Lorentz group on Minkowski spacetime for construction of the two basic principles of plasticity

Hong-Ki Hong*,[a], Chein-Shan Liu[b]

[a]Department of Civil Engineering, National Taiwan University, Taipei 106-17, Taiwan
[b]Department of Mechanical and Marine Engineering, National Taiwan Ocean University, Keelung 202-24, Taiwan

Abstract

We show that a model of plasticity is a necessary consequence of the two basic principles: (1) *causality in the truncated future cone of the Minkowski spacetime* (*or its generalization*) *of augmented states*, and (2) *controllability and non-generativity in a reachable, bounded space of states*. To consider the symmetry switching between $PSO_o(n, 1)$ and $SE(n)$ due to switching on and off the plasticity mechanism, the model is reconstructed as a dynamical system on a composite space, which results from a surgery on Minkowski spacetime. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords: Plasticity; Causality; Controllability; Minkowski spacetime; Lorentz group; Two basic principles

1. Introduction

Plasticity theory of solid materials and structural members under a complicated mechanical environment is a very important topic for engineering science and practice. However, there remain a few fundamental problems yet to be resolved. We name a few among them. What ideas underlie the behavior of plasticity – plastically deformed solids and structures? What attributes should a model of plasticity possess? Why do the present-day models of the (conventional) theory of plasticity gather forms of expressions as diverse as yield conditions, plastic-flow rules, loading–unloading criteria, elastic–plastic decompositions, laws of kinematic hardening, etc.? Indeed, the ingredients of constitutive relations are separately specified and seem to be loosely structured. Curious scholars cannot help asking if there is a backbone idea by which the ingredients become organized. After more than a century of experiments and studies, may we take a step forward to answer these questions? The present paper is supposed to be one of the attempts to face the difficulties encountered above. It investigates internal symmetry, reveals hidden spacetime structure, and derives two basic principles of a plasticity model, which we believe underlies the principal behavior of plasticity in general.

Concerning the questions posed, we call attention to some occasions in the development of the theory of plasticity, which had paved the way for the research of the present paper. First, the flow rule which relates (principal) stresses to (principal) strain increments was suggested by Saint Venant (1870) and by Lévy (1871). Next, the proportionality factor between the (deviatoric) stress components Q^i and the (plastic) strain increment

* Corresponding author. Tel.: + 886-2-2366-1931; fax: + 886-2-2362-2975.

E-mail address: hkhong@ce.ntu.edu.tw (H.-K. Hong).

components dq_i had started becoming more and more appreciated. To illustrate the conceptual progress in this regard, let us write them in progressive order:[1]

$$(1) \quad \frac{dq_1}{Q^1} = \frac{dq_2}{Q^2} = \cdots = \frac{dq_n}{Q^n},$$

$$(2) \quad \frac{dq_1}{Q^1} = \frac{dq_2}{Q^2} = \cdots = \frac{dq_n}{Q^n} = \lambda,$$

$$(3)_a \quad \frac{dq_1}{Q^1} = \frac{dq_2}{Q^2} = \cdots = \frac{dq_n}{Q^n} = d\lambda,$$

$$(3)_b \quad dq_i = Q^i \, d\lambda, \quad i = 1, 2, \ldots, n.$$

Notice that there is neither λ nor $d\lambda$ in (1). While λ in (2) is merely a proportionality factor, the one in $(3)_{a,b}$ appears to signify that an entity with the symbol λ whose increment is $d\lambda$ has been born.

Fourth, the exponential of a dimensionless λ (non-dimensionalized with respect to a characteristic compliance),

$$(4) \quad X^0 := \exp \lambda,$$

has been defined with its evolution explored seriously (see, e.g., [1,2]). Fifth, X^0 has been recognized as a new dimension in the augmented stress space of $(X^1, X^2, \ldots, X^n, X^0)$ defined as

$$(5)_a \quad \mathbf{X} = \begin{bmatrix} X^1 \\ X^2 \\ \vdots \\ X^n \\ X^0 \end{bmatrix} := \frac{X^0}{Q^0} \begin{bmatrix} Q^1 \\ Q^2 \\ \vdots \\ Q^n \\ Q^0 \end{bmatrix}$$

such that the non-linear differential equations of perfect elastoplasticity turn out to be linear in terms of the augmented stresses [3–5]:

$$(5)_b \quad \dot{\mathbf{X}} = \mathbf{A}\mathbf{X}.$$

The augmentation of the states in n dimensions to $(n + 1)$ dimensions is far more than a mere mathematical artifice, for the matrix $\mathbf{A}$ in the resulting Eq. $(5)_b$ possesses such rich properties as to deeply

---

[1] In this introductory section the three different but closely related λ's in (2), $(3)_{a,b}$, (4) are denoted by one and the same symbol for simplicity.

interest us; on this occasion the iceberg of the spacetime structure of plasticity is starting to emerge. We all know that, rather than looking at the tip of the iceberg, it is when one has tried looking beneath the surface of things that one expects to find the deep nature of things as will soon be seen.

