. RAIZER and S.T. STURZHIKOV, *On an observed speed of the optical discharge slow propagation*, Kvantovaya Elektronika, **17**, 7, 937–942, 1990.
2. Z. PERADZYŃSKI, M. DUDECK, J. KURZYNA and Z. SZYMAŃSKI, *Supersonic expansion of the laser sustained plasma*, J. Tech. Phys., 3, 1994.

3. L.M. MILNE-THOMSON, *Theoretical hydrodynamics*, Macmillan and Co. London 1960.
4. J. SŁOMCZYŃSKA, Z. PERADZYŃSKI, *Finite-difference scheme for stationary Navier–Stokes equations with variable coefficients*, R.L. STERNBERG, A.J. KALINOWSKI, J.S. PAPADAKIS [Eds.], [in:] Nonlinear Partial Differential Equations in Engineering and Applied Science, Lecture Notes in Pure and Applied Mathematics, vol. 54, M. Dekker, New York 1980, 343–371.
5. Z. SZYMAŃSKI, Z. PERADZYŃSKI and J. KURZYNA, *Free burning laser-sustained plasma in a forced flow*, J. Phys. D, Applied Physics, (forthcoming).

POLISH ACADEMY OF SCIENCES
INSTITUTE OF FUNDAMENTAL TECHNOLOGICAL RESEARCH

Received August 8, 1994.

Arch. Mech., 46, 6, pp. 893–936, Warszawa 1994

Adiabatic material instabilities in rate-dependent solids

I. DOBOVŠEK (MARIBOR)

A CONTINUUM mechanics framework for the analysis of material instabilities in rate and temperature-dependent solids, based upon a linear stability analysis of the partial differential equations of elastic-thermoviscoplastic flow under the adiabatic conditions, is presented. The derivation is general, three-dimensional, with the effects of rate sensitivity, inertia, temperature and elasticity included. Critical conditions for the formation of two qualitatively different types of material instability are discussed: shear band localization and Hopf bifurcation. The characteristic stability equation is given in explicit form for a general three-dimensional state of stres. It is shown, that the elastic part of the constitutive law based on an additive decomposition of the rate of deformation tensor induces a relative rotation of the instability plane, i.e., the plane with the fastest incipient rate of instability growth with respect to the principal directions of the stress deviator. Comparison with the two limiting cases of material behaviour under the adiabatic conditions, namely the rigidly viscoplastic, and rate-independent ones in particular, is given as well.

1. Introduction

WITHIN THE PAST two decades or so, much attention and considerable efforts have been devoted to the analysis of material instabilities. Examples include Lüders bands, diffuse necking of tensile specimens and shear bands. Diffuse necking may terminate in fracture, but it is often followed by a second instablity process called localized necking or shear banding. In this process, an essentially smooth and continuously varying deformation gives rise to a highly localized deformation pattern which may eventually lead to ductile fracture. A large class of technology processes can be considered as adiabatic when the loading process is dynamic and fast enough so that there is no sufficient amount of time for the heat produced during the viscoplastic irreversible deformation to be conducted away. In such a case, thermal effects may have a significant influence on formation of material instability.

In this work we present a method for predicting the onset of material instability in rate and temperature-dependent solids. The effects of strain hardening or softening, strain rate sensitivity, thermal softening, inertia, and elasticity on the possible mechanisms of material instability are examined in detail. Full three-dimensional analysis is performed within the framework of viscoplasticity theory which in the rate-independent limit corresponds to the standard J_2 flow theory of plasticity with isotropic hardening. The mechanical behaviour of the material is characterized by a general rate-type form of constitutive relation which accounts for the effects of thermal softening, strain rate sensitivity, strain softening, and elasticity.

The theoretical framework for analysing the localization of plastic deformation in rate-independent solids, including bifurcation theory and the analysis of the growth of initial imperfections, is given for example in RICE [17], RUDNICKI and RICE [19], ASARO and RICE [2], and YAMAMOTO [25]. The bifurcation approach is based on the assumption that the phenomenon of shear bands can be explained by seeking the conditions under which the bifurcation from a homogeneous mode of deformation into a localized band mode can occur.

The method of imperfections is the most frequently used method in numerical studies of the same problem, like for example in PAN *et al.* [11]. The origins of the method are due to the widely accepted notion that shear bands nucleate because of the local inhomogeneity or defect causing enhanced local deformation and local heating. Once non-uniform flow begins and straining continues, the deformation becomes unstable. By introducing various geometric and material imperfections, one hopes to induce a representative localization pattern embodied in the form of high deformation gradients in the vicinity of imperfection, which spreads through the specimen in a direction which corresponds approximately to the direction of instantaneous plane of maximum shear. At each stage of the deformation history time integration is performed for the whole specimen, giving the time evolution and spatial resolution of the strain fields which become highly concentrated in the close vicinity of the imperfection. Hence, the evolution of large deformation gradients and other fields that are related to it is highly influenced by the choice of the initial imperfection, and so is the efficiency of the method itself. This makes the method predominantly descriptive. Nevertheless, the method is important in enhancing our understanding of various growth phenomena, and valuable information can be obtained on the evolution and growth of bands triggered by different initial imperfections. Unfortunately, analyses of this type do not provide an adequate answer about the critical orientations of shear bands, which is of great importance in designing effective finite element meshes for localization problems.

To avoid these problems, we study various aspects of material instability by employing a standard linear stability analysis where the formulation incorporates the effects of thermal softening, strain rate sensitivity, strain softening, elasticity, inertia, their mutual interactions and individual influence on the possible mechanisms of material instability. The method has been employed in solid mechanics predominantly in the analysis of one-dimensional problem, BAI [3], FRESSENGEAS and MOLINARI [5], SHAWKI and CLIFTON [21], with a few attempts to generalize it to higer-dimensional problems. A three-dimensional generalization was presented by ANAND *et al.* [1], however they restricted attention to incompressible materials and neglected the effects of elasticity. One of the main objectives of this paper is to present a general method for the linear stability analysis of rate-dependent materials. A plane perturbation wave of an infinitesimal amplitude propagates in an elastic-viscoplastic (rate-dependent) material which is in the state of homogeneous deformation with adiabatically coupled temperature field, in absence of

radiant heat sources. This basically means that in the first law of thermodynamics the term which accounts for heat conduction is neglected. From the perspective of instability, a theoretical model for prediction of two qualitatively different types of material instability is derived, using the basic notions of linear stability analysis in perturbation of the system of equations describing the rate-dependent plastic flow and the first law of thermodynamics. We seek the conditions under which for the material held in a state of pure homogeneous deformation, the superimposed infinitesimal perturbation exhibits an explosive (exponential) growth. Once the infinitesimal disturbance starts to grow, the material may be regarded as unstable. We derive the conditions at which the material system is found on the threshold of instability. We give a detailed derivation of the resulting dispersion relation or stability equation. The problem of material instability is addressed from a broad perspective, where the localization of deformation or shear banding originates as a special case of a general theory. A general form of elastic-thermoviscoplastic constitutive relation is considered, and we emphasize the importance of the elastic portion of the constitutive law and its influence on the determination of the critical orientation of the instability plane, and its evolution in the post-critical regime.

The linear stability analysis can only predict the necessary conditions for the onset of instability. In spite of the limitations of this type of analysis as compared to the full nonlinear analysis, which usually requires the assumption of a special form of the constitutive relation together with the use of numerical techniques, the linear stability analysis predicts the necessary conditions for the onset or formation of instability, and sometimes allows for closed-form analytical results for different classes of constitutive relations. The results of the linear stability analysis are valid only for perturbations whose variations are much faster than those of the homogeneous solution, i.e., the time scale of perturbations is small in comparison to the time scale for the variation of the homogeneous solution. This apparently does not hold for long wavelengths of perturbations where we cannot assume beforehand that the variation of coefficients in the differential equations for perturbations is sufficiently slow, so that the time-dependence of the coefficients can be neglected from the start. Nevertheless, the results for the initial growth of disturbances are quite satisfactory, and comply with the results obtained for the idealized but nonlinear problem. Because of linearization, such an analysis is not expected to characterize completely the stability of the solutions of the nonlinear problem, since the initial growth may or may not lead to instability. This obviously depends on the neglected effects of the higher-order terms and on the nonlinearity of the system. Consequently, the results obtained from the linear stability analysis cannot in general be used for long range instability predictions. Once the instability sets in, the growth of the perturbation makes it invalid to neglect nonlinear terms. However, predictions for the onset of instability obtained by means of linear stability analysis yield results which are in general agreement with numerical solutions of the nonlinear system, and can be consid-

ered quite accurate at the early stages of instability evolution. The linear stability analysis provides important insight into the critical conditions for the onset of instability. From a mathematical point of view, the linear stability analysis gives sufficient conditions for the stability of the solution of the linear perturbation problem, which can also serve as sufficient conditions for the stability of solutions of the nonlinear problem. Moreover, most of the contemporary methods of nonlinear stability analysis are based on a higher-order asymptotic expansion of the governing field equations near the critical state which is determined by means of linear stability analysis. This is one of the reasons why the critical conditions in terms of linear stability should be investigated in detail. The approach based on the same principles will be used in the forthcoming paper, in an attempt to perform the nonlinear stability analysis of the same system of field equations.

The paper is organized as follows. The governing field equations are given in Sec. 2. In Sec. 3 we discuss the kinematics of the perturbation, linearize the equations at the current (intermediate) configuration and derive the characteristic stability equation. In Sec. 4 we display in detail the derivation of a general dispersion relation and expanded stability equation with the characteristic coefficients collected in Tables 1 and 2. Section 5 contains an augmented version of stability criteria that stem from the algebraic structure of the characteristic equation. The analysis of the stationary points of the dispersion relation is followed by a discussion of various aspects and possibilities for the onset of bifurcation. The limiting transformation to recover the constitutive behaviour of a rigid viscoplastic material with the corresponding characteristic coefficients of the dispersion relation is given in Sec. 8. In Sec. 9 we discuss the transition to the rate-independent limit and derive the localization condition for the adiabatic rate-independent plastic deformation.

2. Field equations

Let $\mathbf{x}$ define the current configuration taken as reference. Furthermore, let $\mathbf{v}(\mathbf{x}, t)$ denote the velocity, $\mathbf{L} = \operatorname{grad} \mathbf{v}$ the velocity gradient, $\mathbf{D} = \operatorname{sym} \mathbf{L}$, and $\mathbf{W} = \operatorname{skew} \mathbf{L}$, the rate of deformation and spin tensor, respectively, and $\boldsymbol{\tau}$ the Cauchy (true) stress. Consider a homogeneous solid sustaining stress $\boldsymbol{\tau}$ and temperature field θ. The local balance of linear momentum and energy at $\mathbf{x}$ in the absence of body forces and radiant heat sources, assuming that the irreversible part of deformation is fast enough so that there is no spatial heat exchange with the surroundings, are given as

$$\operatorname{div} \boldsymbol{\tau} = \rho(\mathbf{x}, t) = \dot{\mathbf{v}}(\mathbf{x}, t), \tag{2.1}$$

$$\rho \dot{e} = \boldsymbol{\tau} : \mathbf{L}, \tag{2.2}$$

where $\dot{e}$ is the material time derivative of specific internal energy per unit mass. In what follows, some simplifying assumptions will be used. We assume that the

rate of change of specific internal energy can be taken proportional to the rate of change in temperature field with c_p being the constant of proportionality. Here c_p represents specific heat leading to $c_\theta = \rho c_p$ when measured per unit volume. The elastic part of constitutive relation is decoupled from the temperature field by making use of a scalar parameter χ specifying the fraction of plastic work that is converted to heat. For a comprehensive exposition of thermodynamic theory pertaining to the class of the rate-dependent and rate-independent solids based on quite general grounds, see PERZYNA [13, 14]. For a detailed discussion of the starting assumptions leading to formally the same form of energy balance equation, see DUSZEK and PERZYNA [4]. For metals χ is usually in the range of [0.85 – 0.95]. Consequently, the energy balance equation which now defines the relation between heating and plastic dissipation takes the following form

$$\rho c_p \dot{\theta} = \chi \boldsymbol{\tau} : \mathbf{D}_{vp}. \tag{2.3}$$

The constitutive equations of elastic-viscoplastic solid based on the additive decomposition of the rate of deformation tensor are given by

$$\mathbf{D} = \mathbf{D}_e + \mathbf{D}_{vp}, \tag{2.4}$$

where the elastic and vicoplastic part of the rate of deformation tensor are defined as

$$\mathbf{D}_e = \mathbf{C}^{-1} : \overset{\nabla}{\boldsymbol{\tau}}, \qquad \mathbf{D}_{vp} = \dot{\bar{\varepsilon}}\, \mathbf{p}. \tag{2.5}$$

$\overset{\nabla}{\boldsymbol{\tau}}$ is the Jaumann corotational stress rate, $\mathbf{C}$ a tensor of elastic moduli, $\mathbf{C}^{-1}$ its inverse, $\mathbf{p}$ a plastic flow direction and $\dot{\bar{\varepsilon}}$ an equivalent or effective viscoplastic strain rate,

$$\overset{\nabla}{\boldsymbol{\tau}} = \dot{\boldsymbol{\tau}} + \boldsymbol{\tau} \cdot \mathbf{W} - \mathbf{W} \cdot \boldsymbol{\tau}. \tag{2.6}$$

The elastic response is assumed to be linear isotropic with constant isothermal moduli $\mathbf{C}$, which is appropriate for metal plasticity for stress levels that are small compared to the elastic moduli, and where the elastic strains remain small compared with the plastic strains. Combining Eq. (2.4), (2.5) and solving for $\overset{\nabla}{\boldsymbol{\tau}}$ gives

$$\overset{\nabla}{\boldsymbol{\tau}} = \mathbf{C} : \mathbf{D} - \dot{\bar{\varepsilon}}\, \mathbf{C} : \mathbf{p}. \tag{2.7}$$

The effective viscoplastic strain rate is determined from an experimental curve, and can be written symbolically as a function of effective stress, strain, and temperature

$$\dot{\bar{\varepsilon}} = \dot{\bar{\varepsilon}}\, (\bar{\sigma}, \bar{\varepsilon}, \theta), \tag{2.8}$$

$$\bar{\sigma}^2 = \frac{3}{2}\boldsymbol{\tau}' : \boldsymbol{\tau}', \qquad \boldsymbol{\tau}' = \boldsymbol{\tau} - \frac{1}{3}\mathrm{tr}\,\boldsymbol{\tau}\mathbf{1}. \tag{2.9}$$

The plastic flow direction is derived from a plastic flow potential Q

$$\mathbf{p} = \partial Q/\partial \boldsymbol{\tau}. \tag{2.10}$$

Derivations will be performed for a general class of viscoplastic materials obeying the functional relationship as defined in (2.8). Nevertheless, at certain points during the analysis, we will specialize to the particular form of (2.8) most frequently used in practical analyses, LEMONDS and NEEDLEMAN [8], PIERCE *et al.* [12]:

$$\dot{\bar{\varepsilon}} = \dot{\varepsilon}_0 \left[\frac{\bar{\sigma}}{h(\theta - \theta_0)g(\bar{\varepsilon})}\right]^{-1/m}. \tag{2.11}$$

Here $\dot{\varepsilon}_0$ represents the reference strain rate at which the tensile test was performed, m is the strain rate exponent, $g(\bar{\varepsilon})$ is the strain hardening function, θ_0 is the reference temperature, and the function h specifies thermal softening characteristics of the material. The function $g(\bar{\varepsilon})$ together with

$$\bar{\varepsilon} = \int \dot{\bar{\varepsilon}}\, dt = \int \left(\frac{\mathbf{D}_{vp} : \mathbf{D}_{vp}}{\mathbf{p} : \mathbf{p}}\right)^{1/2} dt \tag{2.12}$$

represents the effective stress versus the effective strain. It may take the form of a simple power law, or it may also incorporate some softening features. Usually, the strain softening is incorporated into the field equations by a proper choice of $g(\bar{\varepsilon})$.

3. Perturbation of field equations

By instability we mean instability with respect to small or linear perturbations that can take the shape of a wave. A perturbation of velocity and temperature field with respect to the current configuration $\mathbf{x}$ as reference is defined as a relative motion of a body $\boldsymbol{\chi} = \boldsymbol{\chi}(\mathbf{x}, t)$, and is characterized by functions $\Delta\mathbf{v}(\mathbf{x}, t)$ and $\Delta\theta(\mathbf{x}, t)$ such that

$$\begin{aligned} \bar{\mathbf{v}}(\mathbf{x}, t) &= \mathbf{v}(\mathbf{x}, t) + \Delta\mathbf{v}(\mathbf{x}, t), \\ \bar{\theta}(\mathbf{x}, t) &= \theta(\mathbf{x}, t) + \Delta\theta(\mathbf{x}, t), \end{aligned} \tag{3.1}$$

$$\begin{aligned} \Delta\mathbf{v}(\mathbf{x}, t) &= \phi\hat{\mathbf{V}} = e^{\omega t + i\mathbf{k}\cdot\mathbf{x}}\hat{\mathbf{V}} = e^{\omega_R t + i\varphi}\hat{\mathbf{V}}, \\ \Delta\theta(\mathbf{x}, t) &= \phi\hat{\Theta} = e^{\omega t + i\mathbf{k}\cdot\mathbf{x}}\hat{\Theta} = e^{\omega_R t + i\varphi}\hat{\Theta}, \end{aligned} \tag{3.2}$$

where $\hat{\mathbf{V}}$ is a constant amplitude with sufficiently small apropriate norm $\|\hat{\mathbf{V}}\| \ll 1$, $\hat{\Theta}$ is the amplitude in the peturbation of temperature, $\varphi = k(ct + \mathbf{n} \cdot \mathbf{x})$ is the phase, $\omega_R = \mathrm{Re}[\omega]$ is the growth rate, $\omega_I = \mathrm{Im}[\omega]$ is the wave frequency, $c = \omega_I/k$ is the wave (phase) speed, $k = 2\pi/\lambda$ is the wave number, and $\mathbf{k} = k\mathbf{n}$ the wave vector. The gradient of velocity increment is thus given as

$$\Delta\mathbf{L} = \Delta\mathbf{v} \overleftarrow{\nabla}_{\mathbf{x}} = (\mathrm{grad}\, \Delta\mathbf{v})^T = ik\Delta\mathbf{v} \otimes \mathbf{n} = \mathbf{g} \otimes \mathbf{n}. \tag{3.3}$$

Consequently, if $\mathrm{Re}[\omega] < 0$, then both $\Delta\mathbf{v}$ and $\Delta\theta$ are exponentially decreasing functions with or without oscillatory modulation according to whether $\mathrm{Im}[\omega]$ is non-vanshing or vanishing. If the wave length of disturbance is short enough, so that the coefficients of linearized equations do not vary significantly over one wave-length ($\lambda = 2\pi/k$), it is permissible in the close neighbourhood of the same state that the coefficients of linearized field equations are considered to be constant. Thus, we may seek the solutions of the system in terms of normal modes, cf. (3.2). The shorter the waves, the stronger the theory. The reference solution ($\Delta\mathbf{v} = 0$, $\Delta\theta = 0$) is reached in the limit as $t \to \infty$, which defines asymptotic stability. Hence, the basic solution given in terms of velocity and temperature fields $\mathbf{v}(\mathbf{x}, t)$, $\theta(\mathbf{x}, t)$ of the current configuration is asymptotically stable. It is assumed that the basic flow defined by the velocity $\mathbf{v}(\mathbf{x}, t)$ is non-periodic. If, on the other hand $\mathrm{Re}[\omega] > 0$, disturbances tend to grow exponentially and the reference (trivial) solution is unstable. The regime which separates the two classes of states, stable from unstable, defines the state of marginal stability. By this definition, the marginal state is the state of neutral stability and $\mathrm{Re}[\omega] = 0$ determines the threshold of instability of the reference state. Two qualitatively different states can be distinguished here. In most cases it is postulated that besides $\mathrm{Re}[\omega] = 0$ one also has $\mathrm{Im}[\omega] = 0$. The eigenvalue ω changes its sign by passing through zero. Some autors refer to this type of critical point as a divergence point. Another possibility is the case of Hopf bifurcation where besides $\mathrm{Re}[\omega] = 0$ one also has $\mathrm{Im}[\omega] \neq 0$. Consequently, we get a bifurcation in a time-periodic flow.

The increment of velocity gradient $\Delta\mathbf{L}$, its symmetric counterpart $\Delta\mathbf{D}$ (the incremental rate of deformation tensor) and skew-symmetric part $\Delta\mathbf{W}$ (the incremental spin tensor), can be expressed as

$$\Delta\mathbf{L} = ik\phi\hat{\mathbf{V}} \otimes \mathbf{n} = \mathbf{g} \otimes \mathbf{n},$$

$$\Delta\mathbf{D} = (\mathbf{g} \otimes \mathbf{n} + \mathbf{n} \otimes \mathbf{g})/2, \tag{3.4}$$

$$\Delta\mathbf{W} = (\mathbf{g} \otimes \mathbf{n} - \mathbf{n} \otimes \mathbf{g})/2, \tag{3.5}$$

$$\Delta\boldsymbol{\tau} = \phi\hat{\boldsymbol{\tau}}, \qquad \|\hat{\boldsymbol{\tau}}\| \ll 1, \tag{3.6}$$

$$\Delta\theta = \phi\hat{\Theta}, \qquad |\hat{\Theta}| \ll 1. \tag{3.7}$$

In the last two expressions (3.6), (3.7) it has been asumed that the kinematics of incremental deformation yields stress and temperature perturbations in the form

depicted, where $\hat{\tau}$ stands for the amplitude tensor of the incremental Cauchy stress with the sufficiently small apropriate norm $\| \cdot \|$, and $\hat{\Theta}$ represents the corresponding amplitude of the temperature field.

Now consider the linearization of the field equations at $\mathbf{x}$ as a starting point for the linear stability analysis and the derivation of dispersion relation in Sec. 4 below. The momentum balance at χ with respect to the intermediate configuration $\mathbf{x}$ as reference can be expressed as

$$\operatorname{div}\left[J\mathbf{F}^{-1} \cdot \bar{\tau}(\mathbf{x}, t)\right] = \rho(\mathbf{x}, t)\, \dot{\mathbf{v}}\,(\mathbf{x}, t). \tag{3.8}$$

The incremental quantities needed for further development are now given as

$$\chi(\mathbf{x}, t) = \mathbf{x} + \Delta\mathbf{u}(\mathbf{x}, t), \tag{3.9}$$

$$\Delta\dot{\mathbf{v}}\,(\mathbf{x}, t) = \partial^2 \Delta\mathbf{u}(\mathbf{x}, t)/\partial t^2, \tag{3.10}$$

$$\bar{\tau}(\mathbf{x}, t) = \tau(\mathbf{x}, t) + \Delta\tau(\mathbf{x}, t). \tag{3.11}$$

The relative deformation gradient $\mathbf{F}$, its inverse $\mathbf{F}^{-1}$, and the corresponding Jacobian of incremental deformation are defined as

$$\begin{aligned} \mathbf{F} = \frac{\partial \chi}{\partial \mathbf{x}} = \mathbf{1} + \mathbf{H}, \qquad \mathbf{F}^{-1} = \mathbf{1} - \mathbf{H}, \\ J = \det(\partial\chi/\partial\mathbf{x}) = 1 + \operatorname{tr}\mathbf{H}, \qquad \mathbf{H}^T = \operatorname{grad}\Delta\mathbf{u}(\mathbf{x}, t), \end{aligned} \tag{3.12}$$

where the higher order terms in expressions for J and $\mathbf{F}^{-1}$ have been neglected. Making use of (3.9)–(3.12), the linearization of (3.8) yields

$$\operatorname{div}\Delta\tau + \operatorname{tr}\mathbf{H}\operatorname{div}\tau - (\operatorname{grad}\tau^T)^T \cdot\cdot\, \mathbf{H} = \rho\Delta\dot{\mathbf{v}}\,. \tag{3.13}$$

By assuming that the strained reference configuration at $\mathbf{x}$ is in homogeneous state of stress, i.e. $\tau(\mathbf{x}, t) = \tau(t)$, Eq. (3.13) reduces to

$$\operatorname{div}\Delta\tau = \rho\Delta\dot{\mathbf{v}}\,. \tag{3.14}$$

Substituting (3.2) and (3.6) into (3.14) gives rise to the momentum jump condition

$$\mathbf{n} \cdot \Delta\tau + \frac{\rho\omega}{k^2}\mathbf{g} = \mathbf{0}, \qquad ik\mathbf{n} \cdot \hat{\tau} - \rho\omega\hat{\mathbf{V}} = \mathbf{0}. \tag{3.15}$$

By analogous considerations with those which have led to (3.15), and by taking into account Eq. (2.2), we can express the equation of energy balance at χ relative to $\mathbf{x}$ as

$$\rho\,\dot{\bar{e}} - \bar{\mathbf{P}} \cdot\cdot\, \dot{\bar{\mathbf{F}}} = 0, \tag{3.16}$$

where $\overline{\mathbf{P}}(\mathbf{x}, t)$ stands for the nominal (first Piola–Kirchhoff) stress tensor. Introducing the incremental relations between the first Piola–Kirchhoff and Cauchy stress tensors in the form

$$\overline{\mathbf{P}}(\mathbf{x}, t) = \mathbf{P}(\mathbf{x}, t) + \Delta\mathbf{P}(\mathbf{x}, t) = \boldsymbol{\tau}(\mathbf{x}, t) + \Delta\boldsymbol{\tau}(\mathbf{x}, t) + \operatorname{tr}\mathbf{H}\,\boldsymbol{\tau}(\mathbf{x}, t) - \mathbf{H}\cdot\boldsymbol{\tau}(\mathbf{x}, t), \tag{3.17}$$

by virtue of relations defined in (3.12), Eq. (3.16) in incremental form yields

$$\rho\Delta\dot{e} - \Delta(\boldsymbol{\tau} \cdot\cdot \mathbf{D}) + \operatorname{tr}\mathbf{H}(\boldsymbol{\tau} \cdot\cdot \mathbf{D}) - \operatorname{tr}(\mathbf{H}\cdot\boldsymbol{\tau}\cdot\mathbf{L}) = 0. \tag{3.18}$$

For a homogeneous thermodynamical process at a basic state at $\mathbf{x}$, the fundamental thermodynamic variables exhibit no spatial dependence, thus in addition to the assumptions that led to (3.14), we may take $\theta(\mathbf{x}, t) = \theta(t)$, and finally, in accordance with the assumptions outlined in Sec. 2, we approximate the left-hand side of (3.18) by taking

$$\rho\Delta\dot{e} - \Delta(\boldsymbol{\tau} \cdot\cdot \mathbf{D}) + \operatorname{tr}\mathbf{H}(\boldsymbol{\tau} \cdot\cdot \mathbf{D}) - \operatorname{tr}\mathbf{H}\cdot\boldsymbol{\tau}\cdot\mathbf{L} \approx c_\theta\Delta\dot{\theta} - \chi\Delta(\boldsymbol{\tau} \cdot\cdot \mathbf{D}_{vp}), \tag{3.19}$$

$$c_\theta\Delta\dot{\theta} - \chi\Delta(\boldsymbol{\tau} \cdot\cdot \mathbf{D}_{vp}) = 0, \tag{3.20}$$

and thus we obtain the following form of the incremental energy balance

$$c_\theta\Delta\dot{\theta} = \chi\Delta(\bar{\sigma}\,\dot{\bar{\varepsilon}}). \tag{3.21}$$

Here the function of the parameter χ has been slightly extended. In addition to the decoupling role which has been originally assigned to χ, the parameter now takes care of the geometric changes due to the differences in incremental relation between the nominal and true stress, leading to a significant simplification of the incremental energy balance equation. A rigorous consideration of geometric changes in incremental energy balance equation complicates the derivation to a large extent, but for all practical purposes when the amplitude of perturbation is infinitesimal, it is permissible either to adopt an extended concept of what used to be originally a Taylor–Quinney coefficient, or to neglect completely the difference between the third and fourth term in (3.19).

By considering the kinematical relations (3.3)–(3.6), the perturbation of the kinematical part of the corotational stress rate gives

$$\Delta\overset{\nabla}{\boldsymbol{\tau}} = (\omega\mathbf{I} + \boldsymbol{\Omega}) : \Delta\boldsymbol{\tau} + \mathbf{S} : \mathbf{n} \otimes \mathbf{g}, \tag{3.22}$$

$$I_{ijkl} = \frac{1}{2}(\delta_{ik}\delta_{jl} + \delta_{jk}\delta_{il}), \tag{3.23}$$

$$\Omega_{(ij)(kl)} = \frac{1}{2}(W_{lj}\delta_{ki} + W_{ki}\delta_{lj} + W_{li}\delta_{kj} + W_{kj}\delta_{li}), \tag{3.24}$$

$$S_{ijkl} = \frac{1}{2}(\tau_{il}\delta_{jk} - \tau_{jk}\delta_{il} + \tau_{jl}\delta_{ik} - \tau_{ik}\delta_{jl}). \tag{3.25}$$

The perturbation increment of the constitutive part of $\overset{\nabla}{\boldsymbol{\tau}}$, cf. (2.7), can be expressed as

$$\Delta \overset{\nabla}{\boldsymbol{\tau}} = \mathbf{C} : \mathbf{n} \otimes \mathbf{g} - \mathbf{X} : \Delta \boldsymbol{\tau}, \tag{3.26}$$

$$\mathbf{X} = f(\omega)\mathbf{C} : \mathbf{p} \otimes \boldsymbol{\tau}' + \dot{\bar{\varepsilon}}\, \mathbf{C} : \mathcal{M}, \tag{3.27}$$

$$\mathcal{M} = \frac{\partial^2 Q}{\partial \boldsymbol{\tau} \partial \boldsymbol{\tau}}, \qquad f(\omega) = \frac{3}{2\bar{\sigma}} \left[\frac{c_\theta A \omega + \chi\, \dot{\bar{\varepsilon}}\, C}{c_\theta \omega - c_\theta B - \chi \bar{\sigma} C} \right], \tag{3.28}$$

where the function $f(\omega)$ was obtained using (3.21) and the following relations:

$$\Delta \dot{\bar{\varepsilon}} = \frac{\partial \dot{\bar{\varepsilon}}}{\partial \bar{\sigma}} \Delta \bar{\sigma} + \frac{\partial \dot{\bar{\varepsilon}}}{\partial \bar{\varepsilon}} \Delta \bar{\varepsilon} + \frac{\partial \dot{\bar{\varepsilon}}}{\partial \theta} \Delta \theta = A \Delta \bar{\sigma} + B \Delta \bar{\varepsilon} + C \Delta \theta, \tag{3.29}$$

$$\Delta \bar{\sigma} = \frac{3}{2\bar{\sigma}} \boldsymbol{\tau}' : \Delta \boldsymbol{\tau}, \qquad \Delta \dot{\bar{\varepsilon}} = \omega \Delta \bar{\varepsilon}, \qquad \Delta \dot{\theta} = \omega \Delta \theta. \tag{3.30}$$

The tensor of elastic moduli in expressed in its standard form in terms of Lamé constants λ, μ

$$C_{ijkl} = \lambda \delta_{ij} \delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{jk}\delta_{il}). \tag{3.31}$$

Finally, a combination of (3.22) and (3.26) yields

$$(\omega \mathbf{I} + \boldsymbol{\Omega} + \mathbf{X}) : \Delta \boldsymbol{\tau} = (\mathbf{C} - \mathbf{S}) : \mathbf{n} \otimes \mathbf{g}. \tag{3.32}$$

If we denote

$$\boldsymbol{\mathcal{A}} = (\omega \mathbf{I} + \boldsymbol{\Omega} + \mathbf{X}), \tag{3.33}$$

$$\boldsymbol{\mathcal{B}} = \mathbf{C} - \mathbf{S}, \tag{3.34}$$

then Eqs. (3.32) and (3.15) give rise to a set of equations

$$\begin{aligned} &\boldsymbol{\mathcal{A}} : \Delta \boldsymbol{\tau} = \boldsymbol{\mathcal{B}} : \mathbf{n} \otimes \mathbf{g}, \\ &\mathbf{n} \cdot \Delta \boldsymbol{\tau} + \frac{\rho \omega}{k^2} \mathbf{g} = \mathbf{0}, \end{aligned} \tag{3.35}$$

which enables us to relate the unknown amplitudes of velocity and stress increments. Multiplying the first of (3.35) by $(\mathbf{n} \cdot \boldsymbol{\mathcal{A}}^{-1} :)$, and substituting into the second yields

$$\left(\mathbf{n} \cdot \boldsymbol{\mathcal{A}}^{-1} : \boldsymbol{\mathcal{B}} \cdot\cdot\, \mathbf{n} + \frac{\rho \omega}{k^2} \mathbf{1} \right) \cdot \mathbf{g} = \mathbf{0}, \tag{3.36}$$

where $\boldsymbol{\mathcal{A}}^{-1}$ is the symmetric inverse of the fourth rank tensor $\boldsymbol{\mathcal{A}}$ such that

$$\boldsymbol{\mathcal{A}}^{-1} : \boldsymbol{\mathcal{A}} = \mathbf{I}. \tag{3.37}$$

By introducing

$$\mathbf{L} = \mathcal{A}^{-1} : \mathcal{B}, \tag{3.38}$$

$$\mathbf{Q} = \mathbf{n} \cdot \mathbf{L} \cdot\cdot \mathbf{n}, \qquad Q_{jl} = n_i L_{ijkl} n_k \,, \tag{3.39}$$

Eq. (3.36) takes the following form

$$\left(\mathbf{n} \cdot \mathbf{L} \cdot\cdot \mathbf{n} + \frac{\rho\omega}{k^2}\mathbf{1}\right) \cdot \mathbf{g} = \mathbf{0}. \tag{3.40}$$

A condition for the nontrivial solution of the homogeneous system (3.40) yields the characteristic equation

$$\det\left[\mathbf{Q} + \frac{\rho\omega}{k^2}\mathbf{1}\right] = 0. \tag{3.41}$$

Expanding the characteristic equation (3.41) gives the dispersion relation of the following form,

$$D(\omega, \mathbf{k}) = \left(\frac{\rho\omega}{k^2}\right)^3 + \left(\frac{\rho\omega}{k^2}\right)^2 \operatorname{tr}\mathbf{Q} + \left(\frac{\rho\omega}{k^2}\right)\frac{1}{2}\left(\operatorname{tr}^2\mathbf{Q} - \operatorname{tr}\mathbf{Q}^2\right) + \det\mathbf{Q} = 0. \tag{3.42}$$

The signs of the real parts of the roots Re[ω] determine the regions of stability and instability. The expansion of (3.41) via (3.42) is by no means simple, since it requires a direct inversion of the tensor $\mathcal{A}$ which explicitly depends on ω representing an unknown eigenvalue of the stability problem. In the following section we give a detailed derivation of the dispersion relation.

4. Derivation of dispersion relation

In what follows, for simplicity, we shall neglect the effect of tensors $\boldsymbol{\Omega}$, $\mathbf{S}$ emerging from the corotational parts of Cauchy stress rate. In this way we limit the nature of perturbation to infinitesimal increments which are almost unidirectional and without significant changes in rotation, i.e., $\Delta \overset{\nabla}{\boldsymbol{\tau}} \approx \Delta \dot{\boldsymbol{\tau}}$.

By introducing two linearly independent, idempotent, symmetric tensors $\mathbf{J}$ and $\mathbf{K}$

$$J_{ijkl} = \frac{1}{3}\delta_{ij}\delta_{kl}, \qquad K_{ijkl} = \frac{1}{2}(\delta_{ik}\delta_{jl} + \delta_{jk}\delta_{il}) - \frac{1}{3}\delta_{ij}\delta_{kl}, \tag{4.1}$$

$$\mathbf{J}^2 = \mathbf{J}, \qquad \mathbf{K}^2 = \mathbf{K}, \qquad \mathbf{J} : \mathbf{K} = \mathbf{K} : \mathbf{J} = \mathbf{0},$$

for plastic flow rule with isotropic hardening and von Mises effective stress, the expressions needed for further development are given as

$$\bar{\sigma}^2 = \frac{3}{2}\boldsymbol{\tau}' : \boldsymbol{\tau}', \qquad \mathbf{p} = \frac{\partial\bar{\sigma}}{\partial\boldsymbol{\tau}} = \frac{3\boldsymbol{\tau}'}{2\bar{\sigma}}, \qquad \mathbf{C} = (3\lambda + 2\mu)\mathbf{J} + 2\mu\mathbf{K}, \tag{4.2}$$

$$\mathcal{M} = \frac{3}{2\bar{\sigma}}\left[\mathbf{K} - \frac{3}{2\bar{\sigma}^2}\boldsymbol{\tau}' \otimes \boldsymbol{\tau}'\right], \qquad \mathbf{X} = \alpha\boldsymbol{\tau}' \otimes \boldsymbol{\tau}' + \beta\mathbf{K}, \tag{4.3}$$

$$\alpha = \frac{9\mu}{2\bar{\sigma}^3}\left[\frac{c_\theta(\bar{\sigma}A - \dot{\bar{\varepsilon}})\omega + c_\theta\,\dot{\bar{\varepsilon}}\,B + 2\chi\bar{\sigma}\,\dot{\bar{\varepsilon}}\,C}{c_\theta(\omega - B) - \chi\bar{\sigma}C}\right], \qquad \beta = 3\mu\frac{\dot{\bar{\varepsilon}}}{\bar{\sigma}}, \tag{4.4}$$

$$\boldsymbol{\mathcal{A}} = \omega\mathbf{J} + (\omega + \beta)\mathbf{K} + \alpha\boldsymbol{\tau}' \otimes \boldsymbol{\tau}', \qquad \boldsymbol{\mathcal{B}} = \mathbf{C} = (3\lambda + 2\mu)\mathbf{J} + 2\mu\mathbf{K}. \tag{4.5}$$

It is a simple matter now to obtain the inverse of the fourth rank tensor $\boldsymbol{\mathcal{A}}$ and, according to (3.38), the expression for $\mathbf{L}$:

$$\begin{aligned}
&\boldsymbol{\mathcal{A}} = \mathbf{E} + \alpha\boldsymbol{\tau}' \otimes \boldsymbol{\tau}',\\
&\mathbf{E} = \omega\mathbf{J} + (\omega + \beta)\mathbf{K}, \qquad \mathbf{E}^{-1} = \frac{1}{\omega}\mathbf{J} + \frac{1}{\omega + \beta}\mathbf{K},\\
&\boldsymbol{\mathcal{A}}^{-1} = \mathbf{E}^{-1} - \frac{\alpha}{1 + \alpha\boldsymbol{\tau}' : \mathbf{E}^{-1} : \boldsymbol{\tau}'}\left[\mathbf{E}^{-1} : \boldsymbol{\tau}' \otimes \boldsymbol{\tau}' : \mathbf{E}^{-1}\right]\\
&\qquad\qquad = \frac{1}{\omega}\mathbf{J} + \frac{1}{\omega + \beta}\mathbf{K} - \alpha^*\boldsymbol{\tau}' \otimes \boldsymbol{\tau}',\\
&\alpha^* = \frac{9\mu}{2\bar{\sigma}^3}\left[\frac{c_\theta(\bar{\sigma}A - \dot{\bar{\varepsilon}})\omega + c_\theta\,\dot{\bar{\varepsilon}}\,B + 2\chi\bar{\sigma}\,\dot{\bar{\varepsilon}}\,C}{(\omega + \beta)(c_\theta\omega^2 + (3\mu c_\theta A - c_\theta B - \chi\bar{\sigma}C)\omega + 3\chi\mu\,\dot{\bar{\varepsilon}}\,C)}\right];
\end{aligned} \tag{4.6}$$

$$\mathbf{L} = \boldsymbol{\mathcal{A}}^{-1} : \mathbf{C} = \frac{(3\lambda + 2\mu)}{\omega}\mathbf{J} + \frac{2\mu}{\omega + \beta}\mathbf{K} - 2\mu\alpha^*\boldsymbol{\tau}' \otimes \boldsymbol{\tau}'. \tag{4.7}$$

Noting that

$$\begin{aligned}
&\mathbf{n} \cdot \mathbf{J} \cdot\cdot\, \mathbf{n} = \frac{1}{3}\mathbf{n} \otimes \mathbf{n}, \qquad \mathbf{n} \cdot \mathbf{K} \cdot\cdot\, \mathbf{n} = \frac{1}{2}\left[\mathbf{1} + \frac{1}{3}\mathbf{n} \otimes \mathbf{n}\right],\\
&\mathbf{n} \cdot \boldsymbol{\tau}' \otimes \boldsymbol{\tau}' \cdot \mathbf{n} = \mathbf{t} \otimes \mathbf{t}, \qquad n_i\tau'_{ij}\tau'_{kl}n_k = \mathbf{t}_j\mathbf{t}_l,
\end{aligned} \tag{4.8}$$

the perturbation wave acoustic tensor $\mathbf{Q}$ admits a decomposition

$$\mathbf{Q} = q_0\mathbf{1} + q_1\mathbf{n} \otimes \mathbf{n} + q_2\mathbf{t} \otimes \mathbf{t}, \qquad Q_{ij} = q_0\delta_{ij} + q_1n_in_j + q_2t_it_j, \tag{4.9}$$

where the scalar parameters q_0, q_1, q_2 which convey the information about the current state of stress and other material parameters being part of the constitutive equation, are given as

$$\begin{aligned}
q_0 &= \frac{\mu}{\omega + \beta},\\
q_1 &= \frac{1}{3}\left(\frac{3\lambda + 2\mu}{\omega} + \frac{\mu}{\omega + \beta}\right),\\
q_2 &= \frac{-2\mu\alpha}{(\omega + \beta)(\omega + \beta + 2\alpha\bar{\sigma}^2/3)}.
\end{aligned} \tag{4.10}$$

To determine the dispersion relation we need to calculate principal invariants of the tensor $\mathbf{Q}$:

$$
\begin{aligned}
I_1 &= \operatorname{tr}\mathbf{Q} = 3q_0 + q_1 + q_2\mathbf{t}^2, \\
I_2 &= (\operatorname{tr}^2\mathbf{Q} - \operatorname{tr}\mathbf{Q}^2)/2 = q_0(3q_0 + 2q_1) + (2q_0 + q_1)q_2\mathbf{t}^2 - q_1q_2\mathbf{t}_n^2, \\
I_3 &= \det\mathbf{Q} = q_0^2(q_0 + q_1) + q_0(q_0 + q_1)q_2\mathbf{t}^2 - q_0q_1q_2\mathbf{t}_n^2.
\end{aligned} \tag{4.11}
$$

By carrying out the expansions, with some rearrangements, (3.42) can be expressed in the following form:

$$
D = d_0 + d_1\mathbf{t}^2 + d_2\mathbf{t}_n^2. \tag{4.12}
$$

Here the coefficients d_0, d_1, d_2 and expresions for $\mathbf{t}^2$ and $\mathbf{t}_n^2$ are given as

$$
\begin{aligned}
d_0 &= \left(\frac{\rho\omega}{k^2} + q_0\right)^2\left(\frac{\rho\omega}{k^2} + q_0 + q_1\right), \\
d_1 &= \left(\frac{\rho\omega}{k^2} + q_0\right)\left(\frac{\rho\omega}{k^2} + q_0 + q_1\right)q_2, \\
d_2 &= -\left(\frac{\rho\omega}{k^2} + q_0\right)q_1q_2,
\end{aligned} \tag{4.13}
$$

$$
\mathbf{t}^2 = \mathbf{t}_i\mathbf{t}_i = n_k{\tau'_{kl}}^2 n_l, \qquad \mathbf{t}_n^2 = (n_i\mathbf{t}_i)^2 = (n_i\tau'_{ij}n_j)^2. \tag{4.14}
$$

Once again, the dispersion relation in explicit form yields

$$
D = d_0 + d_1{\tau'_{kl}}^2 n_k n_l + d_2\tau'_{ij}\tau'_{kl}n_i n_j n_k n_l. \tag{4.15}
$$

Next we observe that the scalar parameters d_0, d_1, d_2 have a common factor which can be factorized from the original expression

$$
\begin{aligned}
&D = D_C D_R, \qquad D_C = \left(\frac{\rho\omega}{k^2} + q_0\right), \\
&D_R = \left(\frac{\rho\omega}{k^2} + q_0\right)\left(\frac{\rho\omega}{k^2} + q_0 + q_1 + q_2\mathbf{t}^2\right) + q_1q_2(\mathbf{t}^2 - \mathbf{t}_n^2).
\end{aligned} \tag{4.16}
$$

This is quite advantageous, since it contributes to a significant simplification, i.e., it lowers the order of polynomial in ω, and we are thus able to solve for ω the two separate equations $D_C = 0$ and $D_R = 0$, respectively. By solving equation $D_C = 0$ first, it yields the negative real parts of the roots. Consequently, the first subset of solutions is always in the stable domain. Therefore, the critical eigenvalues that may eventually cause the instability should be determined from the remaining part of the dispersion relation. It should be kept in mind that the coefficients q_0, q_1, q_2 contain ω as an unknown eigenvalue of the problem as well

as the perturbation wave vector $\mathbf{k} = k\mathbf{n}$, i.e., $q_L = q_L(\omega, \mathbf{k})$; $L = 0, 1, 2$. Carrying out the expansions, after some algebraic manipulations, (4.16) in the expanded form yields

$$D(\omega, \mathbf{k}) = \frac{1}{d}\sum_{k=0}^{N} a_k \omega^{N-k}, \qquad d = \overline{\sigma}^4 \omega(\omega + \beta)^2(c_0\omega^2 + c_1\omega + c_2),$$

$$c_0 = c_\theta, \qquad c_1 = c_\theta(3\mu A - B) - \chi\overline{\sigma}C, \qquad c_2 = 3\chi\mu\,\dot{\bar{\varepsilon}}\,C. \tag{4.17}$$

The order of exponent N in the formula (4.17) depends on the nature of perturbation. We distinguish between the two limiting cases:

i. Quasi-static perturbation. Inertial effects are negligible, and terms containing mass density are neglected. In this case we have $N = 3$ and (4.17) has the form

$$D(\omega, \mathbf{k}) = \frac{1}{d}(a_0\omega^3 + a_1\omega^2 + a_2\omega + a_3). \tag{4.18}$$

ii. Dynamic perturbation. Inertial terms are included. Here we have $N = 7$ and the dispersion relation is of the form

$$D(\omega, \mathbf{k}) = \frac{1}{d}(a_0\omega^7 + a_1\omega^6 + a_2\omega^5 + a_3\omega^4 + a_4\omega^3 + a_5\omega^2 + a_6\omega + a_7). \tag{4.19}$$

In both cases the value of the denominator as defined in (4.17) remains unchanged. The polynomial coefficients $a_k(\mathbf{k})$ which depend on the current state of stress, strain rate, and other material parameters are given in Tables 1 and 2 in the Appendix.

5. Stability criteria revisited

Since the stability is determined by the sign of the real parts of the roots (ω_k) of the numerator in (4.17), one way of checking the sign definiteness of the sequence $\mathrm{Re}[\omega_k]$ is to employ a Routh–Hurwitz criterion (RHC). A necessary and sufficient condition that all the roots have negative real parts which implies asymptotic stability is that all terms forming the sequence of Hurwitz determinants $\{D_k\}$ are positive, i.e., $\{D_1, D_2, \ldots D_N;\ D_k > 0\ \forall k \in [1, N]\}$. The individual members are defined as determinants of the principal minors of

$$D_N = \begin{vmatrix} a_1 & a_0 & 0 & 0 & \ldots & 0 \\ a_3 & a_2 & a_1 & a_0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & a_N & a_{N-1} & a_{N-2} \\ 0 & 0 & 0 & 0 & \ldots & a_N \end{vmatrix} = a_N D_{N-1} > 0.$$

The last two terms of the sequence $\{D_k\}$, namely D_{N-1}, D_N, deserve special attention. According to the RHC we must have $D_{N-1} > 0$ and $D_N = a_N D_{N-1} > 0$. First let us assume that $D_{N-1} > 0$. Thus, for stability we must require that $a_N > 0$. Here index N equals 3 to 7, depending on the type of perturbation we are imposing on the system, quasi-static or dynamic, respectively. By comparing the last two coefficients a_N in the expansion of the dispersion relation $D_R(\omega, \mathbf{k})$ pertaining to the case of quasi-static (coefficient a_3) and dynamic perturbation (coefficient a_7), we observe that in both cases they are the same, regardless of the nature of perturbation. So when a_N changes its sign from $a_N > 0$ to $a_N < 0$, the transition takes place without a direct influence of inertia. When $a_N > 0$ and ω is imaginary, then the stability depends on the behaviour of the term D_{N-1}. The discussion of this particular case is outlined in subsection 2 below.

Since the problem can be viewed as a nonlinear evolution process starting from the initial configuration which presumably is taken to be stable, we are actually interested in determining the conditions under which the state of neutral stability is reached first, and in the nature of the transition from the stable into the unstable domain. By using the Routh–Hurwitz criterion to investigate stability, it turns out that the last two terms in the RH sequence actually dictate the marginal stability behaviour of the system.

Certain peculiarities of the stability criteria can be recognized by inspecting the expanded form of dispersion relation (4.17). Since the characteristic equation takes the form of a rational expression, care should be taken to distinguish between different types of removable and unremovable singularities which may emerge when the conditions for neutral stability are attained and play an important role in defining criteria for the existence of a bifurcation.

5.1. Static loss of stability, divergence point

For $\omega = 0$, ($\mathrm{Re}[\omega] = \mathrm{Im}[\omega] = 0$) one must have $a_N = 0$. However, this does not imply that $D_R(\omega = 0, \mathbf{k})$ vanishes too, because of the ω in the denominator, cf. (4.17). The condition $a_N = 0$ only provides the removable singularity of the characteristic equation at $\omega = 0$. However, for $D_R(0, \mathbf{k}) = 0$ we must also have $a_{N-1} = 0$. Thus, the existence of a nontrivial solution as indicated in (3.41) via (4.18) requires that the following two conditions should be satisfied simultaneously: $a_N = a_{N-1} = 0$. Since both coefficients have to vanish at the same time, they define rather severe conditions under which bifurcation in the form of branching of the solution may set in. There are different ways in which the former two conditions can be satisfied. A detailed study of these particular cases is given in Sec. 7. Theoretically, bifurcation is not excluded, but as it will be seen later on, under certain circumstances, it can be precluded in practice for the class of constitutive relations considered here. Nevertheless, a sign change of the coeficient a_N will always indicate the exchange of stability.

