PHASE DIAGRAM BASED DESCRIPTION OF THE HYSTERESIS BEHAVIOR OF SHAPE MEMORY ALLOYS

A. BEKKER and L. C. BRINSON

1Applied Mathematics Department, Northwestern University, 2145 Sheridan Road, Evanston, IL, U.S.A. and 2Mechanical Engineering Department, Northwestern University, 2145 Sheridan Road, Evanston, IL, U.S.A.

(Received 5 March 1997; accepted 4 December 1997)

Abstract—In this paper, we develop a consistent mathematical description of martensite fraction evolution during athermal thermoelastic phase transformation in a shape memory alloy (SMA) induced by a general thermomechanical loading. The global kinetic law is based on an experimentally defined stress–temperature phase diagram, transformation functions for a one-dimensional SMA body and a novel vector hysteresis model. The global kinetic law provides the phase fraction history given a loading path on the stress–temperature phase diagram and an initial value of martensite fraction. The phase transformation is considered to occur only within transformation strips on the phase diagram and only on loading path segments oriented in the transformation direction. The developed procedure can be used to model a range of different SMA transformation behaviors depending on the choice of transformation functions and local kinetic law algorithms. The phase fraction evolution is examined for a number of characteristic examples, including cyclic loading resulting in oscillatory transformation paths, and internal loops of partial transformation with associated attractor loops. Differences between the various local kinetic law algorithms used in the overall framework are highlighted. The simulation results using a cosine transformation function are found to be in excellent agreement with experimental data.

1. INTRODUCTION

Shape memory alloys (SMAs) are being used in an increasing number of interesting engineering applications. The large recoverable stresses and strains attainable by these materials are due to an athermal thermoelastic martensite phase transformation in the material and depend on the temperature and stress loading history of the material. In order to utilize SMAs effectively in novel applications, accurate and reliable predictions of their hysteresis response are required. An important part of such modeling includes a constitutive law and transformation kinetics relations.

A variety of models have been proposed to describe the transformation hysteresis of shape memory alloys on a macroscale level [1–12], many of which were developed to simulate response induced by a change in temperature only (or stress only) and thus are called scalar models. The scalar models are based on Duhem–Madelung type differential model or the Preisach integral model, then adapted for the case of shape memory phase transformation [6, 7]. One key difference between these two hysteresis models is that the Preisach model is unable to represent dead zones (and dead directions) of transformation or drift of hysteresis loops with cyclic loading during partial transformation. Such behavior has been seen experimentally in NiTi materials [12, 13] and can be captured by Duhem–Madelung models and with the model developed here. Duhem–Madelung models can be separated into those derived from micromechanics [10, 11] and those derived from a general, simplified physical model of the transformation process [1, 3, 6, 8].

Several existing models have formulated SMA kinetics for general thermomechanical loading using the stress–temperature plane and thus are called vector models [1, 3, 8, 12]. However, the vector models available have not provided a consistent procedure for consideration of complex loading paths: calculations are generally performed on an ad hoc basis and the underlying assumptions are not clearly defined. Starting with these phenomenological vector models, in this paper we develop a consistent mathematical description for a macroscale SMA kinetic law based on the geometry of the loading path on an experimentally defined stress–temperature phase diagram. The developed global kinetic law governs application of transformation functions defined for unidirectional monotonic loading (outer loop) to incomplete transformations (internal loops) and enables unambiguous prediction of 1D SMA material response to arbitrary stress–temperature loading. Despite the fact that the procedure has algorithmic recursive character, it is closely related to the Duhem–Madelung differential models, and particularly to the variable hysterons construct [14]. The basic idea of the global kinetic law is to divide a loading path...
into loading segments, in each of which the phase fraction is a single valued function of stress and temperature (local kinetic laws), and then to assemble the chain of local kinetic laws along the loading path. Because of the character of the assembling process the procedure can be called a “chain” model.

In this paper, we assume that the SMA material/ body is trained so that the outer/major hysteresis loop is stable. Tanaka [15] has recently addressed the training/stabilization process of outer loops by introducing a history dependence for the transformation start lines on the phase diagram. In CuZnAl, the transformation start lines for trained materials with stable outer loop strongly depend on the loading history [4, 5, 16]. While we currently present the model for a phase diagram with fixed transformation strips, such a change of transformation lines with loading could be easily accommodated. Cyclic loading data for NiTi materials [13, 15] indicate fixed transformation strips.

In the global kinetic law procedure, we introduce several local kinetic law algorithms with which a variety of popular transformation functions can be used. These distinct algorithms and transformation functions are important features of our kinetic law, allowing the most appropriate to be chosen for a given material.

2. GENERAL DESCRIPTION OF SMA TRANSFORMATION BEHAVIOR

In this Section we describe the transformation behavior of a one-dimensional (1D) SMA polycrystalline body undergoing thermal and uniaxial stress loading. We also introduce a \((\sigma–T)\) phase diagram and show how it facilitates description of the transformation process. The high–temperature phase of the SMA (the parent phase) is austenite (A), which when cooled under zero-stress forms a self-accommodated structure consisting of twin related variants; this low temperature phase is termed twinned martensite (Mt) and there is no associated transformation strain with its formation. A twinned martensitic structure can be detwinned \((M^t \rightarrow M^d)\), or austenite can be converted to detwinned martensite \((A \rightarrow M^d)\) under sufficient stress; the detwinned martensitic structure is associated with a large transformation strain relative to the twinned martensite or austenite state [17].

The natural quantitative measure of the ongoing phase transformation is the volume fraction of martensite, \(\xi\), which is the sum of the twinned and detwinned components. The martensite fraction, \(\xi\), serves as an internal variable in the phenomenological constitutive law (given in Section 5) that can be coupled with the momentum and energy equations to model transformation in 1D SMA body [18].

We assume that the set of variables \((\sigma, T, \xi)\) can fully describe a local thermomechanical state of a 1D SMA body on a macroscale level. The external parameters, \(\sigma\), \(T\), are meant to be time- and space-averaged on some sub-macroscopic level determined by the requirement that \(\xi(x, t)\) must be a representative (smooth enough) function. The length-scale of averaging (size of the constitutive element) is at least an order larger than the average grain diameter. Due to the diffusionless character of martensite transformation, \(\xi\) depends only on current values of \(T, \sigma\) as well as on some features of loading history and is independent of the rate of change of \(T, \sigma\) on a macro-timescale level [3, 12]. A variety of semi-empirical, rate independent kinetic laws for athermal thermoelastic martensite transformations have been proposed [3, 8, 9] in conjunction with a \((\sigma–T)\) phase diagram.