In this treatment X^0 and $(X^1, X^2, \ldots, X^n)$ are no more separable and independent as the usual $(n + 1)$-dimensional vector in Euclidean space $\mathbb{E}^{n+1}$; they are now being organized to an integrated object in Minkowski spacetime $\mathbb{M}^{n+1}$, which is not a simple extension of ordinary Euclidean n-space to $(n + 1)$ dimensions, with X^0 being just one more dimension. Because the corresponding entries in the metric have different signs, -1 versus positive definiteness, the "time" component X^0 is not on the same footing as the n "space" components $(X^1, X^2, \ldots, X^n)$, and the structure built on the spacetime consequently has group properties quite unlike that on Euclidean space.

Although models of plasticity are highly nonlinear, the perfect elastoplastic model has been transformed into a two-phase linear system with an on–off switch. It was made clear that it is a system of linear algebraic equations in the off (i.e. elastic) phase, and in the on phase turns out to be a system of linear differential equations with variable coefficients in the augmented space. Owing to the implicit linearity, a Lorentz group of transformations for the augmented states was discovered and the Minkowski spacetime structure was examined in [5]. In this paper we reverse the line of study, *starting from scratch except for a few general requirements*, then working through mathematical construction, and ending up with an abstract dynamic system, which turns out to be the model of elastoplasticity. The highlight is the discovery of two principles of plasticity.

2. Causality implies the Lorentz group

What we attempt to construct is a causal system in an appropriate spacetime, on which the "time" component X^0 flows forward and causes precede effects. Zeeman [6] was able to show that the causality assumption in Minkowski spacetime

renders a composition of a translation, a dilation, and a proper orthochronous Lorentz transformation. To begin with we will focus on the most essential among the three – the proper orthochronous Lorentz group $SO_o(n, 1)$.

The element $\mathbf{G}$ of the proper orthochronous Lorentz group $SO_o(n, 1)$ satisfies the following properties [7–9]:

$$\mathbf{G}^t\mathbf{g}\mathbf{G} = \mathbf{g}, \tag{1}$$

$$\det \mathbf{G} = 1, \tag{2}$$

$$G_0^0 > 0, \tag{3}$$

where $\mathbf{g}$ is a Minkowskian metric,

$$\mathbf{g} = \begin{bmatrix} \mathbf{g}_{ss} & \mathbf{g}_{s0} \\ \mathbf{g}_{0s} & g_{00} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_n & \mathbf{0}_{n \times 1} \\ \mathbf{0}_{1 \times n} & -1 \end{bmatrix}, \tag{4}$$

and $\mathbf{I}_n$ is the identity tensor of order n.

The Lorentz group $SO_o(n, 1)$ gives uniquely the real Lie algebra $so(n, 1)$ (to within isomorphism). Consider a one-parameter subgroup $\mathbf{G}(t)$ of the Lorentz group $SO_o(n, 1)$, which is a curve passing through the group identity at $t = 0$,

$$\mathbf{G}(0) = \mathbf{I}_{n+1}, \tag{5}$$

and which left acts on the Minkowski spacetime, resulting in a congruence of curves

$$\mathbf{X}(t) = \mathbf{G}(t)\mathbf{X}(0) = [\mathbf{G}(t)\mathbf{G}^{-1}(t_1)]\mathbf{X}(t_1) \tag{6}$$

in the Minkowski spacetime. Owing to the closure property of group, $[\mathbf{G}(t)\mathbf{G}^{-1}(t_1)]$ also belongs to $SO_o(n, 1)$ but in general not to the aforementioned one-parameter subgroup. When t_1 is put very close to t, then $[\mathbf{G}(t)\mathbf{G}^{-1}(t_1)]$ is very close to the identity of $SO_o(n, 1)$; moreover, in view of Eqs. (2) and (3), the group manifold is analytic; hence,

$$\mathbf{A}(t) := \frac{\partial}{\partial t}[\mathbf{G}(t)\mathbf{G}^{-1}(t_1)]\Big|_{t_1 = t}$$
$$= \dot{\mathbf{G}}(t)\mathbf{G}^{-1}(t) \tag{7}$$

defines a string of tangent vectors on the tangent space at the group identity of the group manifold, more precisely, a continuously singly parametrized series of one-dimensional subalgebra of the real Lie algebra $so(n, 1)$ of the Lorentz group $SO_o(n, 1)$. Differentiating Eqs. (1) and (6)$_2$ and setting $t_1 = t$