5.2. Dynamic loss of stability. Transition to a time-period flow. Hopf bifurcation

Here we consider the case in which the dispersion relation has two purely imaginary eigenvalues for a certain value of the control parameter $\Lambda = \Lambda_h$, whereas other eigenvalues have non-vanishing negative real parts. The basic assumption made here is that the eigenvalue is simple, and as under definition at $\Lambda = \Lambda_h$ not only $\mathrm{Re}[\omega] = 0$, but also the branch $\omega_R = \mathrm{Re}[\omega] = \omega(\Lambda)$ crosses the Λ-axis. This is the transversality condition $d\mathrm{Re}[\omega(\Lambda)]/d\Lambda \neq 0$. At $\Lambda = \Lambda_h$ a pair of complex conjugate eigenvalues crosses the imaginary axis with a non-zero velocity. The remaining eigenvalues continue to have negative real parts. For details see MARSDEN and MCCRACKEN [10].

The bifurcation solutions will become periodic in time if $\mathrm{Im}[\omega(\Lambda_h)] \neq 0$. By substituting $\omega^2 = -s^2$ into (4.17), the dispersion relation splits into the real and imaginary part

$$
\begin{gathered}
D_R(\omega, \mathbf{k}) = D_{\mathrm{Re}}(\omega, \mathbf{k}) + iD_{\mathrm{Im}}(\omega, \mathbf{k}) = 0, \\
D_{\mathrm{Re}} = \mathrm{Re}[D_R(\omega, \mathbf{k})] = 0, \qquad D_{\mathrm{Im}} = \mathrm{Im}[D_R(\omega, \mathbf{k})] = 0.
\end{gathered} \tag{5.1}
$$

In the case of a quasi-static perturbation when inertia is negligible, the expressions obtained from (5.1) are relatively simple since we are dealing with a third-order polynomial. Imposing (5.1) upon (4.18) yields the Hopf criterion of the form

$$
H_3 = a_1 a_2 - a_0 a_3 = 0. \tag{5.2}
$$

The preceding equation defines the relation between the coefficients of the dispersion relation which has to be satisfied in order that two complex conjugate roots $\omega^2 + s^2 = 0$ solve (5.1). From (5.2) the relation to RH criteria follows directly. For a general third-order polynomial with real coefficients, the RHC are given as

$$
D_1 = a_1 > 0, \qquad D_2 = a_1 a_2 - a_0 a_3 > 0, \qquad D_3 = a_3 D_2 > 0. \tag{5.3}
$$

By comparing (5.3) and (5.2) it can be observed that the Hopf criterion, $H_3 = 0$, corresponds to $D_2 = 0$. The complete sequence of RH inequalities hence degenerates to

$$
\{D_1, D_2, D_3\} = \{a_1 > 0, 0, 0\}. \tag{5.4}
$$

In the case of a dynamic perturbation, the same procedure gives

$$
\begin{aligned}
a_0 s^6 - a_2 s^4 + a_4 s^2 - a_6 &= 0, \\
a_1 s^6 - a_3 s^4 + a_5 s^2 - a_7 &= 0.
\end{aligned} \tag{5.5}
$$

After introducing a new variable $\xi = s^2$, the problem reduces to eliminating ξ from the two polynomials of the third degree. To eliminate ξ from the last two

equations means to find a relation between the coefficients which must hold if the two equations are both satisfied. In other words, we have to find a necessary condition for the two equations to have a common root. The resultant of the system (5.5) can be represented as a determinant

$$R = \begin{vmatrix} a_0 & -a_2 & a_4 & -a_6 & 0 & 0 \\ 0 & a_0 & -a_2 & a_4 & -a_6 & 0 \\ 0 & 0 & a_0 & -a_2 & a_4 & -a_6 \\ 0 & 0 & a_1 & -a_3 & a_5 & -a_7 \\ 0 & a_1 & -a_3 & a_5 & -a_7 & 0 \\ a_1 & -a_3 & a_5 & -a_7 & 0 & 0 \end{vmatrix} = 0. \tag{5.6}$$

Vanishing of the determinant is a necessary condition for (5.5) to have a common root: a device konwn as Silvester's dialytic method of elimination, WAERDEN [23]. In this case the Hopf criterion can be stated as $H_7 = R = 0$. From (5.6) and from Table 1 it becomes apparent that the HC in this case is rather cumbersome because of the expansion itself and because of the complexity of the coefficients of the dispersion relation. By using one of the determinant expansion methods one can prove that the condition $H_7 = 0$ corresponds to the condition $D_6 = 0$ in the RH sequence for the seventh-order polynomial, implying that the last two terms of the sequence vanish:

$$H_7 = 0 \Rightarrow D_6 = 0, \quad D_7 = a_7 D_6 = 0. \tag{5.7}$$

The Hopf criteria $H_3 = 0$ and $H_7 = 0$ belong to a general class of possible solutions. In the case of a dynamic perturbation, the criterion $H_7 = 0$ is practically unobtainable in explicit form because of the inherent complexity of the individual coefficients a_k. On the other hand, we are not interested in the general class of solutions, but rather in the particular ones, providing the worst case scenario for losing stability once the system in on the threshold of instability. It is more feasible to seek a subset of solutions which defines the most severe conditions under which the system is prone to lose stability in the fastest manner possible. To answer these questions, we take a closer look at the stationary properties of dispersion relation.

6. Stationary point of the dispersion relation

The dispersion relation defines the relationship between the characteristic coefficients as functions of material parameters and eigenvalues ω. To determine the regions of stability, we either have to solve for ω, or use the Routh–Hurwitz criterion to detect when the exchange of stability takes place. For practical use it is especially important to determine the directions in which a given property

of the dispersion relation $D(\omega, \mathbf{k})$ assumes its maximum and minimum values. The reason for this is that expressions which define individual members of the dispersion relation are very complicated from an algebraic standpoint. It may be more feasible to analyze extremal properties of the function D and the relations which follow from them. Since D can be viewed as a function of the eigenvalue ω and the wave vector $\mathbf{k}$, D assumes its extreme value in a certain direction of the wave vector $\mathbf{k} = k\mathbf{n}$ of the imposed disturbance. It is required to find unit vectors $\mathbf{n}$, such that $G(\mathbf{n}) = \mathbf{n} \cdot \mathbf{n} - 1 = 0$, where the function D given by (4.17) assumes minimum value and vanishes for certain values of ω. Mathematically, the problem is to find stationary values of the dispersion relation subject to a constraint. We use the method of Lagrange multipliers seeking extrema of the function

$$F = D - \lambda(\mathbf{n} \cdot \mathbf{n} - 1). \tag{6.1}$$

At the corresponding values of $\mathbf{n}$ the stationary conditions are

$$\begin{aligned} \frac{\partial F}{\partial n_r} &= \frac{\partial D}{\partial n_r} - 2\lambda n_r = 0, \\ \frac{\partial F}{\partial \lambda} &= n_k n_k - 1 = 0. \end{aligned} \tag{6.2}$$

Derivatives in (6.2) can be determined by defining auxiliary quantities

$$\frac{\partial D}{\partial n_r} = 2D_{rk}n_k, \qquad \tau'_{nn} = n_i \tau'_{ij} n_j, \qquad D_{rk} = d_1 \tau'_{rs}\tau'_{sk} + 2d_2 \tau'_{nn}\tau'_{rk}. \tag{6.3}$$

The components n_k and the undetermined Lagrangian multiplier λ are found from (6.2)

$$\lambda = \frac{1}{2}\frac{\partial D}{\partial n_r} n_r = n_r D_{rk} n_k = d_1 \mathbf{t}^2 + 2d_2 \mathbf{t}_n^2, \tag{6.4}$$

$$(D_{rk} - \lambda \delta_{rk}) n_k = 0. \tag{6.5}$$

The introduction of a Lagrange multiplier transforms the minimization problem into a saddle point problem. For fixed λ, D is a minimum at $\lambda = \max \lambda$. This can be easily verified by calculating the second variation of D with respect to $\mathbf{n}$. The geometrical interpretation of the procedure can be illustrated by calculating the gradient vector of the representation surface $D = D(\omega_0 k\mathbf{n})$ of the dispersion function evaluated at some fixed value of ω, say $\omega = \omega_0$. Since we are on the unit sphere, we have

$$\operatorname{grad} D = \frac{\partial D}{\partial \mathbf{n}}\bigg|_{\omega=\omega_0} = 2\lambda \mathbf{n}, \qquad \max |\operatorname{grad} D| = \max \left| \frac{\partial D}{\partial \mathbf{n}} \right| = 2 \max(\lambda). \tag{6.6}$$

By trying to determine $\max(\lambda)$, we are actually searching for the direction that gives the largest value of the gradient on the representation surface, being at the same time the normal gradient of the unit sphere. From this a very powerful conclusion can be drawn. When the necessary conditions for reaching the critical state are met, the instability will evolve at the fastest growth rate in the direction of the gradient with the largest magnitude on the stability surface $D = D(\omega, k\mathbf{n})$. Following the maximum value of the gradient gives the differential direction of the path of the steepest ascent or descent (depending on the sign of gradient) in which the absolute value of perturbation in velocity field exhibits the fastest rate of growth. More directly, by calculating the gradient of $|\Delta\mathbf{v}| = (\Delta\mathbf{v} \cdot \Delta\mathbf{v}^*)^{1/2}$ in the direction of $\mathbf{n}$ we obtain

$$\frac{\partial|\Delta\mathbf{v}|}{\partial n} = \frac{\partial|\Delta\mathbf{v}|}{\partial\mathbf{n}}\mathbf{n}\cdot\operatorname{grad}\omega = |\Delta\mathbf{v}|\mathbf{n}\cdot\frac{\partial\omega}{\partial\mathbf{n}} = \frac{|\Delta\mathbf{v}|}{[\partial D/\partial\omega]_{\omega_0}}\mathbf{n}\cdot\frac{\partial D}{\partial\mathbf{n}} = 2\frac{|\Delta\mathbf{v}|}{D_{\omega_0}}\lambda, \tag{6.7}$$

from where it follows that $\max\lambda$ via $\max|\operatorname{grad} D|$ leads to $\max(\partial|\Delta\mathbf{v}|/\partial n)$, providing that the point under consideration on the representation surface is regular, ie., $D_{\omega_0} = [\partial D/\partial\omega]_{\omega_0} \neq 0$. This is consistent with the starting assumption that the exchange of stability takes place only through an eigenvalue which is simple.

Referring to (6.5) again, the equations can be simplified significantly by taking the system of coordinates constructed on the principal axes of stress tensor. In this case we have

$$\begin{aligned} \tau'_{nn} &= n_1^2\tau'_1 + n_2^2\tau'_2 + n_3^2\tau'_3, \\ \mathbf{t}^2 &= n_1^2\tau_1'^2 + n_2^2\tau_2'^2 + n_3^2\tau_3'^2, \\ \mathbf{t}_n^2 &= (n_1^2\tau'_1 + n_2^2\tau'_2 + n_3^2\tau'_3)^2, \end{aligned} \tag{6.8}$$

and the system of three homogeneous equations

$$\begin{aligned} n_1(d_1\tau_1'^2 + 2d_2\tau'_{nn}\tau'_1 - \lambda) &= 0, \\ n_2(d_1\tau_2'^2 + 2d_2\tau'_{nn}\tau'_2 - \lambda) &= 0, \\ n_3(d_1\tau_3'^2 + 2d_2\tau'_{nn}\tau'_3 - \lambda) &= 0. \end{aligned} \tag{6.9}$$

The systems admits a variety of solutions depending on the stress state and the plane orientation defined by the vector $\mathbf{n}$. In general, no solutions are possible with all three components n_k different from zero and with all three stress deviators different from each other. Other possibilities are summarized in Table 7.

For the sake of convenience, we introduce the following auxiliary variables needed for further development:

$$m^2 = (d_1 + 4d_2)/6d_2, \qquad \ell^2 = (2d_2 - d_1)/6d_2, \tag{6.10}$$

both satisfying the condition $m^2 + \ell^2 = 1$, a rational expression

$$\Xi = d_1/d_2, \tag{6.11}$$

and a normalized value of the Lagrangian multiplier $\overline{\lambda}$, all used in Table 7,

$$\overline{\lambda} = \lambda/d_1. \tag{6.12}$$

To illustrate how Table 7 can be used in the stability analysis, consider the two special states of stress: uniaxial tension and simple shear. Here we assume that only a static exchange of stability takes place during the deformation process. Hence, ω is changing over the real axis, implying that $\mathrm{Im}[\omega(\Lambda)] = 0$ for all values of the control parameter Λ.

6.1. Uniaxial tension

In a physical space of principal stresses the state vector is defined as $\boldsymbol{\tau} = (\tau_1, \tau_2, \tau_3) = X(1,0,0)$ with the corresponding deviatoric counterpart $\boldsymbol{\tau}' = (\tau_1', \tau_2', \tau_3') = (X/3)(2,-1,-1)$. X is the applied external stress in the principal direction 1. According to Table 7 there are two candidates for the solution: Cases 1 and 4A. Since we are interested in the static exchange of stability only, the transition from $\omega < 0$ to $\omega > 0$ takes place through $\omega = 0$. At the critical point we have $\omega = 0$, $\Xi = -1$ and from Table 7 follows the $\max(\overline{\lambda})$ calculated for the Case 1,

$$\overline{\lambda} = 2X^2/9, \qquad \ell^2 = m^2 = 1/2, \qquad n_1^2 = 1/2, \qquad n_2^2 + n_3^2 = 1/2. \tag{6.13}$$

The last two expressions define the direction of the plane with the fastest incipient rate of instability growth. A family of plane envelopes satisfying (6.13) forms a double cone with the apex at the origin, inclined at an angle of $\pi/4$ to the axis of the maximum principal stress deviator. The assertion holds for $\omega = 0$ only, or at least in the limit, when ω tends to zero ($\omega \to 0$). In general, however, when $\omega \neq 0$ then $\Xi(\omega)$ is not constant, implying that there exists a dependence $\mathbf{n}(\omega(\Lambda))$. Here Λ represents the control parameter. The normal vector varies during the deformation and the loading process.

6.2. Simple shear

Without any loss of generality we may take coordinate axes x_1 and x_2 to define a plane of shear with $\tau_3' = 0$. The physical space of principal stress and stress deviators coincide, and the state vector is now given as $\boldsymbol{\tau}' = X(1,-1,0)$, $X = \tau_{12}$. By analogous considerations as those in the case of uniaxial tension we find from Table 7 that the case 6B prevails, giving

$$\overline{\lambda} = X^2, \qquad n_1^2 = 1/2, \qquad n_2^2 = 1/2, \qquad n_3^2 = 0. \tag{6.14}$$

The plane of the fastest instability growth bisects the angle between the smallest and the greatest principal stress. But in this case the eigenvalue dependence $\Xi(\omega)$ is completely eliminated from the start. Regardless of the current value

of the control parameter, once we enter the unstable domain ($\Lambda > \Lambda_c$), the plane orientation with the fastest rate of incipient instability growth is always fixed relative to the current directions of the principal deviators. This is not the case in uniaxial tension, or in any other stress state where the eigenvalue dependence defined by $\Xi(\omega)$ is retained in the governing equations. Here the current directions of the principal stress deviators as well as the direction of the vector $\mathbf{n}$ may change relative to each other. In the case of simple shear $\Xi(\omega)$ cancels out, thus keeping the direction of $\mathbf{n}$ as defined in (6.14) fixed relative to $\boldsymbol{\tau}'$. Whenever the state of material parameters and the critical plane orientation is such that it admits the onset of instability, the instability will evolve in the fastest manner within the plane dividing the planes of max and min stress deviators. The stationary character of the vector $\mathbf{n}$ relative to $\boldsymbol{\tau}'$ may be viewed as an intrinsic property of the stress state of simple shear. Thus, in a way, the stress state of simple shear may have a deceiving character in the stablity analyses of various rate-dependent material models, mostly because of this intrinsic characteristic of the system.

Now we return to the stability analysis under a general stress setting by employing RH criteria. The transition from $\mathrm{Re}[\omega] < 0$ to $\mathrm{Re}[\omega] > 0$ is induced by the sign change of the coefficient a_N:

$$a_N = -9Bc_\theta \dot{\bar{\varepsilon}}^2 \mu^3(3\lambda + 2\mu)\mathbf{t}_s^2 + 3\chi C\bar{\sigma}\dot{\bar{\varepsilon}}^2\mu^3(3\lambda + 2\mu)\mathbf{g}_s^2. \tag{6.15}$$

According to the RHC the system is stable for $a_N > 0$, implying $\mathrm{Re}[\omega] < 0$, and unstable when $a_N < 0$ and $\mathrm{Re}[\omega] > 0$. From the algebraic structure of the coefficient a_N it follows that

$$\chi C\bar{\sigma}\mathbf{g}_s^2 - 3Bc_\theta\mathbf{t}_s^2 > 0. \tag{6.16}$$

To determine possible ways in which the expression (6.16) can change its sign, we first analyze the sign definiteness of the "stress functions" $\mathbf{t}_s^2$ and $\mathbf{g}_s^2$:

$$\begin{aligned} &\mathbf{t}_s^2 \overset{\text{def}}{=} \mathbf{t}^2 - \mathbf{t}_n^2, && \mathbf{g}_s^2 \overset{\text{def}}{=} \bar{\sigma}^2 - 6\mathbf{t}_s^2, \\ &\mathbf{t}^2 = \mathbf{n}\cdot\boldsymbol{\tau}'\cdot\boldsymbol{\tau}'\cdot\mathbf{n}, && \mathbf{t}_n^2 = (\mathbf{n}\cdot\boldsymbol{\tau}'\cdot\mathbf{n})^2. \end{aligned} \tag{6.17}$$

In the coordinate system of principal stresses, functions $\mathbf{t}_s^2$ and $\mathbf{g}_s^2$ can be expressed as quadratic forms

$$\begin{aligned} &\mathbf{t}_s^2 = \boldsymbol{\tau}'^T\cdot\mathbf{A}\cdot\boldsymbol{\tau}' \geq 0, \qquad \mathbf{g}_s^2 = \boldsymbol{\tau}'^T\cdot\mathbf{G}\cdot\boldsymbol{\tau}', \\ &\boldsymbol{\tau}'^T = [\tau_1', \tau_2'], \qquad \mathbf{A} = \begin{bmatrix} a_{11} & a_{12}/2 \\ a_{21}/2 & a_{22} \end{bmatrix}, \qquad \mathbf{G} = \begin{bmatrix} g_{11} & g_{12}/2 \\ g_{21}/2 & g_{22} \end{bmatrix}, \\ &\mathrm{tr}\,\mathbf{A} = 2n_1^2n_2^2 + 5n_2^2n_3^2 + 5n_1^2n_3^2 \geq 0, \qquad \det\mathbf{A} = 9n_1^2n_2^2n_3^2 \geq 0; \end{aligned} \tag{6.18}$$

$$\begin{aligned} \operatorname{tr}\mathbf{G} &= 6(1 - 2n_1^2n_2^2 - 5n_2^2n_3^2 - 5n_1^2n_3^2), \\ \det\mathbf{G} &= 27(1 - 8n_1^2n_2^2 - 8n_2^2n_3^2 - 8n_1^2n_3^2 + 48n_1^2n_2^2n_3^2)/4. \end{aligned} \tag{6.19}$$

From (6.18) a positive semi-definiteness of $\mathbf{t}_s^2$ follows directly, meanwhile (6.19) still needs additional investigation. Since we anticipate that $\mathbf{g}_s^2$ can change its sign irrespective of the state vector of stress deviators for a certain plane orientation with the normal vector $\mathbf{n} = (n_1, n_2, n_3)$, where at least one component in $\mathbf{n}$ is zero, we take a dual pair $n_K^2 + n_L^2 = 1$, with exclusion of the zero component of the normal $\mathbf{n}$, say $n_M = 0$, $(K + L + M = 6)$, and consider the conditions for positive definiteness of the quadratic form $\mathbf{g}_s^2$ in the plane with $n_M = 0$. Hence, the expressions in (6.19) simplify significantly, and yield the condition for positive semi-definiteness of the function $\mathbf{g}_s^2$ in the space of stress deviators of the following form:

$$\begin{aligned} \mathbf{g}_s^2 \geq 0 \quad \forall \Big\{ & n_K^2 \in [(1 - 1/\sqrt{2})/2, (1 + 1/\sqrt{2})/2]; \\ & n_L^2 \in [(1 + 1/\sqrt{2})/2, (1 - 1\sqrt{2})/2]; \quad n_K^2 + n_L^2 = 1 \Big\}. \end{aligned} \tag{6.20}$$

Equation (6.20) transcribed in terms of angles of the plane orientation with the normal $\mathbf{n}$ reads

$$\begin{aligned} & n_K = \cos\Psi, \qquad n_L = \sin\Psi, \qquad n_M = 0, \qquad (K + L + M = 6), \\ & \Psi_1 = 3\pi/8, \qquad \Psi_2 = \pi/8, \qquad \Psi_1 + \Psi_2 = \pi/2. \end{aligned} \tag{6.21}$$

The plane orientation given by the angles Ψ_1 and Ψ_2 defines the threshold for the lack of positive definiteness of the function $\mathbf{g}_s^2$ with respect to the basis of principal stress deviators. If the plane orientation is out of bounds of the interval defined by (6.20), then for some non-trivial linear combination of the components of the state vector of the stress deviators $\boldsymbol{\tau}'$, the value of the function $\mathbf{g}_s^2$ may change its sign passing through zero first. Detecting the critical state of stress and the plane orientation when the function $\mathbf{g}_s^2$ vanishes is of a great theoretical and practical importance. From (6.15) it appears that when $\mathbf{g}_s^2 = 0$, the temperature effects on the material instability in adiabatic thermodynamical process disappear, and the RHC $(a_N > 0)$ in this particular case reduces to that of the isothermal case. Hence, it remains to determine the critical conditions, namely the stress state and the instability plane orientation, when the condition $\mathbf{g}_s^2 = 0$ is first encountered. Towards this end, we express the deviatoric stresses in terms of the effective stress and a single stress-state parameter ζ defined by $\tau'_M = \zeta\bar{\sigma}$, which is zero $(\zeta = 0)$ for simple shear, and attains maximum and minimum values $(\zeta = \pm 1/3)$ for axially-symmetric compression $(\tau'_K = \tau'_M > \tau'_L)$ and extension $(\tau'_K > \tau'_L = \tau'_M)$, respectively. With ζ we can write

$$\frac{\tau'_{K,L}}{\bar{\sigma}} = -\frac{\zeta}{2} \pm \sqrt{\frac{1}{3} - \frac{3\zeta^2}{4}}, \qquad \tau'_M = \zeta\bar{\sigma}. \tag{6.22}$$

The characteristic "stress functions" needed for further development transform accordingly:

$$\frac{\mathbf{t}_s^2}{\bar{\sigma}^2} = \frac{\vartheta}{3}(4 - 9\zeta^2), \qquad \frac{\mathbf{f}_s^2}{\bar{\sigma}^2} = 1 - 3\frac{\mathbf{t}_s^2}{\bar{\sigma}^2} = 1 - \vartheta(4 - 9\zeta^2),$$

$$\frac{\mathbf{g}_s^2}{\bar{\sigma}^2} = 1 - 2\vartheta(4 - 9\zeta^2), \qquad \frac{\mathbf{t}_n^2}{\bar{\sigma}^2} = \left(\frac{\zeta}{2} \mp \sqrt{\left(\frac{1}{3} - \frac{3\zeta^2}{4}\right)(1 - 4\vartheta)}\right)^2, \tag{6.23}$$

$$\vartheta = n_K^2 n_L^2, \qquad n_K^2 + n_L^2 = 1, \qquad n_M = 0,$$
$$\vartheta \in [0, 1/4], \qquad \zeta \in [-1/3, 1/3]. \tag{6.24}$$

In the limit (as $\omega \to 0$), the instability plane orientation parameter ϑ attains its maximum value ($\vartheta = 1/4$) for the plane orientation with the fastest rate of instability growth, meanwhile the minimum value $\vartheta = 0$ corresponds to the first degenerate case of dispersion relation with $\mathbf{t}_s^2 = 0$. From (6.23) we have

$$\mathbf{g}_s^2/\bar{\sigma}^2 = 1 - 2\vartheta(4 - 9\zeta^2) = 0. \tag{6.25}$$

The equation can be solved in two differential ways. First we seek the stress state which satisfies $\mathbf{g}_s^2 = 0$ within the plane of the fastest rate of instability growth. Here we have $\vartheta = 1/4$ and (6.25) yields

$$\zeta = \pm\frac{\sqrt{2}}{3}, \qquad \frac{\tau'_{K,L}}{\bar{\sigma}} = \frac{\pm\sqrt{6} \mp \sqrt{2}}{6}, \qquad \frac{\tau'_M}{\bar{\sigma}} = \pm\frac{\sqrt{2}}{3} \tag{6.26}$$

giving the result for ζ which clearly falls out of bounds of the interval $|\zeta| \leq 1/3$. Moreover, one can show that in the case when $\zeta = \sqrt{2}/3$, the normalized value of Lagrangian multiplier $\bar{\lambda}$ in the plane $n_K = 0$ exceeds the one within the plane $n_M = 0$, i.e. $(-\tau'_L\tau'_M > -\tau'_K\tau'_L)$, leading to the evolution of instability within the plane $n_K = 0$. The same reasoning applies for $\zeta = -\sqrt{2}/3$ where $(-\tau'_K\tau'_M > -\tau'_K\tau'_L)$ implies the instability in the $n_L = 0$ plane, cf. Table 7. Since both solutions violate the starting assumption about the preferred initial orientation of the instability plane defined by $n_M = 0$, they can be disregarded in further development. Previous reasoning leads to the conclusion that in general the equation $\mathbf{g}_s^2 = 0$ cannot be satisfied within the plane of the fastest rate of instability growth. Therefore, instead of seeking the state of stress within the plane ($\vartheta = 1/4$) which satisfies Eq. (6.25), we reverse the task and for a given state of stress characterized by the scalar parameter ζ, try to determine the plane orientation which satisfies (6.25):

$$\vartheta = \frac{1}{2(4 - 9\zeta^2)}, \qquad \vartheta(\zeta = 0) = \frac{1}{8}, \qquad \vartheta\left(\zeta = \pm\frac{1}{3}\right) = \frac{1}{6}. \tag{6.27}$$

Upon using the same angle notation as in (6.21), we can express the angle Ψ in terms of the plane orientation parameter ϑ. The angle Ψ which now defines the plane inclination of isothermal invariance in adiabatic material instability process is thus given by

$$\Psi = \pm \arctan \left(\frac{1 - 2\vartheta \pm \sqrt{1 - 4\vartheta}}{2\vartheta} \right)^{1/2}. \tag{6.28}$$

For the values of $[\vartheta = 1/8,\ \zeta = 0]$, which correspond to the stress state of simple shear, we obtain $[\Psi_{1,2} = 3\pi/8,\ \pi/8]$, which apparently coincides with the threshold for the positive definiteness of the function $\mathbf{g}_s^2$, cf. (6.20). In the case of axisymmetric compression or tension where $[\vartheta = 1/6,\ \zeta = \pm 1/3]$, we have $[\Psi_{1,2} \approx 579\pi/1664,\ 253\pi/1664]$, where the rational approximations for $\Psi_{1,2}$ are accurate to six decimal digits.

We return now to the discussion of the sign-definiteness of the expression (6.16) which in the context of (6.22)–(6.24) can be rewritten in the following form

$$\frac{B}{C} < \frac{\chi\bar{\sigma}}{c_\theta} \left(\frac{1 - 2\vartheta(4 - 9\zeta^2)}{\vartheta(4 - 9\zeta^2)} \right). \tag{6.29}$$

From (6.29), after some rearrangement, we can obtain the upper bound for the general plane orientation for a given state of stress within $|\zeta| < 1/3$ for which RH criterion $a_N > 0$ is still satisfied:

$$\vartheta < \left[\frac{c_\theta B}{\chi\bar{\sigma}C}(4 - 9\zeta^2) + (4 - 18\zeta^2) \right]^{-1}. \tag{6.30}$$

RH criterion in the plane of the fastest rate of instability growth ($\vartheta = 1/4$) yields

$$\frac{B}{C} < -\frac{\chi\bar{\sigma}}{c_\theta} \left(\frac{4 - 18\zeta^2}{4 - 9\zeta^2} \right). \tag{6.31}$$

For our representative form of constitutive relation (2.11) the former inequality transforms into

$$g' > -\frac{\chi\bar{\sigma}gh'}{c_\theta h} \left(\frac{4 - 18\zeta^2}{4 - 9\zeta^2} \right) = \frac{\chi\bar{\sigma}g\gamma}{c_\theta h} \left(\frac{4 - 18\zeta^2}{4 - 9\zeta^2} \right), \tag{6.32}$$

where we have used the paticular form of temperature function

$$\begin{aligned} h(\theta) &= 1 - \gamma(\theta - \theta_0), & \theta_{\mathrm{cr}} &= 1/\gamma + \theta_0, \\ h' &= dh/d\theta = -\gamma, & h(\theta_{\mathrm{cr}}) &= 0, \end{aligned} \tag{6.33}$$

where θ_{cr} defines the critical temperature at which $h(\theta_{\mathrm{cr}})$. From (6.32) it follows that the exchange of the stability can take place when the system is still in the strain-hardening regime. The transition depends on the interaction between the current magnitudes of strain hardening and temperature softening.

7. Some general remarks regarding the possibility of bifurcation

7.1. Steady bifurcation (divergence point, $\omega = 0$)

In accordance with the basic definitions outlined in Sec. 5, the necessary conditions for the onset of bifurcation through $\omega = 0$ are that the coefficients a_N and a_{N-1} vanish simultaneously:

$$a_N = a_{N-1} = 0. \tag{7.1}$$

We are interested in ivestigating the physical relevance of the solutions satisfying Eq. (7.1). From the algebraic structure of the coefficient a_N it follows that there are six possible scenarios which eventually may satisfy Eq. (7.1). We do not impose any restrictions on the behaviour of the material parameters A, B, and C at this point. This in turn may lead to a somehow exotic constiutive behaviour. Nevertheless, particular cases of such behaviour which are not compliant with the starting assumptions regarding individual material properties, like strain-rate sensitivity, strain and thermal softening, can be easily eliminated. Here the discussion is referred to the general case of constitutive relation.

SCENARIO 1. $\mathbf{t}_s^2 \neq 0$, $\mathbf{g}_s^2 \neq 0$

Nontrivial values of stress functions $\mathbf{t}_s^2$ and $\mathbf{g}_s^2$ lead to the system of two linear equations for material parameters B and C. Solving (7.1) yields the critical bifurcation pair expressed in terms of the parameter A

$$\begin{aligned}
B_1 &= -\frac{3\mu \mathbf{g}_s^2}{\overline{\sigma}^2}\left[\frac{\left(A_1\dfrac{\mathbf{f}_s^2}{\mathbf{t}_s^2}+3\dfrac{\dot{\overline{\varepsilon}}}{\overline{\sigma}}\right)}{3\left(\dfrac{\beta}{kc_T}\right)^2-9\widetilde{\nu}\dfrac{\mathbf{t}_n^2}{\mathbf{t}_s^2}-\dfrac{\mathbf{f}_s^2}{\mathbf{t}_s^2}}\right], \\
C_1 &= -\frac{9c_\theta \mu \mathbf{t}_s^2}{\chi\overline{\sigma}^3}\left[\frac{\left(A_1\dfrac{\mathbf{f}_s^2}{\mathbf{t}_s^2}+3\dfrac{\dot{\overline{\varepsilon}}}{\overline{\sigma}}\right)}{3\left(\dfrac{\beta}{kc_T}\right)^2-9\widetilde{\nu}\dfrac{\mathbf{t}_n^2}{\mathbf{t}_s^2}-\dfrac{\mathbf{f}_s^2}{\mathbf{t}_s^2}}\right], \\
\widetilde{\nu} &= \frac{1-2\nu}{2(1+\nu)},
\end{aligned} \tag{7.2}$$

where ν represents Poisson's ratio and $c_T = \sqrt{\mu/\rho}$ is the shear wave speed. Since parameters $[A, B, C]$ evolve on a time scale, we cannot control them directly. So we reverse the task, and for a given triplet $[A, B, C]$ seek: a) the stress state within the plane of the fastest rate of instability growth which can actually bifurcate; and b) the pane orientation for a given state of stress which satisfies this bifurcation

scenario. For the first part we get from (7.2), upon using ($\vartheta = 1/4$), the relation

$$\frac{B_1}{C_1} = \frac{\chi\overline{\sigma}}{c_\theta}\left(\frac{\mathbf{g}_s^2}{3\mathbf{t}_s^2}\right) \quad \Rightarrow \quad \zeta = \pm\frac{2}{3}\sqrt{\frac{\varpi+1}{\varpi+2}}, \qquad \varpi = \frac{c_\theta}{\chi\overline{\sigma}}\left(\frac{B_1}{C_1}\right), \tag{7.3}$$

which, together with the condition $|\zeta| \leq 1/3$, leads to the following lower and and upper bound for the characteristic ϖ: $(-1 \leq \varpi \leq -2/3)$. Whenever the current value of ϖ is within the bounds of the interval $\varpi \in [-1, -2/3]$, the value of the stress parameter ζ will be real, and will remain within the bounds of $\zeta \in [-1/3, 1/3]$. Consequently, we may anticipate that in the plane of fastest rate of instability growth there exist physically admissible stress states which can actually bifurcate according to this scenario. The interval of admissibility for the magnitude of the parameter ϖ helps us to estimate the lower and upper bounds for the allowable values of the material parameters:

$$-\frac{\chi\overline{\sigma}}{c_\theta} \leq \frac{B_1}{C_1} \leq -\frac{2\chi\overline{\sigma}}{3c_\theta}, \tag{7.4}$$

which in a transformed form, upon using our representative form of constitutive relation together with particular choice of temperature function, yields

$$\frac{\chi\overline{\sigma}g\gamma}{c_\theta h} \geq g' \geq \frac{2\chi\overline{\sigma}g\gamma}{3c_\theta h}. \tag{7.5}$$

The last inequality shows that the bifurcation within the plane of the fastest rate of instability growth sets in when the system is still in the strain-hardening regime.

To determine the plane orientation in which, for a given state of stress, the system can actually bifurcate, we fix ζ and from (7.3) obtain

$$\vartheta = \left[(4 - 9\zeta^2)(2 + \varpi)\right]^{-1}, \qquad 0 \leq \vartheta \leq 1/4. \tag{7.6}$$

From (7.6), upon imposing the interval of admissibility for the plane inclination parameter $\vartheta \in [0, 1/4]$, follows the lower bound of the ratio

$$\frac{B_1}{C_1} \geq -\frac{\chi\overline{\sigma}}{c_\theta}\left(\frac{4 - 18\zeta^2}{4 - 9\zeta^2}\right), \tag{7.7}$$

which again, specialized for the representative form of constitutive relation, yields the upper bound estimate for the slope of the strain hardening/softening function:

$$g' \leq \frac{\chi\overline{\sigma}\gamma g}{c_\theta h}\left(\frac{4 - 18\zeta^2}{4 - 9\zeta^2}\right). \tag{7.8}$$

SCENARIO 2. $B = 0,\ C = 0$

Here we analyse the possibilities emerging from the case when both material parameters B and C vanish simultaneously. This scenario is of course precluded for the representative form of constitutive relation with the choice of the temperature function which gives $h' = -\gamma$, thus leading to $C \neq 0$. Nevertheless, for the sake of completeness, we give the results of analysis in the sequel. From (7.1) we obtain

$$A_2 = -\frac{\dot{\bar{\varepsilon}}}{\bar{\sigma}}\left(\frac{3\mathbf{t}_s^2}{\mathbf{f}_s^2}\right). \tag{7.9}$$

But since for this particular selection of the bifurcation pair $[B = C = 0]$ we acquire additional singularity in the denominator, i.e, the free coefficient in the denominator of the dispersion relation vanishes ($c_2 = 0$), hence in addition to Eq. (7.1) we must require $a_{N-2} = 0$ to remove the resulting singularity, which after some algebraic simplification leads to

$$3\left(\frac{\beta}{kc_T}\right)^2 \frac{\mathbf{t}_s^2}{\mathbf{f}_s^2} - 9\tilde{\nu}\frac{\mathbf{t}_n^2}{\mathbf{f}_s^2} - 1 = 0. \tag{7.10}$$

Having taken advantage of the parameters $(\zeta, \bar{\sigma})$, Eq. (7.10) being solved for ζ in the plane with $\vartheta = 1/4$, we obtain

$$\begin{aligned} \zeta &= \pm\frac{2\beta}{3kc_T}\left[\left(\frac{\beta}{kc_T}\right)^2 + \frac{3}{2(1+\nu)}\right]^{-1/2}, \\ c_T &= \sqrt{\frac{\mu}{\rho}}, \qquad \zeta \in [-1/3, 1/3]. \end{aligned} \tag{7.11}$$

Of course, the value of ζ must still reside within the range ($|\zeta| \leq 1/3$) to be consistent with the starting assumption about the initial orientation of the instability plane. For $\zeta = 0$, which corresponds to the stress state of simple shear, it follows that the only way in which (7.11) can be satisfied is the case of quasi-static perturbation ($\rho = 0$). For other stress states (7.11) can be checked accordingly.

SCENARIO 3. $B = 0,\ \mathbf{g}_s^2 = 0$

Under this scenario we seek the critical conditions for the onset of steady bifurcation when $B = \partial\,\dot{\bar{\varepsilon}}/\partial\bar{\varepsilon} = 0$ within the plane of isothermal invariance of adiabatic material instablity process. Hence, we obtain

$$C_3 = \frac{3c_\theta\mu}{2\chi\bar{\sigma}}\frac{(A_3 + \dot{\bar{\varepsilon}}/\bar{\sigma})}{\left[1 + 18\tilde{\nu}\dfrac{\mathbf{t}_n^2}{\bar{\sigma}^2} - \left(\dfrac{\beta}{c_T k}\right)^2\right]}, \tag{7.12}$$

which can be rewritten in the following form

$$\frac{\mathbf{t}_n^2}{\overline{\sigma}^2} = \left(\frac{\zeta}{2} \mp \frac{1}{2}\sqrt{\frac{2}{3} - 3\zeta^2}\right)^2 = \frac{1}{18\widetilde{\nu}}\left[\frac{3\mu c_\theta}{2\chi\overline{\sigma}}\left(\frac{A + \dot{\overline{\varepsilon}}/\overline{\sigma}}{C_3}\right) + \left(\frac{\beta}{c_T k}\right)^2 - 1\right] = \varpi. \tag{7.13}$$

From Eq. (7.13) we can obtain the lower and the upper bound for the parameter ϖ in the interval $|\zeta| \leq 1/3$, namely $\varpi \in [(1/6 - \sqrt{3}/6)^2, 2/9]$. The lower bound is attained at $\zeta_L^2 = 1/9$, and the upper bound is at $\zeta_U^2 = 1/18$. This corresponds to:

$$\begin{aligned} \zeta_L &= \pm\frac{1}{3}, \quad \vartheta_L = \frac{1}{6}, \quad \Psi_{L1,2} \approx \frac{579\pi}{1664}, \frac{253\pi}{1664}, \\ \zeta_U &= \pm\frac{\sqrt{2}}{6}, \quad \vartheta_U = \frac{1}{7}, \quad \Psi_{U1,2} \approx \frac{757\pi}{2082}, \frac{284\pi}{2082}. \end{aligned} \tag{7.14}$$

Here Ψ defines the angle of bifurcation plane inclined with respect to the principal direction of the maximum stress deviator. From the lower and the upper bound for ϖ we derive the inequality

$$1 + \left(\frac{1-\sqrt{3}}{2}\right)^2\left(\frac{1-2\nu}{1+\nu}\right) \leq \frac{3\mu c_\theta}{2\chi\overline{\sigma}}\left(\frac{A_3 + \dot{\overline{\varepsilon}}/\overline{\sigma}}{C_3}\right) + \left(\frac{\beta}{c_T k}\right)^2 \leq \frac{3(1-\nu)}{(1+\nu)}, \tag{7.15}$$

which, specialized for our representative form of constitutive relation, yields

$$1 + \left(\frac{1-\sqrt{3}}{2}\right)^2\left(\frac{1-2\nu}{1+\nu}\right) \leq \frac{3\mu c_\theta(1+m)(\theta_{\mathrm{cr}} - \theta)}{2\chi\overline{\sigma}^2} + \left(\frac{\beta}{c_T k}\right)^2 \leq \frac{3(1-\nu)}{(1+\nu)}. \tag{7.16}$$

For $\nu = 1/2$ the inequalities (7.15) – (7.16) transform into equalities. In the case of quasi-perturbation, the second term in the middle of (7.16) drops out ($\beta/c_T k = 0$) and we are able to make an estimate of the order of the magnitude of the first term in the middle of (7.16). It can be shown that the critical value of the temperature $\theta = \theta_c$ at the onset of bifurcation scenario is quite close to the critical temperature θ_{cr} at which the temperature softening function vanishes, i.e. $h(\theta_{\mathrm{cr}}) = 0$, which leads to the conclusion that this is a rather improbable scenario.

SCENARIO 4. $C = 0$, $\mathbf{t}_s^2 = 0$

The critical condition for the onset of steady bifurcation when $C = \partial\,\dot{\overline{\varepsilon}}/\partial\theta = 0$ within the plane where $\mathbf{t}_s^2 = 0$, which obviously leads to the first degenerate case

of dispersive relation, is given as

$$B_4 = \frac{3\mu A_4}{1 + 9\tilde{\nu}\mathbf{t}_n^2/\bar{\sigma}^2}. \tag{7.17}$$

Again, the scenario is precluded for particular choice of temperature function $h' = -\gamma$, which apparently leads to $C \neq 0$. Moreover, the inspection of the free term in the denominator of the dispersion relation reveals that the term vanishes at $[B_4, C = 0, \mathbf{t}_s^2 = 0]$, i.e. $c_2 = 0$, resulting in a singularity which can be removed by setting $a_{N-2} = 0$, from where after some rearrangements we obtain the supplemented form of the critical condition

$$A_4 = -\frac{\bar{\sigma}\,\dot{\bar{\varepsilon}}}{9\tilde{\nu}^2\mathbf{t}_n^2}\left[\frac{\left(1 + 9\tilde{\nu}\dfrac{\mathbf{t}_n^2}{\bar{\sigma}^2}\right)^2}{\left(4 - 9\dfrac{\mathbf{t}_n^2}{\bar{\sigma}^2}\right)}\right], \qquad B_4 = -\frac{\beta\bar{\sigma}^2}{9\tilde{\nu}^2\mathbf{t}_n^2}\left[\frac{\left(1 + 9\tilde{\nu}\dfrac{\mathbf{t}_n^2}{\bar{\sigma}^2}\right)}{\left(4 - 9\dfrac{\mathbf{t}_n^2}{\bar{\sigma}^2}\right)}\right]. \tag{7.18}$$

To determine the stress state which can bifurcate within the plane $\mathbf{t}_s^2 = 0$, rewrite (7.18) in the form similar to (7.13) with $\vartheta = 0$:

$$\frac{\mathbf{t}_n^2}{\bar{\sigma}^2} = \left(\frac{\zeta}{2} \mp \sqrt{\frac{1}{3} - \frac{3}{4}\zeta^2}\right)^2 = \frac{1}{9\tilde{\nu}}\left(\frac{3\mu A_4}{B_4} - 1\right) = \varpi. \tag{7.19}$$

We solve it for ζ in terms of the characteristic parameter ϖ and obtain

$$\zeta_{1,2}^2 = \frac{1}{2}\left(\frac{2}{3} - \varpi \pm \sqrt{\frac{\varpi}{3}(4 - 9\varpi)}\right). \tag{7.20}$$

From Eq. (7.20) follows the lower and upper bound of ϖ which still renders $|\zeta| \leq 1/3$. We do not go into details of derivation here, let it suffice to say that the interval of admissibility for the parameter ϖ is given by $\varpi \in [1/9, 4/9]$, from where the upper and lower bounds for the ratio of the critical values of material parameters A_4 and B_4 are readily obtained

$$2\mu(1 + \nu) \geq \frac{B_4}{A_4} \geq \mu\frac{(1 + \nu)}{(1 - \nu)}. \tag{7.21}$$

For $\nu = 1/2$ the former inequality reduces to equality: $B_4 = 3\mu A_4$. For a representative form of constitutive relation (7.21) yields

$$-2\mu(1 + \nu)\frac{g}{\bar{\sigma}} \leq g' \leq -\mu\frac{(1 + \nu)g}{(1 - \nu)\bar{\sigma}}, \tag{7.22}$$

which defines a narrow zone of negative slope in the $g = g(\bar{\varepsilon})$ curve which is almost vertical. Consequently, the bifurcation scenario leads to the critical condition with regard to the strain softening function which is very restraining.

SCENARIO 5. $B = 0,\ C = 0,\ \mathbf{t}_s^2 = 0$

A very restrictive scenario from which we cannot expect much. Indeed, the solution of (7.1) upon using $[B = 0, C = 0, \mathbf{t}_s^2 = 0]$ yields $A_5 = 0$, and the critical bifurcation triplet $[A_5 = B_5 = C_5 = 0]$ generates additional singularity of order 2 in the denominator of the dispersion relation where the last two coefficients vanish, $c_1 = c_2 = 0$. Hence, in addition to (7.1), we must require $a_{N-2} = a_{N-3} = 0$, to remove the resulting singularity. A thorough inspection of the algebraic structure of both coefficients shows that in general the coefficients cannot vanish. Hence, the bifurcation under this scenario is precluded.

SCENARIO 6. $B = 0,\ C = 0,\ \mathbf{g}_s^2 = 0$

Here the solution of (7.1) leads to the bifurcation triplet $[A_6 = -\dot{\bar{\varepsilon}}/\bar{\sigma}, 0, 0]$ which apparently cannot be satisfied for the representative form of constitutive relation (2.11), with the additional singularity in the denominator due to $c_2 = 0$. The supplementary condition $a_{N-2} = 0$ leads to the constraint

$$1 + 18\tilde{\nu}\frac{\mathbf{t}_n^2}{\bar{\sigma}^2} - \left(\frac{\beta}{c_T k}\right)^2 = 0. \tag{7.23}$$

In the spirit of previous derivation, we rewrite (7.23), using the relation between the plane orientation parameter ϑ and the stress state ζ which are now constrained by $\mathbf{g}_s^2 = 0$, in the following form:

$$\frac{\mathbf{t}_n^2}{\bar{\sigma}^2} = \left(\frac{\zeta}{2} \mp \frac{1}{2}\sqrt{\frac{2}{3} - 3\zeta^2}\right)^2 = \frac{1}{18\tilde{\nu}}\left[\left(\frac{\beta}{c_T k}\right)^2 - 1\right] = \varpi. \tag{7.24}$$

Equation (7.24) has a similar structure as (7.13), with the value of the characteristic parameter ϖ which is obviously different. Hence, we may use the interval of admissibility of the characteristic parameter $\varpi \in [(1/6 - \sqrt{3}/6)^2, 2/9]$ derived under the scenario 3, from where we obtain, by using the same reasoning as before, the following inequality:

$$\sqrt{1 + \left(\frac{1-\sqrt{3}}{2}\right)^2 \left(\frac{1-2\nu}{1+2\nu}\right)} \leq \frac{\beta}{c_T k} \leq \sqrt{\frac{3(1-\nu)}{(1+\nu)}}. \tag{7.25}$$

The former inequality may prove to be useful in getting the estimate for the lower and upper bound of the wave number which still satisfies the bifurcation criteria imposed by this scenario. In the case of quasi-static perturbation (7.32) cannot be satisfied and the steady bifurcation is precluded. Finally, in the case of elastic incompressibility the former inequality degenerates into equality $\beta = c_T k$.

7.2. Hopf bifurcation ($\mathrm{Re}[\omega] = 0, \quad \mathrm{Im}[\omega \neq 0]$)

The theoretical background and the necessary conditions for the onset of Hopf bifurcation were elaborated in Sec. 5. In general, however, the expressions which define the critical conditions for the onset of HB can be quite complicated. We will try to get some insight into the stability behaviour of the system for the two degenerate cases of the dispersion relation. The crucial part of the dispersion relation that governs the stability behaviour of the system for the first degenerate case with ($\mathbf{t}_s^2 = 0$) reduces to

$$D_{R_1} = \left(\frac{\rho\omega}{k^2} + q_0 + q_1 + q_2\mathbf{t}^2\right)\Big|_{\omega=\pm is} = 0. \tag{7.26}$$

In the expanded form (7.26) yields the third and fifth order polynomials in ω depending on the type of perturbation, quasi-static or dynamic, respectively

$$D_{R_1} = \frac{1}{d}\sum_{k=0}^{N} a_k\omega^{N-k}, \qquad (N = 3,5), \qquad d = \omega(\beta+\omega)(c_0\omega^2+c_1\omega+c_2). \tag{7.27}$$

In the expression above, the coefficients c_0, c_1, c_2 remain the same as in the general case of dispersion relation, cf. (4.17). The coefficients a_k are given in Tables 3 and 4 in the Appendix. The Hopf criteria thus reduce to

$$\begin{aligned} &H_3 = a_1a_2 - a_0a_3 = 0, \qquad s^2 > 0, \\ &H_5 = a_1a_2a_3a_4 + a_0a_2a_3a_5 + 2a_0a_1a_4a_5 - a_0^2a_5^2 - a_1^2a_4^2 \\ &\qquad\qquad - a_0a_3^2a_4 - a_1a_2^2a_5, \qquad s^2 > 0. \end{aligned} \tag{7.28}$$

A derivation stands for $\mathbf{t}_s^2 = 0$, i.e., for the first degenerate case of the dispersion relation: classes 7, 8, 9 and subclass A as defined in Table 7. In the case of simple shear we check the possibilities emerging from the second degenerate case of the dispersion relation defined by $D(\mathbf{t}_n^2 = 0) = 0$. The equation yields

$$D_{R_2} = \left(\frac{\rho\omega}{k^2} + q_0 + q_2\mathbf{t}^2\right)\Big|_{\omega=\pm is} = 0. \tag{7.29}$$

In the expanded form (7.29) yields the second and the fourth order polynomials on ω for quasi-static or dynamic perturbation, respectively:

$$D_{R_2} = \frac{1}{d}\sum_{k=0}^{N} a_k\omega^{N-k}, \qquad (N = 2,4), \qquad d = (\beta+\omega)(c_0\omega^2+c_1\omega+c_2). \tag{7.30}$$

Here again, the coefficients c_0, c_1, c_2 remain the same as before, and the coefficients a_k are given in Tables 5 and 6 in the Appendix. The corresponding Hopf criteria are

$$H_2 = a_1 = 0, \qquad H_4 = a_1a_2a_3 - a_0a_3^2 - a_1^2a_4 = 0, \qquad s^2 > 0. \tag{7.31}$$

From the algebraic structure of the coefficients a_k is evident that H_N can be written as a polynomial in material parameters (A, B, C):

$$H_N = \sum_{K=0}^{\hat{K}} \sum_{L=0}^{\hat{L}} \sum_{M=0}^{\hat{M}} X_{KLM} A^K B^L C^M = 0, \qquad (N = 2,3,4,5),$$
$$(\hat{K} = N - 1, \qquad \hat{L} = \hat{K} - K, \qquad \hat{M} = \hat{L} - L). \tag{7.32}$$

The last expression can be considered as an interaction formula between the material parameters A, B and C. If $H_N = 0$ holds for a certain triplet of $[A_h, B_h, C_h]$ providing that $s^2(A_h, B_h, C_h) > 0$, then we may anticipate the onset of Hopf bifurcation.