The \((\sigma–T)\) phase diagram is of central importance for understanding the process of martensite transformation in SMA materials. Stress–strain curves from isothermal, uniaxial tests are typically used to determine the critical stresses where transformation begins and ends. For polycrystalline, stabilized NiTi materials, it is usually seen that the transformation occurs in transformation strips which are independent of loading history on the phase diagram (see Fig. 1). In an earlier article, the tensile phase diagram shown was discussed at length [18] and here we mention only key features necessary for the present study. Transformation occurs across the strips \([A, M^t, d, t]\) unidirectionally, when a point \((T, \sigma)\) \((A, M^t)\) representing a local state of the SMA system moves across a strip in the direction of transformation, i.e., from the start boundary \([i\] to the finish boundary \([t]\) of the strip \([i]\) as indicated by the director vectors \(n’(i = A, M, d, t)\) on the phase diagram. The unit vectors \(n’\) are normal to the corresponding strips: \(n’ \perp [i]\).

Consequently, the phase diagram is divided into four major regions: in the \(M^t\) region only detwinned martensite can exist; in the \(A\) region only austenite can exist; in the \(M^d\) region both twin
and detwinned martensite can co-exist; in the $M^d_A$ region all phases can co-exist. These regions outside of the transformation strips where no transformation occurs will be referred to as dead zones. The presence of the dead coexistence regions and transformation strips makes it clear that the transformation state of the SMA material is path-dependent.

Limited theoretical [20, 21] and experimental [22] data suggest that the compression phase diagram in the first approximation can be obtained as the mirror reflection of the tension phase diagram with respect to $T$-axis. Then the combined compression–tension phase diagram will have two detwinned martensite zones, $M^+,$ $M^-$ with symmetry related detwinned martensite phases.

In this paper, we use a simplified phase diagram shown in Fig. 2. In this phase diagram only two phases are present (M, A) and correspondingly there are only two strips ([M], [A]). With minor notational changes only, this phase diagram and the ensuing kinetic derivations can be considered to represent (1) the high-stress part of Fig. 1, where transformation is between austenite and detwinned martensite: $A \rightarrow M^d$; (2) the low-stress part of Fig. 1, where transformation is between austenite and twinned martensite: $A \rightarrow M^t$; (3) the low-temperature part of the compression–tension phase diagram, where transformation is between detwinned martensites: $M^t \rightarrow M^d$. For example, to model transformation from austenite to twinned martensite using Fig. 2, M region should be identified with $M^d$ and M everywhere else should be replaced by t: $[M] \rightarrow [t], n^M \rightarrow n^t,$ etc.

If we identify M on Fig. 2 with $M^+$ then the phase diagram (Fig. 2) can be obtained from the relevant part of the theoretical phase diagrams (Fig. 14 [21] and Fig. 11 [20]), by unfolding the boundaries $\sigma_i^d(T), \sigma_i^{M^+}(T)$ of the $AM^+$ metastable region into the transformation strips $[\sigma_i^d(T), \sigma_i^{M^+}(T), \sigma_i^{M^*}(T)]$. We believe that this unfolding is caused by imperfections in the SMA material, which serve as internal constraints, leading to storage of elastic energy during transformation.

In recent work [23], media with “imperfections” were simulated using a random field version of a 3-state spin BEG model which describes the evolution of the 3-state (A, $M^+$, $M^-$) spin variable driven by external “stress” and “temperature” fields. The unfolding of the boundaries of the metastable $AM^-$, $AM^+, M^+M^+$ regions into the transformation strips was observed resulting in a phase diagram very similar to Fig. 1.

Despite the fact that the phase diagram Fig. 1 was constructed using experimental data for Nitinol wires (i.e. SMA system of finite length), results of [20, 21, 23] show it is reasonable to assume that it can be applied locally, for an elementary volume of a 1D SMA transforming body. In Refs [18, 24, 28] we have shown that this phase diagram with appropriate kinetics can be successfully used to model diffuse (two-phase zone migration) and localized
Note that if the [M], [A] strips coincide in Fig. 2, the MA dead region will vanish, rendering the \((s-T)\) phase diagram similar to a \((c-T)\) phase diagram (e.g. Fig. 129f of Ref. [25]), where the [M] ([A]) strip is the counterpart of a solid–liquid two phase region. However, while for solidification the solid phase fraction is a single valued function \(x = x(c, T)\), on the equilibrium \((c-T)\) phase diagram [25], it is impossible to prescribe a unique value of the martensite fraction, \(x\) to all points of the meta-equilibrium \((s-T)\) phase diagram. Therefore even when transformation strips coincide, which is the case for some Cu-based alloys, hysteresis is still present. Hence the development and study of consistent path dependent kinetic laws for SMA materials is necessary for proper thermomechanical modeling.

To address this issue, we first define the thermomechanical loading path at some point of a one-dimensional SMA body as the oriented curve \(L = \{(T(t), \sigma(t))| t \in [t_0, t_E]\}\) (Figs 1 and 2), where \(L\) can have any number of singular points and/or loops. On portions of \(L\) outside of the transformation strips, the martensite fraction is constant. For transformation behavior on portions of \(L\) within the strips \([i]\), \((i = M, A)\), we make the following assumptions, which formalize implicit rules that are often used when describing SMA behavior in conjunction with a phase diagram:

1. Transformation occurs only at the points of \(L\) in \([i]\) where the direction of the loading path has a component in the direction of the director vector, \(n_i\) of the strip \([i]\).
2. The extent of transformation increases with the distance traveled along \(L\) across the transformation strip \([i]\) in the \(n_i\) direction.

In accordance with these assumptions, we define that a segment \(L_j \subset [i]\) in the strip \([i]\) has the transformation direction and denote that as \(L_j^n_i\) if at all points of \(L_j\) the scalar product \(\tau \cdot n_i > 0\), where \(\tau = (T(t), \sigma(t))\) is the tangent vector to \(L_j\). A segment \(L_j\) in the strip \([i]\) is said to have the dead direction, \(L_j^n_i\), if \(\tau \cdot n_i \leq 0\) at any point of the \(L_j\). The martensite fractions change only on the segments \(L_j \subset [i]\) that have the transformation direction.