yields

$$\mathbf{A}^t\mathbf{g} + \mathbf{g}\mathbf{A} = \mathbf{0}, \tag{8}$$

$$\dot{\mathbf{X}}(t) = \mathbf{A}(t)\mathbf{X}(t). \tag{9}$$

The flow generated by such $so(n, 1)$ is the congruence of curves resulting from solving the dynamical system (9). Due to Eqs. (7)$_2$ and (5), $\mathbf{G}(t)$ is the fundamental solution of the system of ordinary differential equations (9). From Eq. (8), $\mathbf{g}\mathbf{A}$ is skew-symmetric; and, therefore, let

$$\mathbf{A} =: \begin{bmatrix} \mathbf{\Omega} & \mathbf{u} \\ \mathbf{u}^t & 0 \end{bmatrix}, \tag{10}$$

where $\mathbf{u} = (u_1, u_2, \ldots, u_n)^t := (A_0^1, A_0^2, \ldots, A_0^n)^t$ and $\mathbf{\Omega}$ is skew-symmetric, i.e., $\mathbf{\Omega}^t = -\mathbf{\Omega}$, in which the entries $\Omega_{ij} := A_j^i$ for $1 \leqslant i < j \leqslant n$.

Using the partition of $\mathbf{G}$

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_s^s & \mathbf{G}_0^s \\ \mathbf{G}_s^0 & G_0^0 \end{bmatrix}, \tag{11}$$

and the transformation formula given in Eq. (6)$_1$ we have

$$X^0(t) = \mathbf{G}_s^0(t)\mathbf{X}^s(0) + G_0^0(t)X^0(0), \tag{12}$$

where $\mathbf{X}^s = (X^1, X^2, \ldots, X^n)^t$. Upon assuming $X^0(0) \neq 0$ the above equation can be written as

$$G_0^0(t) = \frac{X^0(t)}{X^0(0)} - \frac{\mathbf{G}_s^0(t)\mathbf{X}^s(0)}{X^0(0)}. \tag{13}$$

On the other hand, from Eq. (1) and the partition (11) of $\mathbf{G}$ it follows that

$$(\mathbf{G}_0^s)^t(t)\mathbf{G}_0^s(t) - (G_0^0)^2(t) = -1, \tag{14}$$

$$(\mathbf{G}_s^s)^t(t)\mathbf{G}_0^s(t) = G_0^0(t)(\mathbf{G}_s^0)^t(t), \tag{15}$$

$$(\mathbf{G}_s^s)^t(t)\mathbf{G}_s^s(t) - (\mathbf{G}_s^0)^t(t)\mathbf{G}_s^0(t) = \mathbf{I}_n. \tag{16}$$

Substituting Eq. (15) into Eq. (13) for $\mathbf{G}_s^0(t)$ we obtain

$$(G_0^0)^2(t) - \frac{X^0(t)}{X^0(0)}G_0^0(t) + \frac{(\mathbf{G}_0^s)^t(t)\mathbf{G}_s^s(t)\mathbf{X}^s(0)}{X^0(0)} = 0. \tag{17}$$

Using Eqs. (15) and (13) the discriminant of the above quadratic equation for $G_0^0(t)$ is found to be

$$\left[\frac{X^0(t)}{X^0(0)}\right]^2 - \frac{4(\mathbf{G}_0^s)^t(t)\mathbf{G}_s^s(t)\mathbf{X}^s(0)}{X^0(0)}$$

$$= \left[\frac{X^0(t)}{X^0(0)} - \frac{2\mathbf{G}_s^0(t)\mathbf{X}^s(0)}{X^0(0)}\right]^2 \geqslant 0. \qquad (18)$$

Hence the existence of real solutions of Eq. (17) is guaranteed. Now using Eq. (17) we readily deduce that

$$G_0^0(t) > 0 \Leftrightarrow \frac{X^0(t)}{X^0(0)} > 0. \qquad (19)$$

It states that $G_0^0(t) > 0$ is a necessary and sufficient condition for the preservation of the sign of X^0; hence, inequality (3) asserts the preservation of time-orientation.

3. Projection of flow – a primitive model

In view of Eq. (19) we choose to assume $X^0 > 0$ and to define the state n-vector

$$\mathbf{x} = (x_1, x_2, \ldots, x_n)^t := (X^1, X^2, \ldots, X^n)^t/X^0. \qquad (20)$$

Indeed it is a projection. Utilizing this and Eq. (10), we can partition Eq. (9) into

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} X^0\mathbf{x} \\ X^0 \end{bmatrix} = \begin{bmatrix} \mathbf{\Omega} & \mathbf{u} \\ \mathbf{u}^t & 0 \end{bmatrix} \begin{bmatrix} X^0\mathbf{x} \\ X^0 \end{bmatrix}. \qquad (21)$$

Note that it is essential to choose $X^0 > 0$ rather than $X^0 < 0$ for the differential equation of X^0 in the above to be valid. Consequently, the augmented state equation (9) becomes the state equation

$$\dot{\mathbf{x}} = \mathbf{\Omega}\mathbf{x} - (\mathbf{x}^t\mathbf{u})\mathbf{x} + \mathbf{u} \qquad (22)$$

after eliminating X^0. We can identify three vector fields in the nonlinear differential equation: one spinning $\mathbf{\Omega}\mathbf{x}$, one dissipative $(\mathbf{x}^t\mathbf{u})\mathbf{x}$, and one direct $\mathbf{u}$.