In general, however, the expression for H_3 and H_5 can be quite complicated. Even if we specialize the analysis to the specific type of stress state like uniaxial or simple shear, due to the algebraic complexity of the coefficients it is difficult to establish the analytical criteria for the onset of HB following from the condition (7.32). One solution to this problem may be a parametric study of different types of constitutive relations, where one would try to track down numerically whether the criterion can be satisfied for various stress and deformation histories of the body strained beyond the limit point of reversible deformation. We will try to get some insight into a qualitative behaviour of the criterion in the next section, where by neglecting the effects of elasticity, we make the transition to a rigid-viscoplastic constitutive behaviour. Consequently, we obtain the dispersion relation with coefficients which are much simpler and more tractable for analytical considerations. In what follows, we first examine the critical conditions for the onset of HB in the case of uniaxial tension and quasi-static perturbation. For the first degenerate case of dispersion relation with $\mathbf{t}_s^2 = 0$, the stress state of uniaxial tension yields the Hopf criterion of the form

$$H_3 = \mathbf{X}^T \cdot \mathbf{A} \cdot \mathbf{X} + \mathbf{X}^T \cdot \mathbf{b} + X_{000} = 0, \tag{7.33}$$

$$\mathbf{X}^T = [A, B, C], \qquad \mathbf{A} = \begin{bmatrix} X_{200} & X_{110}/2 & X_{101}/2 \\ X_{110}/2 & X_{020} & X_{011}/2 \\ X_{101}/2 & X_{011}/2 & X_{002} \end{bmatrix}, \qquad \mathbf{b}^T = [X_{100}, X_{010}, X_{001}].$$

The expression above defines a quadratic surface in material parameters $[A, B, C]$ considered as independent variables. From (7.33) we can express B in terms of A and C and solve the resulting quadratic equation, which in turn yields the corresponding pair of solutions of the form $B = B(A, C)$,

$$\hat{X}_{020} B^2 + \hat{X}_{010} B + \hat{X}_{000} = 0,$$
$$\hat{X}_{020} = X_{020}, \qquad \hat{X}_{010} = X_{010} + X_{110} A + X_{011} C, \tag{7.34}$$
$$\hat{X}_{000} = X_{000} + X_{100} A + X_{001} C + X_{200} A^2 + X_{101} AC + X_{002} C^2,$$

$$B_{1,2} = -\frac{\hat{X}_{010}}{2\hat{X}_{020}}\left(1 \pm \sqrt{1 - \frac{4\hat{X}_{020}\hat{X}_{000}}{\hat{X}_{010}^2}}\right). \tag{7.35}$$

In Eq. (7.35) the first solution gives a negative value of the square of the Hopf frequency, i.e., $s^2(B_1) < 0$. Hence, HB is precluded and we focus on the solution

$$B_2 = \mu\left(\frac{1+\nu}{1-\nu}\right)A - \frac{\chi\bar{\sigma}}{c_\theta}C, \tag{7.36}$$

$$s^2(B_2) = \left[\frac{\chi\,\dot{\bar{\varepsilon}}\,\mu}{c_\theta}\left(\frac{1+\nu}{1-\nu}\right)\right]C = \frac{\chi\,\dot{\bar{\varepsilon}}\,c_T^2}{c_p}\left(\frac{1+\nu}{1-\nu}\right)C > 0. \tag{7.37}$$

From last inequality it follows that the value of material parameter C must be positive in order to have $s^2 > 0$. For our representative form of constitutive relation, cf. (2.11), (6.33), the Hopf frequency and the corresponding critical value of the slope in the strain hardening/softening function are given by

$$s = \dot{\bar{\varepsilon}}\,c_T\sqrt{\frac{\chi\gamma}{mc_p[1-\gamma(\theta-\theta_0)]}\left(\frac{1+\nu}{1-\nu}\right)}, \tag{7.38}$$

$$g_h' = \left[\frac{\chi\bar{\sigma}\gamma}{c_\theta[1-\gamma(\theta-\theta_0)]} - \left(\frac{1+\nu}{1-\nu}\right)\frac{\mu}{\bar{\sigma}}\right]g(\bar{\varepsilon}). \tag{7.39}$$

Theoretically, g_h' can take positive (strain hardening) and negative (strain softening) values, which depends on the current values of $\bar{\sigma}$ and θ at the onset of HB:

$$g_h'\begin{cases} \geq 0 \quad \text{if} \quad \theta \geq \theta_{\text{cr}} - \chi\left(\dfrac{1-\nu}{1+\nu}\right)\dfrac{\bar{\sigma}^2}{c_\theta\mu}, \\[2ex] < 0 \quad \text{if} \quad \theta < \theta_{\text{cr}} - \chi\left(\dfrac{1-\nu}{1+\nu}\right)\dfrac{\bar{\sigma}^2}{c_\theta\mu}, \end{cases} \qquad \theta_{\text{cr}} = \frac{1}{\gamma} + \theta_0. \tag{7.40}$$

The second degenerate case of dispersion relation, when the perturbation is whithout the influence of inertia and specialized for the stress state of simple shear, leads to the result $s^2 = 0$, which clearly excludes the possibility of HB. If on the other hand, the inertia is accounted for, then the condition $\mathbf{t}_n^2 = 0$ brings additional simplification, namely the order of polynomials in material parameters A and B reduce by one, and we obtain the quadratic equation for the parameter B of the same form as in (7.34), but with different coefficients $\hat{X}_{020}, \hat{X}_{010}$ and $\hat{X}_{000}$. We discard the solution which gives the negative value of the square of the Hopf frequency, i.e., $s^2(B_2) =< 0$.. The other solution yields

$$B_1 = 3\mu A\left(1 + \frac{c_\theta k^2}{3\chi\,\dot{\bar{\varepsilon}}\,\rho C}\right) - \frac{\chi\bar{\sigma}}{c_\theta}C, \qquad s(B_1) = c_T k\sqrt{1 + \frac{3\chi\,\dot{\bar{\varepsilon}}\,\rho}{c_\theta k^2}C}. \tag{7.41}$$

The RH criterion ($a_N > 0$) specialized to the stress state of simple shear yields

$$B + \chi\bar{\sigma}C/c_\theta < 0. \tag{7.42}$$

But from (7.41) we obtain

$$B + \frac{\chi\bar{\sigma}}{c_\theta}C = 3\mu\left(1 + \frac{c_\theta k^2}{3\chi\dot{\bar{\varepsilon}}\rho C}\right)A < 0, \tag{7.43}$$

which can be satisfied without violating $s^2 > 0$ only if

$$(A < 0,\ C > 0) \quad \text{or} \quad \left(A > 0,\ -\frac{c_\theta k^2}{3\chi\rho\dot{\bar{\varepsilon}}} < C < 0\right). \tag{7.44}$$

Hence, the HB may set in only when the values of material parameters A and C are of the opposite sign. In lieu of original relation $\dot{\bar{\varepsilon}}=\dot{\bar{\varepsilon}}\ (\bar{\sigma}, \bar{\varepsilon}, \theta)$, by employing its inverse $\bar{\sigma} = \bar{\sigma}(\bar{\varepsilon}, \dot{\bar{\varepsilon}}, \theta)$, we can see that (7.44) for $A = \partial\dot{\bar{\varepsilon}}/\partial\bar{\sigma} < 0$ and $C = \partial\dot{\bar{\varepsilon}}/\partial\theta > 0$ requires $\partial\bar{\sigma}/\partial\dot{\bar{\varepsilon}} < 0$ and $\partial\bar{\sigma}/\partial\theta > 0$. Hence, the material must be temperature-hardening and strain-rate softening. On the other hand, if $A > 0$ and $C < 0$, then $\partial\bar{\sigma}/\partial\theta > 0$ still requires temperature hardening, and $\partial\bar{\sigma}/\partial\dot{\bar{\varepsilon}} > 0$ leads to strain-rate hardening as well. This clearly eliminates the possibility of HB not just for our representative form of constitutive relation, but for all materials which do not allow temperature hardening. Another intricacy related to (7.41) which merits additional comment can be observed by noting that for $C > 0$, the second relation defined in (7.41) yields the speed of instability wave front which is greater than that of the shear wave. Even in the case when the material is strain-rate hardening, it also has to be in the temperature hardening regime. In a physical sense, the former restrictions may lead to quite a relaxed constitutive behaviour.

8. Transition to the rigid-viscoplastic limit

A constitutive equation representing an elastic-viscoplastic material describes a rather general class of material behaviour. This comprehensive law permits to obtain or recover, with the transition to appropriate limit, various other constitutive models. For example, neglecting the elastic portion of the strain rate tensor in the limit yields a response of a rigid viscoplastic material. In this case the dispersion relation and its coefficients simplify significantly:

$$D(\omega, \mathbf{k}) = \frac{1}{d}(a_0\omega^2 + a_1\omega + a_2) = 0, \qquad \Xi = \text{const} = -1, \tag{8.1}$$

$$a_0 = 3Ac_\theta \left(\overline{\sigma}\,\dot{\overline{\varepsilon}}\right)^2 \rho/k^2, \qquad \mathbf{f}_s^2 = \overline{\sigma}^2 - 3\mathbf{t}_s^2, \qquad \mathbf{g}_s^2 = \overline{\sigma}^2 - 6\mathbf{t}_s^2,$$
$$a_1 = Ac_\theta \overline{\sigma}\,\dot{\overline{\varepsilon}}\,\mathbf{f}_s^2 + 3c_\theta\,\dot{\overline{\varepsilon}}^2\,\mathbf{t}_s^2 + 3\chi\rho C\overline{\sigma}^2\,\dot{\overline{\varepsilon}}^3/k^2,$$
$$a_2 = \chi C\overline{\sigma}\,\dot{\overline{\varepsilon}}^2\,\mathbf{g}_s^2 - 3Bc_\theta\,\dot{\overline{\varepsilon}}^2\,\mathbf{t}_s^2, \qquad d = \omega\left(3\overline{\sigma}\,\dot{\overline{\varepsilon}}\right)^2\left(C\chi\,\dot{\overline{\varepsilon}} + Ac_\theta\omega\right).$$

Since Ξ is constant throughout the loading process, the direction of instability plane remains fixed relative to its initial position. For various stress states it can be simply recovered from the Table 7 by using $\Xi = -1$.

8.1. Steady bifurcation (divergence point, $\omega = 0$)

From the conditions $a_1 = a_2 = 0$ which lead to a linear system of equations for two unknowns B and C expressed in terms of the material parameter A, upon solving the resulting set of equations, we obtain

$$B_c = -\frac{\mathbf{g}_s^2 k^2}{3\rho\overline{\sigma}\,\dot{\overline{\varepsilon}}}\left(1 + \frac{\overline{\sigma}\mathbf{f}_s^2 A}{3\,\dot{\overline{\varepsilon}}\,\mathbf{t}_s^2}\right), \qquad C_c = -\frac{c_\theta \mathbf{t}_s^2 k^2}{\chi\rho\overline{\sigma}^2\,\dot{\overline{\varepsilon}}}\left(1 + \frac{\overline{\sigma}\mathbf{f}_s^2 A}{3\,\dot{\overline{\varepsilon}}\,\mathbf{t}_s^2}\right). \tag{8.2}$$

The result corresponds to the case of dynamic perturbation, where inertia is retained in the leading coefficients defined in (8.1). When the peturbation is quasi-static, equations transform accordingly, and we obtain two independent solutions

$$A_c = -\frac{3\,\dot{\overline{\varepsilon}}\,\mathbf{t}_s^2}{\overline{\sigma}\mathbf{f}_s^2}, \qquad B_c = \frac{\chi\overline{\sigma}\mathbf{g}_s^2}{3c_\theta\mathbf{t}_s^2}C, \tag{8.3}$$

where the last expression which defines the relationship between the material parameters B and C also holds true for the critical values of parameters defined in (8.2). Since the critical value of the material parameter A, according to Eq. (8.3), is always negative in the viscoplastic domain, we may conclude that the steady bifurcation is precluded for the representative form of constitutive relation. On the other hand, the solutions in (8.2) for our representative form of constitutive relation yield

$$g_c' = \frac{\mathbf{g}_s^2 k^2 g}{3\rho\overline{\sigma}\,\dot{\overline{\varepsilon}}^2}\left(m + \frac{\mathbf{f}_s^2}{3\mathbf{t}_s^2}\right), \qquad h_c' = \frac{c_\theta \mathbf{t}_s^2 k^2 h}{\chi\rho(\overline{\sigma}\,\dot{\overline{\varepsilon}})^2}\left(m + \frac{\mathbf{f}_s^2}{3\mathbf{t}_s^2}\right), \tag{8.4}$$

from where it follows, having considered the particular form of temperature softening function (6.33), that the steady bifurcation is precluded for this particular choice of temperature function, because the second equation in (8.4) leads to the contradiction; the right-hand side of the expression is in general positive, meanwhile the left-hand side takes on a negative value, i.e. $h' = -\gamma$, representing the constant negative slope in the temperature softening function $h = h(\theta)$.

8.2. Hopf bifurcation ($\mathrm{Re}[\omega] = 0$, $\mathrm{Im}[\omega \neq 0]$)

The Hopf criterion together with its intrinsic frequency of instability reduces to

$$H_2 = Ac_\theta \bar{\sigma}\, \dot{\bar{\varepsilon}}\, \mathbf{f}_s^2 + 3C\chi(\bar{\sigma}\, \dot{\bar{\varepsilon}})^2\, \dot{\bar{\varepsilon}}\, \rho/k^2 + 3c_\theta\, \dot{\bar{\varepsilon}}^2\, \mathbf{t}_s^2 = 0,$$

$$s^2 = a_2/a_0 = \frac{C\chi\, \dot{\bar{\varepsilon}}^2\, \bar{\sigma}\mathbf{g}_s^2 - 3Bc_\theta\, \dot{\bar{\varepsilon}}\, \mathbf{t}_s^2}{3Ac_\theta(\bar{\sigma}\, \dot{\bar{\varepsilon}})^2\rho/k^2} > 0. \tag{8.5}$$

For the first degenerate case of dispersion relation corresponding to $\mathbf{t}_s^2 = 0$, the Hopf bifurcation is not possible, since from (8.5) we get $s^2(A_h, B_h, C_h) < 0$. The second degenerate case of dispersion relation where $\mathbf{t}_s^2 = \mathbf{t}^2$, $\mathbf{t}_n^2 = 0$, and $\mathbf{t}^2 = \mathbf{n} \cdot \boldsymbol{\tau}' \cdot \boldsymbol{\tau}' \cdot \mathbf{n}$ retains the algebraic structure of the criteria given above, with $\mathbf{t}_s^2$ replaced by $\mathbf{t}^2$. For a general case from (8.5) follows

$$C_h = -\frac{c_\theta \mathbf{t}_s^2 k^2}{3\chi\rho\bar{\sigma}^2\, \dot{\bar{\varepsilon}}}\left(1 + \frac{\mathbf{f}_s^2}{3\mathbf{t}_s^2}\frac{\bar{\sigma}}{\dot{\bar{\varepsilon}}}A\right),$$

$$s^2 = -\frac{\mathbf{t}_s^2 k^2}{\rho\bar{\sigma}^2}\left[\frac{B}{A} + \frac{\mathbf{g}_s^2 k^2}{3\rho\, \dot{\bar{\varepsilon}}^2}\left(\frac{\mathbf{f}_s^2}{3\mathbf{t}_s^2} + \frac{\dot{\bar{\varepsilon}}}{\bar{\sigma}A}\right)\right] > 0. \tag{8.6}$$

The possibility of HB here depends on the particular choice of constitutive functions. In general, the conditions for the onset of HB will be satisfied when

$$A > 0, \qquad B < -\frac{\mathbf{g}_s^2 k^2}{3\rho\, \dot{\bar{\varepsilon}}^2}\left(\frac{\mathbf{f}_s^2}{3\mathbf{t}_s^2}A + \frac{\dot{\bar{\varepsilon}}}{\bar{\sigma}}\right). \tag{8.7}$$

For our representative form of constitutive law we have

$$A = \dot{\bar{\varepsilon}}/m\bar{\sigma}, \qquad B = -\,\dot{\bar{\varepsilon}}\, g'/mg, \qquad C = -\,\dot{\bar{\varepsilon}}\, h'/mh, \tag{8.8}$$

where prime indicates the derivative with respect to the argument, i.e., $g' = dg/d\bar{\varepsilon}$ and $h' = dh/d\theta$. After some algebra one obtains

$$h'_h = \frac{c_\theta h \mathbf{t}_s^2 k^2}{\chi\rho(\bar{\sigma}\, \dot{\bar{\varepsilon}})^2}\left(m + \frac{\mathbf{f}_s^2}{3\mathbf{t}_s^2}\right), \qquad s^2 = \frac{\mathbf{t}_s^2 k^2}{\rho\bar{\sigma}g}\left[g' - \frac{\mathbf{g}_s^2 k^2 g}{3\rho\bar{\sigma}\, \dot{\bar{\varepsilon}}^2}\left(\frac{\mathbf{f}_s^2}{3\mathbf{t}_s^2} + m\right)\right] > 0. \tag{8.9}$$

The requirement $s^2 > 0$ leads to the lower bound for the slope in the strain softening function giving

$$g'_h > \frac{\mathbf{g}_s^2 k^2 g}{3\rho\bar{\sigma}\, \dot{\bar{\varepsilon}}^2}\left(m + \frac{\mathbf{f}_s^2}{3\mathbf{t}_s^2}\right). \tag{8.10}$$

The particular choice of the temperature function, as defined in (6.33) together with $h' = -\gamma$, leads to the contradiction of the first relation in (8.9), since the right-hand side of the expression is always positive. This clearly eliminates the possibility of HB for this particular choice of the function $h = h(\theta)$. For all other types of constitutive relation the criteria (8.6) and (8.7) should be checked accordingly.

9. Transition to the rate-independent limit

The main objective of this section is to analyze how the results obtained from the rate-dependent theory compare to the results from the rate-independent theory. The results from the rate-dependent theory are specialized by taking appropriate limits to recover the rate-independent material behaviour. By letting $\dot{\bar{\varepsilon}} \to 0$ as $\omega \to 0$ in the expressions for the coefficients of the dispersion relation given in Tables 1 and 2, after some simplifications we arrive at

$$\lim_{\omega \to 0}(\omega^2 D) = \frac{1}{(3\mu + H)}\left[H + \frac{9\mu}{\bar{\sigma}^2}\left(\frac{\bar{\sigma}^2}{3} - \mathbf{t}^2 + \frac{1}{2(1-\nu)}\mathbf{t}_n^2\right)\right] = 0, \tag{9.1}$$

where H is expressed as a linear combination of the stress hardening/softening and the thermal hardening/softening moduli, as a result of the following two limits

$$\begin{aligned} H &= h + \frac{\chi\bar{\sigma}}{c_\theta}h_\theta, \qquad h \overset{\text{def}}{=} \lim_{\omega \to 0}\left(-\frac{B}{A}\right) = \frac{\partial\bar{\sigma}}{\partial\bar{\varepsilon}}, \\ h_\theta &\overset{\text{def}}{=} \lim_{\omega \to 0}\left(-\frac{C}{A}\right) = \frac{\partial\bar{\sigma}}{\partial\theta}, \qquad \bar{\sigma} = \bar{\sigma}(\bar{\varepsilon}, \theta). \end{aligned} \tag{9.2}$$

From Eq. (9.1) we find that

$$\frac{\bar{\sigma}^2}{3}\left(1 + \frac{H}{3\mu}\right) - \mathbf{t}^2 + \frac{1}{2(1-\nu)}\mathbf{t}_n^2 = 0, \tag{9.3}$$

which relates to the original form of the dispersion relation by identifying

$$d_0 = \frac{\bar{\sigma}^2}{3}\left(\frac{3\mu + H}{3\mu}\right), \qquad d_1 = -1, \qquad d_2 = \frac{1}{2(1-\nu)}, \tag{9.4}$$

$$\Xi = -1/(2(1-\nu)), \qquad m^2 = (1+\nu)/3, \qquad l^2 = (2-\nu)/3. \tag{9.5}$$

Now we are able to use the expressions from the original analysis once again. For $\nu = 1/2$ we have $m^2 = 1/2$, $l^2 = 1/2$ and the same results are obtained as in the rate-dependent case. In general however, we do not recover results from the rate-dependent theory because of the Eqs. (9.5). Here the direction of the normal vector $\mathbf{n}$ depends on ν(!) Only in two special cases results comply with the ones

from the rate-dependent theory: in the case when $\nu = 1/2$ (elastic incompressibility), and in the case of simple shear where $m^2 = l^2 = 1/2$ holds for both. At first it may seem strange that the result for the orientation of the localization plane does not depend on temperature parameters, neither has the thermal softening any influence on the direction of localization plane. The result is essentially the same as in the case of isothermal deformation. The reason for this is that in constitutive equations the volumetric strains due to temperature changes have not been accounted for, and temperature in the energy balance equation remains coupled only through the deviatoric part of viscoplastic dissipation function. The same result probably applies for all plain J_2 flow theories.

The critical value of hardening modulus can be determined from (9.3) and is defined as

$$\frac{H}{3\mu} = \frac{3}{\bar{\sigma}^2}\left(\mathbf{t}^2 - \frac{1}{2(1-\nu)}\mathbf{t}_n^2\right) - 1. \tag{9.6}$$

Equations (9.5) and (9.6) define the critical orientations of the localization plane and the corresponding critical value of combined hardening modulus in the localization problem of elastic-thermoplastic solid based on the J_2 flow theory of plasticity. In the case of uniaxial tension, the critical value of H corresponding to the critical orientation of the localized plane amounts to

$$h_{\mathrm{cr}} + \frac{\chi\bar{\sigma}}{c_\theta} h_{\theta\mathrm{cr}} = -\frac{\mu(1+\nu)}{2}, \tag{9.7}$$

whereas in the case of simple shear one obtains

$$h_{\mathrm{cr}} + \frac{\chi\bar{\sigma}}{c_\theta} h_{\theta\mathrm{cr}} = 0. \tag{9.8}$$

These results agree with those reported by DUSZEK and PERZYNA [4]. By taking the appropriate limits which define strain and temperature hardening/softening moduli, the results for the rate-independent material are readily obtained as a by-product of the linear stability analysis of a general rate-dependent case. A comparison of the results obtained from the rate-independent bifurcation analysis and the perturbation analysis for the rate-dependent materials leads to the conclusion that there is a difference in the initial inclination of the instability plane in the limit as the growth rate tends to zero, $\omega \to 0$. The perturbation analysis in the rate-dependent case defines the initial inclination of the instability plane as an intrinsic property of the stress state independent of a material parameter of any kind. In the rate-independent limit the normal of the plane depends on the value of Poisson's ratio ν, Eqs. (9.5). This may lead to a notable discrepancy in the initial orientation of the instability plane with the exception of the stress state of simple shear, and in the case of elastic incompressibility ($\nu = 1/2$) where the results coincide.

10. Concluding remarks

The analysis of the influence of thermomechanical coupling under adiabatic conditions on two qualitatively different types of material instability based on the linear stability theory has been presented. The analytical structure of the highly nonlinear governing equations allows for instability responses of a very diversified nature. Physical restrictions imposed on the behaviour of material parameters narrow the family of possible instabilities predicted by the theoretical analysis.

The algebraic structure of the Routh-Hurwitz criterion ($a_N > 0$) reveals the existence of the planes of isothermal invariance in an adiabatic material instability process. Within these planes the influence of temperature under adiabatic conditions vanishes completely, and the RHC transforms to a criterion which is similar to the RH criterion in the isothermal case. The spatial orientation of these planes differs from the plane orientation with the fastest rate of instability growth, hence the material instability scenarios within these planes do not lead to the most catastrophic instability mechanism. Nevertheless, they are interesting from a theoretical point of view.

Elasticity enters the stability equations in a way which is not of a vital importance for the onset of instability. It does not alter the stability properties of the system. The onset of instability depends predominantly on interaction among the critical values of certain material parameters emerging from the viscoplastic part of the constitutive law. This is precisely what we should expect from the start, since the instability behaviour of the system has been constrained to the viscoplastic part of constitutive relation by choosing the incrementally linear hypoelastic law for the elastic part of the rate of deformation tensor. However, the method presented here is quite general, and allows one to incorporate nonlinear effects into the reversible part of constitutive relation without much difficulty. So if one is interested in studying various interaction phenomena between reversible and irreversible parts of the constitutive relation, and their influence on the stability behaviour of the system, then these effects can be readily incorporated into the constitutive relations without significantly altering the method of derivation displayed in the paper. Of course, the material and control parameters must then be redefined accordingly. Even in the case when elasticity retains its standard role representing an incrementally linear and reversible portion of the rate of deformation tensor, which is usually described in terms of Hooke's law, even though the static exchange of stability is not influenced by effects of the elastic portion of the constitutive law, which clearly follows from the analysis, that may not be the case when the dynamic transition takes place. The discussion regarding the onset of bifurcation becomes much easier when elasticity is neglected, and the material can be considered as rigid-viscoplastic. A model of this type can serve quite well in establishing the qualitative predictions arising from bifurcation criteria. For the sake of comparison, the critical conditions describing the onset of HB for two particular stress states, uniaxial and simple shear, were performed for the insta-

bility plane orientations corresponding to $\mathbf{t}_s^2 = 0$, i.e. for the first degenerate case of the dispersion relation, and $\mathbf{t}_s^2 = 0$, i.e. for the second degenerate case of the dispersion relation. In the case of a rigid viscoplastic material, one can explore possibilities arising from other instability plane orientations, since in this case the value of the characteristic parameter Ξ remains fixed ($\Xi = -1$), and we are not confined to the plane orientations corresponding to both degenerate cases of the dispersion relation. Since the inclusion of elasticity causes a substantial increase in algebraic complexity of the leading expressions, especially in the characteristic coefficients of the stability equation, it is not possible to obtain an explicit analytical expression for the criterion describing the onset of dynamic instability. Hence, we are not able to isolate the critical point and the corresponding values of the material parameters in explicit form, where the influence of elasticity and other parameters on the onset of dynamic instability could be inferred directly. Nevertheless, it is possible to formulate the problem in an implicit manner which would enable the detection of the critical points. But this approach comes with one additional difficulty: the complete numerical procedure requires an extensive parametric study. If the representative data are collected, for example, for the constitutive equations most widely used in practical analyses, it is possible to explore the intricacies of each constitutive law individually. Hence, the nature of singularities corresponding to the class of dynamic instabilities under the more general setting still remains to be explored.

Finally, the elastic part of constitutive relation induces a rotation of the plane of the fastest incipient rate of instability growth relative to the principal directions of stress deviators. By neglecting elasticity, this property vanishes completely, and the plane orientation is always fixed relative to the directions of the principal stress deviators throughout the loading process.

Appendix

Table 1. Coefficients of the dispersion relation: dynamic perturbation.

$$a_0 = c_\theta \rho^2 \bar{\sigma}^4 / k^4,$$

$$a_1 = 6 c_\theta \dot{\bar{\varepsilon}}\, \mu \rho^2 \bar{\sigma}^3 / k^4 + 3 A c_\theta \mu \rho^2 \bar{\sigma}^4 / k^4 - B c_\theta \rho^2 \bar{\sigma}^4 / k^4 - C \chi \rho^2 \bar{\sigma}^5 / k^4,$$

$$a_2 = c_\theta \rho \bar{\sigma}^2 \left[9 \dot{\bar{\varepsilon}}^2 \mu^2 \rho + (\lambda + 3\mu) \bar{\sigma}^2 k^2 \right] / k^4 + 18 A c_\theta \dot{\bar{\varepsilon}}\, \mu^2 \rho^2 \bar{\sigma}^3 / k^4 - 6 B c_\theta \dot{\bar{\varepsilon}}\, \mu \rho^2 \bar{\sigma}^3 / k^4 - 3 C \dot{\bar{\varepsilon}}\, \chi \mu \rho^2 \bar{\sigma}^4 / k^4,$$

$$a_3 = c_\theta \dot{\bar{\varepsilon}}\, \mu \rho \bar{\sigma} \left[(6\lambda + 11\mu) \bar{\sigma}^2 + 9 \mu \mathbf{t}^2 \right] / k^2 + 3 A c_\theta \mu \rho \bar{\sigma}^2 \left[9 \dot{\bar{\varepsilon}}^2 \mu^2 \rho + (\lambda + 3\mu) \bar{\sigma}^2 k^2 - 3 \mu \mathbf{t}^2 k^2 \right] / k^4 - B c_\theta \rho \bar{\sigma}^2 \left[9 \dot{\bar{\varepsilon}}^2 \mu^2 \rho + (\lambda + 3\mu) \bar{\sigma}^2 k^2 \right] / k^4 + C \chi \rho \bar{\sigma}^3 \left[9 \dot{\bar{\varepsilon}}^2 \mu^2 \rho - (\lambda + 3\mu) \bar{\sigma}^2 k^2 \right] / k^4,$$

$$a_4 = c_\theta \mu \left[3\mu(3\lambda + 2\mu)\rho\bar{\sigma}^2 \dot{\bar{\varepsilon}}^2 + 27 \dot{\bar{\varepsilon}}^2 \mu^2 \rho \mathbf{t}^2 + (\lambda + 2\mu)\bar{\sigma}^4 k^2\right] / k^2$$
$$+ 3A c_\theta \dot{\bar{\varepsilon}} \mu^2 \rho \bar{\sigma} \left[(6\lambda + 11\mu)\bar{\sigma}^2 - 9\mu \mathbf{t}^2\right] / k^2$$
$$- B c_\theta \dot{\bar{\varepsilon}} \mu \rho \bar{\sigma} \left[(6\lambda + 11\mu)\bar{\sigma}^2 + 9\mu \mathbf{t}^2\right] / k^2$$
$$- C \dot{\bar{\varepsilon}} \chi \mu \rho \bar{\sigma}^2 \left[(3\lambda + 2\mu)\bar{\sigma}^2 k^2 - 27 \dot{\bar{\varepsilon}}^2 \mu^2 \rho + 18\mu \mathbf{t}^2 k^2\right] / k^4,$$

$$a_5 = c_\theta \dot{\bar{\varepsilon}} \mu^2 \bar{\sigma} \left[(3\lambda + 2\mu)\bar{\sigma}^2 + 9(\lambda + 2\mu)\mathbf{t}_s^2 + 9\mu \mathbf{t}_n^2\right]$$
$$+ 3A c_\theta \mu^2 \bar{\sigma}^2 \left[3\mu(3\lambda + 2\mu)\rho \dot{\bar{\varepsilon}}^2 \right.$$
$$\left. + (\lambda + 2\mu)\bar{\sigma}^2 k^2 - 3(\lambda + 2\mu)\mathbf{t}_s^2 k^2 - 3\mu \mathbf{t}_n^2 k^2\right] / k^2$$
$$- B c_\theta \mu \left[3\mu(3\lambda + 2\mu)\rho\bar{\sigma}^2 \dot{\bar{\varepsilon}}^2 + 27 \dot{\bar{\varepsilon}}^2 \mu^2 \rho \mathbf{t}^2 + (\lambda + 2\mu)\bar{\sigma}^4 k^2\right] / k^2$$
$$- C \chi \mu \bar{\sigma} \left[54 \dot{\bar{\varepsilon}}^2 \mu^2 \rho \mathbf{t}^2 - 9\mu(\lambda + 3\mu)\rho \bar{\sigma}^2 \dot{\bar{\varepsilon}}^2 + (\lambda + 2\mu)\bar{\sigma}^4 k^2\right] / k^2,$$

$$a_6 = 9 c_\theta \dot{\bar{\varepsilon}}^2 \mu^3 (3\lambda + 2\mu)\mathbf{t}_s^2 + 3A c_\theta \dot{\bar{\varepsilon}} \mu^3 (3\lambda + 2\mu)\bar{\sigma}(\bar{\sigma}^2 - 3\mathbf{t}_s^2)$$
$$- B c_\theta \dot{\bar{\varepsilon}} \mu^2 \bar{\sigma} \left[(3\lambda + 2\mu)\bar{\sigma}^2 + 9(\lambda + 2\mu)\mathbf{t}_s^2 + 9\mu \mathbf{t}_n^2\right]$$
$$+ C \dot{\bar{\varepsilon}} \chi \mu^2 \bar{\sigma}^2 \left[9\mu(3\lambda + 2\mu)\rho \dot{\bar{\varepsilon}}^2 + 4\mu\bar{\sigma}^2 k^2 - 18(\lambda + 2\mu)\mathbf{t}_s^2 k^2 - 18\mu \mathbf{t}_s^2 k^2\right] / k^2,$$

$$a_7 = -9 B c_\theta \dot{\bar{\varepsilon}}^2 \mu^3 (3\lambda + 2\mu)\mathbf{t}_s^2 + 3C \dot{\bar{\varepsilon}}^2 \chi \mu^3 (3\lambda + 2\mu)\bar{\sigma}(\bar{\sigma}^2 - 6\mathbf{t}_s^2).$$

Table 2. Coefficients of the dispersion relation: quasi-static perturbation.

$$a_0 = c_\theta \mu (\lambda + 2\mu)\bar{\sigma}^4,$$

$$a_1 = c_\theta \dot{\bar{\varepsilon}} \mu^2 \bar{\sigma} \left[(3\lambda + 2\mu)\bar{\sigma}^2 + 9(\lambda + 2\mu)\mathbf{t}_s^2 + 9\mu \mathbf{t}_s^2\right] - B c_\theta \mu (\lambda + 2\mu)\bar{\sigma}^4$$
$$+ 3A c_\theta \mu^2 \bar{\sigma}^2 \left[(\lambda + 2\mu)\bar{\sigma}^2 - 3(\lambda + 2\mu)\mathbf{t}_s^2 - 3\mu \mathbf{t}_n^2\right] - C \chi \mu (\lambda + 2\mu)\bar{\sigma}^5,$$

$$a_2 = 9 c_\theta \dot{\bar{\varepsilon}}^2 \mu^3 (3\lambda + 2\mu)\mathbf{t}_s^2 + 3A c_\theta \dot{\bar{\varepsilon}} \mu^3 (3\lambda + 2\mu)\bar{\sigma}(\bar{\sigma}^2 - 3\mathbf{t}_s^2)$$
$$- B c_\theta \dot{\bar{\varepsilon}} \mu^2 \bar{\sigma} \left[(3\lambda + 2\mu)\bar{\sigma}^2 + 9(\lambda + 2\mu)\mathbf{t}_s^2 + 9\mu \mathbf{t}_n^2\right]$$
$$+ 2C \dot{\bar{\varepsilon}} \chi \mu^2 \bar{\sigma}^2 \left[2\mu\bar{\sigma}^2 - 9(\lambda + 2\mu)\mathbf{t}_s^2 - 9\mu \mathbf{t}_n^2\right],$$

$$a_3 = -9 B c_\theta \dot{\bar{\varepsilon}}^2 \mu^3 (3\lambda + 2\mu)\mathbf{t}_s^2 + 3C \dot{\bar{\varepsilon}}^2 \chi \mu^3 (3\lambda + 2\mu)\bar{\sigma}(\bar{\sigma}^2 - 6\mathbf{t}_s^2).$$

Legend: $A = \dfrac{\partial \dot{\bar{\varepsilon}}}{\partial \bar{\sigma}}$, $B = \dfrac{\partial \dot{\bar{\varepsilon}}}{\partial \bar{\varepsilon}}$, $C = \dfrac{\partial \dot{\bar{\varepsilon}}}{\partial \theta}$, $\mathbf{t}^2 = \mathbf{n} \cdot \boldsymbol{\tau}' \cdot \boldsymbol{\tau}' \mathbf{n}$, $\mathbf{t}_n^2 = (\mathbf{n} \cdot \boldsymbol{\tau}' \cdot \mathbf{n})^2$, $\mathbf{t}_s^2 = \mathbf{t}^2 - \mathbf{t}_n^2$.

Table 3. Coefficients of the dispersion relation: first degenerate case – dynamic perturbation.

$$
\begin{aligned}
a_0 &= c_\theta \rho / k^2, \\
a_1 &= 3c_\theta \dot{\bar{\varepsilon}}\, \mu\rho/\bar{\sigma}k^2 + 3Ac_\theta \mu\rho/k^2 - Bc_\theta \rho/k^2 - C\chi\rho\bar{\sigma}/k^2, \\
a_2 &= c_\theta(\lambda + 2\mu) + 9Ac_\theta \dot{\bar{\varepsilon}}\, \mu^2\rho/\bar{\sigma}k^2 - 3Bc_\theta \dot{\bar{\varepsilon}}\, \mu\rho/\bar{\sigma}k^2, \\
a_3 &= c_\theta \dot{\bar{\varepsilon}}\, \mu \left[(3\lambda + 2\mu)/\bar{\sigma} + 9\mu \mathbf{t}^2/\bar{\sigma}^3\right] + 3Ac_\theta \mu \left[(\lambda + 2\mu) - 3\mu \mathbf{t}^2/\bar{\sigma}^2\right] \\
&\qquad - Bc_\theta(\lambda + 2\mu) + C\chi \left[9\, \dot{\bar{\varepsilon}}^2 \mu^2\rho/\bar{\sigma}k^2 - (\lambda + 2\mu)\bar{\sigma}\right], \\
a_4 &= 3Ac_\theta \dot{\bar{\varepsilon}}\, \mu^2(3\lambda + 2\mu)/\bar{\sigma} - Bc_\theta \dot{\bar{\varepsilon}}\, \mu \left[(3\lambda + 2\mu)/\bar{\sigma} + 9\mu \mathbf{t}^2/\bar{\sigma}^3\right] \\
&\qquad + 2C \dot{\bar{\varepsilon}}\, \chi\mu^2(2 - 9\mathbf{t}^2/\bar{\sigma}^2), \\
a_5 &= 3C\, \dot{\bar{\varepsilon}}^2 \chi\mu^2(3\lambda + 2\mu)/\bar{\sigma}.
\end{aligned}
$$

Table 4. Coefficients of the dispersion relation: first degenerate case – quasi-static perturbation.

$$
\begin{aligned}
a_0 &= c_\theta(\lambda + 2\mu), \\
a_1 &= c_\theta \dot{\bar{\varepsilon}}\, \mu \left[(3\lambda + 2\mu)/\bar{\sigma} + 9\mu \mathbf{t}^2/\bar{\sigma}^3\right] + 3Ac_\theta \mu \left[(\lambda + 2\mu) - 3\mu \mathbf{t}^2/\bar{\sigma}^2\right] \\
&\qquad - Bc_\theta(\lambda + 2\mu) - C\chi(\lambda + 2\mu)\bar{\sigma}, \\
a_2 &= 3Ac_\theta \dot{\bar{\varepsilon}}\, \mu^2(3\lambda + 2\mu)/\bar{\sigma} - Bc_\theta \dot{\bar{\varepsilon}}\, \mu \left[(3\lambda + 2\mu)/\bar{\sigma} + 9\mu \mathbf{t}^2/\bar{\sigma}^3\right] \\
&\qquad + 2C \dot{\bar{\varepsilon}}\, \chi\mu^2(2 - 9\mathbf{t}^2/\bar{\sigma}^2), \\
a_3 &= 3C\, \dot{\bar{\varepsilon}}^2 \chi\mu^2(3\lambda + 2\mu)/\bar{\sigma}.
\end{aligned}
$$

Table 5. Coefficients of the dispersion relation: second degenerate case – dynamic perturbation.

$$
\begin{aligned}
a_0 &= c_\theta \rho / k^2, \\
a_1 &= 3c_\theta \dot{\bar{\varepsilon}}\, \mu\rho/\bar{\sigma}k^2 + 3Ac_\theta \mu\rho/k^2 - Bc_\theta \rho/k^2 - C\chi\rho\bar{\sigma}/k^2, \\
a_2 &= c_\theta \mu + 9Ac_\theta \dot{\bar{\varepsilon}}\, \mu^2\rho/\bar{\sigma}k^2 - 3Bc_\theta \dot{\bar{\varepsilon}}\, \mu\rho/\bar{\sigma}k^2, \\
a_3 &= 9c_\theta \dot{\bar{\varepsilon}}\, \mu^2 \mathbf{t}^2/\bar{\sigma}^3 + 3Ac_\theta \mu^2(1 - 3\mathbf{t}^2/\bar{\sigma}^2) - Bc_\theta \mu + C\chi\mu(9\, \dot{\bar{\varepsilon}}^2 \mu\rho/\bar{\sigma}k^2 - \bar{\sigma}), \\
a_4 &= -9Bc_\theta \dot{\bar{\varepsilon}}\, \mu^2 \mathbf{t}^2/\bar{\sigma}^3 + 3C\, \dot{\bar{\varepsilon}}\, \chi\mu^2(1 - 6\mathbf{t}^2/\bar{\sigma}^2).
\end{aligned}
$$

Table 6. Coefficients of the dispersion relation: second degenerate case – quasi-static perturbaton.

$$a_0 = c_\theta \mu,$$
$$a_1 = 9c_\theta \dot{\bar{\varepsilon}}\, \mu^2 \mathbf{t}^2/\bar{\sigma}^3 + 3Ac_\theta \mu^2(1 - 3\mathbf{t}^2/\bar{\sigma}^2) - Bc_\theta \mu - C\chi\mu\bar{\sigma},$$
$$a_2 = -9Bc_\theta \dot{\bar{\varepsilon}}\, \mu^2 \mathbf{t}^2/\bar{\sigma}^3 + 3C\dot{\bar{\varepsilon}}\, \chi\mu^2(1 - 6\mathbf{t}^2/\bar{\sigma}^2).$$

Legend: $A = \dfrac{\partial \dot{\bar{\varepsilon}}}{\partial \bar{\sigma}}$, $B = \dfrac{\partial \dot{\bar{\varepsilon}}}{\partial \bar{\varepsilon}}$, $C = \dfrac{\partial \dot{\bar{\varepsilon}}}{\partial \theta}$. $\mathbf{t}^2 = \mathbf{n} \cdot \boldsymbol{\tau}' \cdot \boldsymbol{\tau}' \cdot \mathbf{n}$, $\mathbf{t}_n^2 = (\mathbf{n} \cdot \boldsymbol{\tau}' \cdot \mathbf{n})^2$.

Table 7. Stationary points of the dispersion relation $D(\omega, \mathbf{k})$ on the unit sphere $|\mathbf{n}| = 1$.

case	stress	n	λ
1	$\tau_2' = \tau_3'$	$n_1^2 = \ell^2, \quad n_2^2 + n_3^2 = m^2$	$\bar{\lambda} = \tau_1'^2/2$
2	$\tau_3' = \tau_1'$	$n_2^2 = \ell^2, \quad n_1^2 + n_3^2 = m^2$	$\bar{\lambda} = \tau_2'^2/2$
3	$\tau_1' = \tau_2'$	$n_3^2 = \ell^2, \quad n_1^2 + n_2^2 = m^2$	$\bar{\lambda} = \tau_3'^2/2$
4A	$\tau_2' = \tau_3'$	$n_1^2 = 0, \quad n_2^2 + n_3^2 = 1$	$\bar{\lambda} = Z_{\bar{\lambda}}\tau_1'^2/4$
4B	$\tau_2' \neq \tau_3'$	$n_1^2 = 0, \quad n_2^2 = \dfrac{\Xi\tau_1' - 2\tau_3'}{2(\tau_2' - \tau_3')}, \quad n_3^2 = \dfrac{-\Xi\tau_1' + 2\tau_2'}{2(\tau_2' - \tau_3')}$	$\bar{\lambda} = -\tau_2'\tau_3'$
5A	$\tau_3' = \tau_1'$	$n_2^2 = 0, \quad n_1^2 + n_3^2 = 1$	$\bar{\lambda} = Z_{\bar{\lambda}}\tau_2'^2/4$
5B	$\tau_3' \neq \tau_1'$	$n_2^2 = 0, \quad n_1^2 = \dfrac{\Xi\tau_2' - 2\tau_3'}{2(\tau_1' - \tau_3')}, \quad n_3^2 = \dfrac{-\Xi\tau_2' + 2\tau_1'}{2(\tau_1' - \tau_3')}$	$\bar{\lambda} = -\tau_1'\tau_3'$
6A	$\tau_1' = \tau_2'$	$n_3^2 = 0, \quad n_1^2 + n_2^2 = 1$	$\bar{\lambda} = Z_{\bar{\lambda}}\tau_3'^2/4$
6B	$\tau_1' \neq \tau_2'$	$n_3^2 = 0, \quad n_1^2 = \dfrac{\Xi\tau_3' - 2\tau_2'}{2(\tau_1' - \tau_2')}, \quad n_2^2 = \dfrac{-\Xi\tau_3' + 2\tau_1'}{2(\tau_1' - \tau_2')}$	$\bar{\lambda} = -\tau_1'\tau_2'$
7	$\tau_3' \neq 0$	$n_1^2 = 0, \quad n_2^2 = 0, \quad n_3^2 = 1$	$\bar{\lambda} = Z_{\bar{\lambda}}\tau_3'^2$
8	$\tau_2' \neq 0$	$n_1^2 = 0, \quad n_2^2 = 1, \quad n_3^2 = 0$	$\bar{\lambda} = Z_{\bar{\lambda}}\tau_2'^2$
9	$\tau_1' \neq 0$	$n_1^2 = 1, \quad n_2^2 = 0, \quad n_3^2 = 0$	$\bar{\lambda} = Z_{\bar{\lambda}}\tau_1'^2$

Legend: $\ell^2 = (2 - \Xi)/6$, $m^2 = (4 + \Xi)/6$, $Z_{\bar{\lambda}} = [1 + 2/\Xi]$, $\bar{\lambda} = [\mathbf{t}^2 + 2\mathbf{t}_n^2/\Xi]$.

References

1. L. ANAND, K.H. KIM and T.G. SHAWKI, *Onset of shear localization in viscoplastic solids*, J. Mech. Phys. Solids, **35**, 407, 1987.
2. R.J. ASARO and J.R. RICE, *Strain localization in ductile single crystals*, J. Mech. Phys. Solids, **25**, 309, 1977.
3. Y.L. BAI, *Thermo-plastic instability in simple shear*, J. Mech. Phys. Solids, **30**, 195, 1982.
4. M.K. DUSZEK and P. PERZYNA, *The localization of plastic deformation in thermoplastic solids*, Int. J. Solids Structure, **27**, 1419, 1991.
5. C. FRESSENGEAS and A. MOLINARI, *Inertia and thermal effects on the localization of plastic flow*, Acta Metall., **33**, 387, 1985.
6. R. HILL, *Aspects of invariance in solid mechanics*, Advances in Appl. Mech., Vol. 18, Academic Press, N.Y. 1978.
7. H. KOLSKY, *Stress waves in solids*, Dover, N.Y. 1963.
8. J. LEMONDS and A. NEEDLEMAN, *Finite element analyses of shear localization in rate and temperature-dependent solids*, Mech. of Materials, **5**, 339, 1986.
9. Y. LEROY and M. ORTIZ, *Finite element analysis of transient strain localization phenomena in frictional solids*, Int. J. Num. Anal. Methods Geomech., **14**, 93, 1990.
10. J.E. MARSDEN and M. MCCRACKEN, *The Hopf bifurcation and its applications*, Appl. Math. Sciences, 19, Springer Verlag, N.Y. 1976.
11. J. PAN, M. SAJE and A. NEEDLEMAN, *Localization of deformation in rate sensitive porous plastic solids*, Int. J. Fracture, **21**, 261, 1983.
12. D. PEIRCE, R.J. ASARO and A. NEEDLEMAN, *Material rate dependence and localized deformation in crystalline solids*, Acta Metall., **35**, 1951, 1983.
13. P. PERZYNA, *Thermodynamic theory of viscoplasticity*, Advances in Appl. Mech., Vol. 11, Academic Press, N.Y. 1971.
14. P. PERZYNA, *Temperature and rate dependent theory of plasticity of crystalline solids*, Rev. Phys. Appl., **23**, 445, 1988.
15. J.R. RICE, *On structure of stress-strain relation for time-dependent plastic deformation in metals*, J. Appl. Mech., **37**, 728, 1970.
16. J.R. RICE, *Inelastic constitutive relations for solids: an internal variable theory and its application to metal plasticity*, J. Mech. Phys. Solids, **19**, 433, 1971.
17. J.R. RICE, *Localization of plastic deformation*, [in:] Theoretical and Applied Mechanics, Proc. of the 14th IUTAM Congress, Delft, the Netherlands (1976), W. KOITER [Ed.], p. 207, North Holland, N.Y. 1977.
18. J.R. RICE and J. RUDNICKI, *A note on some features of the theory of localization of deformation*, Int. J. Solids Structures, **16**, 597, 1980.
19. J. RUDNICKI and J.R. RICE, *Conditions for the localization of deformation in pressure-sensitive dilatant materials*, J. Mech. Phys. Solids, **23**, 371, 1975.
20. D.G. SCHAEFFER, *Instability and ill-posedness in the deformation of granular materials*, Int. J. Num. Anal. Methods Geomech., **14**, 253, 1990.
21. T.G. SHAWKI and R.J. CLIFTON, *Shear band formation in thermal viscoplastic materials*, Mech. of Materials, **8**, 13, 1989.
22. F. VERHULST, *Nonlinear differential equations in dynamical systems*, Springer Verlag, Berlin 1990.
23. B.L. VAN DER WAERDEN, *Algebra*, vol. 2, Ungar, N.Y. 1970.
24. G.B. WHITHAM, *Linear and nonlinear waves*, John Wiley, N.Y. 1974.
25. H. YAMAMOTO, *Conditions for shear localization in the ductile fracture of void-containing materials*, Int. J. Fracture, **14**, 347, 1978.

DEPARTMENT OF MECHANICAL ENGINEERING,
UNIVERSITY OF MARIBOR, MARIBOR, SLOVENIA.

Received February 15, 1994; *new version August* 29, 1994.

Arch. Mech., 46, 6, pp. 937–951, Warszawa 1994

Kinetic theory and thermocapillarity equations

K. PIECHÓR (WARSZAWA)

IT IS SHOWN that the thermocapillarity equations of liquid-vapour systems can be deduced from the kinetic Enskog-Vlasov equation under the following assumptions: i) the mean free path of hard-sphere molecules and the range of the attractive forces are much shorter than the characteristic macroscopic length; ii) the mean kinetic energy of chaotic motion of molecules is much greater than the potential energy of attractive forces. It turns out that it is necessary to distinguish between kinetic and equilibrium thermodynamics temperatures. Explicit formulae for the internal energy and stress tensor are obtained. Also, orders of magnitude of various effects considered in the thermocapillary theory of liquid-vapour systems are given and their use is discussed.