### 3. DESCRIPTION OF TRANSFORMATION BEHAVIOR ALONG T-AXIS

Before presentation of the kinetics for arbitrary thermomechanical loading paths on the stress–temperature diagram, it is instructive to describe the simplest case of the complete forward \(A \rightarrow M\) and reverse \(M \rightarrow A\) transformation induced by unidirectional monotonic change of temperature from \(+\infty\) to \(-\infty\) and back along the \(T\)-axis (loading paths \(L_-\) and \(L_+\), correspondingly). This will allow us to introduce the fundamental notion of transformation functions. The phase fraction change along \(L = L_- \cup L_+\) can be represented by the two envelope functions \(e_-(T)\), \(e_+(T)\) (lower and upper branches) that define the major hysteresis loop (see Fig. 3):

\[
\xi = \begin{cases} 
    e_+(T), & \text{if } T \in L_+ \\
    e_-(T), & \text{if } T \in L_-
\end{cases}
\] (1)

![Fig. 3. Envelope functions, \(e_\pm(T)\), defining the outer/major hysteresis loops for three different transformation functions, \(f^M, f^A\).](image-url)
where $e_+(T) \geq e_-(T) \forall T \in (-\infty, +\infty)$ and satisfy the saturation conditions:

$$e_+(-\infty) = e_-(\infty) = 1,$$

$$e_+(\infty) = e_-(\infty) = 0$$  \hfill (2)

The transformation functions $f^A, f^M$ are defined in the corresponding transformation intervals on the $T$-axis and represent the change of phase fraction starting from pure M, A phases, respectively:

$$\zeta = f^A(X^A), \text{ if } T \uparrow \text{ in } [T_{S0}^{A}, T_{S0}^{A}]$$

$$\zeta = f^M(X^M), \text{ if } T \downarrow \text{ in } [T_{S0}^{M}, T_{S0}^{M}]$$  \hfill (3)

where dimensionless variables $X^A, X^M \in [0, 1]$ are the normalized distances from the start boundaries of the $[A], [M]$ transformation strips:

$$X^A(T) = \frac{T - T_{S0}^{A}}{T_{S0}^{A} - T_{S0}^{A}}, \quad X^M(T) = \frac{T - T_{S0}^{M}}{T_{S0}^{M} - T_{S0}^{M}}$$  \hfill (4)

Then the envelope functions are given by

$$e_+(T) = \begin{cases} 0, & \text{if } T \uparrow \text{ in } [T_{S0}^A, +\infty) \\ f^A(X^A(T)), & \text{if } T \uparrow \text{ in } [T_{S0}^A, T_{S0}^A] \\ 1, & \text{if } T \uparrow \text{ in } (-\infty, T_{S0}^A] \end{cases}$$

$$e_-(T) = \begin{cases} 0, & \text{if } T \downarrow \text{ in } [T_{S0}^M, +\infty) \\ f^M(X^M(T)), & \text{if } T \downarrow \text{ in } [T_{S0}^M, T_{S0}^M] \\ 1, & \text{if } T \downarrow \text{ in } (-\infty, T_{S0}^M] \end{cases}$$  \hfill (5)

Examples of widely used transformation functions include:

- Algebraic functions,

$$f^A(X^A) = 1 - (X^A)^{\gamma^A}, \quad f^M(X^M) = (X^M)^{\gamma^M}$$  \hfill (7)

with the linear case ($\gamma^A = \gamma^M = 1$) used by many researchers [4, 6, 10, 11].

- Cosine functions:

$$f^A(X^A) = 1 - \frac{1}{2} [1 - \cos(\pi(X^A)^{\gamma^A})],$$

$$f^M(X^M) = \frac{1}{2} [1 - \cos(\pi(X^M)^{\gamma^M})]$$  \hfill (8)

with $\gamma^A = \gamma^M = 1$ proposed by Liang and Rogers [9].

Note, that the saturation conditions hold for the functions (7) and (8) at the start and finish boundaries:

$$f^A(0) = 1,$$

$$f^A(1) = 0$$  \hfill (9)

$$f^M(0) = 0,$$  \hfill (11)

$$f^M(1) = 1$$  \hfill (12)

- Exponential functions:

$$f^A(X^A) = \exp(\pi(X^A)^{\gamma^A}),$$

$$f^M(X^M) = 1 - \exp(\pi(X^M)^{\gamma^M})$$  \hfill (13)

with $\gamma^A = \gamma^M = 1$ proposed by Tanaka [3], $\gamma^A = \gamma^M = 2$ used by Ivshin and Pence [6]. Instead of saturation conditions (10) and (12) the exponential functions require calibration conditions on the finish boundaries:

$$f^A(1) = \Delta, \quad f^M(1) = 1 - \Delta$$  \hfill (14)

where $\Delta \ll 1$ to simulate complete transformation at the finish boundaries of both strips [typically $\Delta = 1\%$ is used, so $\alpha = \ln(0.01)$]. Other types of exponential type transformation functions can be found in Refs [1, 6, 8].

In Fig. 3, the major hysteresis loops for linear, exponential and cosine transformation functions ($\gamma^A = \gamma^M = 1$) are shown. Generally, the cosine function best fits experimental data [26, 27].

4. THE KINETIC LAW

In this Section we formalize the assumptions from Section 2 concerning the process of transformation in shape memory alloys. We create a rigorous kinetic procedure which enables application of the kinetic law unambiguously and consistently to complex loading paths, introducing local kinetic law algorithms which govern application of the transformation functions to incomplete transformations. The global kinetic law is obtained by assembling the chain of local kinetic laws along a loading path.

We address here the general problem: *Given a thermomechanical loading path/history $L = \{(T(t), \sigma(t))| t \in [t_0, t_F]\}$ and the initial value of the martensite fraction, $\xi(t_0) = \xi_{0b}$ determine the transformation path/history $\chi = \{\xi(t)\| t \in [t_0, t_F]\}$. We will utilize the description of the phase diagram to construct the global kinetic law procedure, $\Phi$

$$\xi(t) = \Phi(T(t), \sigma(t); \xi_{0b}, t \in [t_0, t_F])$$  \hfill (15)

which solves the formulated problem*. As postulated above, the kinetics of the transformation is rate independent and time is used here only for parameterization. Therefore whenever possible we will use a time-invariant form for equations. We introduce necessary terminology first, then present the formulation of the local and global kinetic laws.
4.1. Definitions and classification of switching points

- Switching points \(- S_j(T, \sigma), j = 1, 2, \ldots, J\) of a loading path \(\mathcal{L}\) are the points where the direction or sense of transformation changes:
  
  (1) the points where \(\mathcal{L}\) enters or leaves the transformation strip \([i], (i = A, M)\) in the direction of transformation, \(n^i\);
  
  (2) the points inside the strip \([i]\) where \(\mathcal{L}\) changes its direction from \(\mathcal{L}^i[n^i] \) to \(\mathcal{L}^j[n^j]\) or vice versa (see Figs 2 and 4).

Where it is not essential we omit coordinates and indicator of the strip and simply write \(S_j\). The initial point \((T_0, \sigma_0)\) of a loading path \(\mathcal{L}\) we call the zero switching point, \(S_0\).

- Points of a loading path \(\mathcal{L}\) which are not switching points we call regular points. The end point of the loading path \((T_E, \sigma_E)\) may be a regular point.

