MATERIAL LAW FOR NiTi SHAPE MEMORY ALLOYS BASED ON DUAL KRIGING INTERPOLATION

O. Volkov, F. Trochu[*], V. Brailovski, N. Sacépé and P. Terriault

Ecole Polytechnique de Montréal, Mechanical Engineering Department
P.O. Box 6079, Montreal, Quebec, H3C 3A7, Canada

ABSTRACT

A large number of industrial applications could benefit from the remarkable properties of shape memory alloys. The development of a general material law is the first important step before reliable design calculations of shape memory devices can be carried out. This paper presents a new phenomenological constitutive law based on dual kriging, which is a powerful mathematical tool used here as interpolation method to simulate the macroscopic behavior of shape memory alloys. On a set of experimental strain-temperature curves given at constant loads, two surfaces are constructed in the stress, strain and temperature space which describe the cooling and heating behavior of the material at any stress. The response of a specimen subjected to complex thermomechanical loading can be calculated by dual kriging as a general three-dimensional parametric solid anywhere inside the hysteretic domain delimited by the main cooling and heating surfaces. This approach presents the advantage of yielding immediately the explicit equation of any partial cycle inside the main hysteretic thermal cycle, thus allowing a general material law for shape memory alloys to be constructed. Preliminary validations for a set of simple examples demonstrate the potential of this new approach.

1. INTRODUCTION

The mechanical behavior of *Shape Memory Alloys* (SMA) deviates widely from that of usual metals. The thermomechanical state of the material can be austenitic, martensitic or mixed depending on the values of *stress* σ and *temperature T*. The observed responses of strain in function of stress $\varepsilon(\sigma)$ and strain in function of temperature $\varepsilon(T)$ cover various types of behavior which may be divided into three main groups: superelasticity, pseudo-plasticity and shape memory effect (SME).

---

[*] To whom correspondence should be addressed

Each of these remarkable properties of SMA may be characterized by a specific thermomechanical test. The stress-strain curve at constant temperature illustrates well superelasticity and pseudo-plasticity, whereas the strain-temperature diagram obtained in stress-free state or under constant load usually describes the one-way or assisted two-way shape memory effects. Therefore, three important physical parameters interact in the description of SMA: stress, strain and temperature. These variables define a general three-dimensional physical space, in which the complex mechanical behavior of the material could be described. In this article, a new phenomenological model capable of describing in a single formulation superelasticity and shape memory effect is presented. This study breaks up into two parts.

Firstly, we describe how the experimental data related to the behavior of alloys are obtained. For this purpose, a special test bench was constructed to measure strain recovery with temperature. This experimental set-up makes it possible to analyze the assisted two-way shape memory effect by heating of a specimen subjected to constant load. Two types of experimental cycles are studied: (a) a *main thermal cycle* consisting of heating and cooling of the specimen through the zone of phase transformation, so that the material passes from full martensitic state to pure austenite, and conversely: (b) a *partial thermal cycle* consisting of incomplete phase transformations when two phases coexist in the material. The partial cycle belongs to the space delimited by the hysteresis of the main cycle.

Secondly, the experimental data supplied by thermal cycles are interpolated using a new technique called *dual kriging* /1/. In fact, three-dimensional volume parametric kriging /2/ is used here instead of surface parametric kriging /3/ to simulate the behavior of SMA specimens under various thermo-mechanical loading.

The existing phenomenological models describing martensitic transformation can be regrouped in two families:

1. *Models that use the martensite volume fraction as a sole transition parameter that controls the phase transformation* – For example, Amengal *et al.* /4/ use a differential equation based on the fact that the slopes of all higher order partial paths are supposed to be the same as that of first order partial cycles. It is proved that this model describes correctly first order partial cycles for fixed values of stress.

2. *Models based on parametrical description of SMA hysteresis* – For example, Ortin /5,6/ proposed using the Preisach hysteretic model. The strain is given as a function of two stress history parameters necessary for attaining the actual state on a partial cycle curve, where the function is constructed from a set of experimental data. This approach generates stress-strain trajectories for any given temperature, but the precision of the approximation is proved only for first order partial cycles.

Another model based on two-dimensional surface kriging was proposed by Terriault *et al.* /3/. This phenomenological constitute law simulates the material behavior of the principal hysteresis cycles and embedded partial cycles of any order in the three-dimensional stress-strain-temperature space. In the present work, a similar technique is used to create the material model, but it is more general in the

sense that thermomechanical response of the material for any stress, strain or temperature increment is calculated by three-dimensional volume kriging instead of two-dimensional parametric kriging. This approach leads to an easier determination of embedded partial cycles in comparison with the Terriault *et al.* /3/ model, where a new set of loading or unloading subsurfaces must be constructed for each level of partial cycles.