1. Introduction

THERE ARE TWO approaches to liquid-vapour systems. The first one assumes that the liquid phase and the vapour one are separated by a sharp interface endowed with energy and entropy [1].

According to the second concept, the liquid and the vapour are separated by a thin layer, but of finite thickness. Within this layer the flow parameters vary violently but smoothly. The great step in this direction was made by J.D. VAN DER WAALS [2], who proposed one pressure formula suitable both for the liquid and gaseous state of a fluid. His formula reads

$$p = \frac{kn\theta}{1 - bn} - an^2, \tag{1.1}$$

where p is the pressure, k is the Boltzmann constant, n is the number density, θ is the temperature, and a, b are positive constants characterizing the fluid.

Since then many more accurate and, at the same time, more or even very complicated pressure formulae have been proposed. Some of them can be found in [3]. The simplest formulae generalizing that of van der Waals can be written in the form

$$p = kn\theta[1 + bn\,\chi(n)] - an^2\theta^{-\alpha}\psi(n), \tag{1.2}$$

where $\chi(n)$ and $\psi(n)$ are given, positive functions of the density only; α, a, b are positive constants.

The next step was made by J.D. KORTEWEG [4], who contributed to the dynamics of liquid-vapour systems by modifying the Cauchy stress tensor. To model the spatial interaction effects in the transition zone he introduced the first and second order derivatives of the density. However, as shown by DUNN and SERRIN [5], the

Korteweg equations are incompatible with the continuum theory of thermodynamics.

The modern derivation of the capillarity equations can be found in [5, 6] and [7–9]. In [5] a review of earlier papers on the topic is given as well. All of them are based on the phenomenological continuum thermomechanics. However, within this framework one cannot obtain the form of the energy function, nor one can say anything about the order of magnitude of the capillarity terms. To answer such questions an augmented physical description is necessary.

These equations were applied to study film boiling phenomena [8, 10], and shock waves in van der Waals type of media [11–19].

In this paper we deduce the capillarity equations from kinetic theory under a series of assumptions concerning orders of magnitude of some dimensionless parameters. In principle, such assumptions are not demanded by the phenomenological approach.

This resembles the situation very well known from the theory of ideal gases, where the Navier–Stokes equations are derived from the Boltzmann equation under the assumption that within the flow there are no zones of great gradients, such as shock waves. Consequently, the shock wave profile cannot be described correctly by the continuum gasdynamics [20]. In the case of liquid-vapour systems we have not only shock waves but also very thin interfacial zones where the gradients of many flow parameters are even greater. That is why we think that in the case of liquid-vapour systems, a kinetic theory approach should be much more necessary than in the case of ideal gases.

The most fundamental problem we face is the choice of the kinetic equation. It should be suited for liquid-vapour systems and at the same time it should be simple enough to be mathematically tractable. The kinetic equation satisfying our demands is the Enskog–Vlasov equation [24–27]. We have, however, to admit that this equation is far from being satisfactory from a theoretical physicist's, point of view. This is mainly due to the fact that it contains two functions bearing the sense of correlation functions which are chosen to fit, in the equilibrium case, the pressure formula or the transport coefficients known from other theories. Hence, as a matter of fact, the Enskog–Vlasov equation, similarly to the Enskog equation itself, is quasikinetic. Therefore it is deprived of profoundness appurtenant to other statistical theories of liquid-vapour systems. Unfortunately such theories produce equations that are useless for practice.

2. The Enskog–Vlasov equation

We assume that the intermolecular potential is a sum of a hard-sphere core of diameter σ and an attractive part, i.e.

$$\Phi = \begin{cases} \infty & \text{for} \quad r < \sigma, \\ \phi(r) & \text{for} \quad r > \sigma, \end{cases} \tag{2.1}$$

where r is the distance between the centres of molecules.

Consequently, the collisional part of the kinetic equation is a sum of the standard ([24, 25]) or revised ([26, 27]) Enskog operator and a mean-field term.

The Enskog–Vlasov equation written in a dimensionless form reads

$$\frac{\partial f}{\partial t} + v_i \frac{\partial f}{\partial x_i} = -\frac{\partial f}{\partial v_i} F_i + E_\varepsilon(f), \tag{2.2}$$

where $f(\mathbf{x}, \mathbf{v}, t)$ is the (dimensionless) one-particle distribution function; $\mathbf{x} = (x_1, x_2, x_3)$ is the position, $\mathbf{v} = (v_1, v_2, v_3)$ is the molecular velocity, and t is the time.

When reducing $\mathbf{x}$, $\mathbf{v}$, and t to the nondimensional form we referred them in the usual way to a macroscopic length-scale L, thermal speed c, and we used L/c as the time unit. The distribution function f was scaled by referring it to $n_0 c^{-3}$, where n_0 is the characteristic number density.

In (2.2), and through the rest of this paper, the summation is performed over the same indices.

Let Φ_0 be the characteristic value of the attractive potential Φ, and let r_0 be its range.

The self-consistent force of attraction is of the form

$$F = \frac{\Phi_0(n_0 r_0^3)}{c^2} \frac{1}{\tau^4} \int \frac{\mathbf{y} - \mathbf{x}}{|\mathbf{y} - \mathbf{x}|} \phi'\left(\frac{|\mathbf{x} - \mathbf{y}|}{\tau}\right) G(\mathbf{x}, \mathbf{y}) n(\mathbf{y}, t)\, d\mathbf{y}, \tag{2.3}$$

where

$$\tau = \frac{r_0}{L}. \tag{2.4}$$

Φ' is the derivative of Φ, $n(\mathbf{x}, t)$ is the number density

$$n(\mathbf{x}, t) = \int f(\mathbf{x}, \mathbf{v}, t)\, d\mathbf{v}, \tag{2.5}$$

and $G(\mathbf{x}, \mathbf{y}) = G(\mathbf{y}, \mathbf{x}) > 0$ is introduced to take into account correlations between molecules in the regions of strong inhomogeneity.

Finally, in (2.2), E_ε is the revised Enskog operator

$$E_\varepsilon = \frac{1}{\varepsilon} \int d\mathbf{v}_\star \int_{|\mathbf{k}|=1} d\mathbf{k} \Psi\,[\mathbf{k} \cdot (\mathbf{v}_\star - \mathbf{v})] \Big[H(\mathbf{x}, \mathbf{x} + \varepsilon\delta\mathbf{k}) f'(\mathbf{x}) f'_\star(\mathbf{x} + \varepsilon\delta\mathbf{k}) - H(\mathbf{x}, \mathbf{x} - \varepsilon\delta\mathbf{k}) f(\mathbf{x}) f_\star(\mathbf{x} - \varepsilon\delta\mathbf{k}) \Big], \tag{2.6}$$

where

$$\mathbf{v}' = \mathbf{v} + \mathbf{k}[\mathbf{k} \cdot (\mathbf{v}_\star - \mathbf{v})], \qquad \mathbf{v}'_\star = \mathbf{v}_\star - \mathbf{k}[\mathbf{k} \cdot (\mathbf{v}_\star - \mathbf{v})],$$

$\mathbf{k}$ is a unit vector bisecting the angle between the relative velocities $\mathbf{v}_\star - \mathbf{v}$ and $\mathbf{v}'_\star - \mathbf{v}'$; also the common abbreviation is used: $f'(\mathbf{x}) = f(\mathbf{x}, \mathbf{v}', t)$, etc. $\Psi(x)$ is a function of one real variable defined by

$$\Psi(x) = \begin{cases} 0 & \text{for } x < 0, \\ x & \text{for } x \geq 0. \end{cases}$$

The function $H(\mathbf{x}, \mathbf{y}) = H(\mathbf{y}, \mathbf{x}) > 0$ is the local equilibrium pair correlation function. Finally, ε and δ are dimensionless parameters

$$\varepsilon = 1/(n_0 \sigma^2 L), \tag{2.7}$$

$$\delta = n_0 \sigma^3. \tag{2.8}$$

Hence, our equation contains two indefinite functions: G and H. They will be chosen to fit the pressure formula (1.2).

Now, we assume that ε is a small parameter, and treat δ and $\mathbf{F}$ as quantities of order one in ε. In this case we can use the formal Chapman–Enskog procedure as it is described in detail in [28]. For rigorous mathematical results on the hydrodynamical limit of the Boltzmann and Enskog equation see [21–23].

To make calculations simpler we make two assumptions, namely

$$G(\mathbf{x}, \mathbf{y}) = g\left(\frac{\mathbf{x} + \mathbf{y}}{2}, |\mathbf{x} - \mathbf{y}|\right), \tag{2.9}$$

and

$$H(\mathbf{x}, \mathbf{y}) = h\left(\frac{\mathbf{x} + \mathbf{y}}{2}, |\mathbf{x} - \mathbf{y}|\right), \tag{2.10}$$

where g and h are some positive functions.

From (2.10) it follows that

$$H(\mathbf{x}, \mathbf{x} \pm \varepsilon\delta\mathbf{k}) = h\left(\mathbf{x} \pm \frac{1}{2}\varepsilon\delta\mathbf{k}, \varepsilon\delta\right). \tag{2.11}$$

This formula suggests that we can ignore the dependence of h on $|\mathbf{x} - \mathbf{y}|$. We want to explain that assumptions (2.9), (2.10) are not inevitable, and their only role is to simplify the formulae.

To apply the Chapman–Enskog procedure it is enough to treat $\mathbf{F}$ as a given vector field of order one as ε tends to zero. However, to avoid the complicated convolution operator we assume that $\tau \ll 1$, and keep the first two nonvanishing terms. The result is

$$\begin{aligned} F_i = a\left[2g_3\frac{\partial n}{\partial x_i} + n\frac{\partial g_3}{\partial x_i}\right] + A\tau^2\left[3\Delta n\frac{\partial g_3}{\partial x_i} + \frac{3}{2}\Delta g_3\frac{\partial n}{\partial x_i}\right. \\ \left. + \frac{3}{4}n\frac{\partial \Delta g_3}{\partial x_i} + 6g_3\frac{\partial \Delta n}{\partial x_i} + 6\frac{\partial^2 n}{\partial x_i \partial x_j}\frac{\partial g_3}{\partial x_j} + 3\frac{\partial n}{\partial x_j}\frac{\partial^2 g_3}{\partial x_i \partial x_j}\right] + O(\tau^4), \end{aligned} \tag{2.12}$$

where

$$a = \frac{2\pi}{3}\frac{\phi_0 n_0 r_0^3}{c^2}\int_0^\infty y^3\Phi'(y)\,dy, \tag{2.13}$$

and

$$A = \frac{\pi}{15}\frac{\Phi_0 n_0 r_0^3}{c^2}\int_0^\infty y^5\Phi'(y)\,dy, \tag{2.14}$$

are positive constants, and the functions g_ν are defined by

$$g_\nu(\mathbf{x}) = \frac{\int_0^\infty y^\nu g(\mathbf{x}, \tau y)\Phi'(y)\,dy}{\int_0^\infty y^\nu \Phi'(y)\,dy}, \qquad \nu > 0. \tag{2.15}$$

As it is seen from (2.12), (2.15), also the dependence of g on $|\mathbf{x} - \mathbf{y}|$ does not play any important role in this approximation. Therefore, owing to (2.11) and to the above remark, we can assume

$$h = h(\mathbf{x}), \qquad g = g(\mathbf{x}), \tag{2.16}$$

what makes it possible to omit the subscript 3 in g_3.

The hydrodynamic, formal approximation of Eq. (2.2), as $\varepsilon \to 0$, consist of:

a. The mass balance equation

$$\frac{\partial n}{\partial t} + \frac{\partial}{\partial x_i}(nu_i) = 0, \tag{2.17}$$

where n, defined by (2.5), is the number density, and $\mathbf{u} = (u_1, u_2, u_3)$ is the mean velocity vector defined by

$$n\mathbf{u} = \int \mathbf{v} f\,d\mathbf{v}. \tag{2.18}$$

b. The momentum balance equation

$$\frac{\partial}{\partial t}(nu_i) + \frac{\partial}{\partial x_j}(nu_i u_j + p\delta_{ij}) = \varepsilon\frac{\partial}{\partial x_j}D_{ij} + A\tau^2\frac{\partial}{\partial x_j}C_{ij}, \tag{2.19}$$

where δ_{ij} is Kronecker's symbol, p is the pressure

$$p = nT(1 + bnh) - an^2 g, \tag{2.20}$$

T is the kinetic temperature

$$T = \frac{1}{2}\int (\mathbf{v}-\mathbf{u})^2 f\, d\mathbf{v}, \tag{2.21}$$

b is the positive constant defined by

$$b = \frac{2}{3}\pi\delta, \tag{2.22}$$

C_{ij} are the components of the capillarity tensor

$$C_{ij} = \left[6ng\Delta n + \frac{3}{4}n^2\Delta g + 3g\frac{\partial n}{\partial x_k}\frac{\partial n}{\partial x_k} + 6n\frac{\partial n}{\partial x_k}\frac{\partial g}{\partial x_k}\right]\delta_{ij} - \left[6g\frac{\partial n}{\partial x_i}\frac{\partial n}{\partial x_j} + 3n\left(\frac{\partial n}{\partial x_i}\frac{\partial g}{\partial x_j} + \frac{\partial n}{\partial x_j}\frac{\partial g}{\partial x_i}\right)\right], \tag{2.23}$$

and D_{ij} are the components of the dissipation tensor

$$D_{ij} = \kappa\left(\frac{\partial u_k}{\partial x_k}\right)\delta_{ij} + \eta\left[\frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) - \frac{1}{3}\left(\frac{\partial u_k}{\partial x_k}\right)\delta_{ij}\right], \tag{2.24}$$

$\varepsilon\kappa$ and $\varepsilon\eta$ are the bulk and shear viscosity coefficients, respectively. Their explicit form can be found in [28], Sec. 12.5.

c. The energy balance equation

$$\frac{\partial}{\partial t}\left[n\left(\frac{3}{2}T + \frac{1}{2}u^2\right)\right] + \frac{\partial}{\partial x}\left[nu_i\left(\frac{3}{2}T + \frac{1}{2}u^2\right) + pu_i\right] + an^2 g\frac{\partial u_k}{\partial x_k} = \varepsilon\frac{\partial}{\partial x_i}\left(\lambda\frac{\partial T}{\partial x_i}\right) + \varepsilon\frac{\partial}{\partial x_i}(D_{ij}u_j) + A\tau^2 u_i\frac{\partial}{\partial x_j}C_{ij}, \tag{2.25}$$

where $\varepsilon\lambda$ is the coefficient of termal conductivity [28].

Equations (2.19), (2.22), and (2.25) involve two indefinite functions g and h. Their choice will be given in the next section.

3. Consistency with classical thermodynamics

3.1. Inviscid fluid without capillarity

Setting $\varepsilon = 0$, $\tau = 0$ in Eqs. (2.17), (2.19), and (2.25), we deduce the following form of the energy balance equation

$$\left(\frac{3}{2T}\frac{\partial T}{\partial t} - \frac{1+bnh}{n}\frac{\partial n}{\partial t}\right) + u_i\left(\frac{3}{2T}\frac{\partial T}{\partial x_i} - \frac{1+bnh}{n}\frac{\partial n}{\partial x_i}\right) = 0. \tag{3.1}$$

To be more specific, we assume that h is a function of density

$$h(\mathbf{x}) = h[n(\mathbf{x}, t)]. \tag{3.2}$$

Then we obtain

$$\frac{3}{2T}dT - \frac{1 - bnh}{n}dn = dS, \tag{3.3}$$

where

$$S = \frac{3}{2}\ln\frac{T}{T_1} - \int_{n_1}^{n}\frac{1 + b\rho h(\rho)}{\rho}d\rho. \tag{3.4}$$

Here, T_1 and n_1 are some constants.

Owing to (3.3) we can rewrite Eq. (3.1) in the form

$$\frac{\partial S}{\partial t} + u_i\frac{\partial S}{\partial x_i} = 0. \tag{3.5}$$

Inverting (3.4) we get

$$T = T_1 \exp\left\{\frac{2}{3}\left[S + \int_{n_1}^{n}\frac{1 + b\rho h(\rho)}{\rho}d\rho\right]\right\}. \tag{3.6}$$

Next, we assume that

$$g(\mathbf{x}) = g(n(\mathbf{x}, t), T(\mathbf{x}, t)), \tag{3.7}$$

and check easily, using Eq. (3.1), that the following equation holds

$$\frac{\partial}{\partial t}\left[n\left(e_0 + \frac{1}{2}u^2\right)\right] + \frac{\partial}{\partial x_i}\left[nu_i\left(e_0 + \frac{1}{2}u^2 + \frac{p}{n}\right)\right] = 0, \tag{3.8}$$

where

$$e_0(n, S) = \frac{3}{2}T_1 \exp\left\{\frac{2}{3}\left[S + \int_{n_1}^{n}\frac{1 + b\rho h(\rho)}{\rho}d\rho\right]\right\} - a\int_{n_1}^{n} g\left\{\rho, T_1 \exp\left[\frac{2}{3}\left(S + \int_{n_1}^{n}\frac{1 + b\nu h(\nu)}{\nu}d\nu\right)\right]\right\}d\rho, \tag{3.9}$$

and

$$p(n, S) = p(n, T(n, S)), \tag{3.10}$$

with $p(n,T)$ given by (2.20). Additionally, we have

$$p(n,T(n,S)) = n^2 \frac{\partial}{\partial n} e_0(n,S). \tag{3.11}$$

If we denote

$$\theta \equiv \theta(n,S) = \frac{\partial}{\partial S} e_0(n,S), \tag{3.12}$$

then

$$\theta\, dS = de_0 - \frac{p}{n^2} dn. \tag{3.13}$$

Now, Eqs. (3.5), (3.9), (3.11)–(3.13) suggest the following interpretation:

i. S is the local equilibrium entropy. Then, Eq. (3.5) says that the entropy is constant along the streamlines. Let us notice that this equation is a rigorous consequence of (2.17), (2.19), and (2.25) if $\varepsilon = 0$ and $\tau = 0$.

ii. $e_0(n,S)$ is the internal energy. Eq. (3.8) is also a rigorous consequence of (2.17), (2.19), and (2.25). It is of the desired conservative form, contrary to (2.25) even if we set in it $\varepsilon = 0$ and $\tau = 0$.

iii. θ defined by (3.12) is the thermodynamic temperature. Eq. (3.13) expresses the second principle of thermodynamics.

If we perform the differentiation in (3.12) and use (3.4) to eliminate S, we obtain the relation between the thermodynamic temperature θ and the kinetic one T. This relation reads

$$\theta = T - \frac{2}{3} a \int_{n_1}^{n} \exp\left[-\frac{2}{3}\int_{\rho}^{n} \frac{1 + b\nu h(\nu)}{\nu} d\nu\right] \times g'_T\left[\rho, T \exp\left(-\frac{2}{3}\int_{\rho}^{n} \frac{1 + b\nu h(\nu)}{\nu} d\nu\right)\right] d\rho, \tag{3.14}$$

where g'_T is the derivative $\partial g/\partial T$.

We take the following assumption

$$g(n,T) = \hat{g}(n)\, T^{-\alpha}, \qquad \alpha \geq 0, \tag{3.15}$$

where $\hat{g}(n)$ has to be found, and α is the same as in (1.2).

Inserting (3.15) into (3.14) we obtain

$$\theta = T + a\gamma(n) T^{-\alpha}, \tag{3.16}$$

where

$$\gamma(n) = \frac{2}{3}\alpha \int_{n_1}^{n} \hat{g}(\rho) \exp\left[\frac{2}{3}\alpha \int_{\rho}^{n} \frac{1 + b\nu h(\nu)}{\nu} d\nu\right] d\rho. \tag{3.17}$$

In some cases we can solve Eq. (3.16) exactly. The simplest solutions are:
i) $\alpha = 0$, then $T = \theta$;
ii) $\alpha = 1$, then

$$T = \frac{1}{2}\left[\theta + \sqrt{\theta^2 - 4a\gamma}\right], \tag{3.18}$$

where we chose this root which satisfies: $T \propto \theta$ as $a \to 0$.

In the general case we can find only an asymptotic solution to Eq. (3.16). We take the additional assumption

$$a \ll 1, \tag{3.19}$$

and look for solution of (3.16) in the form of a power series in a. The expansion we find is

$$T = \theta\left[1 - \frac{a\gamma}{\theta^{\alpha+1}} - \alpha\left(\frac{a\gamma}{\theta^{\alpha+1}}\right)^2 - \frac{\alpha(3\alpha+1)}{2}\left(\frac{a\gamma}{\theta^{\alpha+1}}\right)^3 + \ldots\right]. \tag{3.20}$$

Assumption (3.19) means that we consider high temperature flows since a is the measure of the ratio of attractive energy to the mean kinetic energy of chaotic motion.

Owing to (3.15) and (3.20) we get the following approximate formula for the pressure

$$p = p(n, T(n,\theta)) = n\theta[1 + bnh(n)] - an^2\left[\hat{g}(n) + \frac{1 + bnh(n)}{n}\right]\theta^{-\alpha} + \ldots . \tag{3.21}$$

Equating (1.2) to the first two terms of (3.21) we obtain

$$h(n) = \chi(n), \tag{3.22}$$

and the integral equation for $\hat{g}$

$$\hat{g}(n) + \frac{2}{3}\alpha\frac{1 + bnh(n)}{n}\int_{n_1}^{n}\hat{g}(\rho)\exp\left[\frac{2}{3}\alpha\int_{\rho}^{n}\frac{1 + b\nu h(\nu)}{\nu}\right]d\rho = \psi(n). \tag{3.23}$$

We differentiate (3.23) with respect to n, and in the obtained equation we use again (3.23) to eliminate the integrals. As the result we get a purely differential equation

$$\frac{d}{dn}\left[\frac{n\hat{g}(n)}{1 + bnh(n)}\right] = \frac{d}{dn}\left[\frac{n\psi(n)}{1 + bnh(n)}\right] - \frac{2}{3}\alpha\psi(n). \tag{3.24}$$

Hence

$$\hat{g}(n) = \psi(n) - \frac{2}{3}\alpha \frac{1 + bnh(n)}{n} \left[\int_{n_1}^{n} \psi(\rho)\, d\rho - C \right], \tag{3.25}$$

where C is an arbitrary constant.

By a direct inspection we check that (3.25) is the solution of Eq. (3.24).

From (2.20), (3.15), (3.16), (3.22) and (3.25) we obtain

$$p(n, \theta) = n\theta[1 + bn\chi(n)] - \frac{an^2\psi(n)}{T^{\alpha}(n, \theta)}. \tag{3.26}$$

This formula coincides with (1.2) only asymptotically as $a \to 0$ (or equivalently as $\theta \to \infty$, or else $T \to \infty$).

Using (3.25) in (3.17) we have

$$\gamma(n) = \frac{2}{3}\alpha \left[\int_{n_1}^{n} \psi(\rho)\, d\rho - C \right]. \tag{3.27}$$

In this way all quantities in Eqs. (3.16) and (3.18), (3.20) have been completely determined.

3.2. Inviscid fluid endowed with capillarity

The second case we consider is that when $\tau > 0$ and $\varepsilon = 0$ in Eqs. (2.17), (2.19), and (2.25). In this case a rigorous consequence of these equations is again Eq. (3.1). Hence, assuming (3.2) we obtain Eq. (3.3) with S given by (3.4). But $e_0(n, S)$ as defined by (3.8) (we keep assumption (3.7)) is no longer the internal energy. Instead, as the internal energy we will treat (see [29])

$$e(n, S) = e_0(n, S) + A\tau^2 \frac{1}{n} \left[(ng)'_n \frac{\partial n}{\partial x_i} \frac{\partial n}{\partial x_i} + (ng)'_S \frac{\partial n}{\partial x_i} \frac{\partial S}{\partial x_i} \right], \tag{3.28}$$

where $e_0(n, S)$ is given by (3.8). Here $g(n, S) = g(n, T(n, S))$, and g'_n and g'_S denote derivatives of $g(n, S)$ with respect to n or, respectively, S.

The energy balance equation (2.25) can be written in the form

$$\frac{\partial}{\partial t} \left[n \left(e + \frac{1}{2}u^2 \right) \right] + \frac{\partial}{\partial x_i} \left[nu_i \left(e + \frac{1}{2}u^2 \right) + \mathfrak{P}_{ij} u_j \right] = \frac{\partial}{\partial x_i} U_i + A\tau^2 \mathcal{S}_i u_i, \tag{3.29}$$

where

$$\mathfrak{P}_{ij} = \left[n^2 \frac{\partial e}{\partial n} - n \frac{\partial}{\partial x_k} \left(n \frac{\partial e}{\partial n_k} \right) \right] \delta_{ij} + n \left(\frac{\partial e}{\partial n_i} n_j + \frac{\partial e}{\partial S_i} S_j \right), \tag{3.30}$$

with

$$n_i \equiv \frac{\partial n}{\partial x_i}, \qquad S_i = \frac{\partial S}{\partial x_i}. \tag{3.31}$$

In (3.30) and in the following formulae the quantities n, n_i, S, S_i are treated as independent variables.

Next,

$$U_i = \dot{n}\, n \frac{\partial e}{\partial n_i} + \dot{S}\, n \frac{\partial e}{\partial S_i}, \tag{3.32}$$

with

$$\dot{n} \equiv \frac{\partial n}{\partial t} + u_i \frac{\partial n}{\partial x_i}, \qquad \dot{S} \equiv \frac{\partial S}{\partial t} + u_i \frac{\partial S}{\partial x_i}. \tag{3.33}$$

And finally

$$\mathcal{S}_i = -3 \frac{\partial}{\partial x_j} \left\{ n^2 \left[\left(g'_n \Delta n + \frac{1}{4} \Delta g \right) \delta_{ij} + \frac{1}{2} \frac{\partial^2}{\partial x_i \partial x_j} g \right] \right\}. \tag{3.34}$$

We rewrite the momentum balance equation (2.19) in the form

$$\frac{\partial}{\partial t}(n u_i) + \frac{\partial}{\partial x}(n u_i u_j + \mathfrak{P}_{ij}) = A \tau^2 \mathcal{S}_i. \tag{3.35}$$

Equations (2.17), (3.29), and (3.35) coincide, modulo the term $\mathcal{S}_i$, with the general thermocapillarity equations as given in [7], [8]. Also the form of the internal energy (3.27) is the same as that obtained in [29]. The only difference is the presence of $\mathcal{S}_i$. A term like that does not appear in any phenomenological approach [5–9], but it occurs in the semi-statistical theory of [29]. In order to get rid of it and to be in agreement with [7–9], i.e. with the phenomenological theories, Gouin used an argument which is not applicable to the present case. We do not see any reason why this term should be eliminated. It is interesting, however, to notice that in the van der Waals case, i.e. if $g \equiv 1$, $\mathcal{S}_i$ vanishes.

3.3. General case

Now we consider the general case of Eqs. (2.17), (2.19), and (2.25) without assuming that $\tau = 0$ and/or $\varepsilon = 0$.

Under assumption (3.2) on h the entropy equation is now

$$\frac{\partial S}{\partial t} + u_i \frac{\partial S}{\partial x_i} = \frac{\varepsilon}{nT} \left[-\frac{\partial}{\partial x_i} \left(\lambda \frac{\partial T}{\partial x_i} \right) + D_{ij} \frac{\partial u_i}{\partial x_j} \right]. \tag{3.36}$$

We rewrite the momentum (2.19) and energy (2.25) balance equations in the form

$$\frac{\partial}{\partial t}(nu_i) + \frac{\partial}{\partial x_j}(nu_iu_j + \mathfrak{P}_{ij}) = \varepsilon\frac{\partial}{\partial x_j}D_{ij} + A\tau^2\mathcal{S}_i\,, \tag{3.37}$$

and

$$\frac{\partial}{\partial t}\left[n\left(e + \frac{1}{2}u^2\right)\right] + \frac{\partial}{\partial x_i}\left[nu_i\left(e + \frac{1}{2}u^2\right) + \mathfrak{P}_{ij}u_j\right] = \frac{\partial}{\partial x_i}U_i + \frac{\varepsilon}{nT}\left[-\frac{\partial}{\partial x_i}\left(\lambda\frac{\partial T}{\partial x_i}\right) + D_{ij}\frac{\partial u_i}{\partial x_j}\right]\left[n\theta - 3A\tau^2(ng)'_n\Delta n - 3A\tau^2(ng)'_S\Delta n\right] + \varepsilon u_i\frac{\partial}{\partial x_j}D_{ij} + A\tau^2u_i\mathcal{S}_i\,, \tag{3.38}$$

where all symbols have the same meaning as previously.

The system of equations (2.17), (3.37), and (3.38) is equivalent to (2.16), (2.19), (2.25), and no approximations have been made.

As we can see, the right-hand side of Eq. (3.38) contains terms which are not usually present in the energy equation with the dissipation included. To give the traditional form to this equation we are forced to make some approximations. First, let us notice that due to the assumption that $\varepsilon \ll 1$ and $\tau \ll 1$ we can reject all terms multiplied by $\varepsilon\tau^2$, as they are small of higher order of magnitude. However, this is not sufficient since after doing that the heat flux term will be of the form

$$-\frac{\varepsilon\theta}{T}\frac{\partial}{\partial x_i}\left(\lambda\frac{\partial T}{\partial x_i}\right),$$

instead of being the divergence of $-\lambda\partial\theta/\partial x_i$. Hence, secondly, to achieve the desired form of this term we assume that a in (3.16) is a small parameter, even if we can solve (3.16) exactly. In this high-temperature approximation $T = \theta + O(a)$. Hence, ignoring terms of order $a\varepsilon$ we can rewrite the energy balance equation in the form

$$\frac{\partial}{\partial n}\left[n\left(e + \frac{1}{2}u^2\right)\right] + \frac{\partial}{\partial x_i}\left[nu_i\left(e + \frac{1}{2}u^2\right) + (\mathfrak{P}_{ij} - \varepsilon D_{ij})u_j\right] - \frac{\partial}{\partial x_i}U_i + \varepsilon\frac{\partial}{\partial x_i}q_i = A\tau^2\mathcal{S}_iu_i\,, \tag{3.39}$$

where

$$q_i = -\lambda\frac{\partial}{\partial x_i}\theta. \tag{3.40}$$

Ignoring terms of order $a\varepsilon$ in (3.36) we write the entropy equation in the standard form

$$\frac{\partial S}{\partial t} + u_i\frac{\partial S}{\partial x_i} = \frac{\varepsilon}{n\theta}\left[-\frac{\partial}{\partial x_i}\left(\lambda\frac{\partial\theta}{\partial x_i}\right) + D_{ij}\frac{\partial u_i}{\partial x_j}\right]. \tag{3.41}$$

In this way we have obtained equations coinciding (modulo the term S_i) with those of [7–9], where a certain analysis of them can be found.

4. Comments

We have shown that the Navier–Stokes equations with capillarity terms can be derived, under some conditions, from the Enskog–Vlasov kinetic equation (2.2). The basic assumptions are:

i.

$$\varepsilon \ll 1, \qquad n_0\sigma^2 = O(1) \quad \text{as} \quad \varepsilon \to 0, \tag{4.1}$$

i.e. the hard-core collisions are the main mechanism leading to the local equilibration of the fluid flow.

ii.

$$\tau \ll 1, \tag{4.2}$$

i.e. the attractive forces are long-ranged. Their range is long as compared to the molecular diameter, but short as compared to the macroscopic length-scale.

Equations (4.1) and (4.2) are the usual assumptions taken when constructing hydrodynamic approximation to the kinetic equations.

iii. Unless the van der Waals case is concerned, we needed an additional stronger assumption that

$$a \ll 1, \tag{4.3}$$

i.e. the temperature is high. More precisely, the condition means that the average kinetic energy of chaotic motion of a molecule is much higher than the attractive potential of forces exerted on it by the surrounding molecules.

The van der Waals case, i.e. the case when $g(\mathbf{x}) \equiv 1$, is exceptional, since then $T = \theta$, and (4.3) is not necessary.

The three parameters ε, τ, and a are independent. Consequently, the terms in Eqs. (3.37), and (3.39) which have ε, a, $A\tau^2$ as multipliers can be of different order of magnitude with respect to each other. The most natural assumption would be

$$a \ll \varepsilon, \qquad A\tau^2 = O(\varepsilon).$$

Then the dissipative and capillarity terms are of the same order of magnitude.

The phenomenological deduction of the capillarity does not use any of our assumptions. Hence, various ordering of the three parameters can be met.

In [8–10] it was assumed that

$$\varepsilon \ll A\tau^2. \tag{4.4}$$

This is not inconsistent with our results if we limit ourselves to Case B (inviscid fluid endowed with capillarity).

On the other hand, in the shock wave problem an assuption opposite to (4.4) was taken. Namely it was assumed in the isothermal case (see [11, 12]) that

$$\varepsilon = O(1), \qquad \text{and} \qquad A\tau^2 = O(1). \tag{4.5}$$

It is important to add that in these papers it was also assumed that $A\tau^2$ cannot be too small. On the other hand, numerical calculations of [19] demanded that the parameter should not exceed an upper bound.

Of course, an assumption like (4.4) is inconsistent with our theory. We have to admit, however, that the necessity of using the capillarity equations in the shock-wave problem is not obligatory. Other approaches were used as well, without mentioning any capillarity (see [30], [31] and [32, 33]). Hence, the use of the kinetic theory in problems like that seems to be even more justified than in the cases of the Boltzmann equation. Unfortunately, to our knowledge, [34] is the only paper where the Enskog equation without attractive forces was used to study the shock waves.

References

1. M.E. GURTIN, *On diffusion in two phase systems; the sharp interface versus transition layer*, [in:] PDEs and Continuum Models of Phase Transitions, M. RASCLE, D. SERRE, M. SLEMROD [Eds.], Lecture Notes in Physics, 344, 99–112, Springer-Verlag, Berlin 1989.
2. J.D. VAN DER WAALS, *On the continuity of the gaseous and liquid states*, Edited with an introductory essay by J.S. ROWLINSON, Studies in Statistical Mechanics, vol. XIV, North-Holland, Amsterdam 1988.
3. P.A. THOMPSON, K.C. LAMBRAKIS, *Negative shock waves*, J. Fluid Mech., **60**, part 1, 187–208, 1973.
4. D.J. KORTEWEG, *Sur la forme que prennent les equations du mouvement des fluides si l'on tient compte des forces capillaires par de variations de densité*, Arch. Nederlandaises des Sci. Exactes et Natur., **28**, 1901.
5. J.E. DUNN, J. SERRIN, *On the thermomechanics of intersticial working*, Arch. Rational Mech. and Anal., **88**, 95–199, 1985.
6. M. E. GURTIN, *On a nonequilibrium thermodynamics of capillarity and phase*, Quarterly of Appl. Math., **47**, 1, 129–145, 1989.
7. P. CASAL, H. GOUIN, *Equations du mouvement des fluides thermocapillaires*, C.R. Acad. Sci. Paris, t. 306, Serie II, 99–104, 1988.
8. P. CASAL, H. GOUIN, *Sur les interfaces liquide-vapour non isothermes*, J. de Mechanique Theorique et Appliquee, 7, 6, 689–718, 1988.
9. P. CASAL, H. GOUIN, *Invariance properties of inviscid fluids of grade N*, [in:] M. RASCLE, D. SERRE, M. SLEMROD [Eds.], PDEs and Continuum Models of Phase Transitions, Lecture Notes in Physics, 344, 85–98, Springer-Verlag, Berlin 1989.
10. H. GOUIN, H.H. FLICHE, *Film boiling phenomena in liquid-vapour interfaces*, [in:] G.E.A. MEIER, P.A. THOMPSON [Eds.], Adiabatic Waves in Liquid-Vapor Systems, IUTAM Symposium, Göttingen 1989, 305–314, Springer-Verlag, Berlin 1990.
11. M. SLEMROD, *Admissing criteria for propagating phase boundaries in a van der Waals fluid*, Arch. Rational Mech. and Anal., **81**, 4, 301–315, 1983.
12. R. HAGAN, M. SLEMROD, *The viscosity-capilllarity criteria for shocks and phase transitions*, Arch. Rational Mech. and Anal., **83**, 333–361, 1983.
13. M. SLEMROD, *Dynamic phase transitions in a van der Waals fluid*, J. Differential Equations, **52**, 1–23, 1984.

14. R. HAGAN, J. SERRIN, *One-dimensional shock layers in Korteweg fluids*, [in:] Phase Transformations and Material Instabilities in Solids, M.E. GURTIN [Ed.], Proc. of a Conference Conducted by the Mathematics Research Center. The University of Wisconsin-Madison, 1983, 113–127, Academic Press, Orlando 1984.
15. M. SLEMROD, *Dynamic of first order phase transitions*, [in:] Phase Transformations and Material Instabilities in Solids, M.E. GURTIN [Ed.], Proc. of a Conference Conducted by the Mathematics Research Center. The University of Wisconsin-Madison, 1983, 163–203, Academic Press, Orlando 1984.
16. M. SHEARER, *Nonuniqueness of admissible solutions of Riemann initial value problems for a system of conservative equations of mixed type*, Arch. Rational Mech. and Anal., **93**, 45–59, 1986.
17. M. SHEARER, *Dynamic phase transitions in a van der Waals gas*, Quarterly of Appl. Math., **46**, 4, 631–636, 1988.
18. M. GRINFELD, *Nonisothermal dynamic phase transitions*, Quarterly of Appl. Math., **47**, 1, 281–296, 1989.
19. M. AFFOUF, R.E. CAFLISH, *A numerical study of Riemann problem solutions and stability for a system of viscous conservation laws of mixed type*, SIAM J. Appl. Math., **51**, 3, 605–634, 1991.
20. W. FISZDON, *The structure of a plane shock wave*, [in:] Rarefied Gas Flows Theory and Experiment, W. FISZDON [Ed.], CISM Courses and Lectures, 224, 447–524, Springer-Verlag, Wien 1981.
21. N. BELLOMO, A. PALCZEWSKI, G. TOSCANI, *Mathematical topics in nonlinear kinetic theory*, World Scientific, Singapore 1988.
22. N. BELLOMO, M. LACHOWICZ, J. POLEWCZAK, G. TOSCANI, *Mathematical topics in nonlinear kinetic theory II. The Enskog equation*, World Scientific, Singapore 1991.
23. M. LACHOWICZ, *Asymptotic analysis of nonlinear kinetic equations: the hydrodynamic limit*, Warsaw University, Dept. of Mathematics, Computer Sciences and Mechanics, Preprint 1/94, Warszawa 1994.
24. M. GRMELA, *Kinetic approach to phase transitions*, J. Statistical Physics, **3**, 3, 347–364, 1971.
25. M. GRMELA, *On the approach to equilibrium in kinetic theory*, J. Math. Phys., **15**, 1, 35–40, 1974.
26. J. KARKHECK, G. STELL, *Kinetic mean-field theories*, J. Chem. Phys., **75**, 3, 1475–1486, 1981.
27. J. KARKHECK, G. STELL, *Maximization of entropy, kinetic equations, and irreversible thermodynamics*, Physical Review A, **25**, 6, 3302–3326, 1982.
28. J. FERZIGER, H.G. KAPER, *Mathematical theory of transport properties in gases*, North-Holland, Amsterdam 1972.
29. H. GOUIN, *Une interpretation moleculaire des fluides thermo-capillaires*, C.R. Acad. Sci. Paris, t. 306, Serie II, 755–759, 1988.
30. M. SHEARER, *The Riemann problem for a class of conservation laws of mixed type*, J. Differential Equations, **46**, 426–443, 1982.
31. R.L. PEGO, *Phase transitions in one-dimensional nonlinear viscoelasticity: admissibility and stability*, Arch. Rational Mech. and Anal., **97**, 4, 353–394, 1987.
32. H. HATTORI, *The Riemann problem for a van der Waals fluid with entropy rate admissibility criterion. Isothermal case*, Arch. Rational Mech. and Anal., **92**, 3, 247–263, 1986.
33. H. HATTORI, *The Riemann problem for a van der Waals fluid with entropy rate admissibility criterion. Nonisothermal case*, J. Differential Equations, **65**, 2, 158–174, 1986.
34. A. FREZOTTI, C. SGARRA, *Numerical analysis of a shock-wave solution to the Enskog equation obtained via a Monte Carlo method*, J. Statistical Physics, **73**, 1/2, 193–208, 1993.

POLISH ACADEMY OF SCIENCES
INSTITUTE OF FUNDAMENTAL TECHNOLOGICAL RESEARCH.

Received June 16, 1994; *new version August* 31, 1994.

rch. Mech., 46, 6, pp. 953–963, Warszawa 1994

Diffracted waves in the problem of electromagnetic diffraction by a screen residing between dissimilar media

A. CIARKOWSKI (WARSZAWA)

THE PAPER IS concerned with electromagnetic wave diffraction by a conducting screen located on the interface between two dielectric media. The aim of the work is to determine amplitudes of the diffracted waves that are generated from the edge of the screen. The analysis is based on the boundary layer method.

. Introduction

HIS WORK DEALS with the problem of electromagnetic plane wave diffraction by a lane, infinitely thin, perfectly conducting screen residing on the interface of two ifferent, homogeneous and isotropic dielectric media. For the screen of arbitrary hape derivation of even an approximate solution is not an easy task. If, however, he edge of the screen is described by a smooth curve and the wavelength is small ompared to the radius of the edge curvature (high frequency approximation), he problem admits asymptotic solution which can be obtained by means of ray echniques.

For the screen of particular shape – the half-plane – the problem considered as an exact solution [1]. With asymptotic analysis of the solution it was shown hat far away from the edge, the total electromagnetic field may be represented in erms of various species of waves. The mechanism of waves generation was studied n [2]. That mechanism can be summarized as follows. In the plane separating wo media the incident plane wave gives rise to two reflected waves, one from the creen surface and the other from the interface, all three waves propagating in ne medium, and to a refracted wave propagating in the opposite medium. These aves constitute the geometrical optics solution. In addition, the incident wave enerates two diffracted waves that emanate from the edge of the screen and ropagate with different velocities in the corresponding media. On the interface hese waves give rise to lateral waves. Propagation of a lateral wave depends trongly on whether the medium it propagates in is optically dense or thin.

In accordance with the ray technique approach, the total field in an arbitrary oint sufficiently distant from the edge is a sum of all waves whose rays reach the oint. On their rays the waves are characterized by phase functions (eikonals) nd amplitude functions. Variation of the phase function along a ray is described y the eikonal equation, and that of the amplitude functions – by the recursive ystem of transport equations. The eikonal and the transport equations are first rder partial differential equations. On the rays, however, they reduce to ordinary

differential equations (cf. [6]). The crucial point in determining the waves on families of their rays is proper determination of initial conditions for the solutions to those equations. In the case of the eikonal equation such a condition follows directly from the principle of phase matching of both the generating and generated waves at the initial manifold (i.e. where the wave generation occurs). For the transport equations, however, choice of the initial conditions is not so simple. In particular, the diffracted rays originate from the screen edge which is a caustic for the diffracted waves. Their amplitudes blow up at the caustic, and hence the initial conditions for the amplitude functions cannot be imposed at the points the diffracted waves are generated from.

In this work it is assumed that the edge of the screen is curvilinear, and the wavelengths admitted are much smaller than the radius of curvature of the edge. In order to find amplitudes of the diffracted waves, an asymptotic approximation will be constructed for the total field in the vicinity of the edge, which in the literature is referred to as a boundary layer. Next, the approximation obtained will be matched to the solution of the transport equations in the region where both field representations remain valid. (Our considerations will be restricted to the leading amplitude function, i.e. the function that satisfies the first equation in the recursive system of transport equations.) As a result, we shall obtain the initial condition necessary for determining the leading term in the asymptotic representation of the diffracted waves. The condition will be recognized as the diffraction coefficient, describing angular dependence of those waves.

The procedure described above is known in the literature as the boundary layer method. BUCHAL and KELLER ([3]) employed it in their study of diffracted waves in a scalar problem of diffraction in a single medium. WITLOX ([4]) adapted the method to the analysis of a vectorial diffraction problem in elastodynamics. The presentation of this paper is based on the latter approach.

2. Formulation of the problem

In the Cartesian coordinate system x, y, z the plane $y = 0$ is the interface of two different homogeneous and isotropic dielectric media. A perfectly conducting screen occupies a portion of this plane. The screen has a curvilinear edge described by a regular curve

$$\mathcal{C}: \quad \mathbf{x} = \mathbf{x}_0(s), \tag{2.1}$$

where s denotes arc length along the curve. The region occupying the halfspace $y > 0$ is assumed to be optically thinner, while the complementary one, filling the halfspace $y < 0$, is optically denser. The wave impedance in the former and the latter regions is equal to $Z_1 = (\mu/\varepsilon_1)^{1/2}$ and $Z_N = (\mu/\varepsilon_N)^{1/2}$, respectively. The quantities ε_1 and ε_N $(= N^2\varepsilon_1)$ are permeabilities in the upper and lower halfspace. μ is a permittivity common to both media. To refer to a particular

medium we shall use the subscript $g = (\varepsilon_g/\varepsilon_1)^{1/2}$, which takes respective values $g = 1$ and $g = N$ in the upper and lower halfspaces. By our earlier assumption $N > 1$.

It is assumed that the electromagnetic plane wave

$$\begin{Bmatrix} \mathbf{E}^i(\mathbf{x}) \\ \mathbf{H}^i(\mathbf{x}) \end{Bmatrix} = \begin{Bmatrix} \mathbf{E}_0 \\ \mathbf{H}_0 \end{Bmatrix} e^{ik_1 S^i(\mathbf{x})} \tag{2.2}$$

propagates in the upper halfspace toward the screen. Its time-dependence is described by $\exp(-i\omega t)$.

Eikonal of this wave is real and equal to

$$S^i(\mathbf{x}) = \nu_1 x - \nu_2 y + \nu_3 z, \qquad \nu_1^2 + \nu_2^2 + \nu_3^2 = 1, \qquad \nu_2 > 0. \tag{2.3}$$

This wave comprises both TM and TE fields with respect to the y-axis. In the following we shall use the subscripts e and m to denote fields TM and TE, respectively([1]). Consequently, the vectorial fields $\mathbf{E}_0$ and $\mathbf{H}_0$ can be expressed as the sums

$$\mathbf{E}_0 = \mathbf{E}_{m0} + \mathbf{E}_{e0}, \qquad \mathbf{H}_0 = \mathbf{H}_{m0} + \mathbf{H}_{e0}, \tag{2.4}$$

where

$$\begin{aligned} \mathbf{E}_{m0} &= k_1^2 Z_1 a_m \hat{\mathbf{y}} \times \nabla S^i, & \mathbf{E}_{e0} &= a_e \nabla S^i \times (\hat{\mathbf{y}} \times \nabla S^i), \\ \mathbf{H}_{m0} &= k_1^2 a_m \nabla S^i \times (\hat{\mathbf{y}} \times \nabla S^i), & \mathbf{H}_{e0} &= -Y_e a_e \hat{\mathbf{y}} \times \nabla S^i \end{aligned} \tag{2.5}$$

and

$$a_e = \frac{E_{0y}}{\nu_1^2 + \nu_3^2}, \qquad a_m = \frac{H_{0y}}{k_1^2(\nu_1^2 + \nu_3^2)}. \tag{2.6}$$

The quantities E_{0y} and H_{0y} are y-components of the fields $\mathbf{E}_0$ and $\mathbf{H}_0$, and $\hat{\mathbf{y}}$ is a unit vector in the direction of the y-axis. a_e and a_m measure TM and TE field contributions in the incident field (2.2).

The total field, by definition being the solution to the problem considered, satisfies in each halfspace the Maxwell equations

$$\nabla \times \mathbf{E}_g = ik_g Z_g \mathbf{H}_g \qquad \text{and} \qquad \nabla \times \mathbf{H}_g = -ik_g Y_g \mathbf{E}_g, \tag{2.7}$$

where $Y_g = Z_g^{-1}$ and $k_g = \omega(\varepsilon_g \mu)^{1/2}$ are the admittance and the wave number in a particular medium. Furthermore, the following conditions should also be satisfied: (i) the continuity condition for the tangent components of the electric and magnetic field in the aperture (i.e. the part of the plane $y = 0$ complementary

([1]) We say that an electromagnetic field is of TM (TE) type with respect to a certain axis if its magnetic (electric) component vanishes along that axis.

to the screen)([2]), (ii) the boundary condition $\mathbf{E}_g \times \hat{\mathbf{y}} = 0$ on the screen([3]), (iii) the edge condition and (iv) the outgoing condition (comp. [6], [1]). The last condition generalizes the Sommerfeld radiation condition to the case of scatterer of infinite extent.

We assume that the radius of curvature of the edge is large as compared to the wavelength. This assumption can be expressed in terms of edge curvature and wave number

$$\frac{|\kappa|}{k_1} \leq \varepsilon, \qquad \varepsilon \ll 1, \tag{2.8}$$

where κ is a "signed" curvature of the edge $\mathcal{C}$. It is understood that $\kappa > 0$ if the edge is locally convex, and $\kappa < 0$ if it is locally concave.

Under these assumptions the asymptotic solution, as $k_g \to \infty$, for the problem considered can be sought using the ray techniques. Underlying assumption in the construction is that the solution for the screen with curvilinear edge comprises the same species of waves that appear in the case of the screen with a straight edge. Naturally, the phase and amplitude functions should account for the curvature of the edge. Such a construction has been carried out in [1]. Another assumption has been adapted there that the field in the vicinity of the curved edge can be locally approximated by the field obtained for the straight edge.

In the present work we reconsider the construction of the diffracted waves, this time not using the last assumption. Instead, we shall construct an approximation to the total field valid in some vicinity of the edge. It then will be used to determine the lacking initial condition needed to determine the amplitude function along the rays. We shall refer to that approximation as the "inner" solution, and the region of its validity as the "inner" region.

A different asymptotic solution, interpretable in terms of propagating waves, we shall call the "outer" solution, and the region where it remains valid – the "outer" region. That solution breaks down in the vicinity of the edge, where diffracted waves have their caustic. Unknown amplitudes of those waves will be sought by matching inner and outer solutions in their common region of validity.

3. The inner solution

At a fixed diffraction point on the edge $\mathcal{C}$ we introduce an orthogonal coordinate system t_1, t_2, t_3 such that the unit vector $\hat{\mathbf{t}}_1(s)$ is orthogonal to the edge and points to the interior of the screen, $\hat{\mathbf{t}}_2 = \hat{\mathbf{y}}$ and $\hat{\mathbf{t}}_3(s) = \hat{\mathbf{t}}_1 \times \hat{\mathbf{t}}_2$. The vector $\mathbf{t}_3$ is tangent to $\mathcal{C}$.

We also define orthogonal curvilinear coordinates ρ, ϕ and s, where ρ and ϕ are, respectively, the distance from the edge and the angle measured from

([2]) If this condition holds then the total field satisfies integral Maxwell equations in regions containing points belonging to both media (comp. [5]).