- The switching moments, \(t_j\), of a loading path \(\mathcal{L}\) are the moments when \(T(t_j) = T_i, \sigma(t_j) = \sigma_i\). The switching moments are numbered in the direction of increasing of time: \(t_0 < t_1 < \ldots < t_j < t_E\) and they divide the total time interval \(I = [t_0, t_E]\) into \((J + 1)\) subintervals, \(I_j\); correspondingly the switching points, \(S_j\), divide \(\mathcal{L}\) into \((J + 1)\) loading segments, \(\mathcal{L}_j = S_j, j = 0, 1, \ldots, J\), with \(\mathcal{L}_j = \mathcal{L}_{j-1}, j = 0, 1, \ldots, J\), so that \(\mathcal{L} = \bigcup_{j=0,1,\ldots,J} \mathcal{L}_j\).

- The \(\tau_j^\pm\) are defined as the tangent vectors \(\tau_j^+(A) = (T'(t_j^\pm A), \sigma'(t_j^\pm A)), A \geq 0\), associated with the switching point, \(S_j\).

\(^\dagger\)Note that if \(S_j\) is a singular point of a curve \(\mathcal{L}\) then \(\lim_{t \to \epsilon} \tau_j^+(A) = \lim_{t \to \epsilon} \tau_j^-(A) = 0\) [see Fig. 5(a)] otherwise \(\lim_{t \to \epsilon} \tau_j^+(A) = \lim_{t \to \epsilon} \tau_j^-(A) = \tau_j^0\) (the right and left tangent vectors in this point are equal).

Using the above definitions, the switching point \(S_j\) of a loading path \(\mathcal{L}\) in the strip \([i]\) can be classified as follows (see also Fig. 4):

- (I) Entrance point, \(S_j^{\text{en}}\): if \(S_j\) lies on the start boundary \([i]\) and there is a small neighborhood, \(0 < t_j - t_0 \leq \delta\), of this point such that \(\tau_j^+(A)n^i > 0\) whenever \(A < \delta\).

- (II) Exit point, \(S_j^{\text{ex}}\): if \(S_j\) leaves the start boundary \([i]\) and there is a small neighborhood, \(0 < t_j - t_0 \leq \delta\), of this point such that \(\tau_j^-(A)n^i > 0\) whenever \(A < \delta\).

- (III) Re-start point of transformation, \(S_j^{\text{rs}}\): if there is a small neighborhood, \(0 < t_j - t_0 \leq \delta\), of this point such that \(\tau_j^+(A)n^i > 0\) whenever \(A < \delta\).

With the exception of \(S_0\), all switching points lie inside or on the boundaries of the transformation strips and any two sequential switching points \(S_j^{\text{en}}, S_j^{\text{ex}}\) of \(\mathcal{L}\) are of different type, \(\tau \neq \beta\). For the initial point \(S_0\), only the second inequality in (II), (III) must be checked.

- The label \(\lambda_j\) of the switching point \(S_j^{\text{ex}}\) is defined as the ordered pair of symbols \(\lambda\) (e.g., for \(S_j^{\text{ex}}\), \(\lambda_j = (\text{MX})\)), where \(\lambda\) has the range \{A, M, D\} and \(\sigma\) has the range \{N, S, R, X, 0\}. If \(S_0\) belongs to a dead zone, we assign the label \(\lambda_0 = DO\) to it.

- Stop and restart points are often called return or reverse points of the transformation [7, 29].

- The base point of a segment \(L_j\) is the switching point \(S_j\) at which the segment begins.

- The memory parameters of a segment \(L_j\) are the parameters \(T_j, \sigma_j, \xi_j\) at the base point of \(L_j\).

- The trace \(\mathcal{F}(\mathcal{L})\) of a path \(\mathcal{L}\) is defined as the string of the \((T, \sigma)\)-memory parameters: \(\mathcal{F}((T_0, \sigma_0, (T_1, \sigma_1), \ldots, (T_j, \sigma_j)).\)
4.2. Local kinetic law

As was shown in Section 2, the martensite fraction is an intrinsically path dependent function of $T, \sigma$: specifically, location and direction of a segment $L_j \subseteq \mathcal{L}$ determines the local kinetic law, $\phi_j$, on it:

$$\xi(t) = \phi_j(T(t), \sigma(t); \xi_j(t), \sigma_j; t \in I_j) \quad (16)$$

where

$$\phi_j(T(t), \sigma_j; \xi_j(t)) = \begin{cases} F^A(T, \sigma; T_j, \sigma_j; \xi_j), & \text{if } L_j \in [A] \text{ and } L_j \uparrow \uparrow n^A \\ F^M(T, \sigma; T_j, \sigma_j; \xi_j), & \text{if } L_j \in [M] \text{ and } L_j \uparrow \uparrow n^M \\ F^D \equiv \xi_j, & \text{otherwise} \end{cases} \quad (17)$$

and the functions $F^A, F^M, F^D$ are the branches of the local kinetic law, expressions for which will be presented shortly. The $F^A, F^M$ branches represent active transformation behavior in the $[A], [M]$ regions in the transformation directions and are single valued functions of all arguments; the $F^D \equiv \xi_j$ represents passive transformation behavior in dead zones and in dead directions in the $[A], [M]$ regions. Note that only the memory parameters at the $f^0$-point are included in the definition of the kinetic law branches.

Using labels, $\lambda_j$, the local kinetic law (equations (16)–(17)) can be written in the simple form

$$\xi = F^{\xi_j}(T, \sigma; T_j, \sigma_j; \xi_j), \quad (T, \sigma) \in L_j \quad (18)$$

with the indicator defined as

$$c_j = c(\lambda_j) = \begin{cases} A, & \text{if } \lambda_j = AN, AR \\ M, & \text{if } \lambda_j = MN, MR \\ D, & \text{otherwise} \end{cases} \quad (19)$$

Thus the type of base switching point of loading path segment $L_j$ determines the branch of local kinetic law prescribed on that segment. The indicator $c_j = c(\lambda_j)$ is also assigned to the corresponding loading segment and time interval: $L_j^\uparrow, t^\uparrow$, so that $L_j^\uparrow \uparrow n^j$ if $c_j = i$.

To define the $F^A, F^M$ branches of the kinetic law we first need to generalize for arbitrary $L_j \subseteq \mathcal{L}$ the $X'$ distances defined on the $T$-axis in Section 3. Here we introduce two distances, $Y'$ and $Z'$, defined along loading path segments that have the transformation direction in the strips $[i], (i = A, M)$.