Finally, note that the present article fulfils the role of extension for the work of Trochu *et al.* /8/ in the sense that it gives more detailed explanation of three-dimensional parametric kriging as well as validation of the method on a simple set of data.

2. EXPERIMENTAL ANALYSIS

This section presents the experimental set-up and the procedure followed to obtain typical strain recovery curves in function of temperature.

2.1. Presentation of the testing bench

The testing bench represented in Figure 1 is designed to subject a SMA wire or ribbon to heating and cooling cycles under constant load. The specimen is heated by an electrical current of given
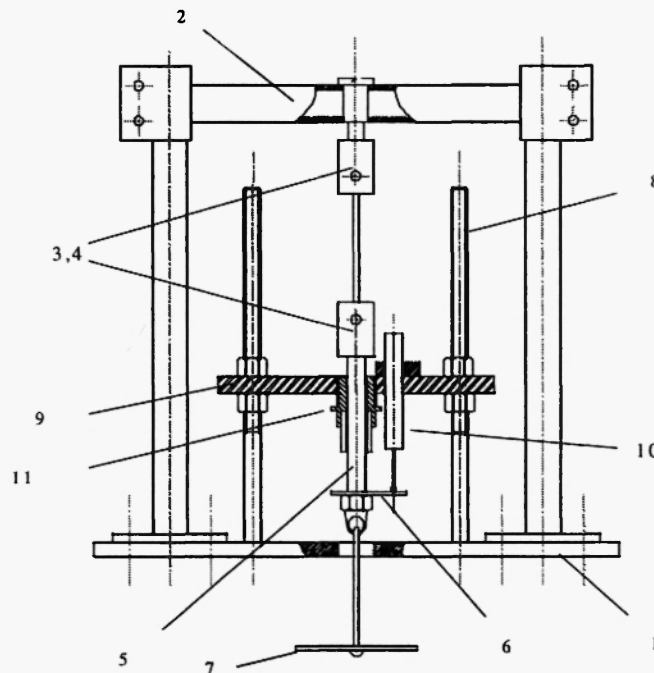

Fig. 1: Assisted two-way SME test bench: (1) base plate; (2) frame; (3,4) grips; (5) guiding rod; (6) thrust block; (7) load bed; (8) column; (9) holding plate; (10) LVDT; (11) cylindrical sliding bearing.

intensity (Joule effect), while the sample cooling is done by convection. A linear variable differential transformer (LVDT) records the length variations of the specimen, and the temperature is measured by a thermocouple fixed directly on the sample. The experimental data are recorded and treated by a specific acquisition system. The testing bench also allows the minimum and maximum temperatures of the specimen to be controlled, so that it is possible to achieve partial transformation.

2.2. Experimental procedure

The testing bench is used to measure during heating and cooling under constant load the elongation of the sample as a function of temperature in case of complete as well as partial assisted two-way shape memory cycle. The procedure is repeated for several values of constant load. At the beginning of each partial cycle, the amount of martensite in the material must be identical for each case of loading, because these partial cycles will be used to generate, by kriging interpolation, heating and cooling surfaces in stress-strain-temperature space. The NiTi specimens used in this analysis are stabilized before test, i.e., the specimens have undergone several *thermal* cycles in order to stabilize their mechanical properties and phase transformation temperatures.

For each case of loading, the experimental procedure may be described as follows:

1) *Main heating and cooling cycle* – As shown in Figure 2, the sample is loaded in martensitic state, then heated to a temperature $T > A_F$ (1-2-3), and finally cooled until it comes back to its martensitic state (3-4-1). The main thermal cycle obtained in this way breaks up into two parts: the higher part of the diagram corresponds to heating, and the lower part to cooling.
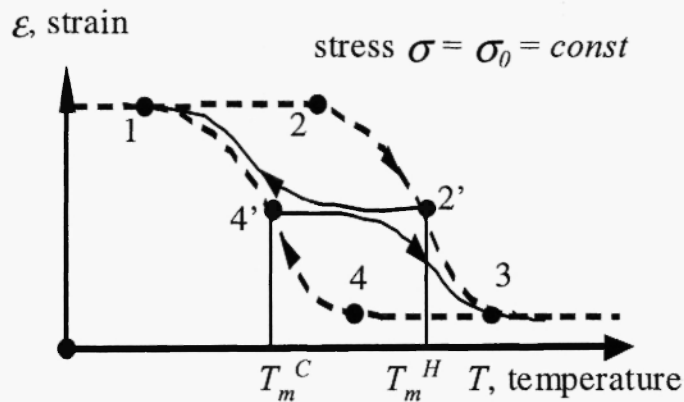

Fig. 2: Main and partial cooling and heating cycles.