For later references we may call the dynamical system derived so far in this section – mainly represented by the nonlinear state equation (22) – as the primitive model. To summarize, the principle of causality in Minkowski spacetime $\mathbb{M}^{n+1}$ implies the Lorentz group $SO_o(n, 1) \ni \mathbf{G}$, which gives uniquely the real Lie algebra $so(n, 1) \ni \mathbf{A}$, which in

turn gives the flow $\mathbf{X}(t)$. Finally, projecting the flow equation $\dot{\mathbf{X}} = \mathbf{A}\mathbf{X}$ yields the primitive model. Note that the primitive model is not the perfect elastoplastic model, but the perfect elastoplastic model [5,7] in the on phase behaves like the primitive model.

4. Controllable non-generative system in reachable bounded state space

The primitive model represents an abstract dynamical system naturally induced by the Lorentz group $SO_o(n, 1)$. Since what we need in addition to the augmented state causality in Minkowski spacetime is, concisely speaking, a controllable, non-generative system with its state space reachable and bounded, we must accordingly place restrictions on its *phase space*. A more effective way to this end is to work on the two-dimensional phase plane $(-\mathrm{div}\,\dot{\mathbf{x}}, \|\mathbf{x}\|)$ rather than directly in the vast $2n$-dimensional phase space $(\mathbf{x}, \dot{\mathbf{x}})$. Here div stands for the divergence with respect to $\mathbf{x}$.

It is known that a continuous dynamical (flow) system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$ is convergent (or dissipative) if $-\mathrm{div}\,\dot{\mathbf{x}} > 0$, divergence-free (or conservative) if $-\mathrm{div}\,\dot{\mathbf{x}} = 0$, and divergent (or generative) if $-\mathrm{div}\,\dot{\mathbf{x}} < 0$, where $\mathbf{f} = (f_1, f_2, \ldots, f_n)^t$ and $-\mathrm{div}\,\dot{\mathbf{x}} = -\sum_{i=1}^n \partial f_i/\partial x_i$. In fact, the scalar quantity $\mathrm{div}\,\dot{\mathbf{x}}$ measures the rate of change of the phase volume. Now in view of our need, restrict the abstract dynamical system (22) to be

(a)$_1$ non-generative, namely $-\mathrm{div}(\dot{\mathbf{x}}) \geqslant 0$, and
(b)$_1$ bounded in the $\mathbf{x}$-state space, say $\|\mathbf{x}\| \leqslant 1$.

The permutation of $-\mathrm{div}\,\dot{\mathbf{x}} = 0, > 0$ and $\|\mathbf{x}\| < 1, = 1$ in the phase plane $(-\mathrm{div}\,\dot{\mathbf{x}}, \|\mathbf{x}\|)$ thus suggests $2^{2 \times 2} = 16$ systems; each system is modeled by the governing equation (22) permitted to work only in an assigned region on the phase plane indicated by the shaded area in one of the chessboards of Fig. 1. But the 16 systems are not all workable, not all meaningful; we need to impose more precise restrictions:

(a)$_2$ The system must be capable of being dissipative; in other words, it should not always be conservative.
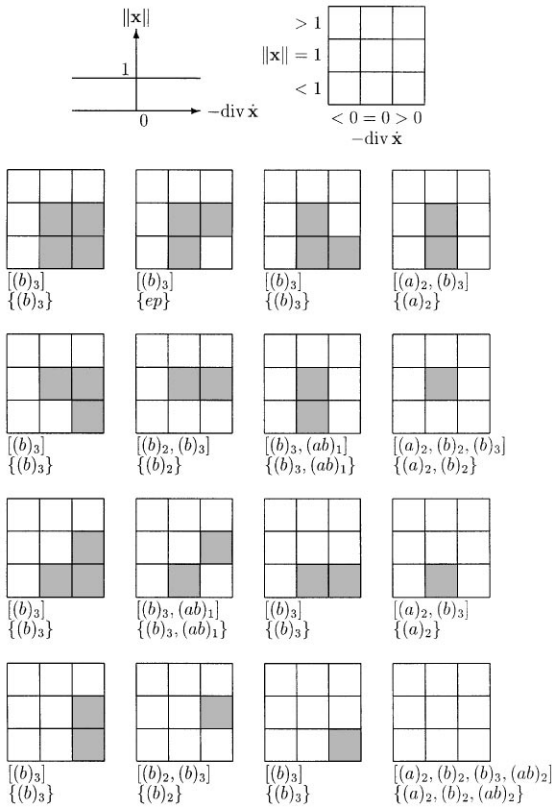
Fig. 1. Phase plane chessboards for $2^{2 \times 2}$ abstract dynamic systems. For example, $[(b)_3]$ means that the system with the governing equation (22) would violate the requirement $(b)_3$ if it were allowed to move in the shaded area on the phase plane $(-\operatorname{div} \dot{\mathbf{x}}, \|\mathbf{x}\|)$, while $\{(b)_3\}$ means that the system with the modified governing equation (22)' would violate the requirement $(b)_3$ if it were allowed to move in the shaded area. The particular system marked with $\{ep\}$, which is governed by Eq. (22)' and allowed to move in the shaded area as indicated, meets all the requirements; indeed, it is nothing but the perfect elastoplastic model.

$(b)_2$ The system can be set free sometimes; in other words, it must contain the free state $\mathbf{x} = \mathbf{0}$ in its state space.

$(b)_3$ The system must be controllable; in other words, any state $\mathbf{x}$ in the admissible closed ball $\|\mathbf{x}\| \leqslant 1$ as required by $(b)_1$ can always be reached under a certain path of controls $\mathbf{u}$ and $\boldsymbol{\Omega}$.

$(ab)_1$ The system should be connected in the phase plane $(-\operatorname{div} \dot{\mathbf{x}}, \|\mathbf{x}\|)$.

$(ab)_2$ The system is not empty.

As indicated within the square brackets $[\cdots]$ in Fig. 1, among the sixteen systems, there are four, four, sixteen, two, one system(s) violating the restrictions $(a)_2$, $(b)_2$, $(b)_3$, $(ab)_1$, $(ab)_2$, respectively. Consequently, none of the 16 systems passes the restrictions. Among these restrictions, $(b)_3$ may need explanation as follows: take the divergence of Eq. (22) with respect to $\mathbf{x}$, giving

$$- \operatorname{div} \dot{\mathbf{x}} = (n + 1)\mathbf{x}^t\mathbf{u}. \qquad (23)$$

Take the inner product of Eq. (22) with $\mathbf{x}$, yielding

$$\frac{\mathrm{d}\|\mathbf{x}\|^2}{\mathrm{d}t} = 2\mathbf{x}^t\mathbf{u}(1 - \|\mathbf{x}\|^2). \qquad (24)$$

By Eqs. (23) and (24) and the restrictions $(a)_1$ and $(b)_1$ we have $\mathrm{d}\|\mathbf{x}\|^2/\mathrm{d}t \geqslant 0$, and hence $\mathrm{d}\|\mathbf{x}\|/\mathrm{d}t \geqslant 0$. In case the initial condition is $\|\mathbf{x}(t_i)\| = 1$, it remains $\|\mathbf{x}(t)\| = 1$ for all $t > t_i$; while in case the initial condition is $\|\mathbf{x}(t_i)\| < 1$, it is always $\mathrm{d}\|\mathbf{x}\|/\mathrm{d}t \geqslant 0$. These indicate an undesirable tendency that the $\mathbf{x}$-path gets more and more deviated from the center $\mathbf{x} = \mathbf{0}$ of the $\mathbf{x}$-state space, and that whenever it is on the hypersphere $\mathbb{S}^{n-1}$, i.e., $\|\mathbf{x}\| = 1$, it will remain there forever. More precisely, the response $\mathbf{x}$-paths never come back inwards. In this sense, the system governed by Eq. (22) is not controllable. It is evident that systems with defective controllability should be excluded.

To remedy the miserable situation that no systems survive the restrictions, one can either relax the restrictions or modify the governing equation (22). Since the aforementioned requirements are reasonable regarding our need, we choose to modify Eq. (22) a little as below:

$$\dot{\mathbf{x}} = \boldsymbol{\Omega}\mathbf{x} - (\mathbf{x}^t\mathbf{u})\mathbf{x} + \mathbf{u} \quad \text{if} -\operatorname{div} \dot{\mathbf{x}} > 0, \qquad (22'_1)$$

$$\dot{\mathbf{x}} = \boldsymbol{\Omega}\mathbf{x} + \mathbf{u} \quad \text{if} -\operatorname{div} \dot{\mathbf{x}} = 0. \qquad (22'_2)$$