The distance $Y' = Y(T, \sigma), (i = A, M)$, represents the normalized distance between a point $(T, \sigma) \in L_j \subseteq \mathcal{L}$ and the start boundary $[i]$ (see Fig. 5):

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig5.png}
\caption{The distances $\rho_0^M, \rho_j^M, \rho_j^M$ and vectors $r_j^M, r_j^M$ in the [M] strip.}
\end{figure}
\[ Y'(T, \sigma) = \frac{g(T, \sigma)}{\psi_0}, \]
\[ g(T, \sigma) = r' \cdot n = n_1(T - T_{s_{0}}) + n_2 \sigma, \]
\[ r' = (T - T_{s_{0}}, \sigma), \]
\[ \psi_0 = g'(T_{j_{0}}, 0) = n_1(T_{j_{0}} - T_{s_{0}}) \] (20)

where \( \psi' \) is the perpendicular distance between the point \((T, \sigma)\) and \([i, \psi_0]\) is the width of the transformation strip \([i]\). Since \(\max_{T, \sigma} g'(T, \sigma) = \psi_0\), the distances \(Y'\) range in value from zero at the start boundary \([i]\) (at the entrance points) to one at the finish boundary \([j]\) (at the exit points) inside the corresponding strips. Note also that on the \(T\)-axis: \(Y'(T, 0) = \psi(T)\).

The distance \(Z' = Z(T, \sigma; T_{j}, \sigma_{j})\), \((i = A, M)\), represents the normalized distance between a point \((T, \sigma)\) on \(Z'_{j}\) \(\in [i]\) and the base point \((T_{j}, \sigma_{j})\) of the segment \(Z'_{j}\) (see Fig. 5):

\[ Z'(T, \sigma; T_{j}, \sigma_{j}) = \frac{g'(T, \sigma) - g'_j}{\psi'_0 - \psi'_j} \]
\[ \psi'_j = g'(T_{j}, \sigma_{j}) = r'_j \cdot n = n_1(T_{j} - T_{s_{0}}) + n_2 \sigma_{j} \] (21)

where \(\psi'_j\) is the distance from the base point \(S_{j}\) of the segment \(Z'_{j}\) to the boundary \([i]\). It follows from this definition that \(Z' = 0\) at the entrance and restart points and \(Z' = 1\) at the exit points.

The distances \(Y', Z', (i = A, M)\) are linear functions of \(T, \sigma\) on \(Z'_{j}\). If \(S_{j}(T_{j}, \sigma_{j}) \in [i]\) then

\[ Y'(T, \sigma) = \frac{\psi(T, \sigma)}{\psi_0} = \frac{n_1(T - T_{s_{0}}) + n_2 \sigma}{n_1(T_{j_{0}} - T_{s_{0}})} \] (22)

which can be written in the notation of Liang and Rogers [9], e.g. for \(i = A:\)

\[ Z^A(T, \sigma; \cdot) = Y^A(T, \sigma) = a^A(T - T_{s_{0}}) + b^A \sigma, \]
\[ a^A = 1/(T_{j_{0}} - T_{s_{0}}), \quad b^A = -a^A/k^A, \]
\[ k^A = -n_1^A/n_2^A, \quad a^A, k^A > 0 \] (23)

From equations (20 and 21) it follows that \(Y', Z'\) are increasing functions of \(t\) on \(Z'_{j}\). The \(Y', Z'\) functions are closely related to the state function \(Z\) introduced by Cory and McNichols [12].

According to the basic assumptions 1, 2 (Section 2) concerning the process of the transformation, the \(F^A, F^M\) branches of the kinetic law can be represented by the following three algorithms using the \(Y = Y(T, \sigma), Y'_j = Y(T_{j}, \sigma_{j}), Z = Z(T, \sigma; T_{j}, \sigma_{j})\) distances:

The Z-algorithm:

\[ F^A = F^A(Z^A; \xi_j) = \xi_j f^A(Z^A), \]
\[ F^M = F^M(Z^M; \xi_j) = \xi_j + (1 - \xi_j) f^M(Z^M) \] (24)

The Y-algorithm:

\[ F^A = F^A(Y^A, Y^A_j; \xi_j) = \xi_j f^A(Y^A - Y^A_j), \]
\[ F^M = F^M(Y^M, Y^M_j; \xi_j) = \xi_j + (1 - \xi_j) f^M(Y^M - Y^M_j) \] (25)

The YF-algorithm:

\[ F^A = F^A(Y^A, Y^A_j; \xi_j) = \xi_j f^A(Y^A) \]
\[ = \xi_j + (1 - \xi_j) f^M(Y^M - f^M(Y^M_j)) \] (26)

with transformation functions \(f^A, f^M\) as defined in Section 3. Equations (24)-(26) generalize the transformation functions for the case when transformation starts from \(\xi = \xi_j\) and allow us to consistently describe incomplete transformations. Since the transformation functions \(f^A, f^M\) are monotone and \(Y', Z'\) are increasing functions of \(t\) on \(Z'_{j}\), the \(F^A, F^M\) are also monotone on \(I'_{j}\).

It is important to note that if the base point of \(Z'_{j}\) lies on the start boundary of the strip, \(S_{j}(T_{j}, \sigma_{j}) \in [i]\), then \(Z(T, \sigma; T_{j}, \sigma_{j}) \equiv Y(T, \sigma)\) and all three algorithms become identical. The YF-algorithm is constructed using a scalar algorithm [6] as a prototype, which is obtained by integration of the Duhem–Madelung model; we believe the Y-algorithm is also related in a similar fashion to some Duhem–Madelung differential model, but this is not the case for the Z-algorithm.

4.3. Global kinetic law

Use of the rules (I-IV) to detect and label switching points along a loading path, \(Z\) (starting from the initial point \(S_{0}\) where \(\xi_{s_{0}}\) is known), in conjunction with the local kinetic law \([\text{equations (16) and (19)}]\) constitutes an algorithm that allows one to propagate the phase fraction, \(\xi\), along a loading path. After the solution is determined for the entire path \(Z = \bigcup_{j=0,1,2,...} Z'_{j}\), the global kinetic law, \(\Phi\), can be written as the sequence of the local kinetic laws along the path:
or using equations (17 and 18)

\[
\xi = \Phi(T(t), \sigma(t); l_{0})_{(T, \sigma) \in \mathcal{L}}
\]

\[
\phi_{i}(T(t), \sigma(t); l_{0}), \quad t \in I_{0}
\]

\[
\phi_{j}(T(t), \sigma(t); l_{j}), \quad t \in I_{j}
\]

(27)

or using equations (17 and 18)

\[
\xi = \Phi(T, \sigma; l_{0})_{(T, \sigma) \in \mathcal{L}}
\]

\[
F_{\xi}(T, \sigma; T_{0}, \sigma_{0}; l_{0}), \quad (T, \sigma) \in \mathcal{L}_{0}
\]

\[
F_{\xi}(T, \sigma; T_{1}, \sigma_{1}; l_{1}), \quad (T, \sigma) \in \mathcal{L}_{1}
\]

\[
\ldots
\]

\[
F_{\xi}(T, \sigma; T_{j}, \sigma_{j}; l_{j}), \quad (T, \sigma) \in \mathcal{L}_{j}
\]

\[
\ldots
\]

\[
F_{\xi}(T, \sigma; T_{J}, \sigma_{J}; l_{J}), \quad (T, \sigma) \in \mathcal{L}_{J}
\]