([3]) The screen itself is a boundary of the region wherein the total field is analyzed.

the screen surface in the plane orthogonal to the edge. The third coordinate s specifies the point on the edge where the plane cuts the edge. Define the inner region by

$$|\kappa|\,\rho \leq \varepsilon. \tag{3.1}$$

In this region we shall find high-frequency asymptotic expression for the total field, which we shall call the inner solution.

Replace the coordinates t_1, t_2 (or ρ, ϕ) and s by the scaled coordinates

$$y_1 = k_1 t_1, \qquad y_2 = k_1 t_2 \qquad (\text{or } r = k_1\rho, \quad \phi), \qquad s. \tag{3.2}$$

In the coordinates y_1, y_2, s the metric coefficients are

$$h_1 = h_2 = \frac{1}{k_1}, \qquad h_3 = 1 - \kappa\rho\cos\phi = 1 + O(\varepsilon) = \frac{k_1 - \kappa y_1}{k_1}. \tag{3.3}$$

Assume that in the boundary layer the total electromagnetic field can be expressed as $\varepsilon \to 0$ in the form

$$\begin{aligned} \mathbf{E}_{g\gamma} &= e^{ik_1 S(s)}\left[\check{\mathbf{e}}_{g\gamma}(y_1, y_2, s) + O(\varepsilon)\right], \\ \mathbf{H}_{g\gamma} &= e^{ik_1 S(s)}\left[\check{\mathbf{h}}_{g\gamma}(y_1, y_2, s) + O(\varepsilon)\right], \end{aligned} \tag{3.4}$$

where $S(s) = S^i(0, 0, s)$ and $\gamma = e, m$.

Also assume that the dependence of $\check{\mathbf{e}}_{g\gamma}$ and $\check{\mathbf{h}}_{g\gamma}$ on y_1 and y_2 is such that their differentiation with respect to both coordinates does not modify their order in ε. Thus we have([4])

$$\begin{aligned} \frac{\partial \mathbf{E}_{g\gamma}}{\partial y_i} &= e^{ik_1 S(s)}\left[\frac{\check{\mathbf{e}}_{g\gamma}(y_1, y_2, s)}{\partial y_i} + O(k_1^{-1})\right], \\ \frac{\partial^2 \mathbf{E}_{g\gamma}}{\partial y_i \partial y_j} &= e^{ik_1 S(s)}\left[\frac{\partial^2 \check{\mathbf{e}}_{g\gamma}(y_1, y_2, s)}{\partial y_i \partial y_j} + O(k_1^{-1})\right]. \end{aligned} \tag{3.5}$$

Furthermore, since $dS/ds = \hat{\mathbf{t}}_3 \cdot \nabla S = \cos\beta_1(s)$, where the angle β_1 is defined with this equality, we obtain

$$\begin{aligned} \frac{\partial \mathbf{E}_{g\gamma}}{\partial s} &= ik_1 \cos\beta_1(s)\mathbf{E}_{g\gamma} + O(1), \\ \frac{\partial^2 \mathbf{E}_{g\gamma}}{\partial s^2} &= [ik_1 \cos\beta_1(s)]^2\,\mathbf{E}_{g\gamma} + O(1). \end{aligned} \tag{3.6}$$

([4]) Since $\kappa = O(1)$, then by virtue of $O(\varphi)O(\psi) = O(\varphi\psi)$ we have $O(\varepsilon) = O(\kappa/k_1) = O(k_1^{-1})$, as $\varepsilon \to 0$ ($k_1 \to \infty$) in the region $\rho < \varepsilon/|\kappa|$. In this region the asymptotic relation $O(k_1^{-1}) = O[(k_1\rho)^{-1}]$ is generally invalid because ρ^{-1} can be arbitrarily large.

Similar formulas can be found for derivatives of the field $\mathbf{H}_{g\gamma}$. Note also that

$$\frac{1}{k_1 - \kappa y_1} = \frac{1}{k_1}\left[1 + \frac{\kappa}{k_1} y_1 + O(k_1^{-1})\right]. \tag{3.7}$$

The Maxwell equations (2.7) lead to

$$k_g^{-2}\nabla \times \nabla \times \mathbf{E}_{g\gamma} - \mathbf{E}_{g\gamma} = 0 \qquad \text{and} \qquad k_g^{-2}\nabla \times \nabla \times \mathbf{H}_{g\gamma} - \mathbf{H}_{g\gamma} = 0. \tag{3.8}$$

Substitute the representations (3.4) into these equations. First notice that in the new coordinates the operation of double curl applied to a doubly differentiable vectorial function $\mathbf{D}$ takes the form

$$\begin{aligned}
\nabla \times \nabla \times \mathbf{D} = k_1^2 \Bigg\{ & \left[\frac{\partial^2 D_2}{\partial y_1 \partial y_2} - \frac{\partial^2 D_1}{\partial y_2^2} + \frac{1}{k_1 - \kappa y_1} \frac{\partial^2 D_3}{\partial y_1 \partial s} - \frac{1}{(k_1 - \kappa y_1)^2} \right. \\
& \left. \times \left(\frac{\partial^2 D_1}{\partial s^2} + \dot{\kappa}\, D_3 + \kappa \frac{\partial D_2}{\partial s} \right) - \frac{\dot{\kappa}\, y_1}{(k_1 - \kappa y_1)^3} \left(\frac{\partial D_1}{\partial s} + \kappa D_3 \right) \right] \hat{\mathbf{t}}_1 \\
& + \left[\frac{\partial^2 D_1}{\partial y_1 \partial y_2} - \frac{\partial^2 D_2}{\partial y_1^2} + \frac{1}{k_1 - \kappa y_1} \left(\frac{\partial^2 D_3}{\partial y_2 \partial s} + \kappa \left(\frac{\partial D_2}{\partial y_1} - \frac{\partial D_1}{\partial y_2} \right) \right) \right. \\
& \left. - \frac{1}{(k_1 - \kappa y_1)^2} \frac{\partial^2 D_2}{\partial s^2} - \frac{\dot{\kappa}\, y_1}{(k_1 - \kappa y_1)^3} \frac{\partial D_2}{\partial s} \right] \hat{\mathbf{t}}_2 \\
& + \left[-\frac{\partial^2 D_3}{\partial y_1^2} - \frac{\partial^2 D_3}{\partial y_2^2} + \frac{1}{k_1 - \kappa y_1} \left(\frac{\partial^2 D_1}{\partial y_1 \partial s} + \frac{\partial^2 D_2}{\partial y_2 \partial s} + \kappa \frac{\partial D_3}{\partial y_1} \right) \right. \\
& \left. + \frac{\kappa}{(k_1 - \kappa y_1)^2} \left(\frac{\partial D_1}{\partial s} + \kappa D_3 \right) \right] \hat{\mathbf{t}}_3 \Bigg\},
\end{aligned} \tag{3.9}$$

where the indices 1, 2 and 3 denote components of the function in the directions of t_1, t_2 and t_3 axes.

Having this in mind and with the use of (3.4)–(3.7) we find

$$\begin{aligned}
& k_1^{-2} e^{-ik_1 S(s)} \nabla \times \nabla \times \mathbf{E}_{g\gamma} \\
& = \left[\frac{\partial^2 \check{e}_{g\gamma 2}}{\partial y_1 \partial y_2} - \frac{\partial^2 \check{e}_{g\gamma 1}}{\partial y_2^2} + i \cos\beta_1 \frac{\partial \check{e}_{g\gamma 3}}{\partial y_1} + \cos^2\beta_1 \check{e}_{g\gamma 1} + O(k_1^{-1}) \right] \hat{\mathbf{t}}_1 \\
& + \left[\frac{\partial^2 \check{e}_{g\gamma 1}}{\partial y_1 \partial y_2} - \frac{\partial^2 \check{e}_{g\gamma 2}}{\partial y_1^2} + i \cos\beta_1 \frac{\partial \check{e}_{g\gamma 3}}{\partial y_2} + \cos^2\beta_1 \check{e}_{g\gamma 2} + O(k_1^{-1}) \right] \hat{\mathbf{t}}_2 \\
& + \left[-\frac{\partial^2 \check{e}_{g\gamma 3}}{\partial y_1^2} - \frac{\partial^2 \check{e}_{g\gamma 3}}{\partial y_2^2} + i \cos\beta_1 \left(\frac{\partial \check{e}_{g\gamma 1}}{\partial y_1} + \frac{\partial \check{e}_{g\gamma 2}}{\partial y_2} \right) + O(k_1^{-1}) \right] \hat{\mathbf{t}}_3.
\end{aligned} \tag{3.10}$$

Define the operator

$$\check{\nabla} = \left[\frac{\partial}{\partial y_1}, \frac{\partial}{\partial y_2}, i \cos\beta_1 \right]. \tag{3.11}$$

Since

$$\check{\nabla}\times\check{\nabla}\times\check{\mathbf{e}}_{g\gamma}=\left[\frac{\partial^2\check{e}_{g\gamma 2}}{\partial y_1\partial y_2}-\frac{\partial^2\check{e}_{g\gamma 1}}{\partial y_2^2}+i\cos\beta_1\frac{\partial\check{e}_{g\gamma 3}}{\partial y_1}+\cos^2\beta_1\check{e}_{g\gamma 1})\right]\hat{\mathbf{t}}_1$$
$$+\left[\frac{\partial^2\check{e}_{g\gamma 1}}{\partial y_1\partial y_2}-\frac{\partial^2\check{e}_{g\gamma 2}}{\partial y_1^2}+i\cos\beta_1\frac{\partial\check{e}_{g\gamma 3}}{\partial y_2}+\cos^2\beta_1\check{e}_{g\gamma 2}\right]\hat{\mathbf{t}}_2$$
$$+\left[-\frac{\partial^2\check{e}_{g\gamma 3}}{\partial y_1^2}-\frac{\partial^2\check{e}_{g\gamma 3}}{\partial y_2^2}+i\cos\beta_1\left(\frac{\partial\check{e}_{g\gamma 1}}{\partial y_1}+\frac{\partial\check{e}_{g\gamma 2}}{\partial y_2}\right)\right]\hat{\mathbf{t}}_3, \tag{3.12}$$

then (3.10) can be written down with the help of this operator in the form

$$k_1^{-2}e^{-ik_1S(s)}\nabla\times\nabla\times\mathbf{E}_{g\gamma}=\check{\nabla}\times\check{\nabla}\times\check{\mathbf{e}}_{g\gamma}+O(k_1^{-1}). \tag{3.13}$$

Substitution of (3.4) to the first equation in (3.8) leads to the equation satisfied by $\check{\mathbf{e}}_{g\gamma}$,

$$g^{-1}\check{\nabla}\times\check{\nabla}\times\check{\mathbf{e}}_{g\gamma}-\check{\mathbf{e}}_{g\gamma}+O(\varepsilon)=0. \tag{3.14}$$

Here, the relation $k_g/k_1=g$ was used.

Additionally, from the boundary condition $\hat{\mathbf{t}}_2\times\mathbf{E}_{g\gamma}=0$ it follows that

$$\hat{\mathbf{t}}_2\times\check{\mathbf{e}}_{g\gamma}+O(\varepsilon)=0 \tag{3.15}$$

holds on the screen.

Now we turn to diffraction of the plane wave (2.2) by the *half-plane* $t_1\geq 0$, $t_2=0$. We assume that the configuration made by the direction of the incident wave propagation and the half-plane edge is the same as that made by this direction and the tangent at the fixed point of diffraction $\mathbf{x}(s)$ to the curvilinear edge. Exact solution to this problem takes the form

$$\widetilde{\mathbf{E}}_{g\gamma}(\mathbf{t})=\widetilde{\mathbf{e}}_{g\gamma}(t_1,t_2)q(t_3)\qquad\text{and}\qquad\widetilde{\mathbf{H}}_{g\gamma}(\mathbf{t})=\widetilde{\mathbf{h}}_{g\gamma}(t_1,t_2)q(t_3), \tag{3.16}$$

where $q(t_3)=\exp[ik_1\cos\beta_1(s)t_3]$. Hence,

$$\frac{\partial\widetilde{\mathbf{E}}_{g\gamma}(\mathbf{t})}{\partial t_i}=k_1q(t_3)\frac{\partial\widetilde{\mathbf{e}}_{g\gamma}\left(\frac{y_1}{k_1},\frac{y_2}{k_1}\right)}{\partial y_i},\qquad i=1,2,$$
$$\frac{\partial\widetilde{\mathbf{E}}_{g\gamma}(\mathbf{t})}{\partial t_3}=ik_1\cos\beta_1(s)q(t_3)\widetilde{\mathbf{e}}_{g\gamma}. \tag{3.17}$$

Since ∇ can be written down as

$$\nabla=\left[\frac{\partial}{\partial t_1},\frac{\partial}{\partial t_2},\frac{\partial}{\partial t_3}\right]=k_1\left[\frac{\partial}{\partial y_1},\frac{\partial}{\partial y_2},i\cos\beta_1(s)\right]=k_1\check{\nabla}, \tag{3.18}$$

then by using the first equation in (3.8) we arrive at the following equaion satisfied by $\tilde{\mathbf{e}}_{g\gamma}$

$$g^{-1}\check{\nabla} \times \check{\nabla} \times \tilde{\mathbf{e}}_{g\gamma} - \tilde{\mathbf{e}}_{g\gamma} = 0. \tag{3.19}$$

On the half-plane surface this function satisfies the boundary condition $\hat{\mathbf{t}}_2 \times \tilde{\mathbf{e}}_{g\gamma} = 0$.

Comparison of the boundary value problems for the functions $\check{\mathbf{e}}_{g\gamma}$ (with fixed s) and $\tilde{\mathbf{e}}_{g\gamma}$ shows that to terms of $O(\varepsilon)$ both problems are the same. Therefore, by restricting our considerations to the leading terms in (3.14) and (3.15), we obtain $\check{\mathbf{e}}_{g\gamma} = \tilde{\mathbf{e}}_{g\gamma}$, and as a consequence,

$$\mathbf{E}_{g\gamma} = e^{ik_1[S(s)-\cos\beta_1(s)t_3]}\left[\tilde{\mathbf{E}}_{g\gamma}(\mathbf{t}) + O(\varepsilon)\right]. \tag{3.20}$$

Similar relation can be found for the magnetic field. Thus the procedure for the approximation of the diffracted field in the vicinity of a curved edge, as used in [1], has been confirmed with the considerations based on the boundary layer method. The solution (3.20) valid in the inner region (3.1) will be referred to as inner solution.

4. Matching of the inner and outer solutions

In order to find the diffraction coefficients for the diffracted waves we shall match inner and outer solutions in the common region of their validity. The outer solution, valid in the outer region $k_1\rho \gg 1$, is given by the sum (vide [1])

$$\begin{Bmatrix} \mathbf{E}_{g\gamma}(\mathbf{x}) \\ \mathbf{H}_{g\gamma}(\mathbf{x}) \end{Bmatrix} = \sum_{\gamma=e,m} \sum_{\beta=i,r1,r2,z,d,b} \begin{Bmatrix} \mathbf{E}^{\beta}_{g\gamma}(\mathbf{x}) \\ \mathbf{H}^{\beta}_{g\gamma}(\mathbf{x}) \end{Bmatrix}, \tag{4.1}$$

where the superscript β indicates the species of waves present in the problem: the incident wave, the wave reflected from the screen and that reflected from the media interface, the refracted wave, the diffracted and lateral waves. Asymptotic expansions for these waves as $k_g \to \infty$ are of the form

$$\begin{Bmatrix} \mathbf{E}^{\beta}_{g\gamma}(\mathbf{x}) \\ \mathbf{H}^{\beta}_{g\gamma}(\mathbf{x}) \end{Bmatrix} \sim \delta^{\beta}_{g}(\mathbf{x}) q^{\beta}(k_g) e^{ik_g S^{\beta}_{g}(\mathbf{x})} \sum_{m=0}^{\infty} \begin{Bmatrix} \mathbf{e}^{\beta}_{g\gamma}(\mathbf{x}) \\ \mathbf{h}^{\beta}_{g\gamma}(\mathbf{x}) \end{Bmatrix} (ik_g)^{-m}, \tag{4.2}$$

where $\delta^{\beta}_{g}(\mathbf{x})$ equals $+1$ in the illuminated region of a particular wave or -1 in its shadow region, and $q^{\beta}(k_g)$ determines its order of magnitude. In particular, the dominant term in the diffracted wave is of the form([5])

$$\mathbf{E}^{d}_{g}(\mathbf{x}) \sim \frac{e^{ik_g S^{d}_{g}}}{\left[k_g \sigma^{d}_{g}\left(1 + \dfrac{\sigma^{d}_{g}}{\rho^{d}_{g}}\right)\right]^{1/2}} \mathbf{D}^{d}_{g}(\varphi, s), \qquad S^{d}_{g} = \frac{k_1}{k_g} S(s^{d}_{g}) + c^{d}_{g}. \tag{4.3}$$

([5]) This is a result of solving the eikonal and transport equations on the diffracted rays.

The vectorial function $\mathbf{D}_g^d(\varphi, s)$ appearing here is the diffraction coefficient of the diffracted wave, which is to be found.

The quantities $\sigma_g^d = R\csc\beta_g(s)$, φ and s are the ray coordinates, specifying a diffracted ray emanating from a diffraction point on the edge and running toward the observation point. The angle $\beta_N(s)$ is related to the angle $\beta_1(s)$ via $k_1\cos\beta_1 = k_N\cos\beta_N$, and R, φ and t_3 are cylindrical coordinates attached at the diffraction point $\mathbf{x}_0(s)$.

The inner and outer solutions are asymptotic ones, valid for large values of k_1. Each solution, however, has a different form and is defined in a different region. Nevertheless, by virtue of (2.8), both solutions have their common region of validity

$$k_1^{-1} \ll \rho \leq \varepsilon/|\kappa|, \tag{4.4}$$

where they can be matched. To this end we expand asymptotically the inner solution with respect to the large parameter $k_1\rho$ (for instance by using the technique of asymptotic evaluation of integrals described in [1]). We thus arrive at the outer expansion of the inner solution. The term corresponding to the diffracted field in this expansion appears to be of the form

$$\begin{aligned}\mathbf{E}_g^d(\mathbf{x}) = -\eta_g\left(\frac{\pi}{2k_g\sin\beta_g\rho}\right)^{1/2} e^{ik_1S(s)+ik_g\sin\beta_g\rho+i\frac{3\pi}{4}} \\ \times\left[\sum_{j=0}^{2}\left\{\left(\csc\frac{\phi-\phi_{jg}}{2}+\csc\frac{\phi+\phi_{jg}}{2}\right)\frac{P_g[w(\phi)]}{P_g[w(\phi_{jg})]}\mathbf{c}_g(\phi)k_g^2Z_gp_j\right.\right. \\ \left.\left.-\left(\csc\frac{\phi-\phi_{jg}}{2}-\csc\frac{\phi+\phi_{jg}}{2}\right)\frac{M_g[w(\phi)]}{M_g[w(\phi_{jg})]}\mathbf{d}_g(\phi)\frac{\eta_gq_j^g}{v_g(w_j)}\right\}\right. \\ \left.+O[(k_g\rho)^{-3/2}]+O(\varepsilon)\right] \\ = e^{ik_g\tilde{S}_g^d}\left\{k_g^{-1/2}\mathbf{E}_{g0}^d[\mathbf{t};\phi^i(s),\beta_1(s)]+O[(k_g\rho)^{-3/2}]\right\}.\end{aligned} \tag{4.5}$$

Here

$$\tilde{S}_g^d(\mathbf{x}) = \frac{k_1}{k_g}S(s)+\sin\beta_g(s)\rho, \tag{4.6}$$

$\mathbf{E}_{g0}^d$ is the principal amplitude of diffracted TE and TM fields in the outer expansion of the inner solution, and

$$\mathbf{c}_g = \hat{\mathbf{y}}\times\nabla S_g^d, \qquad \mathbf{d}_g = \left[\nabla S_g^d\times(\hat{\mathbf{y}}\times\nabla S_g^d)\right]. \tag{4.7}$$

We shall now determine the inner expansion of the outer solution. It can be shown that in the inner region, $R = \rho + O(\varepsilon)$, $\varphi = \phi + O(\varepsilon)$, $\tilde{S}_g^d(\mathbf{x}) = S_g^d(\mathbf{x}) +$

$O(\varepsilon)$, $\sin\beta_g(s_g^d) = \sin\beta_g(s) + O(\varepsilon)$ and $\sigma_g^d/\rho_g^d = O(\varepsilon)$. By using these asymptotic relations in (4.3) we find that the inner expansion of the outer solution is

$$\mathbf{E}_g^d \sim \frac{e^{ik_g S_g^d}}{k_g^{1/2}\left[\dfrac{\rho}{\sin\beta_g(s)}\right]^{1/2}} \mathbf{D}_g^d(\phi, s)\left\{1 + O[(k_g\rho)^{-1}] + O(\varepsilon)\right\}. \tag{4.8}$$

On matching the expansions (4.7) and (4.8) we obtain

$$\mathbf{D}_g^d(\phi, s) = \left(\frac{\rho}{\sin\beta_g(s)}\right)^{1/2} \mathbf{E}_{g0}^d[\mathbf{t}; \phi^i(s), \beta_1(s)]. \tag{4.9}$$

By taking into account the asymptotic relations between both ρ and R and ϕ and φ, we finally arrive at the following expression for the diffraction coefficient of the diffracted waves

$$\begin{aligned}\mathbf{D}_g^d(\varphi, s) = {} & \left(\frac{\pi}{2}\right)^{1/2} \frac{e^{-i\frac{\pi}{4}}}{\sin\beta_g(s)} \\ & \times \sum_{j=0}^{2} \left\{ \left(\csc\frac{\varphi - \varphi_{jg}}{2} + \csc\frac{\varphi + \varphi_{jg}}{2}\right) \frac{P_g[w(\varphi)]}{P_g(w_j)} \mathbf{c_g}(\varphi) k_g^2 Z_g p_j \right. \\ & \left. - \left(\csc\frac{\varphi - \varphi_{jg}}{2} - \csc\frac{\varphi + \varphi_{jg}}{2}\right) \frac{M_g[w(\varphi)]}{M_g(w_j)} \mathbf{d_g}(\varphi) \frac{\eta_g q_j^g}{v_{gj}} \right\},\end{aligned} \tag{4.10}$$

we sought. Here,

$$\begin{aligned} w(\varphi) &= k_g \sin\beta_g \cos\varphi, \qquad v_{g0} = [(k_g \sin\beta_g)^2 - w_0^2]^{1/2}, \\ w_0 &= k_1 \sin\beta_1 \cos\varphi_i, \\ w_1 &= -w_2 = ik_1 \cos\beta_1, \quad v_{g1} = v_{g2} = k_g. \end{aligned} \tag{4.11}$$

The diffraction coefficient $\mathbf{D}_g^d(\varphi, s)$ comprises both TM and TE fields. Substitution of this coefficient into (4.3) concludes the construction of the diffracted waves on families of their rays.

References

1. A. CIARKOWSKI, *Electromagnetic wave diffraction by screens residing on interfaces between dielectric media* [in Polish], Report 4/1994, Inst. Fund. Technol. Research, PAS, 1994.
2. A. CIARKOWSKI, *On the field continuity condition in the problem of electromagnetic plane wave diffraction by a half-plane residing between two media*, Wave Motion, **18**, 227–241, 1993.
3. R.N. BUCHAL and J.B. KELLER, *Boundary layer problems in diffraction theory*, Comm. Pure Appl. Math., XIII, 85–114, 1960.

4. H. WITLOX, *High-frequency diffraction of elastic waves by a plane surface*, Master's Thesis, Eindhoven University of Technology, Dept. Math. and Comp. Sci., 1983.
5. K. BOCHENEK, *Methods of analysis of electromagnetic fields* [in Polish], PWN, Warszawa – Wrocław 1961.
6. D. AHLUWALIA, R. LEWIS and J. BOERSMA, *Uniform asymptotic theory of diffraction by a plane screen*, SIAM J. Appl. Math., **16**, 783–807, 1968.

POLISH ACADEMY OF SCIENCES
INSTITUTE OF FUNDAMENTAL TECHNOLOGICAL RESEARCH.

Received December 30, 1993; *new version September* 9, 1994.

Arch. Mech., 46, 6, pp. 965–988, Warszawa 1994

On the kinematics of the sets of oriented elements

A. BLINOWSKI (WARSZAWA)

A DESCRIPTION of the orientations of the objects in 3-dimensional Euclidean space using the orthogonal tensors is outlined. Local and global geometric properties of the manifold of orientations (tensors of rotation) are discussed. Functions and differential operations on the manifold of orientations are considered, particular attention being paid to the density of orientation (orientation distribution) function. The problems of averaging of the orientation are briefly discussed. The time derivatives of the rotation tensors are examined, their relation with the angular velocity vector being presented. Eulerian and Lagrangian rates of changes of the scalar functions and tangent vectors are defined on the manifold. The local differential law of mass conservation on the manifold is derived and verified for the important example of rigid motion. Co-rotational rates of the Euclidean tensor-valued function on the manifold are discussed as well.

1. Preliminaries

THE NECESSITY of appropriate kinematical description of the behaviour of sets of oriented elements arises almost at every attempt to produce any continual model of the material, particularly in these which are motivated by the considerations on the micro- and/or meso-structural level. The problem of texture evolution in plasticity of polycrystalline materials [2], [3] as well as a complete description of the configuration spaces in thermodynamics of structured polymers [20] can be pointed out here.

Every author dealing with his particular problem solves this problem in his own way, thus it is sometimes difficult not only to verify the correctness of the results, but even to compare the results obtained by different authors. Some results are rather assumed than derived. Even if the particular presumptions were correct (as they probably are) the others could be erroneous. This applies mainly to the descriptions using Euler angles which, in the cases different from the rotation around nearly constant axes with nearly constant angular velocities (like in most rigid body rotation problems [9]), definitely obscure the problems. The intuitive comprehension of the relations derived become almost impossible. Thus even a plain nonsense may be not so easy to detect as it should be using more natural coordinate systems. Moreover, Euler angles taken as the coordinates in the Riemannian space of orientations yield non-diagonal metric tensor, drastically complicating all differential operations [6, 3, 20].

The goal of the present paper consists in the attempt of the unification and simplification of the description of local and global geometry and kinematics of the manifold of orientations. There are several results, among the others that will be quoted here, which the author was not able to find in the earlier sources; he

is however firmly convinced that all of them are known, likely, they simply have not been gathered together yet.

In the course of the considerations which follow, we shall keep in mind as a distant aim the description of the behaviour of structured continuous media. At this stage, however, we have no need to specify neither the *physical nature* of the *oriented* elements under consideration nor even their *spatial localisation*. They can be *gathered together* as, e.g. the set of single crystals constituting the material neighbourhood of the chosen point in polycrystalline material (whatever it exactly means), or they can be quite *dispersed apart* as, for example, the set of honeybees from one hive during a sunny day. It is *the orientation not the localisation* what we are trying to describe.

We shall confine our consideration to such elements for which at any instant t, a set of three orthonormal vectors $\{\boldsymbol{\varepsilon}_i(t)\}$ can be distinctly pointed out, i.e., speaking about *orientation* of the material object we may understand under this term any geometric object capable to specify the actual *directions of three mutually orthogonal vectors*. An *orthogonal tensor* is, in author's opinion, the most appropriate object for this aim – this one, (denoted by $\mathbf{Q}(t)$), which transforms the chosen fixed three orthonormal vectors $\{\mathbf{e}_i\}$ into another orthonormal base vectors $\{\boldsymbol{\varepsilon}_i(t)\}$, describing the orientation of the element under consideration:

$$\boldsymbol{\varepsilon}_i(t) = \mathbf{Q}(t)\mathbf{e}_i \,. \tag{1.1}$$

Therefore we shall say that *the orientation of all elements* $\alpha \in A$ *of the* (finite, countable or continuous) *set* A *is given at any time instant* t *if some* (orthogonal) *tensor-valued function* $\mathbf{Q}(\alpha, t)$ *is specified on the set* A.

In further considerations we shall assume that the determinant of $\mathbf{Q}$ is positive, excluding thus from the considerations all (inappropriate for the description of the behaviour of real material objects) transformations with reflections (inversions).

2. The space of rotations

As it has been pointed out in the previous section, the set of all possible orientations of the material elements can be represented equivalently by the set of all tensors of rotation (the proper orthogonal group).

In the properly chosen orthonormal basis, for which the unit vector of the axis of rotation $\mathbf{n}$ is taken as the third base vector, the representation of any tensor of rotation assumes the following form [7]

$$\mathbf{Q} = \begin{bmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{2.1}$$

where φ is the angle of rotation of the vectors orthogonal to the axis of rotation. It is not difficult to notice that the form of this matrix does not depend on the

choice of the first and the second base vector in the plane of rotation. Taking into account that the determinant of the tensor **Q** is positive (equal to one), it is evident that any tensor of the form (2.1) can be expressed in the absolute "dyadic" notation as follows

$$\mathbf{Q} = \mathbf{n} \otimes \mathbf{n} + (\mathbf{1} - \mathbf{n} \otimes \mathbf{n}) \cos\varphi + \boldsymbol{\varepsilon}\mathbf{n} \sin\varphi, \tag{2.2}$$

where **1** is the unit tensor, $\boldsymbol{\varepsilon}$ denotes Levi-Civitta skew-symmetric tensor. Representation of tensor **Q** in arbitrary orthonormal basis takes the following form:

$$Q_{ij} = n_i n_j + (\delta_{ij} - n_i n_j) \cos\varphi + n_k \varepsilon_{kij} \sin\varphi. \tag{2.3}$$

Expressing the unit normal vector **n** as follows:

$$\mathbf{n} = \sin\theta \cos\psi \mathbf{e}_1 + \sin\theta \sin\psi \mathbf{e}_2 + \cos\theta \mathbf{e}_3\,, \tag{2.4}$$

and examining simultaneously expressions (2.3) and (2.4), one easily concludes that the set of rotation tensors is a smooth three-parametric compact subset of the nine-dimensional space of the second rank tensors (the tensorial "square" of the three-dimensional Euclidean vector space $\mathbf{V}^3 \otimes \mathbf{V}^3$). These three parameters $(\psi, \theta, \varphi,)$ will be used as the coordinates at the three-dimensional manifold of the rotation tensors (or – what is equivalent – the manifold of orientations).

For the further considerations we shall need the definition of the scalar product of the second rank tensors. We shall adopt the simplest and most standard one:

$$\mathbf{A} \circ \mathbf{B} \equiv \operatorname{tr}\left(\mathbf{A}\mathbf{B}^{\mathrm{T}}\right) = A_{ij} B_{ij}. \tag{2.5}$$

This product bears the following norm in nine-dimensional space:

$$\|A\| \equiv \left(\operatorname{tr} \mathbf{A}\mathbf{A}^{\mathrm{T}}\right)^{1/2} = \sqrt{A_{ij} A_{ij}}\,. \tag{2.6}$$

One can easily see that, in the sense of this norm, the submanifold under consideration is a subset of eight-dimensional hypersphere of the radius $\sqrt{3}$. It is also not difficult to show that the manifold of rotations cannot be embedded in any Euclidean proper subspace of the nine-dimensional space. If it were not true then it would exist a non-trivial (different from zero) tensor being orthogonal to all rotation tensors. Let us take arbitrary tensor **T** and choose the basis of eigenvectors of the symmetric part of **T**. In such a basis for all non-diagonal terms we have $T_{ij} = -T_{ji}$ $(i \neq j)$. We can immediately point out representations of such six rotation tensors that bear the linear combination equal to **T**:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

If all these tensors were orthogonal to $\mathbf{T}$ then we would have $T_{ij} = 0$ $(i = j)$ for diagonal terms and $T_{ij} = T_{ji}$ $(i \neq j)$ for the others, but this is possible only if these term also vanish. This proves our assertion. Other details concerning the global geometry of the manifold will be discussed later.

We shall examine now the local geometric properties. To this end we need the derivatives of the vector $\mathbf{n}$ with respect to ψ and θ. Differentiating (2.4) we obtain the following relations:

$$\frac{\partial \mathbf{n}}{\partial \psi} = -\sin\theta \sin\psi \mathbf{e}_1 + \sin\theta \cos\psi \mathbf{e}_2 = \sin\theta \mathbf{e}_\psi , \tag{2.7}$$

$$\frac{\partial \mathbf{n}}{\partial \theta} = \cos\theta \cos\psi \mathbf{e}_1 + \cos\theta \sin\psi \mathbf{e}_2 - \sin\theta \mathbf{e}_3 = \mathbf{e}_\theta . \tag{2.8}$$

The meaning of the two unit vectors $\mathbf{e}_\psi$ and $\mathbf{e}_\theta$ (tangent to the unit sphere in 3-dimensional space) is clear from the context of the Eq. (2.7) and (2.8).

For clarity we shall introduce the following convention: we shall sometimes treat all vectors from nine-dimensional space $\mathbf{V}^3 \otimes \mathbf{V}^3$ like any other vectors, forgetting about their tensorial nature, thus in order to avoid the misunderstandings we have to distinguish between these objects and the usual "three-dimensional" vectors. Thus, we shall denote them by capital boldfaced symbols and, for the amplification of this distinction, capital letters will be used also for indices. Writing the representations of nine-dimensional vectors in the three-dimensional bases of the subspaces of the nine-dimensional space, we shall use single (capital) indices, while the representations of the same objects treated as tensors from $\mathbf{V}^3 \otimes \mathbf{V}^3$ will be written using pairs of indices.

Differentiation of the second order tensors with respect to scalar parameters (coordinates) bears of course objects of the same kind. Making use of relations (2.7) and (2.8) and denoting parameters $\{\psi, \theta, \varphi\}$ by $\{\xi^1, \xi^2, \xi^3\}$, respectively, we are able to obtain the following relations for the vectors (from nine-dimensional space!), tangent to the three-parametric manifold of orientations [16]:

$$\mathbf{G}_1 \equiv \frac{\partial \mathbf{Q}}{\partial \psi} = (\mathbf{n} \otimes \mathbf{e}_\psi + \mathbf{e}_\psi \otimes \mathbf{n})(1 - \cos\varphi)\sin\theta - \varepsilon \mathbf{e}_\psi \sin\varphi \sin\theta, \tag{2.9}$$

$$\mathbf{G}_2 \equiv \frac{\partial \mathbf{Q}}{\partial \theta} = (\mathbf{n} \otimes \mathbf{e}_\theta + \mathbf{e}_\theta \otimes \mathbf{n})(1 - \cos\varphi) - \varepsilon \mathbf{e}_\theta \sin\varphi, \tag{2.10}$$

$$\mathbf{G}_3 \equiv \frac{\partial \mathbf{Q}}{\partial \varphi} = -(\mathbf{1} - \mathbf{n} \otimes \mathbf{n})\sin\varphi - \varepsilon \mathbf{n} \cos\varphi. \tag{2.11}$$

These three vectors $\mathbf{G}_K$ constitute a basis of the three-dimensional space $\mathbf{T}^3$, tangent to the manifold at a given point (for given values of the coordinates). In the

space $\mathbf{T}^3$ a scalar product can be introduced giving rise to its Euclidean structure; it is quite natural to use the same definition (2.5) as in the nine-dimensional space containing $\mathbf{T}^3$ as a subspace. Space $\mathbf{T}^3$, although isomorphic to $\mathbf{V}^3$ space, should not be confused with the latter; we shall point out later at least two important reasons for this.

The adopted definition of the scalar product operation makes it possible to define covariant representation G_{IJ} of the metric tensor $\mathbf{G}$ at the manifold:

$$
\begin{aligned}
\mathbf{G}_1 \circ \mathbf{G}_1 &= G_{11} = 4\sin^2\theta(1-\cos\varphi), \\
\mathbf{G}_2 \circ \mathbf{G}_2 &= G_{22} = 4(1-\cos\varphi), \\
\mathbf{G}_3 \circ \mathbf{G}_3 &= G_{33} = 2, \\
\mathbf{G}_I \circ \mathbf{G}_K &= G_{IK} = 0 \qquad \text{for} \quad I \neq K.
\end{aligned} \tag{2.12}
$$

Due to the diagonality of the metric tensor, the expressions for co-vectors and the contravariant representation of the metric tensor are almost trivial:

$$
\left.
\begin{aligned}
G^{II} &= \frac{1}{G_{II}}, \qquad G^{IK} = 0 \qquad \text{for} \quad I \neq K \\
\mathbf{G}^K &= \mathbf{G}_K G^{KK}
\end{aligned}
\right\} \quad \text{no summation!} \tag{2.13}
$$

We shall not specify these terms explicitly here. It is worthwhile, however, to quote the representations of the base vectors $\mathbf{G}^K$ (considered as the second order tensors) using the basis $\{\mathbf{n}, \mathbf{e}_\psi, \mathbf{e}_\theta\}$ of the 3-dimensional "physical" space $\mathbf{V}^3$.

$$
\begin{aligned}
\mathbf{G}_1 &= 2\sin\theta\sin\frac{\varphi}{2}\begin{bmatrix} 0 & \sin\frac{\varphi}{2} & \cos\frac{\varphi}{2} \\ \sin\frac{\varphi}{2} & 0 & 0 \\ -\cos\frac{\varphi}{2} & 0 & 0 \end{bmatrix}, \\
\mathbf{G}_2 &= 2\sin\frac{\varphi}{2}\begin{bmatrix} 0 & -\cos\frac{\varphi}{2} & \sin\frac{\varphi}{2} \\ \cos\frac{\varphi}{2} & 0 & 0 \\ \sin\frac{\varphi}{2} & 0 & 0 \end{bmatrix}, \\
\mathbf{G}_3 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\sin\varphi & -\cos\varphi \\ 0 & \cos\varphi & -\sin\varphi \end{bmatrix}.
\end{aligned} \tag{2.14}
$$

Taking into account that for the basis under consideration the tensor $\mathbf{Q}$ (this "point" of the manifold for which we determined the tangent vectors $\mathbf{G}_K$) has the

particular representation (2.1), one can rewrite expressions (2.14) in the following form

$$\mathbf{G}_K = -\mathbf{W}^{(K)}\mathbf{Q} \qquad (K = 1,2,3), \tag{2.15}$$

where $\mathbf{W}^{(K)}$ are a skew-symmetric second rank tensors (from $\mathbf{V}^3 \otimes \mathbf{V}^3$). The product operation at the right-hand side of the last relation has the usual sense of matrix multiplication (simple contraction) $(W^{(K)}_{ij}Q_{jl})$. For $\mathbf{W}^{(K)}$ the following expressions hold true:

$$\begin{aligned} \mathbf{W}^{(1)} &= \sin\theta\boldsymbol{\varepsilon}\,[\mathbf{e}_\psi \sin\varphi - \mathbf{e}_\theta(1-\cos\varphi)]\,, \\ \mathbf{W}^{(2)} &= \boldsymbol{\varepsilon}\,[\mathbf{e}_\psi(1-\cos\varphi) + \mathbf{e}_\theta \sin\varphi]\,, \\ \mathbf{W}^{(3)} &= \boldsymbol{\varepsilon}\mathbf{n}. \end{aligned} \tag{2.16}$$

Relations (2.15) and (2.16) suggest the natural mapping of the tangent space onto the Euclidean three-dimensional "physical" space $\mathbf{V}^3$, according to the following scheme:

$$\mathbf{S} \to \mathbf{V}^3 : \mathbf{A} \to \boldsymbol{\varepsilon} : \left(\mathbf{A}\,\mathbf{Q}^{\mathrm{T}}\right) \quad \left(\mathbf{A} \to \varepsilon_{pqr}A_{qs}Q_{rs}\mathbf{e}_p \text{ in orthogonal base}\right), \tag{2.17}$$

where by A_{qs} we denote tensorial representation of the tangent vector $\mathbf{A}$. It is not difficult to see that such a mapping can be defined for all points of the manifold, moreover this mapping is linear and preserves the structure of the Euclidean linear space. Nevertheless one should not identify the tangent space with the "physical" Euclidean space – the main difference is based upon the fact that the mapping depends on the "radius-vector" $\mathbf{Q}$ (i.e. on the coordinates on the manifold), and therefore there is no decomposition of the vectors from the tangent space corresponding to one radius-vector in the basis of the tangent space belonging to different $\mathbf{Q}$ (different radius-vector). It can be easier to accept this fact intuitively, if one realises that the tangent vectors corresponding to the radius-vector $\mathbf{Q}$ are orthogonal to it in the sense of scalar product in nine-dimensional space (compare (2.1) and (2.14)). The comprehension of these differences is quite important for everyone who wishes to avoid the fatal errors when differentiating the vector (tensor) fields defined on the manifold.

Calculating the Riemann–Christoffel symbols according to well-known formula([1]) [7], [16]:

$$\Gamma^K_{IJ} = \frac{1}{2}G^{KL}\left(\frac{\partial G_{LI}}{\partial \xi^J} + \frac{\partial G_{LJ}}{\partial \xi^I} - \frac{\partial G_{IJ}}{\partial \xi^L}\right), \tag{2.18}$$

([1]) Using nothing but the well-known definitions $G_{IK} \equiv \mathbf{G}_I \circ \mathbf{G}_K$, $\mathbf{G}_I \circ \mathbf{G}^K \equiv \delta^K_I$, etc., one can readily rewrite (2.18) in the following, equivalent form:

$$\Gamma^I_{KL} = \frac{\partial \mathbf{G}_K}{\partial \xi^L} \circ \mathbf{G}^I = -\frac{\partial \mathbf{G}^I}{\partial \xi^L} \circ \mathbf{G}_K\,. \tag{2.18}'$$

This interpretation will be sometimes useful in the course of the forthcoming considerations.

we arrive at the following expressions for these coefficients of Riemannian connection which are not equal to zero:

$$
\begin{aligned}
&\Gamma^1_{12} = \operatorname{ctg}\theta, \qquad \Gamma^1_{13} = \Gamma^2_{23} = \frac{1}{2}\operatorname{ctg}\frac{\varphi}{2}, \\
&\Gamma^2_{11} = -\sin\theta\cos\theta, \\
&\Gamma^3_{22} = -\sin\varphi, \qquad \Gamma^3_{11} = -\sin^2\theta\sin\varphi.
\end{aligned} \tag{2.19}
$$

Knowing the representation of metric tensor and the Riemann-Christoffel coefficients one can calculate the components of the tensor of curvature according to the following relation:

$$
\begin{aligned}
R_{LK,IJ} = \frac{1}{2}\left(\frac{\partial^2 G_{LJ}}{\partial\xi^K\partial\xi^I} - \frac{\partial^2 G_{LI}}{\partial\xi^K\partial\xi^J} + \frac{\partial^2 G_{KJ}}{\partial\xi^L\partial\xi^I} - \frac{\partial^2 G_{KI}}{\partial\xi^L\partial\xi^J}\right) \\
+ G_{PQ}\left(\Gamma^P_{LJ}\Gamma^Q_{KI} - \Gamma^P_{KJ}\Gamma^Q_{LI}\right).
\end{aligned} \tag{2.20}
$$

Performing differentiation one obtains:

$$
\begin{aligned}
&R_{12,12} = 2\sin^2\theta(1-\cos\varphi)^2, \\
&R_{13,13} = \sin^2\theta(1-\cos\varphi), \\
&R_{12,12} = (1-\cos\varphi).
\end{aligned} \tag{2.21}
$$

All other components are either equal to zero, or can be obtained from the symmetry conditions, which must be obeyed by any tensor of curvature:

$$
R_{IJ,KL} = -R_{JI,KL} = R_{JI,LK} = -R_{JI,KL} \ldots . \tag{2.22}
$$

Bearing in mind relations (2.12) one can write the following important identity:

$$
R_{PQ,RS} = \frac{1}{8}\left(G_{PR}G_{QS} - G_{PS}G_{QR}\right). \tag{2.23}
$$

Relation (2.23) defines the constant curvature K of Riemannian space [16, 19] where

$$
K = 1/8. \tag{2.24}
$$

The quantity r defined by the following relation

$$
r = \frac{1}{\sqrt{K}} \tag{2.25}
$$

is known as a radius of curvature.

Another tensor of curvature called Ricci or Einstein tensor can be obtained from the following relation:

$$R_{IK} = G^{PQ} R_{IP,QK} . \tag{2.26}$$

In the case under consideration we obtain:

$$R_{IK} = -\frac{1}{4} G_{IK} . \tag{2.27}$$

The following quantity R:

$$R = G^{IK} R_{IK} \tag{2.28}$$

is called the scalar curvature, in our case

$$R = -3/4. \tag{2.29}$$

Both curvatures are scalar invariants characterising the deviation of local geometric properties of the manifold from the Euclidean ones, for Euclidean spaces both are equal to zero (in general, these quantities can depend on the coordinates).

It has been proved [19], [13] that all n-dimensional spaces of constant curvature are isometric to the factor-spaces S^n/Γ (sets of orbits with respect to Γ [13, 11]), where S^n is a n-dimensional sphere and Γ denotes a discontinuous group of isometry acting freely (having no fixed points). Such spaces, in the case of positive **K**, are called spherical space forms [13].

Closer examination of such subsets (c.f. [14]) of the manifold of the orthogonal transformations seems to be promising from the viewpoint of the description of the spaces of orientations for the oriented elements having non-trivial symmetries, e.g. crystal lattices. It is of course a well recognised item (among the people involved in the texture evolution problems), that, e.g. for the case of cubic or hexagonal crystals the set of all rotations is too wide. Nevertheless it seems that far not always, as it has been pointed out in [4], proper clarity reigns in the considerations concerning that point. We shall, however, for the time being, leave this question open for further studies.

The simplest spherical form, called elliptic space (projective sphere), is generated by the two-element group Γ consisting of the inversion and identity transformations. Orbits of such a group are two-point sets of opposite poles of the sphere. One can perceive such a form as the sphere with opposite points identified.

Elliptic space is the only – besides the full sphere – spherical space form preserving maximal homogeneity, i.e. possessing the group of isometric transformations of the maximum possible number of parameters: $n(n+1)/2$, in our case — 6 [13, 19]. One can point out easily the six-parametric group of isometries of the space of rotations, e.g. the one generated by two three-parametric subgroups:

left translations $(\mathbf{Q}) : \mathbf{Q} \to \mathbf{L}\mathbf{Q}$, where $\mathbf{L}$ is any rotation tensor, and rotations $\mathbf{R}(\mathbf{Q}) : \mathbf{Q} \to \mathbf{R}\mathbf{Q}\mathbf{R}^{\mathrm{T}}$, where $\mathbf{R}$ is any orthogonal tensor. Another possible isometric transformations like right translations can be obtained as the superpositions of the former two. On the other hand, we have essential reasons to assert that the form under consideration is not a sphere, namely: the orthogonal elements of opposite sign $-\mathbf{Q}$ are not rotations and do not belong to the manifold under consideration.

Owing to the simplicity of the manifold, we are able to point out some features of its global geometry. Let us consider for this aim the equations of geodesics

$$\frac{d^2\xi^K}{ds^2} = \Gamma^K_{IJ}\frac{d\xi^I}{ds}\frac{d\xi^J}{ds} = 0. \tag{2.30}$$

Using (2.19) one obtains:

$$\begin{aligned}
\frac{d^2\psi}{ds^2} &= -\cot\theta\frac{d\psi}{ds}\frac{d\theta}{ds} - \frac{1}{2}\cot\frac{\varphi}{2}\left(\frac{d\psi}{ds}\right)^2, \\
\frac{d^2\theta}{ds^2} &= -\frac{1}{2}\cot\frac{\varphi}{2}\frac{d\theta}{ds}\frac{d\varphi}{ds} + \sin\theta\cos\theta\left(\frac{d\theta}{ds}\right)^2, \\
\frac{d^2\varphi}{ds^2} &= \sin\varphi\left(\frac{d\theta}{ds}\right)^2 + \sin^2\theta\sin\varphi\left(\frac{d\psi}{ds}\right)^2.
\end{aligned} \tag{2.31}$$

On the other hand we have:

$$(ds)^2 = G_{KL}d\xi^K d\xi^L. \tag{2.32}$$

Thus, taking

$$\theta = \text{const}, \qquad \psi = \text{const}, \qquad |\varphi| = \frac{s}{\sqrt{2}} = 2\frac{s}{r}, \tag{2.33}$$

we fulfil equations (2.31) defining, at the same time, along the geodesic line, the length parameter s congruent with the local metric.

One can see now that any orbit of the identity transformation $\mathbf{I}$ with respect to the subgroup of the rotations around the fixed axis

$$\mathbf{Q}(s) = \mathbf{R}(s)\mathbf{I}, \tag{2.34}$$

where

$$\begin{aligned}
\mathbf{R} &= (\mathbf{a}\otimes\mathbf{a})\left(1 - \cos\frac{2s}{r}\right) + \mathbf{I}\cos\frac{2s}{r} + \mathbf{a}\boldsymbol{\varepsilon}\sin\frac{2s}{r}, \\
\mathbf{a} &= \text{const}, \qquad |\mathbf{a}| = 1
\end{aligned} \tag{2.35}$$

defines a geodesic line.

In virtue of the uniformity of the space we can claim that an orbit of any element with respect to arbitrarily chosen subgroup of rotations around the fixed axis defines some geodesic line. Taking into account the known fact of stationarity (in this case minimality) of the path along the geodesic lines, one can define a global notion of the distance between two orientations.

Let two orientations $\mathbf{Q}_1$ and $\mathbf{Q}_2$ be given; we shall denote by $\mathbf{R}$ the rotation tensor transforming by the left-hand translation the orientation $\mathbf{Q}_1$ into $\mathbf{Q}_2$:

$$\mathbf{R} = \mathbf{Q}_2\mathbf{Q}_1^{\mathrm{T}}. \tag{2.36}$$

Let us denote by α the angle of rotation associated with the orthogonal tensor $\mathbf{R}$([2]). Then we have

$$\cos\alpha = \frac{\operatorname{tr}\mathbf{Q} - 1}{2}, \tag{2.37}$$

thus, according to (2.33), we can write

$$s = \sqrt{2}\arccos\left(\frac{\operatorname{tr}\left(\mathbf{Q}_2\mathbf{Q}_1^{\mathrm{T}}\right) - 1}{2}\right). \tag{2.38}$$

For small angles this definition of the distance between the two orientations is equivalent to the global metrics in nine-dimensional space of the second rank tensors.

$$\lim_{\alpha\to 0}\frac{|\mathbf{R}\,\mathbf{Q} - \mathbf{Q}|}{\sqrt{2}\arccos\left(\dfrac{\operatorname{tr}\left(\mathbf{Q}_2\mathbf{Q}_1^{\mathrm{T}}\right) - 1}{2}\right)} = \lim_{\alpha\to 0}\frac{\sqrt{6 - 2(1 + \cos\alpha)}}{\sqrt{2}\alpha} = 1. \tag{2.39}$$

One can evaluate the maximal possible distance between two orientations, corresponding to the rotation angle equal to π:

$$d_{\max} = \sqrt{2}\pi = \frac{\pi r}{2}, \tag{2.40}$$

while the maximal possible distance between two rotation tensors in nine-dimensional space equals $2\sqrt{2}$. It is not difficult to notice that one can not point out uniquely the element most distant from the given one, there are many of them; the set of all most distant elements is often called the polar plane.