Take the divergence of Eq. (22)' with respect to $\mathbf{x}$, giving

$$- \operatorname{div} \dot{\mathbf{x}} = (n + 1)\mathbf{x}^t\mathbf{u} \quad \text{if} -\operatorname{div} \dot{\mathbf{x}} > 0, \qquad (23'_1)$$

$$- \operatorname{div} \dot{\mathbf{x}} = 0 \quad \text{if} -\operatorname{div} \dot{\mathbf{x}} = 0. \qquad (23'_2)$$

Take the inner product of Eq. (22′) with $\mathbf{x}$, yielding

$$\frac{\mathrm{d}\|\mathbf{x}\|^2}{\mathrm{d}t} = 2\mathbf{x}^t\mathbf{u}(1 - \|\mathbf{x}\|^2) \quad \text{if } -\mathrm{div}\,\dot{\mathbf{x}} > 0, \qquad (24'_1)$$

$$\frac{\mathrm{d}\|\mathbf{x}\|^2}{\mathrm{d}t} = 2\mathbf{x}^t\mathbf{u} \quad \text{if } -\mathrm{div}\,\dot{\mathbf{x}} = 0. \qquad (24'_2)$$

Equation (23′$_2$) may appear trivial, but its harmony reflects that the modification in Eq. (22′$_2$) is essential for the system to be truly conservative in the divergence-free phase, since by this subtle modification in the said phase we have disconnected the unreasonable link between $\mathbf{x}^t\mathbf{u}$ and $-\mathrm{div}\,\dot{\mathbf{x}}$ such that a vanishing divergence of the system no longer absurdly relies on the perpendicularity of the response $\mathbf{x}$ and the control $\mathbf{u}$ and thereby the control $\mathbf{u}$ is free so that the system in the divergence-free phase is controllable as can be seen by considering Eq. (24′$_2$).

Again apply the restrictions to the new systems governed by Eq. (22′). First, among the $2^{3\times3} = 512$ systems obtainable from the permutation of $-\mathrm{div}\,\dot{\mathbf{x}} < 0, = 0, > 0$ and $\|\mathbf{x}\| < 1, = 1, > 1$ in the whole phase plane $(-\mathrm{div}\,\dot{\mathbf{x}}, \|\mathbf{x}\|)$, there are $2^{2\times2} = 16$ systems which meet the requirements (a)$_1$ and (b)$_1$. Then, considering the remaining 16 systems, we find that there are, respectively, four, four, nine, two, one system(s) violating the restrictions (a)$_2$, (b)$_2$, (b)$_3$, (ab)$_1$, (ab)$_2$, as indicated within the braces $\{\cdots\}$ in Fig. 1, and that, remarkably, there is exactly one system (marked by $\{ep\}$ in Fig. 1) which meets all the requirements (a)$_1$, (b)$_1$, (a)$_2$, (b)$_2$, (b)$_3$, (ab)$_1$, and (ab)$_2$. Moreover, it is very easy to see that the shaded area shown in Fig. 1 for the $\{ep\}$ system can be readily obtained by directly imposing conditions (a)$_1$, (b)$_1$ and

(ab)$_3$ $-\mathrm{div}\,\dot{\mathbf{x}}\|\mathbf{x}\| = -\mathrm{div}\,\dot{\mathbf{x}}$

in the phase plane $(-\mathrm{div}\,\dot{\mathbf{x}}, \|\mathbf{x}\|)$. That is to say, imposing the seven restrictions (a)$_1$, (b)$_1$, (a)$_2$, (b)$_2$, (b)$_3$, (ab)$_1$, and (ab)$_2$ altogether on Eq. (22′) is equivalent to imposing the three conditions (a)$_1$, (b)$_1$, and (ab)$_3$ on Eq. (22′). Furthermore, Eq. (22′) combined with (a)$_1$, (b)$_1$, and (ab)$_3$ is equivalent to

$$\dot{\mathbf{x}} = \mathbf{\Omega}\mathbf{x} - u_0\mathbf{x} + \mathbf{u}, \qquad (25)$$

$$u_0 \geqslant 0, \qquad (26)$$

$$\|\mathbf{x}\| \leqslant 1, \qquad (27)$$

$$\|\mathbf{x}\|u_0 = u_0, \qquad (28)$$

where

$$u_0 := \frac{-\mathrm{div}\,\dot{\mathbf{x}}}{n+1} = \frac{\dot{X}^0}{X^0}. \qquad (29)$$

As a consequence, we have established in a natural and logical way a system governed by Eq. (25) and the complementary trios (26)–(28), which is nothing other than a normalized, dimensionless version of the perfect elastoplastic model [3–5,7].