(28)

where

\[
\xi_{j} = F_{\xi}(T_{j}, \sigma_{j}; T_{j-1}, \sigma_{j-1}; l_{j-1}), \quad j = 1, 2, \ldots, J
\]

(29)

The memory parameters \(T_{j-1}, \sigma_{j-1}, l_{j-1}\) of loading path segment \(\mathcal{L}_{j-1}\) determine the memory parameter \(\xi_{j}\) in the base point of segment \(\mathcal{L}_{j}\). Thus, through recursive relation (29) for \(\xi_{j}\) coupling between two sequential local kinetic laws is realized and memory is transported from one loading segment to another. Note also that from equations (18) and (24)-(26) it follows

\[
\xi_{j} = F_{\xi}(T_{j}, \sigma_{j}; T_{j-1}, \sigma_{j-1}, l_{j-1}), \quad j = 1, 2, \ldots, J
\]

(30)

which ensures continuous matching of the local kinetic laws at the switching points. The formulated global kinetic law \(\xi(t) = \Phi(T(t), \sigma(t); l_{E})\) is a single valued, piecewise monotone, continuous function of \(t\) and almost everywhere differentiable on \(I\). In contrast the function \(\xi = \Phi(T, \sigma; l_{E})_{(T, \sigma) \in \mathcal{L}}\) can be multivalued when \(\mathcal{L}\) has loops.

The form of equation (28) emphasizes that the structure of the global kinetic law (i.e., the type and sequence of local kinetic laws) is determined by the signature of the loading path, \(S(\mathcal{L})\). This structure is very sensitive to pinning perturbations of a loading path which result in SR-points and thus change the signature. However, the sensitivity of \(\xi\) to such perturbations depends on the choice of algorithm for the \(F^{A}, F^{M}\) branches and the choice of the transformation function (e.g. the \(Y\)-algorithm leads to \(\Phi\) which is insensitive to SR-points). Physically, sensitivity to SR-points implies a difference between moving across a transformation strip continuously or with stops. This issue is also of concern for numerical implementations of an SMA kinetic law.

Given the trace \(S(\mathcal{L})\), the signature \(S(\mathcal{L})\) is easily found and the trace of the transformation path \(\Xi(l) = \{\xi_{0}, \xi_{1}, \ldots, \xi_{j}\}\) is determined by the recursive relation (29); thus the global kinetic law [equation (28)] is completely specified. Therefore, if two loading paths with the same trace have the same phase fraction, \(\xi_{0}\), in the initial point then they have equal phase fractions in the intersection/common points.

To discuss the memory properties of the kinetic law, note that the phase fraction at the end point \((\cdot)\mathcal{E}(T(t), \sigma(t))\) of a path \(\mathcal{P} = \bigcup_{j=0,1,2,\ldots} \mathcal{P}_{j}\) is

\[
\xi(t) = F_{\xi}(T(t), \sigma(t); T_{J}, \sigma_{J}; l_{J}) = \mathcal{E}(T(t), \sigma(t); l_{E})
\]

(31)

where

\[
\xi_{j} = F_{\xi}(T_{j}, \sigma_{j}; T_{j-1}, \sigma_{j-1}, l_{j-1}), \ldots, \xi_{k} = F_{\xi}(T_{k}, \sigma_{k}; T_{k-1}, \sigma_{k-1}, l_{k-1}), \ldots, \xi_{l} = F_{\xi}(T_{l}, \sigma_{l}; T_{l-1}, \sigma_{l-1}, l_{l-1})
\]

(32)

Note that if a switching point \(S_{k}^{i}\) \((0 < k < J)\) is an exit point, the value of \(\xi_{k}\) is known \((\xi_{k} = 0, 1\) for \(i = A, M\), respectively) and memory of the previous history, \(j < k\), is erased. Therefore,

(1) If \(\mathcal{L}\) does not have exit points then phase fraction at the end point of \(\mathcal{L}\) remembers the gross behavior (trace \(S\)) of the entire loading path and the initial phase fraction, \(\xi_{0}\).

(2) If \(\mathcal{L}\) has exit points then the phase fraction, \(\xi_{E}\), remembers only the gross behavior (trace \(S_{E}\)) of the loading path to the nearest exit point, \(S_{E}\); in the exit points the memory is erased.

Thus, the formulated kinetic law has discrete memory so that only a finite amount of information about the loading path is needed.

Note that in this paper we restricted memory of the \(F^{A}\)-local kinetic law, \(F^{A}\) to the set of memory parameters \((T_{0}, \sigma_{0}, l_{0})\) at the base point of \(\mathcal{L}\) only, however, if one needs to describe hysteresis behavior such that the internal loops are closed (as for some copper based alloys) one must expand memory of the \(F^{A}, F^{M}\) branches and include memory parameters at previous switching points.

In summary, the solution of the basic problem (determining \(\xi\) along a loading path \(\mathcal{L}\)) can be viewed from two perspectives. First, the solution \(\xi\) can be propagated along the loading path by using rules (I-IV) at each step to identify switching points and then applying the local kinetic law (equation (18)). Alternatively, if the trace (and thus signature) of the loading path is known \(a\ priori\), the transformation path/history can be written im-

†With the exception of stop points and possibly other switching points, dependent upon choice of transformation function. For example, the linear transformation function (equation (3.7)) leads to \(\Phi\) which is not differentiable at any switching point.
mediately as $\chi = \{ \Phi(T(t), \sigma(t); \xi_0) | t \in P_{(T,s) \in \mathbb{R}} \}$. In either solution technique, experimental data for a given material is used to determine the best fitting transformation function and most appropriate kinetic law algorithm.

5. NUMERICAL EXAMPLES

In this Section we present results of calculations for characteristic examples of SMA hysteresis behavior and their qualitative comparison with experimental data. In each case, a loading path is chosen in advance and the developed kinetic law is employed to determine the transformation path. The global kinetic law procedure can be conveniently implemented in a computer program and is used for the calculations. Several popular SMA transformation functions (exponential, linear and cosine) are used with the three different local kinetic law algorithms ($Z$, $Y$, $YF$) and results are compared.

The kinetic law can be coupled with a phenomenological, rate independent SMA constitutive law $[2, 3, 18]$

$$\varepsilon = \sigma/E + \varepsilon_{\text{e}}, \xi$$

(33)

Results can then be presented for both phase fraction, $\xi$, evolution and stress–strain ($\sigma$–$\varepsilon$) response. In equation (33), the phase fraction $\xi$ is the fraction of detwinned martensite, as that is the component related to macroscopic strain in the SMA material. Consequently, for the stress–strain results in this section we consider the transformation between austenite and detwinned martensite. In the calculations, the thermal strain has been neglected.