2) *Identification of the average temperature* – The average temperature T_m corresponds to the medium point of the deformation either in heating (T_m^H) or in cooling (T_m^C).

3) *Partial heating or cooling cycles* – The sample is loaded and heated up to a temperature $T > A_F$ (1-2-3), then cooled down to temperature T_m^C (3-4-4') as illustrated in Figure 2. When this state is

reached, the alloy is heated again until $T > A_F$ (4'-3). Finally, it is cooled until it reaches its final martensitic state (see in Figure 3 lower subcycles inside each main hysteresis curve). Partial cooling cycles have also been represented in Figure 3 (upper subcycles in the main hysteresis curves).
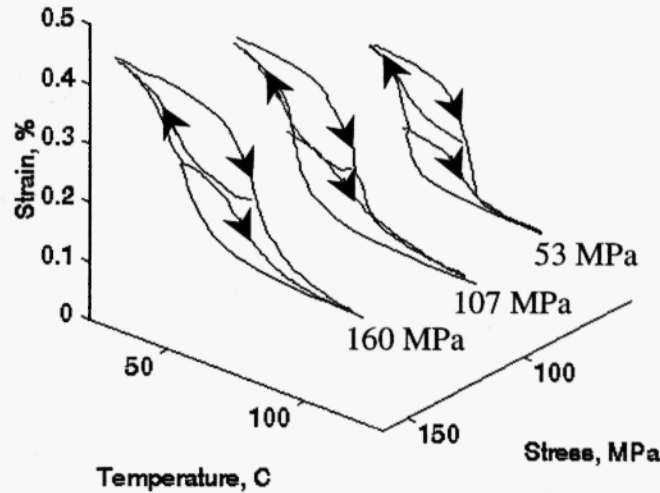

Fig. 3: Main and partial thermal cycles obtained experimentally for three cases of loading: 53, 107 and 160 MPa. (Partial cycles correspond here to about 50% of martensite content.)

3. PARAMETRIC MODELING USING THREE-DIMENSIONAL KRIGING

This section presents the shape memory model based on dual kriging /1/, a general interpolation method which contains spline and Bezier curves or surfaces as particular cases and least square methods as a limit case. The thermomechanical behavior of stabilized SMA samples is totally defined in a three-dimensional space of strain, stress and temperature. The response of the material is described by a series of curves and surfaces. At each point M inside the hysteresis volume defined by the main thermal cycle, only one downward surface corresponding to *heating* and only one ascending surface corresponding to *cooling* may exist.

In order to define a thermomechanical heating or cooling trajectory, it is necessary to introduce supplementary curvilinear coordinates in the model. This will be done using parametric dual kriging.

3.1. Parameterization

The basic requirement for natural parameters to be used in the dual kriging interpolation is the necessity to confine them to an explicit physical meaning. The parameters introduced here are the following: as shown in Figure 4, ε_H is the normalized strain on heating, ε_C the normalized strain on cooling, and σ denotes the normalized stress. The minimum and maximum values of each variable are
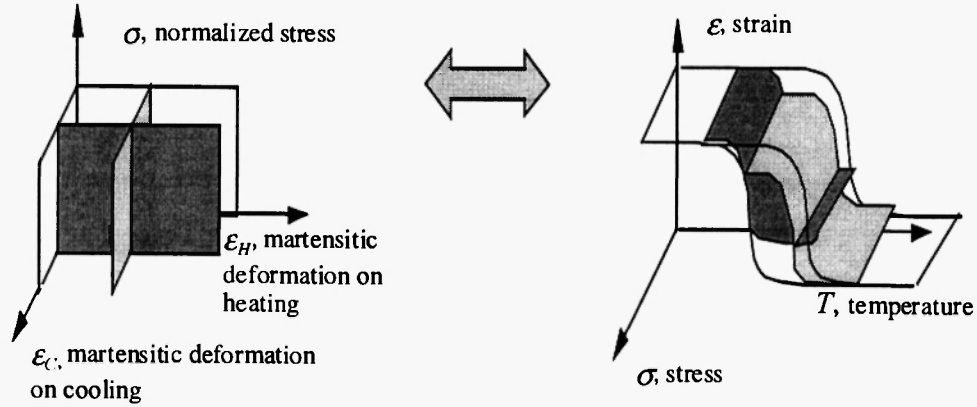
Fig. 4: A plane in the parameter space $(\varepsilon_H, \varepsilon_C, \sigma)$ associated with a single surface in the space (ε, T, σ).