([2]) Note please that this is not the angle of the rotation of the tensors in three-dimensional Euclidean space. Indeed, let the orthogonal tensor $\mathbf{P}$ generate a rotation, mapping $\mathbf{Q}_1$ onto $\mathbf{Q}_2$: $\mathbf{P}\mathbf{Q}_1\mathbf{P}^{\mathrm{T}} = \mathbf{Q}_2$; then the corresponding value of $\mathbf{R}$ is equal to $\mathbf{Q}_2\mathbf{Q}_1^{\mathrm{T}} = \mathbf{P}\mathbf{Q}_1\mathbf{P}^{\mathrm{T}}\mathbf{Q}_1^{\mathrm{T}}$. Now, if e.g. $\mathbf{P}$ and $\mathbf{Q}_1$ commute, then this yields $\mathbf{R} = \mathbf{I}$ *ie.* $\alpha = 0$, independently of the angle of rotation corresponding to $\mathbf{P}$.

The elementary volume element in the space of orientations is defined by the usual relation

$$dV = \sqrt{G}\, d\varphi d\theta d\psi , \tag{2.41}$$

where G denotes the determinant of the matrix $[G_{ij}]$. This gives rise to the following expression for the volume of the whole space

$$V = \int_V dV = \int_0^{2\pi}\int_0^{\pi}\int_0^{\pi} \sqrt{G}\, d\varphi d\theta d\psi \tag{2.42}$$

$$= \int_0^{2\pi}\int_0^{\pi}\int_0^{\pi} \sqrt{32 \sin^2\theta(1-\cos\varphi)}\, d\varphi d\theta d\psi = 16\sqrt{2}\pi^2 = \pi r^3 .$$

The angle φ varies here from 0 to π (instead of, say, from $-\pi$ to π) because the change of its sign is equivalent to the change of the sign of $\mathbf{n}$ (compare (2.1), (2.2)), while all possible directions of $\mathbf{n}$ have already been taken into account (cf. (2.42)) by the suitable choice of the variation range for ψ and θ.

All relations like (2.40) and (2.42), coincide with the results obtained in the framework of Riemann's non-Euclidean geometry [13] (not to be confused with the geometry of Riemannian spaces).

3. Functions and operators on the set of orientations

Assigning an element from some set to every orientation (every rotation tensor) one defines a function on the manifold, which can be effectively represented as a function of three coordinates $\{\xi^1, \xi^2, \xi^3\}$. Functions of two different kinds are worthwhile to be considered here: scalar functions and the functions assigning to every orientation (every point at the manifold of orientations) a vector from the tangent space taken at the same point. Any other reasonable cases can be reduced to, or obtained (as a generalisation) from, the two cases mentioned above.

Among the scalar functions most useful in practical applications one should mention the function of the (mass) density of orientation (orientation distribution) which, for the case of material elements, can be defined as such a function $\mu(\xi^i)$ for which the following relation holds true for every domain Ω of the manifold of orientations:

$$m(\Omega) = \int_\Omega \mu(\mathbf{Q})\, dV = \int\int\int_\Omega \mu(\varphi, \theta, \psi)\sqrt{32 \sin^2\theta(1-\cos\varphi)}\, d\varphi d\theta d\psi, \tag{3.1}$$

where $m(\Omega)$ denotes the total mass of all elements whose orientations fall into domain Ω. We assume of course that the set of oriented elements is large

enough to justify such an approach and the proper regularity demanded by the Radon-Nikodym theorem can be assured.

We should mention here that for practical applications the question of the interpretation and, hence, of the range of validity of the orientation density concept can be more important than in the case of material density in the "physical" space. This applies particularly to these cases when we consider the "localized" sets of oriented elements such as "material particles" in polycrystalline materials. In some problems of plastic deformations the characteristic length of the problem, e.g. the radius of curvature of the tool, may be equal to several grain diameters. Under these conditions, speaking about the "continual distribution of crystal orientations in every material particle" sounds somewhat odd.

The same problem of contradiction between "continuity" and "smallness" of the material particle arises of course in the classical continuum mechanics of real (discrete on the atomic level) matter. It is, in that case, however, suppressed immediately by the powerful Avogadro number. The portion of the material can be considered simultaneously to be very small when compared to any characteristic length of the macroscopic problem and large enough to contain such a great number of molecules which makes sensible all the local state variables of continuum mechanics and thermodynamics. Unfortunately there is no analogous great number in the meso-scale considerations, for which the concept of orientation distribution is usually applied.

It seems that, if the probabilistic interpretation were involved, then the rigorous concept of the orientation density, sensible even for not very numerous sets, could be defined. We shall, however, leave aside this question in the present paper. We shall, nevertheless, come back to the problem of continuity in Sec. 4 in the context of the velocity field definition at the orientation manifold.

It is not difficult to define the mean orientation density – averaged over a chosen domain – as $\overline{\mu} \equiv m(\Omega)/V(\Omega)$, where $V(\Omega)$ denotes the volume of the domain Ω. Such a quantity (particularly when divided by the total mass of whole set) can be possibly useful for semi-quantitative estimation of crystallographic textures of polycrystalline materials. The author cannot however point out any serious advantages of this approach when compared to the use of the X-ray orientation distribution diagrams (pole figures) based on the slightly different concepts [2], which are at present widely used for this purpose.

Here, the related question arises: is it possible to define *mean orientation*, for a domain (or for the whole set of orientations), if the orientation density function is known everywhere? If such a notion made sense then one would expect to get from it some preliminary information, concerning, for example, the rough approximation of the orientation of the orthotropy axes of polycrystalline material.

The difficulty lies upon the fact that the set of the orthogonal tensors is not closed with respect to summation, thus the usual averaging procedures can not be used even in the case of the finite sets of orthogonal tensors: $(\mathbf{Q}_1 + \mathbf{Q}_2)/2$ is

not an orthogonal tensor and there is no orientation associated with this quantity. Definitely we should look for other concepts. Let us notice that, in the case of mass distribution in three-dimensional space, we can define the centre of mass of the material body (or of the set of material points) as such a point for which the moments of inertia with respect to any axis passing through this point are minimal (in the set of moments of inertia with respect to all parallel axes – compare the well known Steiner theorem [18], [9]). Recalling that the trace of the inertia tensor is invariant with respect to frame rotation we can define the mass centre $\mathbf{c}$ as such a point for which the following condition holds

$$\int_{\Omega} |\mathbf{c} - \mathbf{x}|^2 \rho(\mathbf{x})\, dV = \text{Min}. \tag{3.2}$$

Differentiating the integral in (3.2) with respect to Cartesian components of $\mathbf{c}$ one obtains, from the conditions of vanishing of the partial derivatives, well known formulas for the coordinates of the mass centre.

Proceeding in the same way, we can define the mean orientation as described by such a rotation tensor $\mathbf{R}$, for which the following condition is fulfilled:

$$\int_{\Omega} \left[\arccos \left(\frac{\operatorname{tr}\left(\mathbf{R}\,\mathbf{Q}^{\mathrm{T}}\right) - 1}{2} \right) \right]^2 \mu(\mathbf{Q})\, d = \text{Min}. \tag{3.3}$$

Conditions of minimum of this functional contain the nonlinear functions of unknown tensor $\mathbf{R}$ under the integral sign and, most probably, are of no practical significance for the analytic considerations, thus we shall not even quote them here. It seems that the problem, for every domain and every density distribution, should be separately solved numerically.

The problem becomes much simpler if we take the concept of distance from the space of second rank tensors, then the corresponding minimality condition takes the following form:

$$\int_{\Omega} |\mathbf{R} - \mathbf{Q}|^2 \mu(\mathbf{Q})\, dV = 2 \int_{\Omega} (3 - \mathbf{R} \circ \mathbf{Q}) \mu(\mathbf{Q})\, dV = \text{Min}\,. \tag{3.4}$$

Condition (3.4) is evidently equivalent to the following:

$$\mathbf{R} \circ \int_{\Omega} \mathbf{Q} \mu(\mathbf{Q})\, dV = \mathbf{R} \circ \mathbf{A} = \text{Max}\,. \tag{3.5}$$

Thus we have reduced our problem to finding the "nearest" rotation tensor for a given tensor $\mathbf{A}$.

One may ask which definition of the mean orientation is "better". The answer to this question lies beyond the scope of the author's present knowledge. Some two-dimensional analogies can suggest that the use of the definition (3.4) may sometimes yield rather surprising, from the viewpoint of common sense, results.

Beginning from this point, up to the end of the paper, we shall assume that the set of oriented elements is large enough and the functions defined are as smooth as required for the use of differential calculus.

The field of the gradient ∇f of the scalar function f on the manifold of orientations is defined in a usual way:

$$\nabla f(\mathbf{Q}) = \frac{\partial f(\mathbf{Q})}{\partial \xi^K} \mathbf{G}^K . \tag{3.6}$$

We remind here once again that the value of the gradient of the scalar function for given $\mathbf{Q}$ (given coordinate point (ξ^1, ξ^2, ξ^3)), is (as any vector tangent to the manifold of the orientations) at the same time a second order tensor from $\mathbf{V}^3 \otimes \mathbf{V}^3$. This tensor can be represented as $\mathbf{\Lambda Q}$, where the skew-symmetric tensor $\mathbf{\Lambda}$ is a linear combination of the tensors $\mathbf{W}^{(I)}$ defined by the relations (2.16):

$$\mathbf{\Lambda} = -\sum_{I=1}^{3} \frac{\partial f}{\partial \xi^I} \mathbf{W}^{(I)} . \tag{3.7}$$

Three functions

$$\xi^I = \xi^I(s) \tag{3.8}$$

define the curve at the manifold:

$$\mathbf{Q} = \mathbf{Q}(s). \tag{3.9}$$

Vector $\boldsymbol{\tau}(s)$ defined by the following relation

$$\boldsymbol{\tau}(s) = \frac{d\xi^K}{ds} \mathbf{G}^K \tag{3.10}$$

will be called a vector tangent to the curve $\mathbf{Q}(s)$. If this vector along the whole curve is a unit vector, then the parameter s will be called natural. We shall omit here the proof of the rather obvious fact that the linear part of the increment of the function f corresponding to the increment of the parameter s by the value h can be expressed as follows:

$$f(s+h) = f(s) + \frac{\partial f(\xi^I)}{\partial \xi^K} \frac{\partial \xi^K}{\partial s} + \mathcal{O}(h) = f(s) + \nabla f \circ \boldsymbol{\tau} + \mathcal{O}(h), \tag{3.11}$$

where $\mathcal{O}(h)$ denotes, as usual, terms of higher order in h : $\lim_{h \to 0} \mathcal{O}(h)/h = 0$. And, obviously the following relation for the derivative along the curve holds true:

$$\frac{df(\mathbf{Q})}{ds} = \nabla f(\mathbf{Q}) \circ \boldsymbol{\tau}. \tag{3.12}$$

There is no need to re-derive all the known results of the differential geometry of Riemannian spaces [13, 16, 17], we shall just quote them for the completeness. Thus, if $\mathbf{a}(\xi^I)$ denotes the function ascribing to every point (radius-vector) $\mathbf{Q}$ of

the manifold a vector from the tangent space at **Q**, then the second order tensor $\nabla\mathbf{a}$ from the tensorial product of the tangent space by itself $\mathbf{T}^3 \otimes \mathbf{T}^3$, defined by the following expression:

$$\nabla\mathbf{a} = \left(\frac{\partial a^K}{\partial \xi^I} + a^J \Gamma^K_{JI}\right)\left(\mathbf{G}_K \otimes \mathbf{G}^I\right) \tag{3.13}$$

will be called a gradient of the tangent vector field $\mathbf{a}(\mathbf{Q})$. Such a tensor can, of course, be interpreted also as a fourth-order tensor from $\mathbf{V}^3 \otimes \mathbf{V}^3 \otimes \mathbf{V}^3 \otimes \mathbf{V}^3$. The term in parentheses represents covariant derivative([3]) of the field $a^K(\xi^I)$:

$$a^K_{,I} = \frac{\partial a^K}{\partial \xi^I} + a^J \Gamma^K_{JI}\,. \tag{3.14}$$

The corresponding expression for covariant derivatives of the covariant representation is the following:

$$a_{K,I} = \frac{\partial a_K}{\partial \xi^I} - a_J \Gamma^J_{KI}\,. \tag{3.15}$$

We shall not quote explicitly all the expressions for these derivatives, and confine our attention to the trace of this tensorial expression – the divergence of the tangent vector field:

$$\begin{aligned}\operatorname{div}\mathbf{a} = a_{K,L}G^{KL} &= \left(\frac{\partial a_K}{\partial \xi^L} - a_J \Gamma^J_{KL}\right) G^{KL} \\ &= \frac{\partial a_1}{\partial \xi^1}G^{11} + \frac{\partial a_2}{\partial \xi^2}G^{22} + \frac{\partial a_3}{\partial \xi^3}G^{33} - a_2\Gamma^2_{11}G^{11} - a_3\left(\Gamma^3_{11}G^{11} + \Gamma^3_{22}G^{22}\right) \\ &= \frac{\partial a_1}{\partial \psi}\frac{1}{4\sin^2\theta(1-\cos\varphi)} + \frac{\partial a_2}{\partial \theta}\frac{1}{4(1-\cos\varphi)} + \frac{\partial a_3}{\partial \varphi}\frac{1}{2} \\ &\quad - a_2\frac{\operatorname{ctg}\theta}{4(1-\cos\varphi)} - a_3\frac{\sin\varphi}{2(1-\cos\varphi)}\,.\end{aligned} \tag{3.16}$$

Relation (3.16) immediately yields the following expression for the Laplace operator of the scalar function on the manifold of orientations:

$$\begin{aligned}\Delta f(\mathbf{Q}) = f_{,KL}G^{KL} &= \frac{1}{4(1-\cos\varphi)} \\ &\times \left(\frac{\partial^2 f}{\partial \psi^2}\frac{1}{\sin^2\theta} + \frac{\partial^2 f}{\partial \theta^2} + 2\frac{\partial^2 f}{\partial \varphi^2}(1-\cos\varphi) - \frac{\partial f}{\partial \theta}\operatorname{ctg}\theta - 2\frac{\partial f}{\partial \varphi}\sin\varphi\right).\end{aligned} \tag{3.17}$$

([3]) Usually for the explanation of the geometric sense of the covariant derivative, the notion of parallel displacement of the tangent vector along a curve at the manifold is introduced. In our case, when we can point out the Euclidean nine-dimensional space in which all the tangent spaces can be embedded, we are able to propose the straightforward interpretation of the covariant derivative as the tangent component of the partial derivative of the vector. Recalling (2.18)′ we are able to write

$$\frac{\partial \mathbf{a}}{\partial \xi^K} \circ \mathbf{G}^L = \frac{\partial}{\partial \xi^K}(a^I \mathbf{G}_I) \circ \mathbf{G}^L = \frac{\partial a^I}{\partial \xi^K}\delta^L_I + a^I \frac{\partial \mathbf{G}_I}{\partial \xi^K} \circ \mathbf{G}^L = \frac{\partial a^L}{\partial \xi^K} + a^I \Gamma^L_{IK}\,,$$

and similarly for the covariant components.

4. Time derivative of orientation

Let for every instant t a rotation tensor $\mathbf{Q}(t)$ be given. We shall assume that the function $\mathbf{Q}(t)$ is smooth enough to make sensible all subsequent differential operations. We shall not specify, for the time being, how our time-dependent tensor is connected with the motion of material bodies or media. It can describe either the sequence of the orientations of one chosen material element or, on the contrary, the change of the orientation at the chosen point of the (considered as remaining in rest) "physical" space, which can be occupied by different (oriented) material particles at different time instants.

Let, then, the tensor-valued function of time under consideration be specified by the following three scalar functions:

$$\xi^i = \xi^i(t), \qquad \Big(\psi = \psi(t), \quad \theta = \theta(t), \quad \varphi = \varphi(t)\Big). \tag{4.1}$$

Relations (4.1), on the other side, define a time-parametrised curve (a trajectory) in the manifold, thus the time derivative $d[\mathbf{Q}(t)]/dt$ is a tangent vector and we are able to write

$$\frac{\partial \mathbf{Q}}{dt} = \frac{\partial \mathbf{Q}}{d\xi^K}\frac{d\xi^K}{dt} = \frac{d\xi^K}{dt}\mathbf{G}_K = \mathbf{W}\mathbf{Q}, \tag{4.2}$$

where skew-symmetric second rank tensor $\mathbf{W}$ from $\mathbf{V}^3 \otimes \mathbf{V}^3$ is defined as follows:

$$\mathbf{W} = -\sum_{I=1}^{3} \mathbf{W}^{(I)}\frac{d\xi^I}{dt}. \tag{4.3}$$

Hence, we are able to write

$$\frac{d\mathbf{Q}}{dt} = -\boldsymbol{\omega}\,\boldsymbol{\varepsilon}\,\mathbf{Q}, \tag{4.4}$$

where $\boldsymbol{\omega}$ is a vector from $\mathbf{V}^3$ dual to the tensor $\mathbf{W}$, $\omega_k = \frac{1}{2}W_{ij}\varepsilon_{ijk}$. Vectors dual to tensors $\mathbf{W}^{(I)}$ form (after normalisation) an orthonormal basis $\{\boldsymbol{\eta}_i\}$ in $\mathbf{V}^3$:

$$\begin{aligned} \boldsymbol{\eta}_1 &= \mathbf{e}_\psi \cos\frac{\varphi}{2} - \mathbf{e}_\theta \sin\frac{\varphi}{2}, \\ \boldsymbol{\eta}_2 &= \mathbf{e}_\psi \sin\frac{\varphi}{2} + \mathbf{e}_\theta \cos\frac{\varphi}{2}, \\ \boldsymbol{\eta}_3 &= \mathbf{n}. \end{aligned} \tag{4.5}$$

The representation of vector $\boldsymbol{\omega}$ in this basis

$$\boldsymbol{\omega} = \omega_i \boldsymbol{\eta}_i \tag{4.6}$$

has the following form:

$$\begin{aligned} \omega_1 &= 2\sin\theta\sin\frac{\varphi}{2}\frac{d\psi}{dt}, \\ \omega_2 &= 2\sin\frac{\varphi}{2}\frac{d\theta}{dt}, \\ \omega_1 &= \frac{d\varphi}{dt}, \end{aligned} \tag{4.7}$$

and for the square of its length we obtain

$$\omega^2 = (1-\cos\varphi)\sin^2\theta\left(\frac{d\psi}{dt}\right)^2 + (1-\cos\varphi)\left(\frac{d\theta}{dt}\right)^2 + \left(\frac{d\varphi}{dt}\right)^2. \tag{4.8}$$

It is not difficult to notice that in the case of "infinitesimal rotation" vector $\boldsymbol{\omega}$ is directed along the axis of rotation, or, more rigorously:

$$\lim_{\varphi\to 0}\frac{\boldsymbol{\omega}}{|\boldsymbol{\omega}|} = \mathbf{n}. \tag{4.9}$$

Let a rotating vector $\mathbf{a}(t)$ (from $\mathbf{V}^3$) have the form

$$\mathbf{a}(t) = \mathbf{Q}(t)\mathbf{a}_0, \tag{4.10}$$

Then for its time derivative we can write

$$\frac{d}{dt}\mathbf{a}(t) = \frac{d}{dt}\mathbf{Q}(t)\mathbf{a}_0 = -\boldsymbol{\omega}\,\varepsilon\,\mathbf{Q}\,\mathbf{a}_0 = \boldsymbol{\omega}\,\varepsilon\,\mathbf{a} = \boldsymbol{\omega}\times\mathbf{a}, \tag{4.11}$$

thus one can see on this example that we have properly chosen a symbol $\boldsymbol{\omega}$ for the vector under consideration which had turned out to be exactly equal to the well-known vector of angular velocity.

This is a proper place here for two remarks:

The first – one should not confuse two related, but entirely different vectors: vector $d\mathbf{Q}/dt$ from the space tangent to the manifold at $\mathbf{Q}$, and vector $\boldsymbol{\omega}$ from $\mathbf{V}^3$; it is not difficult to notice that the mapping $d\mathbf{Q}/dt \to \boldsymbol{\omega}$ depends on the coordinates, thus these two vector fields have different properties, especially with respect to differential operations.

The second – we have not assumed (nor obtained) any relation between our skew-symmetric tensor $\mathbf{W}$ and the skew-symmetric part of the velocity gradient in the motion of the continuous medium. We have not even assumed that our oriented elements are embedded in any continuous material. Moreover, even if we considered such a situation, we do not see any reason for identification of these two tensors, with one exception, probably – the case of rigid motion.

5. Velocity field on the manifold of orientations and the mass conservation law

The main scope of this section consists in the obtaining of the local mass conservation law (evolution rule of orientation density) on the manifold of orientations. Before we start this main task, we have to come back to the fundamental problems of continuity already discussed in the Sec. 3.

In the dynamics of continuous media the fundamental property of the (topologic) continuity of motion is assumed. The very notion of material fibre and material surface is invariant with respect to the motion (except, maybe, some sets of zero three-dimensional measure). One realises of course that, from the physical point of view, this assumption is incorrect; if it were true, then no diffusion process could be possible in the material. Nevertheless, due to the different time and space scales of the dynamic and diffusion problems, nobody takes care of this inconsistency – free paths of molecules even in gases (wit the exception of extremely rarefied) are, due to the already mentioned Avogadro number, much less than any characteristic length of the usual boundary value problem. In the case of the manifold of orientations, as it has been already mentioned, all these problems are not so evident.

Let the countable set of the oriented elements be given. There is no conceptual difficulty to ascribe the velocity vector from the tangent space to every element, simply we take as $\mathbf{Q}(t)$ the orientation of the given element and take a time derivative of this function denoting it (as analogue to material time derivative in the "physical" space) by $\dot{\mathbf{Q}}$. Thus we have the following situation: at the countable set of points of the orientation manifold (at the points occupied by the oriented material elements) two functions are defined: the first – a finite mass of the element and the second – a velocity vector $\dot{\mathbf{Q}}$ from the tangent space $\mathbf{T}^3$.

This situation looks like in the large scale description of the set of all stars of the galaxy, or a portion of rarefied gas. There are however some essential differences: there is no physical reason for *different elements* forbidding them to occupy *the same place* in the manifold of orientations, also no physical reasons can be pointed out to ban *different velocities* being assigned to the elements having *the same orientation*. It is not difficult in the "physical" Euclidean space to define the mean velocity of several material points summing their momenta an dividing by the total mass. For the obvious reasons such an operation can not be performed in the case of rotating elements in the space of orientations. As it has been already mentioned, the number of oriented elements is usually much smaller than e.g. the number of molecules in "small" neighbourhood of any point in real materials.

All this creates a lot of doubts about the sensibility of the averaging and smoothing of the rotation velocity $\dot{\mathbf{Q}}$ distributions defined on the manifold of rotations. It is however the author's good hope, that the probabilistic approach is capable of eliminating these doubts in future, supplying, at the same time, some cues concerning the limits of applicability of the continuity assumptions

with respect both to the orientation distribution and to the functions defined on the sets of oriented elements.

It is quite possible that a new approach, permitting the simultaneous existence of the continuous velocity spectrum for every orientation, would be more adequate for the description of the density of orientations evolution. The present author, however, prefers rather to leave this unrecognised path to the future students of the problem, confining the present considerations to the simplest model.

Thus we shall assume, that a unique velocity field $\dot{\mathbf{Q}}(\mathbf{Q}, t)$ is given for some domain of the orientations and some time interval $<t_1, t_2>$. We shall also assume that at some time instant $t \in <t_1, t_2>$ the orientation density function $\mu(\mathbf{Q}, t)$ is known. We assume at last that the field $\dot{\mathbf{Q}}(\mathbf{Q}, t)$ is regular enough (e.g. certain Liepshitz conditions are satisfied [15]) to provide the existence of well-defined integral curves, making possible to introduce such a system of "Lagrangian" coordinates $\{\Xi^1, \Xi^2, \Xi^3\}$ (i.e., in fact, the functions $\Xi^I(\xi^K, t)$) in the set of the oriented elements which satisfy the following conditions:

$$\det\left[\frac{\partial \Xi^I\left(\xi^1, \xi^2, \xi^3, t\right)}{\partial \Xi^K}\right] > 0, \tag{5.1}$$

$$\dot{\mathbf{Q}} = \frac{\partial \mathbf{Q}}{\partial \xi^K} \frac{\partial \xi^K\left(\Xi^1, \Xi^2, \Xi^3, t\right)}{\partial t} = \mathbf{G}_K \frac{\partial \xi^K\left(\Xi^1, \Xi^2, \Xi^3, t\right)}{\partial t}. \tag{5.2}$$

The first condition (5.1) is necessary for the relation (5.2) to be sensible([4]).

Like in the "physical" space, we can introduce here two different kinds of the time derivatives: the local (Eulerian)

$$\frac{\partial f}{\partial t} \equiv \frac{\partial f(\xi^K, t)}{\partial t}, \tag{5.3}$$

and the Lagrangian (we shall rather not call it "material" here):

$$\overset{\circ}{f} \equiv \frac{\partial f(\Xi^K, t)}{\partial t}. ([5]) \tag{5.4}$$

([4]) The situation here is resembling the one in the theory of mixtures, where, despite the diffusion, some "material in the averaged sense" or "quasimaterial" coordinates can be introduced [1].

([5]) If our set of oriented elements is "localised", i.e. constitutes a material particle in Euclidean space then this notation can be confusing: we would not know if the Lagrangian or Eulerian approach in "physical" space is considered. It is however beyond the author's mental capacity to give an interpretation of such a derivative which is "Eulerian" in the Euclidean space, being at the same time "Lagrangian" at the manifold of orientations. For consistency it would be better, probably, to denote the time derivative of the orientation of a given material element by $\overset{\circ}{\mathbf{Q}}$ rather than by $\dot{\mathbf{Q}}$, for the reasons of tradition, however, we shall preserve the notation used in the previous sections.

There is a simple (expected from the analogy with the "physical" space) relation between these two quantities. To show it, we transform the expression (5.4) in the following way:

$$\overset{\circ}{f} = \frac{\partial}{\partial t} f\left(\xi^I\left(\Xi^K, t\right), t\right) = \frac{\partial f}{\partial \xi^I}\frac{\partial \xi^I\left(\Xi^K, t\right)}{\partial t} + \frac{\partial f\left(\xi^I, t\right)}{dt} \tag{5.5}$$

$$= \left(\frac{\partial f}{\partial \xi^I}\mathbf{G}^I\right) \circ \left(\mathbf{G}_K \frac{\partial \xi^K\left(\Xi^J, t\right)}{\partial t}\right) + \frac{\partial f}{\partial t} = \nabla f \circ \dot{\mathbf{Q}} + \frac{\partial f}{\partial t}.$$

In the forthcoming considerations concerning the formulation of the local mass conservation principle, we shall consider the orientation density function as a continuous approximation of the real discrete distribution of the orientations of the elements of a very numerous set. It will be also assumed, that all the elements of the same orientation have the same rate of the orientation change (rotation velocity) $\dot{\mathbf{Q}}$. The author believes that for such a restricted case the Reader will agree with these considerations. As regards the general situation – he has the right to treat all reasoning as heuristic only.

Let us choose a surface $\mathbf{S}$ in the manifold of orientations given by the following equation:

$$F(\mathbf{Q}) = 0. \tag{5.6}$$

Let us assume that, for a given oriented element II with Lagrangian coordinates $(\Phi_0, \Theta_0, \Psi_0)$ and a varying in time orientation $\mathbf{O}(II, t) = \mathbf{O}(\Phi_0, \Theta_0, \Psi_0, t)$, the following inequality holds true at some time instant t_0:

$$F(\mathbf{O}(II, t_0)) < 0. \tag{5.7}$$

For the Lagrangian derivative we can write

$$\overset{\circ}{F} = \frac{dF}{d\mathbf{O}} \circ \dot{\mathbf{O}}. \tag{5.8}$$

Trajectory $\mathbf{O}(II, t)$ of the element II in the orientation space will intersect the surface $\mathbf{S}$ at a time instant from the interval $(t_0, t_0 + h)$ if and only if the mean value of the time derivative $\overset{\circ}{F}{}^*$ satisfies the following inequality:

$$\overset{\circ}{F}{}^* h \geq -F(\mathbf{O}(\mathbf{P}, t_0)). \tag{5.9}$$

Denoting by d the distance between the orientation $\mathbf{O}(\mathbf{P}, t_0)$ and the surface $\mathbf{S}$ at time instant t_0 and making use of the property (3.11) of the gradient of the function F, we can write:

$$h\nabla F \cdot \dot{\mathbf{Q}} \geq |\nabla F(\mathbf{O}(II, t_0 + h))| d + \mathcal{O}(d). \tag{5.10}$$

Introducing the following notation for the unit vector normal to $\mathbf{S}$([6])

$$\mathbf{N} \equiv \frac{\nabla F}{|\nabla F|} \tag{5.11}$$

and taking into account that the mean value of derivative over the time interval $(t_0, t_0 + h)$ differs from its value at any instant inside the interval at most by the terms of the order of h, we are able to formulate the following condition:

During the time interval of duration h orientations of only these elements will cross the surface $\mathbf{S}$ *which are contained in the layer of thickness d, adjacent to the surface* $\mathbf{S}$, *where the thickness d is given* (with the accuracy to the terms linear in h and/or d) *by the following relation:*

$$\frac{d}{h} = \mathbf{N} \circ \dot{\mathbf{O}} \,. \tag{5.12}$$

The mass of all elements contained in the layer (per unit area of $\mathbf{S}$), is equal to μh, thus, under the assumption that the mass of every element remains constant, we are able to write, for the arbitrary domain Ω of the manifold, the following integral mass balance law:

$$\frac{\partial m(\Omega, t)}{\partial t} = -\int_{\partial\Omega} \mu \, \dot{\mathbf{Q}} \circ \mathbf{N} \, dS. \tag{5.13}$$

Taking into account the definition (2.1) of density μ and making use of the Gauss–Ostrogradskii formula [17, 10] one can rewrite expression (5.13) in the form of the following volume integral:

$$\int_{\Omega} \left\{ \frac{\partial \mu(\mathbf{Q}, t)}{\partial t} + \mathrm{Div} \left[\mu(\mathbf{Q}, t) \, \dot{\mathbf{Q}} \, (\mathbf{Q}, t) \right] \right\} dV = 0. \tag{5.14}$$

By virtue of the arbitrariness of the domain Ω, the last integral equality is equivalent to the following local condition

$$\frac{\partial \mu}{\partial t} + \mathrm{Grad}\, \mu \circ \dot{\mathbf{Q}} + \mu \mathrm{Div}\, \dot{\mathbf{Q}} = 0, \tag{5.15}$$

where operators Grad and Div should be interpreted in terms of the differential calculus on the manifold, thus e.g. the field $\mathrm{Grad}\, \mu$ is a field of the second rank tensors tangent to the manifold. Using relation (5.5) one can readily rewrite (5.15) in another form:

$$\overset{\circ}{\mu} + \mu \mathrm{div}\, \dot{\mathbf{Q}} = 0 \tag{5.16}$$

resembling the most popular form of the expression for the local mass conservation law in the "physical" space.

([6]) One should remember that N, being normal to the surface S is a *tangent* vector with respect to the *whole manifold.*

If all the elements of the set rotate with the same angular velocity ω, one should expect vanishing $\overset{\circ}{\mu}$ and the Eulerian rate of the change of μ to be governed by the sole convective term

$$\frac{\partial \mu}{\partial t} = -\nabla \mu \circ \dot{\mathbf{Q}} \qquad \text{(in rigid motion)}. \tag{5.17}$$

Indeed let $\boldsymbol{\omega} = \text{const}$, then by virtue of (4.4) one can write:

$$\frac{\partial}{\partial \xi^K}\left(\dot{\mathbf{Q}}\,\mathbf{Q}^{\mathsf{T}}\right) = \frac{\partial \dot{\mathbf{Q}}}{\partial \xi^K}\mathbf{Q}^{\mathsf{T}} + \dot{\mathbf{Q}}\,\frac{\partial \mathbf{Q}^{\mathsf{T}}}{\partial \xi^K} = \mathbf{0}. \tag{5.18}$$

Multiplying (5.18) (in the sense of a simple contraction in three-dimensional space) from the right-hand side by $\mathbf{Q}$ one obtains:

$$\frac{\partial \dot{\mathbf{Q}}}{\partial \xi^K} + \dot{\mathbf{Q}}\,\frac{\partial \mathbf{Q}^{\mathsf{T}}}{\partial \xi^K}\mathbf{Q} = \frac{\partial \dot{\mathbf{Q}}}{\partial \xi^K} - \dot{\mathbf{Q}}\,\mathbf{Q}^{\mathsf{T}}\frac{\partial \mathbf{Q}}{\partial \xi^K} = \frac{\partial \dot{\mathbf{Q}}}{\partial \xi^K} - \dot{\mathbf{Q}}\,\mathbf{Q}^{\mathsf{T}}\mathbf{G}_K = \mathbf{0}. \tag{5.19}$$

For better clarity of further considerations we shall denote contravariant components of $\dot{\mathbf{Q}}$ in the basis of the tangent space by ϑ^K:

$$\dot{\mathbf{Q}} = \vartheta^K \mathbf{G}_K\,. \tag{5.20}$$

Then, taking into account (2.18)′ one can rewrite (5.19) as follows:

$$\left(\frac{\partial \vartheta^L}{\partial \xi^K} + \vartheta^J \Gamma^L_{JK}\right)\mathbf{G}_L = \mathbf{W}\mathbf{G}_K\,, \tag{5.21}$$

where $\mathbf{W}$ denotes a skew-symmetric tensor defined by (4.2). Multiplying both sides of (5.21) by $\mathbf{G}^K$ (in the sense of scalar product in nine-dimensional space, in accordance with the summation convention), one obtains the following equality:

$$\left(\frac{\partial \vartheta^L}{\partial \xi^K} + \vartheta^J \Gamma^L_{JK}\right)\delta^K_L = (\mathbf{W}\mathbf{G}_K) \circ (\mathbf{G}^K), \tag{5.22}$$

which, by virtue of the definitions of the divergence and the scalar product in $\mathbf{V}^3 \otimes \mathbf{V}^3$, taking into account diagonality of the matrix $[G_{KL}]$, (compare (2.13)) can be rewritten as follows:

$$\operatorname{div}\mathbf{Q} = \mathbf{W} \circ \left(\mathbf{G}^K \mathbf{G}^{\mathsf{T}}_K\right) = \mathbf{W} \circ \left(\frac{\mathbf{G}_1\mathbf{G}_1^{\mathsf{T}}}{G_{11}} + \frac{\mathbf{G}_2\mathbf{G}_2^{\mathsf{T}}}{G_{22}} + \frac{\mathbf{G}_3\mathbf{G}_3^{\mathsf{T}}}{G_{33}}\right) = 0. \tag{5.23}$$

The last term equals zero due to the symmetry of the second order tensor in parenthesis and the skew symmetry of $\mathbf{W}$. This result, proving our assertion concerning the vanishing of the Lagrangian rate of change of the orientation density

for the case of rigid rotation, supplies an indirect evidence of the sensibility of the introduced rates and differential operations. One has to understand the last sentence as a declaration that we would be rather inclined to reject the theory violating this condition.

Expression (5.4) gives a rule for the transition from the Eulerian to Lagrangian rate of change of any scalar function defined on the manifold of orientations. This rule can be generalised for tangent vectors and tensors of any rank from the corresponding tensorial products of n copies of the tangent space $\mathbf{T}^3 \otimes \mathbf{T}^3 \otimes \ldots \otimes \mathbf{T}^3$. The present author, however, does not see any direct application of such quantities in continuum mechanics.

Much more useful may turn out to be the quantities describing the rate of change of the functions ascribing Euclidean tensors from $\mathbf{E}^3 \otimes \mathbf{E}^3 \otimes \ldots \otimes \mathbf{E}^3$ to the oriented elements and/or to the points of the manifold of orientations. The problem exists with respect to both Lagrangian and Eulerian rates of change. We shall consider it here, however, for the individual element, the interpretation of the obtained results in terms of Lagrangian rates will be meaningful in the same sense as the very notion of the Lagrangian coordinates at the manifold.

Let some time-dependent tensorial object $\mathbf{A}$ be assigned to the rotating element (e.g. vector of external force or moment, second rank tensor of dielectric permeability, fourth rank tensor of elasticity etc.). It may be useful to define such a rate of change of a tensor $\mathbf{A}$, which is equal to zero if the representation of a tensor in the basis rotating together with the element under consideration remains constant, and equals the usual time derivative (e.g. Lagrangian in the "physical" space) if the element does not rotate with respect to the "laboratory frame of reference". In other words, we try to define such a rate of change $\overset{*}{\mathbf{A}}$ as the one measured by the observer rotating together with the element and "not knowing" about this rotation.

Thus let $\{\boldsymbol{\varepsilon}_k\}$ denote the time-dependent orthonormal basis in $\mathbf{V}^3$: $\boldsymbol{\varepsilon}_k(t) = \mathbf{Q}(t)\boldsymbol{\varepsilon}_k^0$, where $\mathbf{Q}$ describes the orientation of the element. Let $\mathbf{A} = A_{i\ldots s}(t)(\boldsymbol{\varepsilon}_i \otimes \ldots \otimes \boldsymbol{\varepsilon}_s)$ be a tensor of any rank, then the following quantity:

$$\overset{*}{\mathbf{A}} \equiv \overset{*}{A}_{i\ldots s}\,(\boldsymbol{\varepsilon}_i \otimes \ldots \boldsymbol{\varepsilon}_s) \tag{5.24}$$

will be called a co-rotational derivative of tensor $\mathbf{A}$ with respect to the rotating element of orientation $\mathbf{Q}$.

We can express this quantity in the form independent of the choice of the base $\{\boldsymbol{\varepsilon}_k\}$. To this end we write

$$\overset{\circ}{\mathbf{A}} = \overset{\circ}{A}_{i\ldots q}\,(\boldsymbol{\varepsilon}_i \otimes \ldots \otimes \boldsymbol{\varepsilon}_q) + A_{i\ldots q}\left(\dot{\mathbf{Q}}\,\mathbf{e}_i \otimes \ldots \otimes \mathbf{Q}\,\mathbf{e}_q\right) + \ldots + A_{i\ldots q}\left(\mathbf{Q}\,\mathbf{e}_i \otimes \ldots \otimes \dot{\mathbf{Q}}\,\mathbf{e}_q\right), \tag{5.25}$$

or

$$\overset{\circ}{\mathbf{A}} = \overset{*}{\mathbf{A}} + A_{i\ldots q}\,(\mathbf{W}\boldsymbol{\varepsilon}_i \otimes \ldots \otimes \boldsymbol{\varepsilon}_q) + \ldots + A_{i\ldots q}\,(\boldsymbol{\varepsilon}_i \otimes \ldots \otimes \mathbf{W}\boldsymbol{\varepsilon}_q), \tag{5.26}$$

and finally we obtain

$$\overset{*}{\mathbf{A}} = \overset{\circ}{\mathbf{A}} - \sum_{i=1}^{n} \mathbf{W} \overset{(2,i)}{\circ} \mathbf{A}, \tag{5.27}$$

where by $\mathbf{W} \overset{(2,i)}{\circ} \mathbf{A}$ we understand a simple contraction of the second index of tensor $\mathbf{W}$ with the i-th index of tensor $\mathbf{A}$.

This terminates, for the time being, our considerations on the kinematics on the manifold of orientations.

Acknowledgements

This paper contains results obtained in the framework of the research program *Mathematical Description and Forming of the Local Properties of the Aggregated Media* headed by J. RYCHLEWSKI, supported by grant No 3 P404 027 05 of the Committee of Scientific Research (Poland).

References

1. A. BLINOWSKI, *Gradient model of capillary phenomena* [in Polish], IFTR Reports, **42**, 1974.
2. H.J. BUNGE, *Texture analysis in materials science*, Butterworths, London 1982.
3. P.H. DŁUŻEWSKI, *Crystal orientation spaces and remarks of the modelling of polycrystal anisotropy*, J. Mech. Phys. Solids, **39**, 5, 651–661, 1991.
4. P.H. DŁUŻEWSKI, *Some remarks on the integration domain of the slip orientation*, Arch. Mech., **39**, 419–422, 1987.
5. J. GANCARZEWICZ, *Differential geometry* [in Polish], PWN, Warszawa 1982.
6. I.M. GELFAND, R.A. MINLOS and Z.YA. SHAPIRO, *Representation of the group of rotations and the Lorentz group; their applications* [in Russian], Gos. Izd. Fiz-Mat. Lit., Moskva 1958.
7. S. GOŁĄB, *Tensor calculus* [in Polish], PWN, Warszawa 1966.
8. A. GOETZ, *Differential geometry* [in Polish], PWN, Warszawa 1965.
9. H. GOLDSTEIN, *Classical mechanics*, Addison-Wesley, Cambridge Mass. 1950.
10. L. LANDAU and E. LIFSHITS, *Field theory* [in Polish], PWN, Warszawa 1958.
11. J. LUBARSKI, *Group theory and its applications in physics* [in Polish], PWN, Warszawa 1961.
12. J. MANDEL, *Directors vectors and constitutive equations for plastic and viscoplastic media*, [in:] Problems of Plasticity, A. SAWCZUK [Ed.], 135–43, Nordhoff, Groningen, 1974.
13. *Encyclopaedia of Mathematics* vol. 1–5, "Sovetskaya Enciklopedia", Moskva 1977–1985.
14. W. NOWACKI, Comment. Math. Helv., **7**, 81–93, 1934.
15. I.G. PETROVSKII, *Lectures on ordinary differential equations* [in Russian], "Nauka", Moskva 1964.
16. P.K. RASZEWSKII, *Riemann geometry and tensor analysis* [in Polish], PWN, Warszawa 1958.
17. J. SYNGE and A. SHIELD, *Tensor calculus*, Toronto 1949.
18. J. SYNGE, *Classical dynamics*, [in:] Encyclopaedia of Physics, S. FLÜGGE [Ed.], vol. III/1, Springer Verlag, Berlin-Göttingen-Heidelberg, 1960.
19. J.A. WOLF, *Spaces of constant curvature*, Univ. of California, Berkeley, California 1972.
20. A. ZIABICKI, *Configurational space for clusters in the theory of nucleation*, Arch. Mech., **42**, 6, 703–715, 1990.

POLISH ACADEMY OF SCIENCES
INSTITUTE OF FUNDAMENTAL TECHNOLOGICAL RESEARCH

Received July 29, 1994; *new version September* 14, 1994.

Arch. Mech., 46, 6, pp. 989–1004, Warszawa 1994

On some one-to-one constitutive correspondences in plasticity

Q.-C. HE and A. CURNIER (LAUSANNE)

ON THE BASIS of non-smooth convex analysis, it is shown that: (i) when the elastic domain is a convex set containing the origin, a gauge yield function (a non-negative positively homogeneous convex function) can be constructed so that the elastic domain is uniquely related to this function; (ii) if the normality rule is further postulated for plastic flow, the polarity operation can be used to establish a one-to-one symmetric correspondence between the dissipation and gauge yield functions. Several examples are given to illustrate the results.

1. Introduction

IN A PREVIOUS paper, HILL [8] has shown from geometric considerations that if the elastic domain is convex and the normality rule applies, a one-to-one symmetric relationship exists between the positively homogeneous convex (i.e. sublinear) yield and dissipation functions. The present paper proposes to reexamine this duality correspondence in the light of *non-smooth convex analysis* (MOREAU [14], ROCKAFELLAR [18]).

Two observations motivate this work. Firstly, the dissipation function for rate-independent plasticity or its conjugate, the indicator function of the elastic domain, is not differentiable but subdifferentiable and hence taking it or its conjugate, as a constitutive potential makes sense only when the notion of subdifferential of (non-smooth) convex analysis is introduced. Secondly, although the correspondences between closed convex sets and closed sublinear functions on the one hand, and between the characteristic functions of a closed convex set and its polar on the other hand, are well established and classical in convex analysis (HÖRMANDER [9], MOREAU [14], ROCKAFELLAR [18]), yet they are far from being well-known in mechanics and their application in the mathematical theory of plasticity, going back to a work of PRAGER [17], remains fragmental (FRIAÂ [2], HALPHEN *et al.* [4], EVE *et al.* [1], HE *et al.* [6], ROMANO *et al.* [19]). A systematic and detailed exposition of this application is now called for. It is the aim of the present paper to provide one, centering on the duality correspondence investigated by HILL [8]. In some sense, this work contributes to further structuring the classical mathematical theory of plasticity.

In Sec. 2, after introducing some basic concepts of convex analysis, Moreau's pseudo-potential (or super-potential) formalism for plasticity is briefly recalled. In Sec. 3, it is shown and illustrated that: (i) provided the elastic domain is a convex set containing the origin, a gauge yield function can be constructed for

the elastic domain to be uniquely related to it; (ii) if the normality rule is further postulated for plastic flow, the polarity operation can be used to establish a duality correspondence between the dissipation and gauge yield functions, the inequality given by YANG [21] appearing as a by-product. For the purpose of this paper, consideration of rigid-perfectly plastic materials undergoing small deformations is sufficient.

2. Pseudo-potential formalism for plasticity

As in classical theories of plasticity (HILL [7], LUBLINER [12]), we assume the existence of a *convex* elastic domain in stress space and postulate the *normality rule* for plastic flow. For any plasticity theory based on these two hypotheses, convex analysis represents a suitable and powerful mathematical tool for formulating its constitutive equations (MOREAU [15, 16], GERMAIN [3], HALPHEN and NGUYEN [4], ZEIDLER [22], MAUGIN [13]). Before entering into the title problem of the paper, we first recall some elements of convex analysis and then give a brief review of Moreau's pseudo-potential formalism which will be used as the starting point of our further development.

2.1. Some basic notions of convex analysis

Let us designate by $\mathcal{V}$ a normed real linear space and by $\mathcal{U}$ its *dual*. A function f from $\mathcal{V}$ to $\overline{\mathcal{R}} := \mathcal{R} \cup +\infty \cup -\infty$ is said to be *convex* if

$$f[\lambda \mathbf{v} + (1-\lambda)\mathbf{v}'] \leq \lambda f(\mathbf{v}) + (1-\lambda) f(\mathbf{v}') \tag{2.1}$$

holds for each $\lambda \in [0,1]$ and all $\mathbf{v}, \mathbf{v}' \in \mathcal{V}$. The set $\operatorname{dom}(f) := \{\mathbf{v} \in \mathcal{V} | f(\mathbf{v}) < +\infty\}$ is called the *effective domain* of f. When $\operatorname{dom}(f) \neq \emptyset$, i.e. f is not identically equal to $+\infty$, we say that f is *proper*.

One of the most often used notions in convex analysis is the *lower semi-continuity*. A function $f : \mathcal{V} \to \overline{\mathcal{R}}$ is *lower semi-continuous* (l.s.c) at $\mathbf{v}_0$ if, for every $\varepsilon > 0$, there is a $\delta > 0$ such that $\|\mathbf{v} - \mathbf{v}_0\| < \delta$ implies that $f(\mathbf{v}_0) - \varepsilon \leq f(\mathbf{v})$. When f verifies this condition at each $\mathbf{v}_0 \in \mathcal{V}$, f is called l.s.c. on $\mathcal{V}$.

A subset C of $\mathcal{V}$ is said to be *convex* if the line segment $[\mathbf{v}, \mathbf{v}'] := \{(1-\lambda)\mathbf{v} + \lambda \mathbf{v}' | \ 0 \leq \lambda \leq 1\}$ is included in C whenever $\mathbf{v}, \mathbf{v}' \in C$. The *epigraph* of f, defined as the subset

$$\operatorname{epi}(f) := \{(\mathbf{v}, \alpha) \in \mathcal{V} \times \mathcal{R} \mid f(\mathbf{v}) \leq \alpha\}, \tag{2.2}$$

is a constructive notion in convex analysis. Indeed, a function $f : \mathcal{V} \to \overline{\mathcal{R}}$ is convex if and only if $\operatorname{epi}(f)$ is a convex subset of $\mathcal{V} \times \mathcal{R}$; in addition, f is l.s.c. on $\mathcal{V}$ if and only if $\operatorname{epi}(f)$ is *closed* in $\mathcal{V} \times \mathcal{R}$ [18].

Given a convex set C of $\mathcal{V}$, the function $\sigma_C : \mathcal{U} \to \overline{\mathcal{R}}$, defined by

$$\sigma_C(\mathbf{u}) := \sup_{\mathbf{v} \in C} \{\mathbf{v}.\mathbf{u}\}, \tag{2.3}$$

is called the *support function* of C. It can be interpreted geometrically by Fig. 1, in which the *supporting hyperplane* P to C is perpendicular to $\mathbf{u}$ and has $\mathbf{v}_0$ as the supporting point. We can easily show that σ_c is *positively homogeneous* (of degree one), i.e. $\sigma_c(\lambda\mathbf{u}) = \lambda\sigma_c(\mathbf{u})$ for $\lambda \geq 0$, and convex.

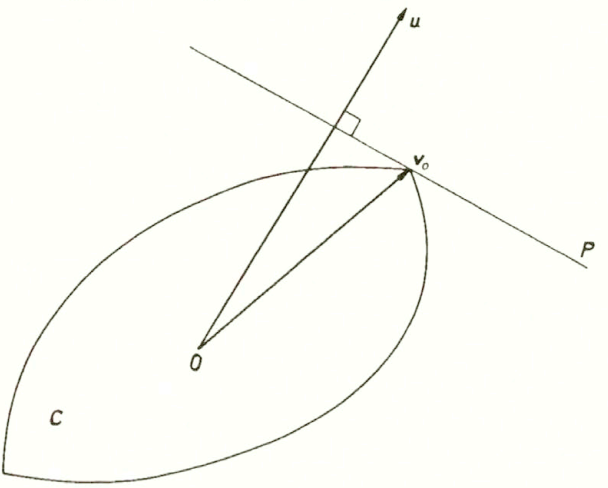


FIG. 1. Geometric interpretation of the support function of C: $\sigma_c(\mathbf{u}) = \mathbf{u}.\mathbf{v}_0$.