The phase fraction, $\xi$, plays the role of an internal variable in equation (33). However, if we substitute in equation (33) the expressions (31)–(32), it becomes clear that the actual internal variables of the model are the set of memory parameters, $(\sigma_{fr}, T_{fr}, \xi_0)$, or $(\mathcal{F}, \xi_0)$. The challenge of developing thermodynamics formalism for the constitutive model (33) and kinetic law (31) and (32) with these discrete internal variables remains open.

First we present a simple case of thermoinduced transformation under zero stress for which a qualitative comparison to experiments is possible, (see Figs 6 and 7). The major hysteresis loop is shown and also several simple inner paths are depicted for heating/cooling from various initial martensite fractions. Figure 7 is experimental data from Paskal and Monasevich [13]; Fig. 6 contains our kinetic law predictions with cosine transformation function employed and start/finish transformation temperatures determined from the data in Fig. 7. The results demonstrate an excellent qualitative agreement with the experimental data. Note that the temperatures at which the transformations begin are the same for the outer loop and all the inner paths in Fig. 7, indicating that for the Nitinol material a phase diagram with dead zones, dead directions and fixed transformation strips is appropriate.

For the remaining examples, the material properties are for a Nitinol SMA, with experimental data for the phase diagram taken from the literature [9, 19] listed in Table 1. The loading paths for the examples I, II, …, V are shown on Fig. 8, with the switching points indicated.
In Example I, the temperature is cycled under constant stress, $\sigma_1 = 200$ MPa, between two fixed points: C in [A], and D in the [M] strip, as shown on Fig. 8. The $j^{th}$-loading cycle can be represented by the four switching points, $S_j^{MS}$, $S_{j+1}^{AN}$, $S_{j+2}^{AS}$, $S_{j+3}^{MN}$. Results are depicted in Figs 9(a)–(d) where the transformation regions and switching points are indicated. For the first few cycles, the inner loops drift upward, but within several cycles converge to an attractor limit cycle. Such behavior has been found previously characteristic for Nitinol [6, 12]. In Figs 9(a) and (b) the cosine function is used and the difference is only in the location of the point D in the strip [A]; in Fig. 9(a), D is the midpoint of [A]; in Fig. 9(b), D is close to the start boundary, [A]. Note that as the point D nears the start boundary, [A, the drift of the internal loop upward is quite pronounced with cycling. Conversely, as the D nears [A, the drift is very small and the attractor loop is attained almost immediately.

Figures 9(a, c and d) demonstrate the effect of the different transformation functions. In each of these cases, the temperature is cycled between the midpoints of the strips and the transformation functions used are the cosine [equation (8)], linear [equation (7)], and exponential [equation (13)] respectively. The results using the cosine and linear functions exhibit similar behavior, while the exponential function case shows very little drift and the attractor loop is very near to the major hysteresis loop. This feature is due to the initial rapid change of the exponential function such that half-way through [A], the martensite phase fraction has already decreased to 0.1 (where in the cosine or linear cases the magnitude is 0.25). Since the loading path has no restart points in these four cases of Example I, the distances $Z critical temperature

In Example II, the SMA body is subjected to isothermal ($T_H = 55^\circ C$) stress cycling with gradually diminishing range of stress, shown on Fig. 8 and the inset of Fig. 10(a). Again the $j^{th}$-loading cycle can be represented by the four switching points $S_j^{MS}$, $S_{j+1}^{AN}$, $S_{j+2}^{AS}$, $S_{j+3}^{MN}$, ($j = 0, 1, 2, \ldots$). As $j \to \infty$ the stress range is decreasing such that $\sigma_1 \to \sigma_{j+1} \to \sigma_{j+2} \to \sigma_{j+3} \to \sigma_\infty \to \sigma_0 \to T_H$. The attractor is a collapsed limit cycle—the straight line segment, $[\sigma_1 (T_H), \sigma_\infty (T_H)]$. Figure 10(a) shows the stress–strain response using the cosine transformation function. Figure 10(b) shows the martensite fraction hysteresis corresponding to Fig. 10(a).

A similar calculation is presented in Fig. 11(a) for thermal cycling with a decreasing temperature range under zero-stress, which also produces the collapsed limit cycle. This example shows excellent qualitative agreement with experimental results shown in Fig. 11(b) [13]. The close correspondence between the experimental data and the kinetic law prediction in this example is one more confirmation that the dead directions and dead zones are physically realistic for NiTi materials.

The isostressed loading path for Example III is shown in Fig. 8, where a number, $N$, of SR-points are introduced equi-spaced along the loading path. The results for different transformation functions using the Z-algorithm are shown in Figs 12(a)–(c). Such a loading path is of interest since use of the SMA kinetic law in numerical computation schemes could inadvertently create SR-points via an incremental procedure. At the SR-points the memory parameters are reloaded and the Z and Y-distances are recalculated using the new base point. The introduction of SR-points on the loading path will

<table>
<thead>
<tr>
<th>$A$ slope, $k^A$</th>
<th>13.8 MPa/°C</th>
<th>$M$-slope, $k^M$</th>
<th>8 MPa/°C</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$-start critical temperature, $T^M_{0}$</td>
<td>34.5 °C</td>
<td>$M$-start critical temperature, $T^M_{0}$</td>
<td>5.9 °C</td>
</tr>
<tr>
<td>$A$-finish critical temperature, $T^M_{1}$</td>
<td>49 °C</td>
<td>$M$-finish critical temperature, $T^M_{1}$</td>
<td>−2.85 °C</td>
</tr>
<tr>
<td>Young’s modulus, $E$</td>
<td>5x10^4 MPa</td>
<td>Maximum transformation strain, $\epsilon_M$</td>
<td>0.067</td>
</tr>
</tbody>
</table>

Table 1. Material parameters used for calculations [9, 19]
not change the results in the case of the $YF$-algorithm, however introduction of $SR$-points make a dramatic difference in the results for the $Z$-algorithm. Although not shown here, use of the $Y$-algorithm leads to similar for the $Z$-algorithm results with increasing numbers of $SR$-points.

As seen in Fig. 12(a), the effect of $SR$-points with the cosine function is dramatic, causing the transformation path to lie outside of the major hysteresis loop. As the number of $SR$-points increases, the phase fraction change slows and is delayed until the neighborhood of the finish boundary, $M$, eventually attaining a sharp corner instead of the smooth cosine shape. Physically, this result can be interpreted that stopping then restarting the transformation has a retarding effect on the process.

Fig. 8. Loading paths of examples illustrated on phase diagram.