For the abstract dynamic system (25)–(28), it is now rather straightforward to deduce the on–off switching criteria

$$u_0 = \begin{cases} \mathbf{x}^t\mathbf{u} > 0 & \text{if } \|\mathbf{x}\| = 1 \text{ and } \mathbf{x}^t\mathbf{u} > 0, \\ 0 & \text{if } \|\mathbf{x}\| < 1 \text{ or } \mathbf{x}^t\mathbf{u} \leqslant 0. \end{cases} \qquad (30)$$

When the power of dissipation $u_0 > 0$, the mechanism of dissipation (or plasticity or plastic irreversibility) is on and the system is said to be in the on (or dissipative or elastoplastic or irreversible) phase, while when $u_0 = 0$, the mechanism of dissipation is off and the system is said to be in the off (or conservative or elastic or reversible) phase.

In summary, we have modified the primitive model a little bit and imposed the requirements of non-generativity and controllability on the modified model and also placed the restrictions of boundedness and reachability on the admissible state space, resulting in an abstract dynamic system, which turned out to be the model of perfect elastoplasticity.

5. Symmetry switching between SE(n) and PSO$_o$(n,1)

In the foregoing section the consideration of the phase plane suggested that the validity of the SO$_o(n, 1)$ symmetry be restricted to the cone, when the dynamic system is in the on phase. While in the off phase, the state equation becomes

$$\dot{\mathbf{x}} = \mathbf{\Omega}\mathbf{x} + \mathbf{u}, \qquad (31)$$

as from Eqs. (25) and (30)$_2$. It is clear from the previous derivation that $\mathbf{\Omega}$ belongs to the real Lie

algebra so(n) of the n-dimensional special ortho-gonal (or proper rotation) group SO(n), which is the group of rotations of the n-space $(x_1, x_2, \ldots, x_n)$ around its origin $(0, 0, \ldots, 0)$. Hence in the off phase, one has

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_s^s & \mathbf{G}_0^s \\ \mathbf{0}_{1 \times n} & 1 \end{bmatrix} \tag{32}$$

with

$$\dot{\mathbf{G}}_s^s = \boldsymbol{\Omega}\mathbf{G}_s^s,$$

$$\dot{\mathbf{G}}_0^s = \mathbf{u},$$

where $\mathbf{G}_s^s \in \mathrm{SO}(n)$, $\mathbf{G}_0^s \in T(n)$, and $\mathbf{G} \in \mathrm{SE}(n)$, and the dynamic system has an internal symmetry charac-terized by the special Euclidean (or proper motion) group SE(n), which is the semi-direct product of the translation group $T(n)$ with the proper rotation group SO(n).

This inspires us to remould the spacetime by removing the interior of the cone and gluing a con-tinuously infinite number of stacking Euclidean closed n-discs (i.e. closed n-balls) to the cone. The surgery results in a composite space endowed with the Minkowski metric (4) on the cone and the Euclidean metric $\mathbf{I}_n$ on the closed discs. As a result, the dynamic system in the state space of $\mathbf{x}$ has symmetry switching between the special Euclidean group SE(n) acting on the closed n-ball of admiss-ible states and the projective proper orthochronous Lorentz group $\mathrm{PSO}_o(n, 1)$ acting on the yield hyper-sphere.

6. Truncated future hyperboloids

Now, let us further examine the flow $\mathbf{X}(t)$ induced by the real Lie algebra so($n, 1$) of the Lorentz group $\mathrm{SO}_o(n, 1)$. By virtue of Eqs. (4), (9) and (8) we have

$$\sum_{i=1}^{n} X^i \dot{X}_i - X^0 \dot{X}^0 = \mathbf{X}^t\mathbf{g}\dot{\mathbf{X}} = \mathbf{X}^t\mathbf{g}\mathbf{A}\mathbf{X}$$

$$= -\mathbf{X}^t\mathbf{A}^t\mathbf{g}\mathbf{X} = -\dot{\mathbf{X}}^t\mathbf{g}\mathbf{X} = 0. \tag{33}$$

After integration,

$$\sum_{i=1}^{n} (X^i)^2 - (X^0)^2 = \mathbf{X}^t\mathbf{g}\mathbf{X} = \text{constant} =: -r. \tag{34}$$

It is the equation of a hyperboloid – an n-dimen-sional pseudo-Riemannian submanifold of con-stant curvature which admits the Minkowski metric (4). The argument leading to Eq. (34) is highly remarkable in that it asserts that the hyper-boloid is a homogeneous manifold under $\mathrm{SO}_o(n, 1)$ because the Lorentz group $\mathrm{SO}_o(n, 1)$ left acts transi-tively on the hyperboloid according to Eq. (34). If the constant $r = 0$, Eq. (34) is *the cone*. If $r > 0$, it represents two copies of Minkowskian spheres in *the interior*. If $r < 0$, it is a hyperbolic space in *the exterior*. In Section 2 we assumed $X^0 \neq 0$; further, in Section 3 we choose to assume $X^0 > 0$; further-more in Section 4, Eqs. (29) and (26) amounted to requiring that

$$X^0 \geqslant 1. \tag{35}$$

Eqs. (34) and (35) together indicate that *truncated future hyperboloids* fulfill our need in modeling plasticity.