Consider any proper function $f : \mathcal{V} \to \overline{\mathcal{R}}$. Then the function $f^* : \mathcal{U} \to \overline{\mathcal{R}}$ defined through the Legendre–Fenchel transformation

$$f^*(\mathbf{u}) := \sup_{\mathbf{v}\in\mathcal{V}}\{\mathbf{v}.\mathbf{u} - f(\mathbf{v})\} \tag{2.4}$$

is called the *conjugate* of f and the function $f^{**} : \mathcal{V} \to \overline{\mathcal{R}}$, defined as

$$f^{**}(\mathbf{v}) := \sup_{\mathbf{u}\in\mathcal{U}}\{\mathbf{v}.\mathbf{u} - f^*(\mathbf{u})\}, \tag{2.5}$$

is the *biconjugate* of f. In the case where f is from $\mathcal{R}$ to $\overline{\mathcal{R}}$, a geometric method of constructing its conjugate f^* is given in Fig. 2. Let us recall the important result that f is convex and l.s.c. if and only if $f^{**} = f$ (see e.g. [18]).

The *subdifferential* of a convex function is another key concept necessary for the presentation of Moreau's pseudo-potential formalism. Letting $f : \mathcal{V} \to \overline{\mathcal{R}}$ be a proper convex function, the subset

$$\partial f(\mathbf{v}) := \{\mathbf{u} \in \mathcal{U} \mid f(\mathbf{v}') \geq f(\mathbf{v}) + (\mathbf{v}' - \mathbf{v}).\mathbf{u}, \quad \forall \mathbf{v}' \in \mathcal{V}\} \tag{2.6}$$

is called the subdifferential of f at $\mathbf{v}$. The elements of $\partial f(\mathbf{v})$ are called the *subgradients* of f at $\mathbf{v}$. If f is not finite at $\mathbf{v}$, we define $\partial f(\mathbf{v}) = \emptyset$. In the case of $\partial f(\mathbf{v}) \neq \emptyset$, f is said to be *subdifferentiable* at $\mathbf{v}$. In geometric terms, $\partial f(\mathbf{v}_0)$ is composed of all the slopes a supporting hyperplane to epi(f) at $(\mathbf{v}_0, f(\mathbf{v}_0))$ can take (Fig. 2). Clearly, $\partial f(\mathbf{v})$ reduces to $\nabla f(\mathbf{v})$ when f is differentiable at $\mathbf{v}$.

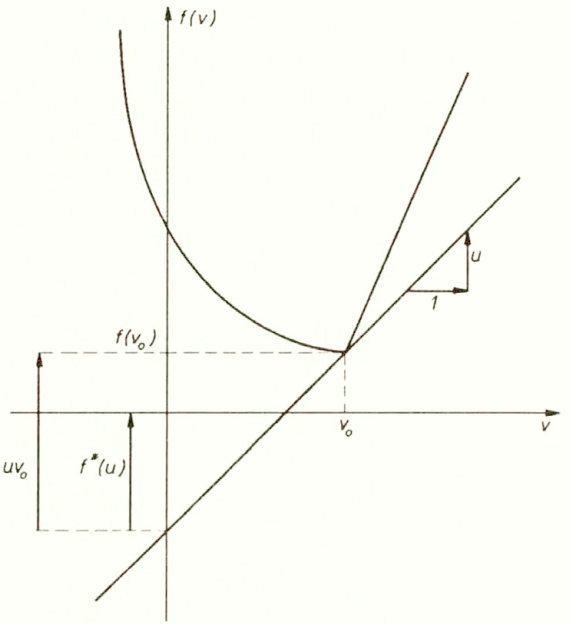


FIG. 2. Geometric interpretation of the conjugate of f : $f^*(u) = uv_0 - f(v_0)$.

2.2. Pseudo-potential formalism for plasticity

Let $\mathcal{V}$ be identified with the symmetric stress tensor space $\mathcal{S}$ and $\mathcal{U}$ with the symmetric plastic strain rate tensor space $\dot{\mathcal{E}}$. These two spaces are placed in duality by introducing the bilinear form: $\mathbf{S} : \dot{\mathbf{E}} := \operatorname{tr}(\mathbf{S}\dot{\mathbf{E}}^T) = \operatorname{tr}(\mathbf{S}\dot{\mathbf{E}})$ for $\mathbf{S} \in \mathcal{S}$ and $\dot{\mathbf{E}} \in \dot{\mathcal{E}}$, where tr represents the trace and $\dot{\mathbf{E}}^T$ the transpose of $\dot{\mathbf{E}}$.

Let C be a given *closed convex* set in $\mathcal{S}$, that contains the origin $\mathbf{0}$. Suppose the yield surface corresponds to its boundary ∂C and the elastic domain to its interior $\operatorname{int}(C)$. By definition, the *indicator function* $I_c(\mathbf{S})$ of C is

$$I_c(\mathbf{S}) := \begin{cases} 0, & \text{if } \mathbf{S} \in C, \\ +\infty, & \text{if } \mathbf{S} \notin C. \end{cases} \tag{2.7}$$

On account of the conditions imposed on C, $I_c(\mathbf{S})$ is proper, convex and l.s.c.. Applying the definition (2.4) to $I_c(\mathbf{S})$ and using (2.3), we get

$$I_c^*(\dot{\mathbf{E}}) = \sigma_c(\dot{\mathbf{E}}), \tag{2.8}$$

i.e. the conjugate of $I_c(\mathbf{S})$ coincides with the support function $\sigma_c(\dot{\mathbf{E}})$ of C, which is convex and positively homogeneous. Furthermore, owing to the convexity and lower semi-continuity of I_c, we can write [18]

$$\sigma_c^*(\mathbf{S}) = (I_c^*)^*(\mathbf{S}) = I_c(\mathbf{S}). \tag{2.9}$$

It is seen from Eqs. (2.8) and (2.9) that *the Legendre–Fenchel transformation* (2.4) *induces a one-to-one symmetric correspondence between* $I_c(\mathbf{S})$ *and* $\sigma_c(\dot{\mathbf{E}})$.

Following Moreau's pseudo-potential formalism for plasticity [10, 11], the plastic law takes the form

$$\dot{\mathbf{E}} \in \partial I_c(\mathbf{S}). \tag{2.10}$$

With the definitions (2.6) and (2.7) in mind, we can easily verify that

$$\partial I_c(\mathbf{S}) = \begin{cases} \emptyset, & \text{if } \mathbf{S} \notin C, \\ N_c(\mathbf{S}), & \text{if } \mathbf{S} \in C, \end{cases} \tag{2.11}$$

where $N_c(\mathbf{S})$ represents the *outward normal cone* to the convex C at $\mathbf{S} \in C$ (Fig. 3):

$$N_c(\mathbf{S}) := \{\dot{\mathbf{E}} \in \dot{\mathcal{E}} \mid (\mathbf{S}' - \mathbf{S}) : \dot{\mathbf{E}} \leq 0, \quad \forall \mathbf{S}' \in C\}, \tag{2.12}$$

i.e., $\dot{\mathbf{E}}$ is an element of $N_c(\mathbf{S})$ if $\dot{\mathbf{E}}$ does not make an "acute angle" with any $\delta\mathbf{S}$ which has $\mathbf{S}$ as starting point and is in C. In particular,

$$N_c(\mathbf{S}) = \{\mathbf{0}\} \quad \text{when} \quad \mathbf{S} \in \text{int}\,(C), \tag{2.13}$$

and $N_c(\mathbf{S})$ reduces to the outward normal half-line to ∂C when $\mathbf{S}$ is a regular point of ∂C (Fig. 3). Therefore, the classical normality rule of plasticity is included in

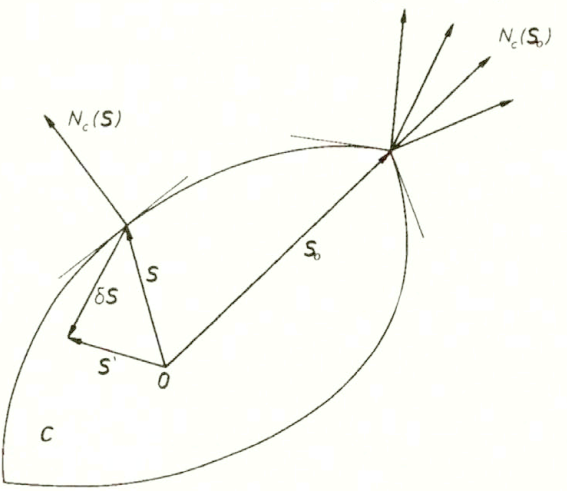


FIG. 3. Outward normal cone to the convex C.

(2.10). A justification for (2.13) can be given as follows. Suppose $\mathbf{S} \in \text{int}\,(C)$ and consider $\dot{\mathbf{E}} \in N_c(\mathbf{S})$. Then, for any $\delta\mathbf{S}$ of sufficiently small norm such that $\delta\mathbf{S}$ with $\mathbf{S}$ as starting point is in C, we have $\delta\mathbf{S} : \dot{\mathbf{E}} \leq 0$ and $-\delta\mathbf{S} : \dot{\mathbf{E}} \leq 0$. Hence, it is necessary that $\dot{\mathbf{E}} = \mathbf{0}$.

It is well-known in convex analysis [14, 18] that the inclusion (2.10) is equivalent to either one of the following two conditions:

$$\mathbf{S} \in \partial I_C^*(\dot{\mathbf{E}}) = \partial \sigma_C(\dot{\mathbf{E}}), \tag{2.14}$$

$$\mathbf{S} : \dot{\mathbf{E}} = I_C(\mathbf{S}) + I_C^*(\dot{\mathbf{E}}) = I_C(\mathbf{S}) + \sigma_C(\dot{\mathbf{E}}), \tag{2.15}$$

in which (2.14) acts as the inverse of (2.10) in the sense of multivalued mappings, i.e. $\mathbf{S} \in \partial I_C^*(\dot{\mathbf{E}})$ if and only if $\dot{\mathbf{E}} \in \partial I_C(\mathbf{S})$. In (2.15), the term $\mathbf{S} : \dot{\mathbf{E}}$ represents the plastic dissipation power density since $\dot{\mathbf{E}}$ and $\mathbf{S}$ are implicitly related by $\dot{\mathbf{E}} \in \partial I_C(\mathbf{S})$ or $\mathbf{S} \in \partial I_C^*(\dot{\mathbf{E}})$. As $I_C(\mathbf{S}) = 0$ for $\mathbf{S} \in C$, it follows from (2.15) that $\sigma_C(\dot{\mathbf{E}}) = \mathbf{S} : \dot{\mathbf{E}}$ when $\mathbf{S} \in C$ and hence $\sigma_C(\dot{\mathbf{E}})$ plays the role of the *plastic dissipation function*. Moreover, (2.12) shows that the plastic law (2.10) or its inverse (2.14) is consistent with the maximum dissipation principle of plasticity.

3. Some one-to-one correspondences

3.1. Gauge yield function

How to analytically describe the convex elastic domain C is not specified in the precedent section. However, this specification is necessary for the explicit expression of the plastic law (2.10).

With Moreau's pseudo-potential formalism, we have seen that the dissipation function of plasticity is identified with the support function $\sigma_C(\dot{\mathbf{E}})$ of the elastic domain. According to convex analysis [18], the support function $\sigma_C(\dot{\mathbf{E}})$ of a closed convex set C is *characteristic* in the sense that C is completely determined by $\sigma_C(\dot{\mathbf{E}})$, i.e.

$$C = \{\mathbf{S} \in \mathcal{S} \mid \mathbf{S} : \dot{\mathbf{E}} \leq \sigma_C(\dot{\mathbf{E}}), \quad \forall \mathbf{S} \in \mathcal{S}\}. \tag{3.1}$$

Physically, this means that the elastic domain C of a (rigid-perfectly) plastic material obeying the normality rule can be determined by measuring its plastic dissipation power. The fact that $\sigma_C(\dot{\mathbf{E}})$ is a characteristic function of C is due to its convexity and positive homogeneity properties. More precisely, *any positively homogeneous closed proper convex function can be regarded as the support function of a certain closed convex set and a one-to-one correspondence exists between them* [18]. In the context of plasticity, the idea of characterizing the elastic domain by the plastic dissipation function was initially introduced by PRAGER [17]. In a work by TELEGA [20], the same idea has been used for dry friction. On studying certain specific cases LUBLINER [10, 11] has remarked that if the plastic dissipation function is positively homogeneous and non-smooth, then a yield criterion follows.

Describing the elastic domain C by the plastic dissipation function is not the usual way employed in plasticity theories. Indeed, a scalar function f, defined over the stress space $\mathcal{S}$ and called the yield function, is commonly used to describe C:

$$C = \{\mathbf{S} \in \mathcal{S} \mid f(\mathbf{S}) \leq 0\}. \tag{3.2}$$

In such a way, a countless number of (convex continuous) functions $f(\mathbf{S})$ can be utilized in (3.2) to get the same convex closed set C. Hence, it is natural to ask whether among these yield functions there exists a *characteristic* one in the sense that it is associated with C in a *one-to-one* manner. Let us illustrate this problem by a simple example.

EXAMPLE 3.1. For von Mises criterion, the following two yield functions of different forms are often used in the literature:

$$(a) \qquad f(\mathbf{S}) = J_2 - k^2; \qquad (b) \qquad f(\mathbf{S}) = \sqrt{J_2} - k,$$

where $J_2 = (\mathbf{S}^d : \mathbf{S}^d)/2$, $\mathbf{S}^d$ denoting the deviator of $\mathbf{S}$ and k being material elastic limit in simple shear stress. It is obvious that the two convex sets obtained by Eq. (3.2) together with (a) and (b) are the same.

Now we come to the construction of the characteristic yield function. Geometrically, remark that every convex set C in $\mathcal{S}$ can be considered as a cross-section of certain convex cone K in $\mathcal{S} \times \mathcal{R}$. Indeed, if we define the convex cone K from C by

$$K := \{(\lambda \mathbf{S}, \lambda) \in \mathcal{S} \times \mathcal{R} \mid \mathbf{S} \in C, \ \lambda \geq 0\}, \tag{3.3}$$

then the intersection of K with the hyperplane $P := \{\mathbf{S}, \lambda) \mid \mathbf{S} \in \mathcal{S}, \ \lambda = 1\}$ in $\mathcal{S} \times \mathcal{R}$ is precisely C. A one-dimensional example is given in Fig. 4. Let us introduce $g(\mathbf{S})$ by

$$g(\mathbf{S}) := \inf\{\lambda \mid (\mathbf{S}, \lambda) \in K\}. \tag{3.4}$$

This function can be shown to be closed, convex and positively homogeneous and its epigraph coincides with K [18].

Substituting (3.3) into (3.4), we obtain

$$g(\mathbf{S}) = \inf\{\lambda \geq 0 \mid \mathbf{S} \in \lambda C\} =: \gamma_c(\mathbf{S}) \tag{3.5}$$

which identifies $g(\mathbf{S})$ with the *gauge* or *Minkowski distance function* $\gamma_c(\mathbf{S})$ of the convex C. It is readily shown that $\gamma_c(\mathbf{S})$ *is a non-negative positively homogeneous convex function with* $\gamma_c(\mathbf{0}) = 0$. A geometrical interpretation for this $\gamma_c(\mathbf{S})$ is given in Fig. 5. Let $\mathbf{S} \neq \mathbf{0}$ and $\mathbf{S}_0$ be the point at which the ray vector from the origin and along $\mathbf{S}$ pierces the boundary of C. (If C is not bounded in the direction $\mathbf{S}$, $\mathbf{S}_0$ is a point at infinity). According to (3.5), we have $\mathbf{S} = \gamma_c(\mathbf{S})\mathbf{S}_0$ and so

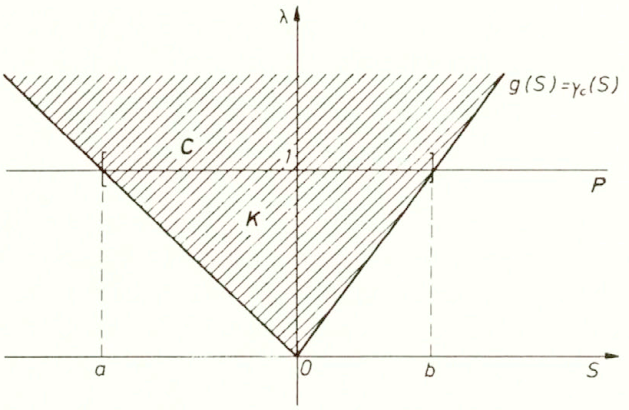


FIG. 4. Cone K and gauge function $\gamma_c(\mathbf{S})$ associated with the line segment $C = [a, b]$.

$\gamma_c(\mathbf{S}) = \|\mathbf{S}\|/\|\mathbf{S}_0\|$ with $\|.\|^2 := (.) : (.)$. Thus, the value of the function $\gamma_c(\mathbf{S})$ can be regarded as the distance (relative to the convex C) from the origin to the point $\mathbf{S}$ if $\|\mathbf{S}_0\|$ is taken as unit of measure. When the convex C is not a ball centered at the origin, the unit measure $\|\mathbf{S}_0\|$ depends on the orientation of $\mathbf{S}$. In addition, $\gamma_c(\mathbf{S}) = 0$ if the ray vector along $\mathbf{S}$ does not intersect the boundary of C.

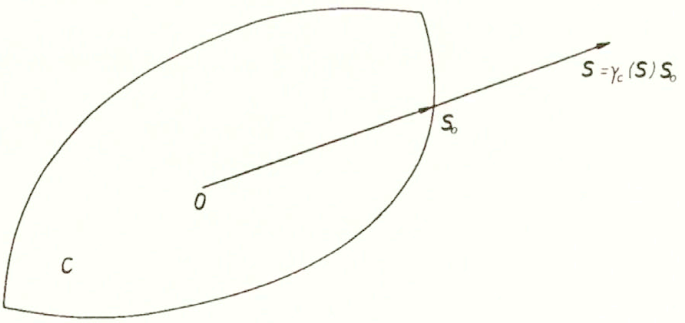


FIG. 5. Geometrical interpretation of $\gamma_c(\mathbf{S})$.

In particular, if, in addition to the conditions (closedness, convexity and $\mathbf{0} \in \text{int}(C)$) previously imposed on C, we assume it to be *bounded* and *centrally symmetric with respect to the origin* $\mathbf{0}$ (i.e., $-\mathbf{S} \in C$ for all $\mathbf{S} \in C$), its gauge function $\gamma_c(\mathbf{S})$ has the properties of a norm [14, 18]:

(a) $\gamma_c(\mathbf{S}) \geq 0$ for all $\mathbf{S} \in \mathcal{S}$;

(a′) $\gamma_c(\mathbf{S}) = 0$ if and only if $\mathbf{S} = \mathbf{0}$;

(b) $\gamma_c(\alpha\mathbf{S}) = |\alpha|\gamma_c(\mathbf{S})$ for $\mathbf{S} \in \mathcal{S}$ and all $\alpha \in \mathcal{R}$;

(c) $\gamma_c(\mathbf{S} + \mathbf{S}') \leq \gamma_c(\mathbf{S}) + \gamma_c(\mathbf{S}')$ for all $\mathbf{S}, \mathbf{S}' \in \mathcal{S}$.

In plasticity, the boundedness of the elastic domain C is often not satisfied; for instance, the elastic domains of most metal materials are pressure-independent. In this case, $\gamma_c(\mathbf{S})$ reduces to a *semi-norm* from lack of property (a'). However, it

is interesting to observe that certain gauge functions $\gamma_c(\mathbf{S})$ which are semi-norms in $\mathcal{S}$ behave as norms in some subspaces of $\mathcal{S}$.

EXAMPLE 3.2. For a von Mises material, the gauge function of its elastic domain C has the form of

$$\gamma_c(\mathbf{S}) = \frac{1}{2}\sqrt{J_2} = \frac{1}{\sqrt{2}k}\|\mathbf{S}^d\| =: \gamma_c|_{\mathcal{S}^d}(\mathbf{S}^d).$$

In $\mathcal{S}$, γ_c is only a semi-norm but its restriction $\gamma_c|_{\mathcal{S}^d}$ to the deviatoric stress tensor space $\mathcal{S}^d$ is a norm in $\mathcal{S}^d$.

EXAMPLE 3.3. For a Tresca material, we have

$$\gamma_c(\mathbf{S}) = \frac{1}{2k}\sup_{i,j}\{|S_i - S_j|\} = \frac{1}{2k}\sup_{i,j}\{|S_i^d - S_j^d|\} =: \gamma_c|_{\mathcal{S}^d}(\mathbf{S}^d),$$

where S_i and S_i^d $(i, j = 1, 2, 3;\ i \neq j)$ are respectively the principal stresses of $\mathbf{S}$ and $\mathbf{S}^d$. As in Example 3.2, γ_c is a semi-norm in $\mathcal{S}$ whereas $\gamma_c|_{\mathcal{S}^d}$ is a norm in $\mathcal{S}^d$.

EXAMPLE 3.4. In the plane stress case, the gauge functions for von Mises and Tresca materials are given respectively by

$$\text{(i)} \qquad \gamma_c(\mathbf{S}) = \frac{1}{k}\sqrt{J_2},$$

$$\text{(ii)} \qquad \gamma_c(\mathbf{S}) = \frac{1}{2}\sup\{|S_1|, |S_2|, |S_1 - S_2|\},$$

where $\mathbf{S}$ belongs to the two-dimensional stress tensor space $\mathcal{S}^2$. These two functions act as norms in $\mathcal{S}^2$.

Nevertheless, if C has no central symmetry with respect to the origin, $\gamma_c(\mathbf{S})$ is even not a semi-norm due to the fact that $\gamma_c(-\mathbf{S}) = \gamma_c(\mathbf{S})$ cannot hold for every $\mathbf{S} \in \mathcal{S}$, and hence the property (b) necessary for a semi-norm is not verified. In order to illustrate this remark, let us consider the following example:

EXAMPLE 3.5. When Drucker-Prager criterion applies, the gauge function $\gamma_c(\mathbf{S})$ for the corresponding convex elastic domain C has the form of

$$\gamma_c(\mathbf{S}) = \frac{1}{k}\left(\sqrt{J_2} + \frac{\mu}{\sqrt{6}}I_1\right),$$

where $I_1 = \operatorname{tr}(\mathbf{S})$ and μ is a material constant. Obviously, this function is not centrally symmetric relative to the origin.

In agreement with the definition (3.5), the convex C is given by

$$C = \{\mathbf{S} \in \mathcal{S} \mid \gamma_c(\mathbf{S}) \leq 1\}. \tag{3.6}$$

According to convex analysis [18], given a *closed* convex set C, a *unique* closed gauge function $\gamma_c(\mathbf{S})$ is defined through (3.4) or (3.5); conversely, C is *uniquely* determined by $\gamma_c(\mathbf{S})$. Up to this stage, we have resolved the problem of finding a characteristic function (defined over $\mathcal{S}$) for C such that a one-to-one correspondence is established between this function and C. From now on, it is natural to refer to $\gamma_c(\mathbf{S})$ as the *gauge yield function* and to $\gamma_c(\mathbf{S}) \leq 1$ as the *gauge yield criterion*.

3.2. Plastic flow rule with the gauge yield function

Firstly, let us specify (2.10) in the case where C is given in the form of (3.2):

$$\begin{aligned} &\dot{\mathbf{E}} \in \lambda \partial f(\mathbf{S}), \\ &f(\mathbf{S}) \leq 0, \quad \lambda \geq 0, \quad \lambda f(\mathbf{S}) = 0, \end{aligned} \tag{3.7}$$

which is obtained by aplying the Kuhn-Tucker theorem of convex analysis [18] to (2.11) and (3.2). Then, putting $f(\mathbf{S}) = \gamma_c(\mathbf{S}) - 1$ in (3.7), we get

$$\begin{aligned} &\dot{\mathbf{E}} \in \lambda \partial \gamma_c(\mathbf{S}), \\ &\gamma_c(\mathbf{S}) \leq 1, \quad \lambda \geq 0, \quad \lambda \gamma_c(\mathbf{S}) = \lambda, \end{aligned} \tag{3.8}$$

the expression of the plastic flow rule relative to the gauge yield function $\gamma_c(\mathbf{S})$.

In the study of differentiable homogeneous functions, a remarkable fact is Euler's identity. Let $h(\mathbf{S})$ be a *differentiable homogeneous* function (of degree one) from $\mathcal{S}$ to $\mathcal{R}$. Then Euler's identity reads $h(\mathbf{S}) \equiv \nabla h(\mathbf{S}) : \mathbf{S}$; additionally, $\nabla h(\mathbf{S}) = \nabla h(\mathbf{0})$ for all $\mathbf{S} \in \mathcal{S}$. In the following, a similar result is established for any *subdifferentiable* and *positively homogeneous* function $h(\mathbf{S})$. This allows us to show some properties of the multiplier λ in (3.8), due to the utilization of $\gamma_c(\mathbf{S})$.

Let $h(\mathbf{S})$ be a *subdifferentiable positively homogeneous* function. Suppose $\dot{\mathbf{E}} \in \partial h(\mathbf{S})$. By definition, $h(\mathbf{S}') \geq h(\mathbf{S}) + \dot{\mathbf{E}} : (\mathbf{S}' - \mathbf{S})$, $\forall \mathbf{S}' \in \mathcal{S}$. Taking $\mathbf{S}' = \alpha \mathbf{S}$ with $\alpha - 1 > 0$ and using the assumption that $h(\mathbf{S})$ is positively homogeneous, we have $(\alpha - 1)h(\mathbf{S}) \geq (\alpha - 1)\dot{\mathbf{E}} : \mathbf{S}$, i.e. $h(\mathbf{S}) \geq \dot{\mathbf{E}} : \mathbf{S}$. On the other hand, if $\mathbf{S}' = \alpha \mathbf{S}$ with $\alpha \in [0, 1[$, it follows that $h(\mathbf{S}) \leq \dot{\mathbf{E}} : \mathbf{S}$. Consequently,

$$h(\mathbf{S}) = \dot{\mathbf{E}} : \mathbf{S} \quad \text{for} \quad \dot{\mathbf{E}} \in \partial h(\mathbf{S}); \tag{3.9'}$$

in addition,

$$\partial h(\mathbf{S}) \subseteq \partial h(\mathbf{0}), \quad \forall \mathbf{S} \in \mathcal{S} \tag{3.9''}$$

since $\dot{\mathbf{E}} \in \partial h(\mathbf{S})$ implies that $h(\mathbf{S}') \geq h(\mathbf{S}) + \dot{\mathbf{E}} : (\mathbf{S}' - \mathbf{S}) = \dot{\mathbf{E}} : \mathbf{S} = h(\mathbf{0}) + \dot{\mathbf{E}} : (\mathbf{S}' - \mathbf{0})$ for all $\mathbf{S}' \in \mathcal{S}$, namely $\dot{\mathbf{E}} \in \partial h(\mathbf{0})$.

We now return to (3.8). In order to apply (3.9) to (3.8), we write $\dot{\mathbf{E}} \in \lambda \partial \gamma_c(\mathbf{S})$ as $\dot{\mathbf{E}} = \lambda \mathbf{N}$ with $\mathbf{N} \in \partial \gamma_c(\mathbf{S})$. Thus, $\dot{\mathbf{E}} : \mathbf{S} = \lambda \mathbf{N} : \mathbf{S} = \lambda \gamma_c(\mathbf{S})$. Due to the fact that, in (3.8), $\lambda \geq 0$ when $\gamma_c(\mathbf{S}) = 1$ and $\lambda = 0$ when $\gamma_c(\mathbf{S}) < 1$, we have

$$\lambda = \dot{\mathbf{E}} : \mathbf{S} = \sigma_c(\dot{\mathbf{E}}). \tag{3.10}$$

HILL [8] refers to λ as a *"work-equivalent measure of strain-rate"*, since the plastic dissipation power is the same for all $\dot{\mathbf{E}}$ with a common value λ of $\sigma_c(\dot{\mathbf{E}})$.

EXAMPLE 3.6. The plastic flow associated to the gauge yield criterion of Drucker–Prager is

$$\dot{\mathbf{E}} = \frac{\lambda}{\sqrt{2}k} \left[\frac{\mathbf{S}^d}{\|\mathbf{S}^d\|} + \frac{\mu}{\sqrt{3}} \mathbf{I} \right],$$
$$\gamma_c(\mathbf{S}) \leq 1, \quad \lambda \geq 0, \quad \lambda \gamma_c \mathbf{S} = \lambda.$$

A simple calculation shows that

$$\lambda = \dot{\mathbf{E}} : \mathbf{S} = \sigma_c(\dot{\mathbf{E}}) = \frac{k\sqrt{2\dot{\mathbf{E}} : \dot{\mathbf{E}}}}{\sqrt{1 + \mu^2}}.$$

3.3. Duality correspondence between $\sigma_c(\dot{\mathbf{E}})$ and $\gamma_c(\mathbf{S})$

Prior to showing the duality relationship between $\sigma_c(\dot{\mathbf{E}})$ and $\gamma_c(\mathbf{S})$, we introduce the polar of a convex set, a very useful notion in the duality theory of convex sets and functions [14, 18]. By definition, the polar C^0 of the convex C is given by

$$C^0 := \{\dot{\mathbf{E}} \in \dot{\mathcal{E}} \mid \mathbf{S} : \dot{\mathbf{E}} \leq 1, \quad \forall \mathbf{S} \in C\}. \tag{3.11}$$

It is straightforward to verify that C^0 is a convex set *containing the origin*. As C is assumed to be as *closed* convex set containing the origin, the following classical result of convex analysis [18]

$$C^{00} = C \tag{3.12}$$

holds and implies that a *one-to-one* correspondence exists between C and its polar C^0. In Fig. 6, the polars to the elastic domains of von Mises and Tresca materials subjected to plane stress (see example 3.4) are shown as an illustration. (Note that von Mises and Tresca criteria are here matched in shear rather than traction).

Let $\sigma_{c^0}(\mathbf{S})$ and $\gamma_{c^0}(\dot{\mathbf{E}})$ represent the support and gauge functions of C^0:

$$\sigma_{c^0}(\mathbf{S}) := \sup_{\dot{\mathbf{E}} \in C^0} \{\mathbf{S} : \dot{\mathbf{E}}\}, \tag{3.13}$$

$$\gamma_{c^0}(\dot{\mathbf{E}}) := \inf\{\lambda \geq 0) \mid \dot{\mathbf{E}} \in \lambda C^0\}. \tag{3.14}$$

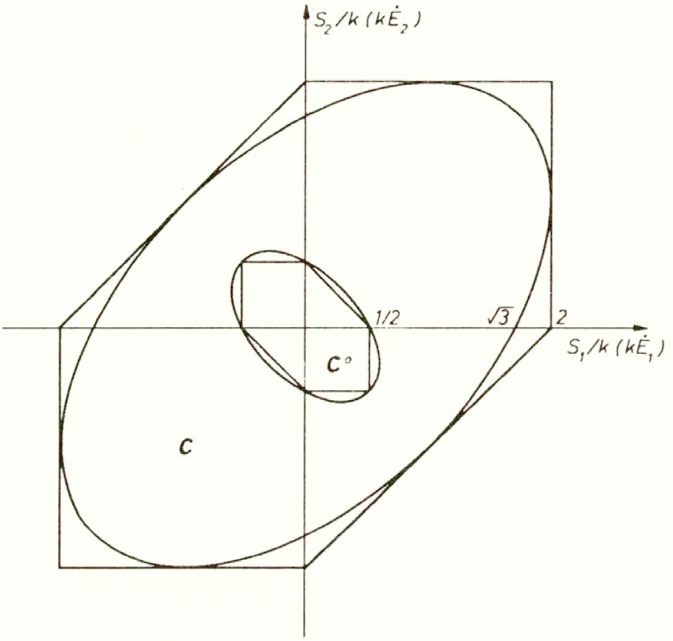


FIG. 6. Polars C^0 to the elastic domains of von Mises and Tresca for plane stress.

In convex analysis (see e.g. [14, 18]), we have the following important relations:

$$\gamma_{C^0}(\dot{\mathbf{E}}) = \sigma_C(\dot{\mathbf{E}}); \tag{3.15}$$

$$\gamma_{C^{00}}(\mathbf{S}) = \sigma_{C^0}(\mathbf{S}) = \gamma_C(\mathbf{S}). \tag{3.16}$$

A simple demonstration is now given. First, in view of definitions (3.14) and (3.11),

$$\begin{aligned}\gamma_{C^0}(\dot{\mathbf{E}}) &= \inf\{\lambda \geq 0 \,|\, \dot{\mathbf{E}} \in \lambda C^0\} \\ &= \inf\{\lambda \geq 0 \,|\, \dot{\mathbf{E}} : \mathbf{S} \leq \lambda, \ \forall \mathbf{S} \in C\} \\ &= \sup_{\mathbf{S}\in C}\{\dot{\mathbf{E}} : \mathbf{S}\} \\ &= \sigma_C(\dot{\mathbf{E}})\end{aligned}$$

which justifies (3.15); next, replacing C^0 by C and invoking (3.12), we obtain (3.16). Taken together, formulae (3.15) and (3.16) show that *the polarity operation* (3.11) *induces a unique symmetric correspondence between the support and gauge functions of* C (*or* C^0).

In order to describe this fact in an *analytical* way, ROCKAFELLAR [18] defines the polar g^0 of a gauge or support function g as

$$g^0(\dot{\mathbf{E}}) := \inf\{\lambda \geq 0 \,|\, \dot{\mathbf{E}} : \mathbf{S} \leq \lambda g(\mathbf{S}), \ \forall \mathbf{S} \in \mathcal{S}\}. \tag{3.17}$$

With this definition, the following relations

$$\gamma_c^0(\dot{\mathbf{E}}) = \gamma_{c^0}(\dot{\mathbf{E}}) = \sigma_c(\dot{\mathbf{E}}), \tag{3.18}$$

$$\gamma_c^{00} = \gamma_{c^{00}}(\mathbf{S}) = \sigma_c(\mathbf{S}) = \sigma_{c^0}(\mathbf{S}) = \sigma_c^0(\mathbf{S}) \tag{3.19}$$

hold [18] and hence the duality relation between $\gamma_c(\mathbf{S})$ and $\gamma_c(\dot{\mathbf{E}})$ is established without involving the polar C^0 of C.

When $g(\mathbf{S})$ is *positive* for $\mathbf{S} \neq 0$ and *finite* for every $\mathbf{S} \in \mathcal{S}$, the definition (3.17) is equivalent to the following

$$g^0(\dot{\mathbf{E}}) = \sup_{\substack{\mathbf{S}\in\mathcal{S} \\ \mathbf{S}\neq\mathbf{0}}} \frac{\mathbf{S} : \dot{\mathbf{E}}}{g(\mathbf{S})}. \tag{3.20'}$$

For example, this formula is applicable to $\gamma_c(\mathbf{S})$ or $\sigma_c(\dot{\mathbf{E}})$ if the convex C is *compact* (closed and bounded) with $\mathbf{0} \in \text{int}(C)$. By using the positive homogeneity of $g(\mathbf{S})$, we can write (3.20') as

$$g^0(\dot{\mathbf{E}}) = \sup_{\|\mathbf{S}\|=1} \frac{\mathbf{S} : \dot{\mathbf{E}}}{g(\mathbf{S})}. \tag{3.20''}$$

If $g(\mathbf{S})$ is furthermore assumed to be symmetric with respect to the origin, i.e. $g(-\mathbf{S}) = g(\mathbf{S})$, it is a norm and $g^0(\dot{\mathbf{E}})$ in the form of (3.20'') can easily be recognized as its *dual norm*. For (3.20') or (3.20'') to be applicable to von Mises or Tresca criterion, it is sufficient to replace $\mathbf{S}$, $\dot{\mathbf{E}}$, $\mathcal{S}$ and $\dot{\mathcal{E}}$, respectively, by $\mathbf{S}^d$, $\dot{\mathbf{E}}^d$, $\mathcal{S}^d$ and $\dot{\mathcal{E}}^d$.

As a direct result of (3.17), we have the inequality

$$\mathbf{S} : \dot{\mathbf{E}} \leq g(\mathbf{S})g^0(\dot{\mathbf{E}}), \quad \forall \mathbf{S} \in \mathcal{S} \text{ and } \forall \dot{\mathbf{E}} \in \dot{\mathcal{E}}. \tag{3.21}$$

Cauchy–Schwarz and Hölder inequalities are two examples of this. In fact, given a gauge function $g(\mathbf{S})$, $g^0(\dot{\mathbf{E}})$ represents the *smallest* of the functions $q(\dot{\mathbf{E}})$ satisfying the condition $\mathbf{S} : \dot{\mathbf{E}} \leq g(\mathbf{S})q(\dot{\mathbf{E}})$ for all $\mathbf{S} \in \mathcal{S}$ and $\dot{\mathbf{E}} \in \dot{\mathcal{E}}$. If (3.21) is applied to $\gamma_c(\mathbf{S})$ or $\sigma_c(\dot{\mathbf{E}})$ while taking account of (3.18) and (3.19), we get

$$\mathbf{S} : \dot{\mathbf{E}} \leq \gamma_c(\mathbf{S})\gamma_c^0(\dot{\mathbf{E}}) = \gamma_c(\mathbf{S})\sigma_c(\dot{\mathbf{E}}), \quad \forall \mathbf{S} \in \mathcal{S} \text{ and } \forall \dot{\mathbf{E}} \in \dot{\mathcal{E}}. \tag{3.22}$$

This inequality for plasticity has been obtained recently by YANG [21] with the help of a generalized Hölder inequality that can be regarded as a particular form of (3.21). As a matter of fact, (3.21) reduces to Yang's generalized inequality when $g(\mathbf{S})$ is a differentiable norm.

In the case of $\dot{\mathbf{E}} \in \partial I_c(\mathbf{S})$, the inequality (3.22) is attained, i.e.

$$\mathbf{S} : \dot{\mathbf{E}} = \gamma_c(\mathbf{S})\sigma_c(\dot{\mathbf{E}}) \quad \text{if} \quad \dot{\mathbf{E}} \in \partial I_c(\mathbf{S}). \tag{3.23}$$

This is immediate from the formula (3.22) together with (3.8) and (3.10).

Below, two examples are given to illustrate the polarity operation (3.17) and the duality correspondence between $\sigma_c(\dot{\mathbf{E}})$ and $\gamma_c(\mathbf{S})$ in detail.

EXAMPLE 3.7. It is proposed to construct the dissipation function $\sigma_c(\dot{\mathbf{E}})$ for a von Mises material from its gauge yield function $\gamma_c(\mathbf{S})$ by introducing $g = \gamma_c(\mathbf{S}) = \sqrt{J_2}/k$ into (3.17):

$$\sigma_c(\dot{\mathbf{E}}) = \gamma_c^0(\dot{\mathbf{E}}) = \inf\{\lambda \geq 0 \mid \sqrt{2}k\,\dot{\mathbf{E}} : \mathbf{S} \leq \lambda\sqrt{\mathbf{S}^d : \mathbf{S}^d}, \quad \forall \mathbf{S} \in \mathcal{S}\}.$$

We need to distinguish two cases. If $\operatorname{tr}\dot{\mathbf{E}} \neq 0$, $\sigma_c(\dot{\mathbf{E}}) = +\infty$; to see this, it is enough to take $\mathbf{S} = p\mathbf{I}$ ($p \in \mathcal{R}$). Next, if $\operatorname{tr}\dot{\mathbf{E}} = 0$, the space $\mathcal{S}$ can be replaced by $\mathcal{S}^d$ and $\sigma_c(\dot{\mathbf{E}})$ is nothing else than the dual norm of $\gamma_c(\mathbf{S})$ in $\mathcal{S}^d$, i.e., $\sigma_c(\dot{\mathbf{E}}) = k(2\dot{\mathbf{E}}^d : \dot{\mathbf{E}}^d)^{1/2}$. In sum, $\sigma_c(\dot{\mathbf{E}})$ can be expressed in the form of

$$\sigma_c(\dot{\mathbf{E}}) = k\left(\sqrt{2\,\dot{\mathbf{E}}^d : \dot{\mathbf{E}}^d} + I_{\{0\}}(\operatorname{tr}\dot{\mathbf{E}})\right).$$

Here the presence of the indicator function $I_{\{0\}}(\operatorname{tr}(\dot{\mathbf{E}})$ is due to the plastic incompressibility of a von Mises material.

EXAMPLE 3.8. Consider the *anisotropic* gauge yield function

$$\gamma_c(\mathbf{S}) = \sqrt{\mathbf{S} : H\mathbf{S}},$$

where H is a *positive definite three-dimensional tensor of fourth order* with the usual symmetries $H_{ijkl} = H_{jikl} = H_{klij}$ $(i, j, k, l = 1, 2, 3)$. We apply (3.20′) to calculating the polar $\sigma_c(\dot{\mathbf{E}})$ of $\gamma_c(\mathbf{S})$; that is

$$\sigma_c(\dot{\mathbf{E}}) = \sup_{\mathbf{S}\neq\mathbf{0}} \pi(\dot{\mathbf{E}}, \mathbf{S}) \quad \text{with} \quad \pi(\dot{\mathbf{E}}, \mathbf{S}) := \frac{\dot{\mathbf{E}} : \mathbf{S}}{\sqrt{\mathbf{S}.H\mathbf{S}}}.$$

As H is positive definite, the function $\pi(\dot{\mathbf{E}}, \mathbf{S})$ is differentiable provided $\mathbf{S} \neq \mathbf{0}$. For a given $\dot{\mathbf{E}}$ and with $\mathbf{S} \neq \mathbf{0}$, the condition necessary for $\pi(\dot{\mathbf{E}}, \mathbf{S})$ to attain the maximum takes the form

$$\mathbf{0} = \frac{\partial\pi(\dot{\mathbf{E}}, \mathbf{S})}{\partial\mathbf{S}} = \frac{\dot{\mathbf{E}}}{\sqrt{\mathbf{S} : H\mathbf{S}}} - \frac{(\dot{\mathbf{E}} : \mathbf{S})H\mathbf{S}}{(\mathbf{S} : H\mathbf{S})^{3/2}}, \qquad \text{viz.} \quad H^{-1}\dot{\mathbf{E}} = \frac{(\dot{\mathbf{E}} : \mathbf{S})\mathbf{S}}{\mathbf{S} : H\mathbf{S}}.$$

If this condition is satisfied, we have

$$\dot{\mathbf{E}} : H^{-1}\dot{\mathbf{E}} = \frac{(\dot{\mathbf{E}} : \mathbf{S})^2}{\mathbf{S} : H\mathbf{S}} = \left[\frac{(\dot{\mathbf{E}} : \mathbf{S})}{\sqrt{\mathbf{S} : H\mathbf{S}}}\right]^2.$$

Then, it is not difficult to see that

$$\sigma_c(\dot{\mathbf{E}}) = \sup_{\mathbf{S}\neq\mathbf{0}} \pi(\dot{\mathbf{E}}, \mathbf{S}) = \sqrt{\dot{\mathbf{E}} : H^{-1}\dot{\mathbf{E}}}.$$

Note that, when the function $g(\mathbf{S})$ in (3.20′) is non-differentiable at a point other than $\mathbf{S} = \mathbf{0}$, the calculation of $g^0(\dot{\mathbf{E}})$ may become much less straightforward.

4. Conclusion

In the light of some fundamental notions and results of non-smooth convex analysis, an attempt has been made to give a comprehensive account of several one-to-one constitutive correspondences in plasticity. In keeping with Moreau's pseudo-potential formalism, this should contribute to getting a deeper insight into the intricacy of plasticity and to rendering the classical mathematical theory of plasticity more structured.

Acknowledgement. The authors are grateful to the Referee for his critical remarks leading to improvement of the manuscript.

References

1. R.A. EVE, B.D. REDDY and R.T. ROCKAFELLAR, *An internal variable theory of elastoplasticity based on the maximum plastic work inequality*, Q. Appl. Math., **48**, 59–83, 1990.
2. A. FRIAÂ, *Les matériaux de Norton-Hoff généralisés et ses applications en analyse limite*, C.R. Acad. Sci. Paris, **A-286**, 953–956, 1978.
3. P. GERMAIN, *Cours de Mécanique des Milieux Continus*, vol. I, Masson, Paris 1973.
4. B. HALPHEN and Q.S. NGUYEN, *Sur les matériaux standard généralisés*, J. Mécanique, **14**, 39–63, 1975.
5. B. HALPHEN and J. SALENÇON, *Elasto-plasticité*, Presses de l'Ecole Nationale des Ponts et Chaussées, Paris 1987.
6. Q.-C. HE and A. CURNIER, *On a constitutive duality correspondence in plasticity*, Internal Report, EPFL–DME, 1–12, 1992.
7. R. HILL, *The mathematical theory of plasticity*, Oxford University Press, 1950.
8. R. HILL, *Constitutive dual potentials in classical plasticity*, J. Mech. Phys. Solids, **35**, 23–33. 1987.
9. L. HÖRMANDER, *Sur la fonction d'appui des ensembles convexes dans un espace localement convexe*, Arkiv för Mat., **3**, 181–186, 1954.
10. J. LUBLINER, *Non-smooth dissipation functions and yield criteria*, Acta Mechanica, **22**, 289–293, 1975.
11. J. LUBLINER, *On thermodynamic yield criteria*, Acta Mechanica, **37**, 259–263, 1980.
12. J. LUBLINER, *Plasticity theory*, MacMillan Publishing Company, New York 1990.
13. G.A. MAUGIN, *The thermomechanics of plasticity and fracture*, Cambridge University Press, 1992.

14. J.J. MOREAU, *Fonctionnelles convexes*, [in:] Séminaire sur les Equations aux Dérivées Partielles II, pp. 1–108, Collège de France, Paris 1966–1967.
15. J.J. MOREAU, *Sur les lois de frottement, de plasticité et de viscosité*, C. R. Acad. Sci. Paris, **A–271**, 608–611, 1970.
16. J.J. MOREAU, *On unilateral constraints, friction and plasticity*, [in:] New Variational Techniques in Mathematical Physics, G. CAPRIZ and G. STAMPACCHIA [Eds.], pp. 175–322, CIME, Edizioni Cremonese, Roma 1973.
17. W. PRAGER, *Théorie générale des états limites d'équilibre*, J. Math. pures Appl., **34**, 395–406, 1955.
18. R.T. ROCKAFELLAR, *Convex analysis*, Princeton University Press, 1970.
19. G. ROMANO, L. ROSATI and F. MAROTTI DE SCIARRA, *An internal variable theory of inelastic behaviour derived from the uniaxial rigid-perfectly plastic law*, Int. J. Engng. Sci., **31**, 1105–1120, 1993.
20. J.J. TELEGA, *Topics on unilateral contact problems of elasticity and inelasticity*, [in:] Nonsmooth Mechanics and Applications, J.J. MOREAU and P.D. PANAGIOTOPULOS [Eds.], pp. 341–462, Springer Verlag, Wien 1988.
21. R.T. YANG, *On generalized Hölder inequality*, Nonlinear Analysis, Theory, Methods and Applications, **16**, 489–498, 1991.
22. E. ZEIDLER, *Nonlinear functional analysis and its applications – IV. Applications to mathematical physics*, Springer Verlag, New York 1988.

LABORATOIRE DE MÉCANIQUE APPLIQUÉE, DÉPARTEMENT DE GENIE MÉCANIQUE
ECOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE, SWITZERLAND.

Received June 10, 1994; *new version September* 15, 1994.

Arch. Mech., 46, 6, pp. 1005–1027, Warszawa 1994

Outlooks in Saint Venant theory
I. Formal expansions for torsion of Bredt-like sections

F. dell'ISOLA and G. C. RUTA (ROMA)

STARTING FROM a regular curve (middle line), we define a section of a Saint Venant cylinder transporting the curve along its normal; this we call a Bredt-like section. A formal series expansion for Prandtl stress flow function valid for these sections is given. Starting from it, and using standard relationships, formal expansions for warping, shear stress vector field and torsional stiffness are obtained. Although the expansions are affected by degeneration in regularity, at least where curvature of the middle line is not smooth or at its ends, we obtain all traditional results of the theory of strength of materials and are able to give an estimate of the error associated with them. Moreover, some interesting open problems are formulated, concerning both the problems of regularizing the expansions and of generalizing them for sections with variable thickness.

1. Introduction

IT IS WELL KNOWN [1–7] that in the Saint Venant linearized theory of torsion there are two different ways to find the tangential stress vector field, each leading to a boundary value problem. Both ways are equivalent; one is based on the search of a potential function for the stress, sometimes called the Saint Venant function, which is also proportional to the warping; the other leads to the so-called Prandtl stress flow function. If we look at the general expression of the problem, one has to find that vector field $\mathbf{t}$, defined over the section $\mathcal{D}$ of a Saint Venant cylinder, which obeys the Cauchy balance equations

$$\operatorname{div}\mathbf{t} = 0 \quad \text{in } \mathcal{D}^o, \tag{1.1}$$

$$\mathbf{t}\cdot\mathbf{n}_{\partial\mathcal{D}} = 0 \quad \text{along } \partial\mathcal{D}, \tag{1.2}$$

and the integrability condition

$$\oint_{\mathcal{L}} \mathbf{t}\cdot\mathbf{l}_{\mathcal{L}} = 2G\tau A_{\mathcal{L}}, \qquad \forall\mathcal{L}. \tag{1.3}$$

Here $\mathbf{n}_{\partial\mathcal{D}}$ denotes the outer normal unit vector to the boundary of the domain $\mathcal{D}^o$, where $\mathcal{D}^o$ denotes the interior of $\mathcal{D}$, τ is the angle of twist, G is the modulus of elasticity in shear and $\mathcal{L}$ is an arbitrary circuit (closed curve) in $\mathcal{D}$. By $\mathbf{l}_{\mathcal{L}}$ we denote the unit vector tangent to $\mathcal{L}$ while $A_{\mathcal{L}}$ is the area confined by $\mathcal{L}$, with sign depending on the orientation of the cycle with respect to the orientation of the plane $\mathcal{P}$ containing $\mathcal{D}$.

The first method to solve system (1.1)–(1.3) starts from a particular function which satisfies (1.3):

$$\tilde{\mathbf{t}} = G\tau * \mathbf{r}, \tag{1.4}$$

where $*$ is the Hodge operator, which rotates a vector contained in $\mathcal{P}$ by $\pi/2$ in the positive orientation of $\mathcal{P}$. By $\mathbf{r}$ we denoted the position vector of a typical point q in $\mathcal{P}$. Even if the domain is not simply linearly connected (we will suppose, that it is monoconnected), from (1.4) and (1.3) it follows that

$$\oint_{\mathcal{L}} (\mathbf{t} - \tilde{\mathbf{t}}) \cdot \mathbf{l}_{\mathcal{L}} = 0 \tag{1.5}$$

for any circuit $\mathcal{L}$ included in $\mathcal{D}$. Equation (1.5) is equivalent to

$$\mathbf{t} - \tilde{\mathbf{t}} = G\tau \operatorname{grad} \phi, \tag{1.6}$$

where $\phi : \mathcal{D} \to \mathbb{R}$ is a potential function for the field $\mathbf{t} - \tilde{\mathbf{t}}$, sometimes referred to as the Saint Venant warping function. By substitution of (1.6) in (1.1)–(1.2), one obtains

$$\Delta\phi = 0 \qquad \text{in } \mathcal{D}^o, \tag{1.7}$$

$$\operatorname{grad}\phi \cdot \mathbf{n}_{\partial\mathcal{D}} = - * \mathbf{r} \cdot \mathbf{n}_{\partial\mathcal{D}} \qquad \text{along } \partial\mathcal{D}, \tag{1.8}$$

where Δ is the Laplace operator. The problem of finding tangential stress vector field is now reduced to a boundary value problem which is in the standard Dini–Neumann form.