Fig. 9. Example I: Attractor loops formed by cycling between [M] and [A] strips with loading path I from Fig. 8. Except for (b), $Z_c = Y_c = 1/2$ in [M] and $Z_D = Y_D = 1/2$ in [A]. Transformation regions highlighted. (a) cosine transformation function; (b) $Z_c = Y_c = 1/2$ in [M] and $Z_D = Y_D = 1/7$ in [A], cosine transformation function; (c) linear transformation function; (d) exponential transformation function.
until the neighborhood of $M$ where the system suddenly jumps to the martensite phase. For moderate numbers of $SR$-points, the transformation paths have a staircase character which is reminiscent of the experimental results [30, 31] obtained for the motion of a single interface in monocrystals. As shown in the inset of Fig. 12(a), it was found that if a specimen was partly transformed, then held at constant temperature for some time, the interface became stabilized, requiring overheating/undercooling in order to move the interface further.

Figure 12(b) shows results for the exponential function, in which the opposite trend is seen: introduction of hold points accelerates the transformation process, causing the transformation path to lie inside of the major hysteresis loop. For the exponential case the result is not nearly as dramatic: the limit curve ($N = \infty$) lies close to the major loop.

Figure 12(c) depicts results for the algebraic transformation functions. In the linear case ($\gamma = 1$), paths for all values of $N$ lie along the major hyste-
esis loop, while the case with $\gamma = 2$ exhibits results similar to those for the cosine function; the transformation path for the $\gamma = 1/2$ case lies inside the major hysteresis loop as in the exponential case, but in the limit $N = \infty$ again approaches an instantaneous transformation (here at the start boundary, [M]). Note that the tendency of SR points to retard or accelerate transformation depends upon the behavior of the transformation function near the start boundary of the strip ($Z = Y = 0$): slowly increasing functions (such as $Z^2$ and cosine) retard transformation as $N$ increases, while rapidly increasing functions (such as $Z^{1/2}$ and exponential) accelerate transformation.

In Example IV, the isostressed loading path begins at the start boundary of the transformation strip [M] with $\xi_0 = 0$. Then the temperature is cycled between two fixed points C, D that are stop and restart points, $S_j^M$ and $S_j^{M*}$ ($j = 1, 2, \ldots$), with $\rho_j^{M*} = \xi\rho_0^{M}$ and $\rho_j^M = \xi\rho_0^M$ as shown in Fig. 8. The results for the $Z$, $Y$, $YF$-algorithms are shown for the cosine function in Fig. 13. The transformation path in each case has oscillatory character and the phase fraction is seen to accumulate and eventually reach the saturation level $\xi = 1$ at the major loop. Such an accumulating or ratcheting phenomenon (so-called lack of vibro-stability) is well-known in the hysteresis literature [6, 14] and is inherent to Duham–Madelung models, but has not yet been addressed experimentally for SMA materials.

The loading path for the final Example V is a closed rectangular loop on the phase diagram, as illustrated in Fig. 8. The two vertices of the rectangular transformation path in the transformation strips have been taken as SR-points for the results depicted in Fig. 14. Again drift to an attractor loop in stress–strain–temperature space is clear.
Fig. 12. Increasing numbers of SR-points placed in [M] strip, for Example III. (a) Cosine transformation function used with $Z$-algorithm. Inset Fig. 11 from Ref. [44]. Left: (b) Exponential transformation function used with $Z$-algorithm. Right: (c) Algebraic transformation functions in equation (7) used with $Z$-algorithm
Calculations without the SR-points at the vertices of the rectangular loop were also performed, but the results were not remarkably different from those shown.

6. CONCLUSIONS

In this paper, we have developed a consistent mathematical framework for the phase diagram based macroscale kinetics which can be used to calculate phase fraction evolution in a one-dimensional SMA body due to thermomechanical loading. Typically such calculations have been performed in an ad hoc manner, with no clearly defined procedure for complex loading paths. While ad hoc calculations are often sufficient for very simple stress or temperature monotone loadings of SMA, the results here emphatically illustrate that

![Figure 13](image-url)

Fig. 13. Results for loading path IV from Fig. 8; here $\varphi_D = \frac{1}{2} \varphi_0$ and $\varphi_C = \frac{1}{2} \varphi_0$. Cosine transformation function used in $Z$, $Y$ and $YF$ algorithms. At the stop point $Z_C = 0.4375$ while $Y_C = 0.5$; at the restart point $Z_D = 0$ while $Y_D = 0.1111$.

![Figure 14](image-url)

Fig. 14. Loading path V of Fig. 8 with SR points at two vertices. Results calculated using cosine transformation function and $Z$-algorithm. Convergence to attractor loop in stress–temperature–strain space is shown.
for more complex loading, certain important issues regarding the description of transformation process (loading path) must be addressed.

The rigorous kinetic law procedure was developed based on a few clearly defined and physically reasonable assumptions concerning the transformation process. The procedure is easy to apply in numerical calculations; it involves primarily tracking and classifying the switching points of a loading path on the phase diagram and using an appropriate local kinetic law algorithm with an associated transformation function. To tune the procedure for a given material, experiments are needed to identify the most appropriate transformation function and local kinetic law algorithm, since, as shown, in some cases the differences between their predictions are striking.

Results of calculations have been presented for several characteristic loading paths using a variety of transformation functions suggested in the literature. It was shown that even simple isothermal (iso-stressed) loading paths justify employment of the developed framework, since dramatically different transformation paths can result from relatively minor changes in the description of a loading path. It was shown that introduction of a number of hold (SR-) points on a simple loading path across a transformation strip can cause the phase fraction change to be accelerated or retarded, depending on the choice of local kinetic law algorithm and transformation function. These results imply that care must be taken in implementation of SMA kinetics in incremental computational schemes. Based on experimental results [22, 30] we believe that staircase transformation paths obtained that lie outside of major loop are physically realistic at least for some SMA materials. More experiments are required to clarify this issue.

For cyclic loading paths between the transformation strips it was shown that drift of the phase fraction internal loops is predicted by all algorithms and transformation functions, a phenomenon which has not been well addressed yet experimentally for SMA materials. Limited experimental data on the thermoinduced transformation agrees very well qualitatively with the results obtained.

For cyclic loading paths inside the transformation strips, it was shown that accumulation of the phase fraction is universal for all algorithms and transformation functions; for SMAs this type of behavior also has not yet been investigated experimentally. Due to the popularity of the exponential transformation function, we should mention that it produces the most dissimilar response compared to other functions (when used for the same loading paths with identical algorithms) and does not fit available experimental data well.

Although the kinetic law procedure presented here is for a trained SMA material in which a stable outer loop has already been achieved and the start transformation boundaries do not depend on the loading history, such a change of the boundaries with loading could be easily accommodated into the model by introducing the dependence of $T_{f0}$ on $\chi$.

Acknowledgements—The authors would like to recognize the National Science Foundation for support of this work.

REFERENCES