In the present paper we restrict ourselves in the on phase to the cone and hence to *perfect* plasticity. However, it is interesting to know that the other two choices, the Minkowski sphere and the hyper-bolic space, will lead to hardening and softening, respectively.

7. Two principles of plasticity

Based on the study of these two sections, the whole – seemingly sophisticated and diversified – ideas which govern the behavior of plasticity crystallize into two basic principles:

(1) causality in an n-dimensional truncated future hyperboloid which admits the Minkowski met-ric (or an appropriate nonconstant-curvature pseudo-Riemannian submanifold) and is em-bedded in $(n + 1)$-dimensional Minkowski spacetime (or an appropriate pseudo-Rieman-nian manifold) of augmented states;
(2) controllability and non-generativity in a reach-able, bounded n-space of states of generalized stresses, which are projections of the aug-mented states.

We emphasize that the model of plasticity is a necessary consequence of the two principles, as

has been shown. This seems to have clarified the controversies surrounding the existence of yield surface and the feasibility of switching off the mechanism of plasticity.

Since the two principles are quite simple and transparent in the phenomenological sense and make clear why nature is the way it is, we expect that a systematic generalization from the composite (pseudo-Euclidean) space to composite (pseudo-Riemannian) manifolds will create more organized and self-consistent models of plasticity capable of modeling varieties of constitutive relations with reasonable accuracy for various engineering materials and structural members.

8. Conclusions

On the basis of the above study we conclude that the ideas underlying the behavior of plasticity are "*causality in a truncated future hyperboloid in the Minkowski spacetime* (or its appropriate generalization) *of augmented states*, whose projections are the states of generalized stresses" and "*controllability and non-generativity in a reachable, bounded space of states*," and that the models of plasticity are the necessary consequences of these two principles. To facilitate the modeling of plasticity we devised a constructive methodology, whose legitimacy is rooted in the study in Section 4, especially the two principles of plasticity and our investigation of totally $2^{3 \times 3}$ abstract dynamic systems in their reduced phase spaces.

This paper dealt only with the (isotropic) *perfect* elastoplastic models without considering hardening, softening, anisotropy, *etc.* However, we believe that the two principles crystallized in Section 7 will apply generally and the methodology developed in Section 4 will help create more powerful and faithful models which are all organized under the same basic principles of plasticity. As an immediate outlook, one may further consider the composition of all Zeeman's three (an $SO_o(n, 1)$, a translation and a dilation), and even more sophisticated group actions on pseudo-Riemannian manifolds, say causal group and conformal group. Thus making an explicit use of the powerful group-theoretic and dynamics-theoretic methods of Sections 2 and 3, one would expect to end up with various models of plasticity, such as bilinear elastoplasticity [8], mixed hardening, viscoplasticity, and more.

References

[1] H.-K. Hong, H.-S. Lan, A new point-of-view at plastic materials of the Prandtl-Reuss type, in: A.S. Khan, M. Tokuda (Eds.), Advances in Plasticity 1989, Proceedings of the Second International Symposium on Plasticity and its Current Applications, Pergamon, Oxford, 1989, pp. 95–98.

[2] H.-K. Hong, J.-K. Liou, Integral-equation representations of flow elastoplasticity derived from rate-equation models, Acta Mech. 96 (1993) 181–202.

[3] H.-K. Hong, C.-S. Liu, Prandtl-Reuss elastoplasticity: on–off switch and superposition formulae, Int. J. Solids Struct. 34 (1997) 4281–4304.

[4] H.-K. Hong, C.-S. Liu, On behavior of perfect elastoplasticity under rectilinear paths, Int. J. Solids Struct. 35 (1998) 3539–3571.

[5] H.-K. Hong, C.-S. Liu, Internal symmetry in the constitutive model of perfect elastoplasticity, Int. J. Non-Linear Mech. 35 (2000) 447–466.

[6] E.C. Zeeman, Causality implies the Lorentz group, J. Math. Phys. 5 (1964) 490–493.

[7] H.-K. Hong, C.-S. Liu, Lorentz group $SO_o(5, 1)$ for perfect elastoplasticity with large deformation and a consistency numerical scheme, Int. J. Non-Linear Mech. 34 (1999) 1113–1130.

[8] H.-K. Hong, C.-S. Liu, Internal symmetry in bilinear elastoplasticity, Int. J. Non-Linear Mech. 34 (1999) 279–288.

[9] J.F. Cornwell, Group Theory in Physics, Vol. 2, Academic Press, London, 1984.