The second method to find the field $\mathbf{t}$ starts with observing that (1.2) implies

$$\oint_{\partial\lambda} \mathbf{t} \cdot \mathbf{n}_{\partial\lambda} = \oint_{\partial\lambda} * \mathbf{t} \cdot \mathbf{l}_{\partial\lambda} = 0. \tag{1.9}$$

With this result, we may now deduce from (1.1) that, for any cycle $\mathcal{L}$,

$$\oint_{\mathcal{L}} * \mathbf{t} \cdot \mathbf{l}_{\mathcal{L}} = \oint_{\partial\mathcal{S}} * \mathbf{t} \cdot * \mathbf{n}_{\partial\mathcal{S}} = \oint_{\partial\mathcal{S}} \mathbf{t} \cdot \mathbf{n}_{\partial\mathcal{S}} = \int_{\mathcal{S}} \operatorname{div} \mathbf{t} = 0, \tag{1.10}$$

$\mathcal{S} \subset \mathcal{D}$ being a surface whose boundary is $\mathcal{L} \bigcup \partial\lambda$. Equation (1.10) assures that

$$* \mathbf{t} = G\tau \operatorname{grad} \psi \quad \Rightarrow \quad \mathbf{t} = -G\tau * \operatorname{grad} \psi, \tag{1.11}$$

where $\psi : \mathcal{D} \to \mathbb{R}$ is known as Prandtl stress flow function [8]. If we substitute this result in (1.1)–(1.3), we obtain another boundary value problem, this time

in a Dirichlet form:

$$\Delta\psi + 2 = 0 \qquad \text{in } \mathcal{D}^o, \tag{1.12}$$

$$\psi = \overline{\psi}_i \qquad \text{along } \partial\mathcal{D}_i, \tag{1.13}$$

$$\oint_{\mathcal{L}} \operatorname{grad}\psi \times \mathbf{l}_{\mathcal{L}} = -2A_{\mathcal{L}} \qquad \forall \text{ closed line } \mathcal{L}. \tag{1.14}$$

In (1.14) the $\partial\mathcal{D}_i$ are all the connected components of $\partial\mathcal{D}$ and the ψ_i' in (1.13) are uniquely determined by (1.14) up to an additive constant. The constant can be chosen in such a way that $\psi = 0$ along the outher connected part of the boundary.

Both systems (1.7)–(1.8) and (1.12)–(1.14) in general do not have a solution in analytical form: such a solution is known only for a few domains, that is, for sections whose boundary is a circle, an ellipse, a square or a triangle with equal sides; in the last two cases the solution is given in terms of a series and is due to Saint Venant himself.

There always is, of course, the possibility of searching solutions of (1.7)–(1.8) and (1.12)–(1.14) by a numerical approach, but this would lead to list of numbers giving no qualitative understanding of the problems. Qualitative results are a good tool for the designer: they put into evidence which parameters are of importance in the design of a structure undergoing torsion.

Among these there are also "technical" results derived to searching solutions sufficiently close to the exact one by approximating the functions of interest by means of linear functions; such an approximation is valid only when the section is made up by thin stripes. These results are usually attributed to VLASOV [9] for open sections and to BREDT [10] for closed, hollow (tubular) sections: Vlasov starts from a constrained material (the section is rigid in its own plane and the middle line of the domain is free of shear), while Bredt uses a particular geometrical construction of the domain which lets him attain his result quite easily.

Our aim is, starting from Bredt's way to imagine the geometry of sections, to develop a different approach to the problem: we will use a "rational" series expansion for the functions of interest (warping, stress, torsional stiffness) and a perturbation technique [11]. This is a well known standard procedure that will provide a set of equations, valid for thicker domains. We will prove that this method will let us recover "technical" results and give us further qualitative results for sections which cannot be considered "thin", both for open and closed sections.

We choose to study a section whose geometry can be described starting from its "middle line", trying to follow Bredt in his analysis of torsion of "thin" closed hollow sections. Let us choose in $\mathcal{P}$ a line Γ of length $\bar{l}$, which can be either open or closed, and parametrize it through an abscissa s:

$$\Gamma(s) := \left\{ q \in \mathcal{P} \,:\, q - o = \mathbf{r}_0(s), \;\; s \in [0, \bar{l}] \right\}, \tag{1.15}$$

where $\mathcal{P}$ is the plane to which the section belongs and o is one of its points, chosen as the origin. It is possible to define a Frenet basis over Γ:

$$\mathbf{l}(s) := \frac{d\mathbf{r}_0(s)}{ds} = \mathbf{r}_0'(s), \tag{1.16}$$

$$\mathbf{m}(s) := \frac{d\mathbf{r}_0'(s)}{ds} = \kappa(s) * \mathbf{l}(s), \tag{1.17}$$

where $\kappa(s)$ is the curvature of Γ at s. After this, let us build the section $\mathcal{D}$ by "thickening" the middle line along $\mathbf{m}$; more precisely, we say that the domain we consider is

$$\mathcal{D}_{\Gamma,\delta} := \left\{ p \in \mathcal{P} : \; p - o = \mathbf{r}_0(s) + z * \mathbf{r}_0'(s), \;\; z \in \left[-\frac{\delta(s)}{2}, \frac{\delta(s)}{2} \right] \right\}, \tag{1.18}$$

$\delta(s)$ being section's thickness at s; the section will be said to be open or closed, depending on Γ being open or closed. We can also write:

$$\mathcal{D} := \left\{ p \in \mathcal{P} : \; p - o = \mathbf{r}_0(s) + z\frac{\delta(s)}{2} * \mathbf{r}_0'(s), \;\; s \in [0, \bar{l}], \;\; z \in [-1, 1] \right\}; \tag{1.19}$$

we choose the orientation of the boundary of $\mathcal{D}$, in order to have always outer normals expressed by the equation

$$\mathbf{n}_{\partial\mathcal{D}} = -\mathbf{m}. \tag{1.20}$$

In this paper we will confine ourselves to the case of sections with constant thickness. This will imply that the curvilinear coordinate system $\mathcal{C} = \{(s, z) \backslash s \in [0, \bar{l}], \; z \in [-1, 1]\}$ (see (1.19)) is orthogonal and therefore the Laplace operator can be represented without use of covariant derivatives, i.e. Christoffel symbols. The choice of such a coordinate system does not lead to a well behaving asymptotic expansion (for a discussion of this problem see [11] Sec. 2.5): indeed, although the first terms of our expansions perfectly match the results already known and supply their generalization, some regularity problems are left open in the case of middle lines whose curvature is not analytic. Moreover, in the case of open sections, there occurs a problem of matching inner and outer solutions of an edge boundary layer. These phenomena most likely arise from the fact that the particular expansion chosen is singular: the starting elliptic operator is turned into a hierarchy of ordinary differential equations in which either a *shift in singularity* or a *coordinate singularity* appears (see [11] Sec. 2.4 and following).

2. Representation of Laplace operator in system $\mathcal{C}$

One of the most interesting features of torsion is that, in general, sections cannot remain plane and undergo a deformation traditionally known as *warping*.

Such a strain is strictly linked with section's shape and resistance and so we will begin our study analyzing it. If we wish to know the warping of a section undergoing torsion, we should remember the following relation between warping and the potential function ϕ:

$$w(p) = \tau\phi(p), \quad p \in \mathcal{D}. \tag{2.1}$$

To find warping we have now a boundary value problem similar to (1.7)–(1.8):

$$\Delta w = 0 \qquad \text{in } \mathcal{D}^o, \tag{2.2}$$

$$\operatorname{grad} w \cdot \mathbf{n}_{\partial\mathcal{D}} = -\tau * \mathbf{r} \cdot \mathbf{n}_{\partial\mathcal{D}} \qquad \text{along } \partial\mathcal{D}. \tag{2.3}$$

Problem (2.2)–(2.3) is well-posed, that is to say, as

$$\oint_{\partial\mathcal{D}} \operatorname{grad} w \cdot \mathbf{n}_{\partial\mathcal{D}} = 0. \tag{2.4}$$

Condition (2.4) assures that there exists a unique solution of the system (2.2)–(2.3) and that this solution is the only one. However, it is well known that in order to determine this solution for sections of interest in technical applications, there are two possible approaches:

a) numerical methods of integration;

b) the so-called "technical" theories for cylinders with "thin" sections, i.e. "thin-walled beams" (for example [9, 10]). These theories give solutions under the hypotheses usually based on the fact that if the section is "thin", one may approximate functions involved in the problem by first order polynomials.

Let us now use curvilinear coordinates system $\mathcal{C}$, that is to say, abscissa s along the middle line Γ and ordinate z along the direction of Frenet normal $\mathbf{n}$; let us also suppose $\delta(s)$ to be constant with respect to s. A basis for $\mathcal{P}$ can then be obtained considering vectors tangent to the coordinate lines:

$$\mathbf{i}_1(s,z) := \frac{\partial p(s,z)}{\partial s} = \mathbf{r}_0'(s)\left(1 - z\kappa(s)\frac{\delta}{2}\right), \tag{2.5}$$

$$\mathbf{i}_2(s,z) := \frac{\partial p(s,z)}{\partial z} = \frac{\delta}{2} * \mathbf{r}_0'(s); \tag{2.6}$$

such a basis is orthogonal. To avoid tedious notation let us omit the dependence of the functions of interest on the coordinates, where this does not lead to confusion. If we wish to have an orthonormal basis, it is sufficient to divide (2.5)–(2.6) by their norm, and to obtain (cf. Fig. 1)

$$\mathbf{e}_1 = \frac{\mathbf{i}_1}{\|\mathbf{i}_1\|} = \mathbf{r}_0', \tag{2.7}$$

$$\mathbf{e}_2 = \frac{\mathbf{i}_2}{\|\mathbf{i}_2\|} = *\mathbf{r}_0'. \tag{2.8}$$

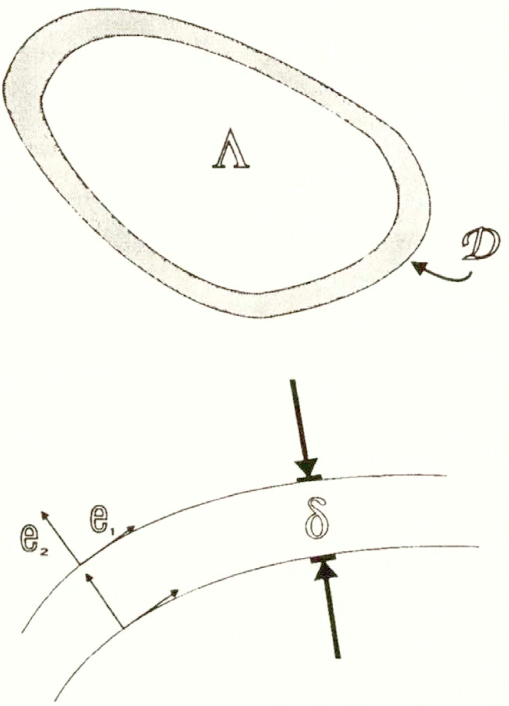


FIG. 1.

We may therefore formulate the problem (2.2)–(2.3) in the following way [12, 13, 14]:

$$\frac{1}{\|\mathbf{i}_1\|\,\|\mathbf{i}_2\|}\left[\frac{\partial}{\partial s}\left(\frac{\|\mathbf{i}_2\|}{\|\mathbf{i}_1\|}\frac{\partial w}{\partial s}\right)+\frac{\partial}{\partial z}\left(\frac{\|\mathbf{i}_1\|}{\|\mathbf{i}_2\|}\frac{\partial w}{\partial z}\right)\right]=0 \qquad \text{in } \mathcal{D}^o, \tag{2.9}$$

$$\left(\frac{1}{\|\mathbf{i}_1\|}\frac{\partial w}{\partial s}\mathbf{r}_0'+\frac{1}{\|\mathbf{i}_2\|}\frac{\partial w}{\partial z}*\mathbf{r}_0'\right)\cdot *\mathbf{r}_0'=-\tau\mathbf{r}_0\cdot\mathbf{r}_0' \qquad \text{along } \partial\mathcal{D}. \tag{2.10}$$

In our case it is easy to see that Eq. (2.9) assumes the form

$$\frac{1}{\delta\left(1-z\kappa\frac{\delta}{2}\right)}\left[\frac{\partial}{\partial s}\left(\frac{\frac{\delta}{2}}{1-z\kappa\frac{\delta}{2}}\frac{\partial w}{\partial s}\right)+\frac{\partial}{\partial z}\left(\frac{1-z\kappa\frac{\delta}{2}}{\frac{\delta}{2}}\frac{\partial w}{\partial z}\right)\right]=0; \tag{2.11}$$

we have to assume that

$$\delta<\sup_{s\in[0,\bar{l}]}\left|\frac{2}{\kappa}\right|, \tag{2.12}$$

to be sure that the Jacobian of $\mathcal{C}$ is not singular.

To solve problem (2.9)–(2.10) by means of an "asymptotic" analysis, let us suppose to have a set of domains built exactly as $\mathcal{D}$, but with different thickness, constant for each domain. It is then obvious that warping depends on δ as well as on s, z: we suppose that it is possible to write a formal series expansion for w in terms of δ:

$$w(s,z,\delta) = \sum_{n=0}^{\infty} w_n(s,z)\delta^n; \tag{2.13}$$

with (2.13), system (2.9)–(2.10) becomes now

$$\begin{aligned}\sum_{n=0}^{\infty}\Bigg\{\delta^{n+3}\left[-\frac{z\kappa}{8}w_{n,ss} - \frac{z^3\kappa^3}{8}w_{n,zz} + \frac{z\kappa'}{8}w_{n,s} - \frac{z^2\kappa^3}{8}w_{n,z}\right]\\ +\delta^{n+2}\left[\frac{1}{4}w_{n,ss} + \frac{3}{4}z^2\kappa^2 w_{n,zz} + z\frac{\kappa^2}{2}w_{n,z}\right]\\ +\delta^{n+1}\left[-\frac{3}{2}z\kappa w_{n,zz} - \frac{\kappa}{2}w_{n,z}\right] + \delta^n[w_{n,zz}]\Bigg\} = 0 \qquad \text{in } \mathcal{D}^o,\end{aligned} \tag{2.14}$$

$$\sum_{n=0}^{\infty} w_{n,z}\delta^n = -\tau\frac{\delta}{2}\mathbf{r}_0\cdot\mathbf{r}_0' \qquad z=\pm 1, \forall s; \tag{2.15}$$

here $g_{,x}$ stands for the derivative of g with respect to x. In particular, (2.15) implies

$$w_{n,z} = 0, \qquad n \neq 1, \tag{2.15}'$$

$$w_{n,z} = -\frac{\tau}{2}\mathbf{r}_0\cdot\mathbf{r}_0', \qquad n = 1. \tag{2.15}''$$

To (2.14)–(2.15) we shall also add, for closed sections, the condition of continuity

$$w(\bar{l}) = w(0), \qquad \forall z. \tag{2.16}$$

Thus we have transformed (2.2)–(2.3) into a hierarchy of systems, one for each power of δ. One could expect that, since the originary problem was well-posed, so should also be each problem of the hierarchy; unfortunately, it is not so.

This may be proved by induction: one can at once see that the zeroth step of the hierarchy gives us an equation for $w_{0,zz}$ so that we know the dependence of w_0 on z; however, to have an equation for $w_{0,ss}$, one must wait for the third step of the hierarchy. It is clear though that, in order to solve the same third step, one is supposed to know also the exact expressions for the other terms that appear in that equation; in particular that implies one should know $w_{1,z}$. This is possible because of the very simple relations one has at the first two steps; but proceeding further this implies overwhelming difficulties.

In general, at the k-th step one has a relation between $w_{k,zz}$ and the derivatives of w at previous steps. Unfortunately, boundary conditions (2.15) do not completely determine w_k, leaving in it an undetermined function of s. This implies that fundamental information for that step are at our disposal only by solving the following steps, whereas for this solution one has to know all the preceding solutions. It seems to be impossible to solve the system (2.14)–(2.15) using the proposed procedure: even though (2.2)–(2.3) was a well-posed problem, its analysis through a perturbation technique leads to a ill-posed set of successive problems.

We may give a justification of this phenomenon by observing that the expression of Laplace operator as given by (2.11) is affected by the possible singularities occurring when (2.12) is not verified: moreover, the regularity of coordinates C is C^1, while for a theorem due to KELLOGG [19] the solution belongs to the class $C^2(\overline{\mathcal{D}}) \cap C^\omega(\mathcal{D}^o)$. As a matter of fact, a singularity (see [11]) occurs in expansion (2.13) whose terms have a regularity lower than that of the exact solution. Such an effect is in our case enhanced by the circumstance that warping is given as the solution of a Dini–Neumann problem, i.e. a boundary value problem with data on normal derivatives, whose regularity is even weaker.

Our idea is then to look for a solution of the warping problem by an analogous technique, starting from basic torsion problem expressed in terms of Prandtl stress function.

3. Asymptotic expansion for Prandtl function

We choose an expansion for Prandtl function of the form

$$\psi(s, z, \delta) = \sum_{n=0}^{\infty} \psi_n(s, z)\delta^n; \tag{3.1}$$

in this way, system (1.12)–(1.14) is turned into

$$\begin{aligned} \sum_{n=0}^{\infty} \Big\{ & \delta^{n+3} \left[-\frac{z\kappa}{8}\psi_{n,ss} - \frac{z^3\kappa^3}{8}\psi_{n,zz} + \frac{z\kappa'}{8}\psi_{n,s} - \frac{z^2\kappa^3}{8}\psi_{n,z} \right] \\ & + \delta^{n+2} \left[\frac{1}{4}\psi_{n,ss} + \frac{3}{4}z^2\kappa^2\psi_{n,zz} + z\frac{\kappa^2}{2}\psi_{n,z} \right] \\ & + \delta^{n+1} \left[-\frac{3}{2}z\kappa\psi_{n,zz} - \frac{\kappa}{2}\psi_{n,z} \right] + \delta^n[\psi_{n,zz}] \Big\} \\ & = -\frac{\delta^2}{2}\left(1 - \frac{3}{2}z\kappa\delta + \frac{3}{4}z^2\kappa^2\delta^2 - \frac{1}{8}z^3\kappa^3\delta^3\right) \qquad \text{in } \mathcal{D}^o, \end{aligned} \tag{3.2}$$

$$\sum_{n=0}^{\infty} \psi_n\delta^n = \sum_{n=0}^{\infty} \overline{\psi}_n\delta^n \qquad \text{along } \partial\mathcal{D}_i, \tag{3.3}$$

$$\oint_{z=1} 2\sum_{n=0}^{\infty} \psi_{n,z}\delta^n - \oint_{z=1} \kappa \sum_{n=0}^{\infty} \psi_{n,z}\delta^{n+1} = 2\overline{A}\delta - \overline{I}\delta^2 + \frac{\pi}{4}\delta^3 \qquad \text{along } \partial\lambda. \tag{3.4}$$

If the section is simply connected (in our case, if it is *open*), its boundary is only one closed curve, and Eqs. (3.2) and (3.3) supply a hierarchy of well-posed problems, as will be proved below. If the section is not simply connected (in our case, if it is *closed*, with a hollow), we must consider, in addition to (3.2) and (3.3), Eq. (3.4) as a condition of integrability; $\overline{A}$ is the area enclosed by the middle line and $\partial\lambda$ is the boundary of the hollow, which has opposite orientation to that of the same curve regarded as part of the domain boundary.

We must observe, as it was remarked at the end of the previous section, that also asymptotic expansion (3.2) is affected by a degeneration in singularity phenomenon; however, as the boundary value problem is in a Dirichlet form, the sequence of ψ_n allows an evaluation of all quantities of interest in torsion, at least when Γ is analytic. However, although we do not prove here that this is really the case for expansion (3.2), we believe that at least its first steps reasonably approximate the *true* solutions; we limit ourselves to refer to [11] for a more detailed discussion on this subject.

We now look for the solution of system (3.2)–(3.4), step by step, both for closed and open sections; we remember that one may always choose ψ so that at the outer boundary ($z = -1$) it equals zero.

3.1. Closed sections

Solution at step δ^0

Considering in system (3.2)–(3.4) the coefficient of the zeroth power of δ, we have

$$\psi_{0,zz} = 0 \;\Rightarrow\; \psi_0(s,z) = z\psi_{0:1}(s) + \psi_{0:0}(s); \tag{3.5}$$

substituting (3.5) into (3.3) we obtain

$$\psi_0(s,-1) = 0 \;\Rightarrow\; \psi_0(s,z) = (z+1)\psi_{0:0}(s), \tag{3.6}$$

$$\psi_0(s,1) = \text{const} \;\Rightarrow\; \psi_0(s,z) = \psi(z) = (z+1)\psi_{0:0}, \tag{3.7}$$

where $\psi_{0:0}$ is a constant independent both of s and z. It is now possible to use (3.4) and (3.6):

$$2\oint_{z=1} \psi_{0,z} = 0 \;\Rightarrow\; \psi_{0:0} = 0 \;\Rightarrow\; \psi_0(z) \equiv 0. \tag{3.8}$$

Solution at step δ^1

Let us go on with our solution; keeping in mind (3.8), the second step of Eq. (3.2) gives us

$$\psi_{1,zz} = 0 \;\Rightarrow\; \psi_1(s,z) = z\psi_{1:1}(s) + \psi_{1:0}(s); \tag{3.9}$$

again applying (3.3) it results that

$$\psi_1(s,-1) = 0 \;\Rightarrow\; \psi_1(s,z) = (z+1)\psi_{1:0}(s), \tag{3.10}$$

$$\psi_1(s,1) = \text{const} \;\Rightarrow\; \psi_1(s,z) = \psi_1(z) = (z+1)\psi_{1:0}, \tag{3.11}$$

$\psi_{1:0}$ being a constant; another relation must be used, that is, Eq. (3.4); substituting in it (3.11), we have

$$2\oint_{z=1}\psi_{1,z} = 2\overline{A} \;\Rightarrow\; \psi_{1:0} = \frac{\overline{A}}{\overline{l}} \;\Rightarrow\; \psi_1(z) = \frac{\overline{r}}{2}(z+1), \tag{3.12}$$

where $\overline{r} := 2\dfrac{\overline{A}}{\overline{l}}$ is some sort of "equivalent radius" for the middle line Γ. It is worth noting that expression (3.12) is the same as that which can be found in "technical" books (for example [15, 16, 17]) under the assumption that the cylinder is "thin", that is to suppose the Prandtl function t be linear along z.

Solution at step δ^2

If we take (3.2) at its third step, Eqs. (3.8) and (3.12) lead to

$$\psi_{2,zz} - \frac{\kappa}{2}\frac{\overline{r}}{2} = -\frac{1}{2} \;\Rightarrow\; \psi_2(s,z) = \frac{z^2}{4}\left(\kappa\frac{\overline{r}}{2} - 1\right) + z\psi_{2:1}(s) + \psi_{2:0}(s). \tag{3.13}$$

Equation (3.13) and the boundary conditions (1.13) in the form of (3.3) give

$$\psi_2(s,-1) = 0 \;\Rightarrow\; \psi_2(s,z) = \frac{1}{4}(z^2-1)\left(\kappa\frac{\overline{r}}{2} - 1\right) + (z+1)\psi_{2:1}(s), \tag{3.14}$$

$$\psi_2(s,1) = \text{const} \;\Rightarrow\; \psi_2(s,z) = \frac{1}{4}(z^2-1)\left(\kappa\frac{\overline{r}}{2} - 1\right) + (z+1)\psi_{2:1}, \tag{3.15}$$

where $\psi_{2:1}$ is a constant. As usual, we now turn to apply (3.15) to (3.4):

$$2\oint_{z=1}\psi_{2,z} - \oint_{z=1}\kappa\psi_{1,z} = -\bar{l} \;\Rightarrow\; \psi_{2:1} = 0 \tag{3.16}$$

$$\Rightarrow\; \psi_2(s,z) = \frac{1}{4}(z^2-1)\left(\kappa\frac{\overline{r}}{2} - 1\right).$$

Solution at step δ^3

Equations (3.8), (3.12) and (3.16) give us

$$\psi_{3,zz} = z\frac{\kappa}{4}(\kappa\bar{r} - 1) \quad \Rightarrow \quad \psi_3(s,z) = z^3\frac{\kappa}{24}(\kappa\bar{r} - 1) + z\psi_{3:1}(s) + \psi_{3:0}(s). \tag{3.17}$$

Equation (3.17) and boundary conditions (3.3) lead us to the results

$$\psi_3(s,-1) = 0 \quad \Rightarrow \quad \psi_3(s,z) = \frac{\kappa}{24}(z^3 + 1)(\kappa\bar{r} - 1) + (z + 1)\psi_{3:1}(s), \tag{3.18}$$

$$\psi_3(s,1) = c \quad \Rightarrow \quad \psi_3(s,z) = \frac{\kappa}{24}(z^3 - z)(\kappa\bar{r} - 1) + c(z + 1), \tag{3.19}$$

where c is a constant. If we apply (3.19) to (3.4), we obtain

$$2\oint_{z=1}\psi_{3,z} - \oint_{z=1}\kappa\psi_{2,z} = \oint_{z=1}\frac{\kappa}{4} \quad \Rightarrow \quad c = \frac{1}{\bar{l}}\oint_{z=1}\frac{\kappa}{24}(\kappa\bar{r} - 1) = \frac{1}{\bar{l}}(a\bar{r} - 2\pi) \tag{3.20}$$

$$\Rightarrow \quad \psi_3(s,z) = \frac{\kappa}{24}(z^3 - z)(\kappa\bar{r} - 1) + \frac{1}{24\bar{l}}(a\bar{r} - 2\pi)(z + 1),$$

where

$$a := \oint_{z=1}\kappa(\kappa\bar{r} - 1). \tag{3.21}$$

Solution at step δ^4

It is a straightforward calculation which gives us (κ'' is the second-order derivative of κ)

$$\psi_{4,zz} = (1 - 9z^2)\frac{\kappa^2(\kappa\bar{r} - 1)}{48} + (1 - 3z^2)\frac{\kappa''\bar{r}}{32} + \frac{\kappa(a\bar{r} - 2\pi)}{48\bar{l}}; \tag{3.22}$$

and

$$\psi_4(s,z) = (1 + 2z^2 - 3z^4)\frac{\kappa^2(\kappa\bar{r} - 1)}{192} - (5 - 6z^2 + z^4)\frac{\kappa''\bar{r}}{384} + (z^2 - 1)\frac{\kappa(a\bar{r} - 2\pi)}{48\bar{l}}. \tag{3.23}$$

Without going on to further steps, we observe that every ψ_i is uniquely determined. We may prove this by induction: having proved that we are able to find a complete solution at the first four steps, it is easy to see that at each step i one has:

a. One equation, derived from (3.2), which gives us a relationship between $\psi_{i,zz}$ and an algebraic sum of derivatives of ψ_j with $j < i$; as these constitute a

homogeneous polynomial, it turns out that $\psi_i(s, z)$ is a polynomial of degree i in z and it depends on two functions of s alone;

b. Two equations, implied by (3.3), which define the form of the polynomial $\psi_i(s, z)$ and give us some information on the functions of s above mentioned; more precisely, we obtain one of these functions in terms of the other. Therefore we are left with only one function to determine, about which we have some information, of the kind: it is a constant, or equals a known expression plus a constant;

c. Such information is employed by the equation derived from (3.4) which finally lets us solve the problem and find the expression of $\psi_i(s, z)$. One interesting property of $\psi_i(s, z)$ is that it presents only odd or even powers of z, depending on whether i is odd or even.

We may say that this way of seeking a solution of the torsion problem for hollow cylinders is rather succesful (see Conclusions); let us then turn to examine the case of open sections.

3.2. Open sections

As already said, an open section is simply connected, which implies that the problem is now ruled by Eqs. (3.2) and (3.3) alone; in particular, if $\partial\mathcal{D}$ is connected, we can say that Eq. (3.3) reduces to

$$\psi = 0 \qquad \text{along } \partial\mathcal{D}. \tag{3.24}$$

Let us now look for the solution of system (3.2)–(3.24). It should be reminded that the solution which will be given here is an *outer* solution (see [11]), that is, it coincides with the exact solution only outside a thin region near the shorter edges of the section. Indeed, in closed sections an edge effect layer arises, since solutions given in this section are not capable to satisfy *all* boundary conditions. They need to be matched with an *inner* solution valid in the edge layer which satisfies boundary condition at $s = 0$, $s = \bar{l}$. It will be a further step of our research to investigate what happens in the edge layer and to provide complete matched asymptotic expansions. We remark explicitly here that the outer solution we supply coincides with those found in the literature [4, 17].

Solution at step δ^0

It is quite easy to see in this case that (3.2) implies

$$\psi_{0,zz} = 0, \tag{3.25}$$

which means that ψ_0 is affine in z; but, as (3.24) tells us that ψ_0 vanishes both at $z = -1$ and $z = 1$, we may at once conclude that

$$\psi_0 = 0. \tag{3.26}$$

Solution at step δ^1

At this step results (3.25) and (3.26) show us that we have exactly the same situation as in the previous section concerning step δ^0; we may thus conclude that

$$\psi_1 = 0. \tag{3.27}$$

Solution at step δ^2

Equations (3.2), (3.26) and (3.27) tell us that now we have

$$\psi_{2,zz} = -\frac{1}{2} \quad \Rightarrow \quad \psi_2(s,z) = -\frac{z^2}{4} + z\psi_{2:1}(s) + \psi_{2:0}(s), \tag{3.28}$$

while the restriction of Eq. (3.24) to the lines $z = \pm 1$ gives us

$$\psi_2(s,z) = \psi_2(z) = \frac{1-z^2}{4}. \tag{3.29}$$

The remaining part of the boundary condition corresponding to the lines $s = 0$ and $s = \bar{l}$ cannot be fulfilled. This is a well-known phenomenon in singular perturbation problems. We will return to this problem in the final section. We can observe that this is the solution one usually finds in ordinary handbooks on the strength of materials.

Solution at step δ^3

This time we have

$$\psi_{3,zz} = -\frac{z\kappa}{4} \quad \Rightarrow \quad \psi_3(s,z) = -\frac{z^3\kappa}{24} + z\psi_{3:1}(s) + \psi_{3:0}(s), \tag{3.30}$$

which, with the help of (3.24), becomes

$$\psi_3(s,z) = (1-z^2)\frac{z\kappa}{24}. \tag{3.31}$$

Solution at step δ^4

This step is also easy to solve; we obtain from (3.2), (3.27), (3.29), (3.31),

$$\psi_{4,zz} = \frac{\kappa^2}{48}(1-9z^2) \quad \Rightarrow \quad \psi_3(s,z) = \frac{\kappa^2}{192}(2z^2-3z^4) + \psi_{4:1}(s)z + \psi_{4:0}(s) \tag{3.32}$$

and this, with the help of (3.24), becomes

$$\psi_4(s,z) = \frac{\kappa^2}{192}(2z^2-3z^4+1). \tag{3.33}$$

We can stop now and repeat what we said about the solution of system (3.2)–(3.4) for closed sections: we have, as a matter of fact, the same structure at each step as that observed there; we can conclude that also for open sections the torsion problem in terms of Prandtl function may be solved in this way. Let us now look for the relations between warping and ψ.

4. Asymptotic expansion for warping

An equation which lets us link the warping and Prandtl functions is given directly by (1.5), (1.7), (1.13), (2.1):

$$\operatorname{grad} w = -\tau * (\operatorname{grad} \psi + \mathbf{r}) \tag{4.1}$$

where each function is a field defined over the domain. We may now explicitely express the gradients in (4.1) in terms of the coordinates s and z, use expansions (2.13) and (3.1) for w and ψ and write as well the complete expression for $\mathbf{r}$ given by (1.19). We get

$$\frac{1}{2}\sum_{n=0}^{\infty} w_{n,s}\delta^{n+1} = \tau\left[\sum_{n=0}^{\infty}\psi_{n,z}\delta^{n} - z\frac{\kappa}{2}\sum_{n=0}^{\infty}\psi_{n,z}\delta^{n+1} + \left(\frac{\delta}{2} - z\kappa\frac{\delta^2}{4}\right)\left(z\frac{\delta}{2} - \mathbf{r}_0 \times \mathbf{r}_0'\right)\right], \tag{4.2}$$

$$\sum_{n=0}^{\infty} w_{n,z}\delta^{n} - z\frac{\kappa}{2}\sum_{n=0}^{\infty} w_{n,z}\delta^{n+1} = -\tau\left[\frac{1}{2}\sum_{n=0}^{\infty}\psi_{n,s}\delta^{n+1} + \left(\frac{\delta}{2} - z\kappa\frac{\delta^2}{4}\right)\mathbf{r}_0 \cdot \mathbf{r}_0'\right]; \tag{4.3}$$

having proved in the previous section that we are able to find the complete expression for ψ_k for each k, we can as well find complete expressions for w_k at each step, what we were not able to do following the method described in Sec. 2, both for closed and open sections.

4.1. Closed sections

Solution at step δ^0

It follows directly from (4.2) and (4.3) that

$$\tau\psi_{0,z} = 0, \tag{4.4}$$

$$w_{0,z} = 0. \tag{4.5}$$

It is worth noting that (4.4) corresponds to (3.8) and (4.5) could be derived also from an attempt to solve (2.2)–(2.3) at the first step; in this case we may note that at the first step the shift in singularity phenomenon did not occur, so that also system (2.2)–(2.3) provides a correct solution.

Solution at step δ^1

Again starting from (4.2)–(4.3) and keeping in mind (3.8), (3.12) we have

$$w_{0,s} = 2\tau\frac{\overline{r}}{2} - \tau\mathbf{r}_0 \times \mathbf{r}_0', \tag{4.6}$$

$$w_{1,z} = -\frac{\tau}{2}\mathbf{r}_0 \cdot \mathbf{r}_0'. \tag{4.7}$$

Simple integration of (4.5) and (4.6) gives

$$w_0(s,z) = w_0(s) = 2\tau\left(\frac{\bar{r}}{2}s - \Omega(s)\right) + w_0(0), \tag{4.8}$$

where $\Omega(s)$ is the area enclosed by the arc comprised by the values 0 and s of the abscissa on the middle line and the two radii $\mathbf{r}(0)$ and $\mathbf{r}(s)$. Equation (4.8) is the same which is possible to find in "technical" books under assumptions of "thin" section (see the already quoted books [9] and [18]).

Solution at step δ^2

If we now substitute (3.8), (3.12), (3.16) and (4.8) into (4.2), (4.3), we have

$$w_{1,s} = -z\frac{\tau}{2}(1 + \kappa\mathbf{r}_0' \times \mathbf{r}_0), \tag{4.9}$$

$$w_{2,z} = 0. \tag{4.10}$$

By integration we find

$$w_1(s,z) = -z\frac{\tau}{2}\mathbf{r}_0 \cdot \mathbf{r}_0', \tag{4.11}$$

plus an inessential constant; this is a new result which enables us to predict the warping of a hollow section of the kind described in Sec. 2 when the wall is not so "thin"; we may see that warping is not constant and varies with z, such a variation might be called *tilting*. Equation (4.10) corresponds to the boundary conditions (2.3).

Solution at step δ^3

At this step, with the same procedure as that used in previous sections, we find

$$w_{2,s} = \frac{\tau}{12}\left[\kappa(1 - \kappa\bar{r}) + \frac{a\bar{r} - 2\pi}{\bar{l}}\right], \tag{4.12}$$

$$w_{3,z} = -\frac{\tau}{16}(z^2 - 1)\kappa'\bar{r}, \tag{4.13}$$

and integrating we obtain

$$w_2(s,z) = w_2(s) = \frac{\tau}{12}\left[\frac{a\bar{r} - 2\pi}{\bar{l}} - \int_0^s \kappa(\kappa\bar{r} - 1)\right]. \tag{4.14}$$

Solution at step δ^4

We have now

$$w_{3,s} = \frac{\tau}{48}(3z^3 - z^3)\kappa''\bar{r} \tag{4.15}$$

which implies

$$w_3(s,z) = \frac{\tau}{48}(3z^3 - z^3)\kappa'\bar{r}. \tag{4.16}$$

4.2. Open sections

In this section we will apply the same procedure as that used in Sec. 4.1, starting from Eqs. (4.2) – (4.3); we shall make use, of course, of expression for the Prandtl (outer) function for open domains, that is to say, Eqs. (3.26), (3.27), (3.29), (3.31), (4.33).

Solution at step δ^0

We have immediately

$$\tau\psi_{0,z} = 0, \tag{4.17}$$

$$w_{0,z} = 0, \tag{4.18}$$

about which we may make the same remarks as those concerning Eqs. (4.4), (4.5).

Solution at step δ^1

Going on to the next step we have

$$w_{0,s} = -\tau\mathbf{r}_0 \times \mathbf{r}_0', \tag{4.19}$$

$$w_{1,z} = -\frac{\tau}{2}\mathbf{r}_0 \cdot \mathbf{r}_0'. \tag{4.20}$$

From (4.18) and (4.19), one obtains

$$w_0(s,z) = w_0(s) = -\tau\mathbf{r}_0 \times \mathbf{r}_0', \tag{4.21}$$

which is the well known Vlasov equation of warping for "thin-walled beams" of open section [9].

Solution at step δ^2

This time we have

$$w_{1,s} = -z\frac{\tau}{2}(1 - \kappa\mathbf{r}_0 \times \mathbf{r}_0'), \tag{4.22}$$

$$w_{2,z} = 0, \tag{4.23}$$

integration of (4.20) and (4.22) gives

$$w_1(s,z) = -z\frac{\tau}{2}\mathbf{r}_0 \cdot \mathbf{r}_0', \tag{4.24}$$

plus a constant; this is a new result concerning sections which are not so "thin"; note the *tilting* already mentioned concerning (4.11).

Solution at step δ^3

If we want to follow the procedure for further steps, we have

$$w_{2,s} = \tau \frac{\kappa}{12}, \tag{4.25}$$

$$w_{3,z} = 0, \tag{4.26}$$

we may integrate (4.23) and (4.25) to obtain

$$w_2(s, z) = w_2(s) = \frac{\tau}{12} \int_0^s \kappa, \tag{4.27}$$

which is another new result.

Solution at step δ^4

Let us examine the last step:

$$w_{3,s} = 0, \tag{4.28}$$

$$w_{4,z} = -\tau \frac{\kappa'}{48}(z - z^3); \tag{4.29}$$

one integration of (4.26), (4.28) gives us

$$w_3(s, z) = \text{const}, \tag{4.30}$$

which is a new and interesting result that shows us that there is a gap between the third and fifth step in this calculation.

We remark that at all steps the cross-derivatives of w_k are equal, which is an important test of the validity of our derivations.

So we may say to have found a kinematical solution of the torsion problem for these sections. Let us now examine some interesting statical properties.

5. Tangential stress field

Let us suppose to have a formal series expansion of the tangential stress field $\mathbf{t}$:

$$\mathbf{t}(s, z, \delta) = \sum_{n=0}^{\infty} \mathbf{t}_n(s, z)\delta^n; \tag{5.1}$$

as we know each coefficient of the formal series expansion (3.1), we may use (1.11) and (1.6) to find an equation for $\mathbf{t}$. In particular, we have

$$\sum_{n=0}^{\infty} \mathbf{t}_n \delta^{n+1} - \frac{z\kappa}{2} \sum_{n=0}^{\infty} \mathbf{t}_n \delta^{n+2} = 2G\tau \left[\left(\sum_{n=0}^{\infty} \psi_{n,z} \delta^n - \frac{z\kappa}{2} \sum_{n=0}^{\infty} \psi_{n,z} \delta^{n+1} \right) \mathbf{r}_0' - \frac{1}{2} \sum_{n=0}^{\infty} \psi_{n,s} \delta^{n+1} (*\mathbf{r}_0') \right]; \tag{5.2}$$

so, operating as in the previous sections, we can solve system (6.2)–(6.3) step by step both for the closed and open sections.

5.1. Closed sections

Solution at step δ^0

We immediately have, on the basis of (5.2),

$$\psi_{0,z} = 0, \tag{5.3}$$

conciding with a result already obtained, cf. Eq. (3.8).

Solution at step δ^1

We have now

$$\mathbf{t}_0 = G\tau\bar{\tau}\mathbf{r}_0' = \frac{2G\tau\overline{A}}{\bar{l}}. \tag{5.4}$$

Equation (5.4) can be found in technical literature: due to the assumption of "thin" section, the state of stress is plane, directed along the middle line, and besides one may neglect terms of higher than the second order in a power series expansion of Prandtl function, and obtain a constant distribution of tangential stress.

Solution at step δ^2

Simply applying (5.2) and (5.4), we get

$$\mathbf{t}_1 = G\tau z\left(\kappa\frac{\overline{r}}{2} - 1\right)\mathbf{r}_0'; \tag{5.5}$$

equation (5.5) is a new result that, as already said in connection with (4.11), takes into account the fact that the section is not so "thin".

Solution at step δ^3

If we go on with our simple procedure, it is quite easy to prove that

$$\mathbf{t}_2 = G\tau\left\{\left[\frac{1}{4}(\kappa\overline{r} - 1)\left(z^2\kappa - \frac{1}{3}\right) + \frac{1}{12}(a\overline{r} - 2\pi)\right]\mathbf{r}_0' + \left[\frac{\kappa'\overline{r}}{16}(1 - z^2)\right](*\mathbf{r}_0')\right\}, \tag{5.6}$$

which shows that this is the first step in which a component of shear stress along $\mathbf{m}$ appears.

Solution at step δ^4

Our last result for this case is

$$\mathbf{t}_3 = z\kappa\mathbf{t}_2 + G\tau + \frac{z\kappa'}{24}(2\kappa\overline{r} - 1)(1 - z^2)(*\mathbf{r}_0'), \tag{5.7}$$

which is another new result. Let us examine what happens in the case of an open section, keeping in mind that we have at our disposal only a partial knowledge of the complete solution.

5.2. Open sections

Solution at step δ^0

We immediately have from (5.2)

$$\psi_{0,z} = 0 \tag{5.8}$$

about which we may repeat the remark made before, concerning (5.3).

Solution at step δ^1

We have now the possibility to write tangential stress vector field at the first step:

$$\mathbf{t}_0(s,z) = 0, \tag{5.9}$$

this is a result already known: stress distribution in an open section cannot be determined before arriving at the third step of Prandtl function.

Solution at step δ^2

We get now from (5.2), (5.8), (5.9),

$$\mathbf{t}_1(s,z) = -G\tau z\mathbf{r}_0'. \tag{5.10}$$

Solution at step δ^3

We have at this step, remembering the results of previous steps,

$$\mathbf{t}_2(s,z) = G\tau\frac{\kappa}{12}(1-3z^2)\mathbf{r}_0', \tag{5.11}$$

which is the first of our new results for open sections.

Solution at step δ^4

Equations (5.2), (5.8), (5.9), (5.10) and (5.11) yield

$$\mathbf{t}_3(s,z) = G\tau\left[z\frac{\kappa^2}{8}\left(\frac{1}{3}-z^2\right)\mathbf{r}_0' + z\frac{\kappa'}{24}(z^2-1)*\mathbf{r}_0'\right], \tag{5.12}$$

which is another new result. We see that, starting from this step, the effect of tangential stress along the normal to middle line becomes appreciable.

6. Higher order torsional stiffnesses

Usually the torsional stiffness is defined as the ratio of torque divided by unit torsion angle, that is to say,

$$K = \frac{T}{\tau}; \tag{6.1}$$

since it is possible to express tangential stresses in two ways, either by means of a potential function or through Prandtl function; it is also possible to derive two different equations for torque, each starting from a different point of view. That implies we have also two different formulae for torsional stiffness:

$$K = G(J - D), \tag{6.2}$$

$$K = 2G\left(\int_{\mathcal{D}} \psi + A_\lambda \overline{\psi}_\lambda\right); \tag{6.3}$$

in (6.2) J is the section's polar moment of inertia and D is the so-called Dirichlet integral, which is given by

$$D = \int_{\mathcal{D}} \operatorname{grad} \phi \times \mathbf{r}. \tag{6.4}$$

In (6.3) A_λ is the area of the hole. To simplify the calculation, we use (6.3). Let us suppose for K a formal power series expansion similar to those already used:

$$K(\delta) = \sum_{n=0}^{\infty} K_n \delta^n; \tag{6.5}$$

we may then substitute (6.5) into (6.3), as well as the formal series expansion (3.1) for Prandtl function. We then have the following results.

6.1. Closed sections

We may see at once that

$$K_0 = 0 \tag{6.6}$$

which means that at the zero order, that is to say, when thickness tends to zero, the considered section exhibits no resistance to torque, which is physically reasonable. Going on with our calculation we have

$$K_1 = 2G\left(\int_{\mathcal{D}} \psi_0 + \overline{A}\frac{2\overline{A}}{\overline{l}}\right) = \frac{4G\overline{A}^2}{\overline{l}}, \tag{6.7}$$

which coincides with the well-known Bredt formula that one may found in every book on the strength of materials. At next step we find

$$K_2 = 2G\left(\frac{1}{2}\int_D \psi_1 - \overline{A}\right) = 0 \tag{6.8}$$

which is a new result and gives us a mathematical reason for the validity of Bredt formula: since there are no immediate corrections to the torsional stiffness of a Bredt-like section, its approximate value has proved to be good enough for some applications. If we go on, we find

$$K_3 = G\left[\frac{\bar{l} - \pi\bar{r}}{3} + \frac{a\bar{r}}{12}\right], \tag{6.9}$$

which gives us a third-order correction for Bredt formula. At the next step we find

$$K_4 = 0. \tag{6.10}$$

6.2. Open sections

It is absolutely trivial to show that

$$K_0 = K_1 = K_2 = 0; \tag{6.11}$$

it is also easy to see that

$$K_3 = \frac{G\bar{l}}{3} \tag{6.12}$$

which is the usually known result for "thin" sections; going on we have

$$K_4 = 0 \tag{6.13}$$

which is a new result, interesting because it enables us to see that for open sections there are corrections for K only at higher orders of δ. However, we have to remark that we have not developed a matching of Prandtl function close to the ends of the section to account for the edge effects. This seems not to affect (6.12), but it could change the result given by (6.13). Further investigations will clear up this point.

7. Conclusion and open problems

The only exact solution available in literature with which we can compare our results is that of a circular ring. Our asymptotic expansions supply exact solutions

after three steps. All sections belonging to the class $\mathcal{B}$ we have studied share the property of having a vanishing second order torsional stiffness. We remark that the only other solution in analytical form for our problem which may be found in literature is that of an elliptic shaft [4] and that, unfortunately, this section does not belong to $\mathcal{B}$: indeed, its second order torsional stiffness does not vanish. Further investigations will be developed in the following directions:

a. To find an expansion valid for Bredt-like sections with variable thickness;

b. To develop a matched asymptotic procedure to handle the edge layer effect appearing in open sections;

c. To cure the loss of regularity arising in present expansions using the methods proposed in [11];

d. To determine mathematically the range of validity of the proposed expansions using the potential theory, as first done by KELLOG [19];

e. To estimate rigorously the errors associated with the Bredt results in some cases interesting from the point of view of applications.

In our opinion, the expansions proposed in this paper show a new example of a well-known phenomenon occurring in asymptotic analysis: first terms supply an accurate approximation of true solutions, while starting from a given step, a loss of regularity may be observed.

Acknowledgement

The authors thank Professor ANTONIO DI CARLO for having pointed out the existence of problems dealt in this paper.

References

1. A. CLEBSCH, *Théorie de l'élasticité des corps solides (Traduite par MM. Barré de Saint-Venant et Flamant, avec des Notes étendues de M. Barré de Saint-Venant)*, Dunod, Paris 1883; Reprinted by Johnson Reprint Corporation, New York 1966.
2. I.S. SOKOLNIKOFF, *Mathematical theory of elasticity*, McGraw-Hill, New York 1946.
3. A.E.H. LOVE, *A treatise on the mathematical theory of elasticity*, Dover, New York 1949.
4. S. TIMOSHENKO, J.N. GOODIER, *Theory of elasticity*, McGraw-Hill, New York 1951.
5. V.V. NOVOZHILOV, *Theory of elasticity* [in Russian], Sudpromgiz, Leningrad 1958; English translation: *Theory of elasticity*, Pergamon, London 1961.
6. Y.C. FUNG, *Foundations of solid mechanics*, Prentice-Hall, Englewood Cliffs, 1965.
7. B.M. FRAEIJS DE VEUBEKE, *A course in elasticity*, Springer-Verlag, New York 1979.
8. L. PRANDTL, *Zur Torsion von prismatischen Stäben*, Phys. Zeits., **4**, 758, 1903.
9. V.Z. VLASOV, *Thin-walled elastic rods* [in Russian], Fizmatgiz, Moskwa 1959; English translation: *Thin-walled elastic beams*, Israel Program for Scientific Translations, Jerusalem 1961.
10. R. BREDT, *Kritische Bemerkungen zur Drehungselastizität*, Zeits. Ver. Deutsch. Ing., **40**, 815, 1896.
11. A. NAYFEH, *Perturbation methods*, John Wiley and Sons, New York 1973.
12. V.I. ARNOLD, *Mathematical methods of classical mechanics* [in Russian], MIR, Moskva 1979; Italian translation: *Metodi matematici della meccanica classica*, Editori Riuniti, Roma 1992.
13. A.N. TICHONOV, A.A. SAMARSKIJ, *Equations of mathematical physics*, MIR, Moskva 1977; Italian translation: *Equazioni della fisica matematica*, Editori Riuniti, Roma 1981.

14. C.E. WEATHERBURN, *An introduction to Riemannian geometry and the tensor calculus*, Cambridge University Press, 1963.
15. V. FEODOSYEV, *Strength of materials*, MIR, Moskva 1968; Italian translation: *Resistenza dei materiali*, Editori Riuniti, Roma 1977.
16. J. CHASE, A.H. CHILVER, *Strength of materials and structures*, Edward Arnold, London 1971.
17. R. BALDACCI, *Scienza delle costruzioni*, UTET, Torino 1970.
18. A. GJELSVIK, *The theory of thin-walled bars*, John Wiley and Sons, New York 1981.
19. O.D. KELLOGG, *Potential theory*, Dover, New York 1929.

DIPARTIMENTO DI INGEGNERIA STRUTTURALE E GEOTECNICA,
UNIVERSITÀ DI ROMA "LA SAPIENZA", ROMA, ITALIA.

Received July 8, 1994; *new version September* 16, 1994.