normalized here between 0 and 1. By this parameterization of SMA behavior, a volume is generated between two surfaces corresponding respectively to heating and cooling. The variation of parameters $(\varepsilon_H, \varepsilon_C, \sigma)$ is represented in a unit cube. Indeed, by fixing one of the coordinates at a constant value, while varying the two others in a unit interval, a single plane perpendicular to one parametric axis is obtained in the parameter space.

Two sets of parameterized surfaces are constructed: the first set represents the variation of the parameter ε_H, and the corresponding surfaces in the space $(\varepsilon_H, \varepsilon_C, \sigma)$ describe partial *heating* cycles in the stress-strain-temperature space (see Figure 5). The second set represents the variation of ε_C, and the corresponding surfaces in the space $(\varepsilon_H, \varepsilon_C, \sigma)$ describe partial *cooling* cycles. As illustrated in Figure 4, these surfaces can be represented as the two sets of planes parallel to $(0, \varepsilon, \sigma)$ and $(\varepsilon_H, 0, \sigma)$ in the unit cube $(\varepsilon_H, \varepsilon_C, \sigma)$. The advantage of this representation is that we can control any path in the stress-strain-temperature space including any type of partial cycle by knowing only three coordinates $(\varepsilon_H, \varepsilon_C, \sigma)$.

4.2. Kriging results

By using three-dimensional volume parametric kriging /1/, the hysteretic heating and cooling volumes are completely described by following parametric equations (for the sake of completeness, a more detailed description of the method is outlined in the Appendix):

$$\varepsilon_{\text{heating}} = \varepsilon_{\text{heating}}(\varepsilon_H, \varepsilon_C, \sigma)$$
$$T'_{\text{heating}} = T'_{\text{heating}}(\varepsilon_H, \varepsilon_C, \sigma) \tag{1}$$

and

$$\varepsilon_{\text{cooling}} = \varepsilon_{\text{cooling}}(\varepsilon_H, \varepsilon_C, \sigma)$$
$$T_{\text{cooling}} = T'_{\text{cooling}}(\varepsilon_H, \varepsilon_C, \sigma) \tag{2}$$

4.3. Verification

The first experimental set of thermal cycles shown in Figure 3 was treated numerically to define the heating volumes. Figure 5 shows the results of calculated heating thermal cycles.
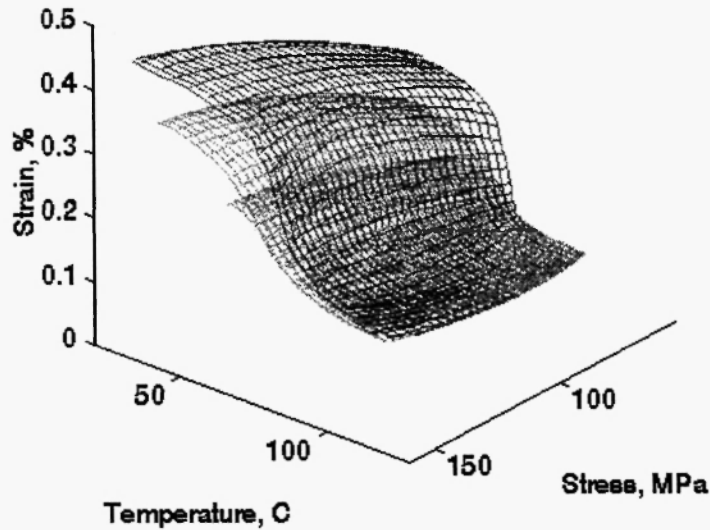

Fig. 5: Heating thermal cycles corresponding to an initial content of 40, 70 and 100% of martensite: the hysteretic volume is created using the three sampling sets of Figure 3 obtained for loads of 53, 107 and 160 MPa.

Figure 6 shows the comparison between calculated trajectories and the experience. We verify here that the approximated curves match the partial heating cycle accurately; this demonstrates the stability of the model with respect to the initial data used to generate the kriged volume. The next step will be to validate this approach by comparison with a stress-strain-experimental curve at a given temperature.
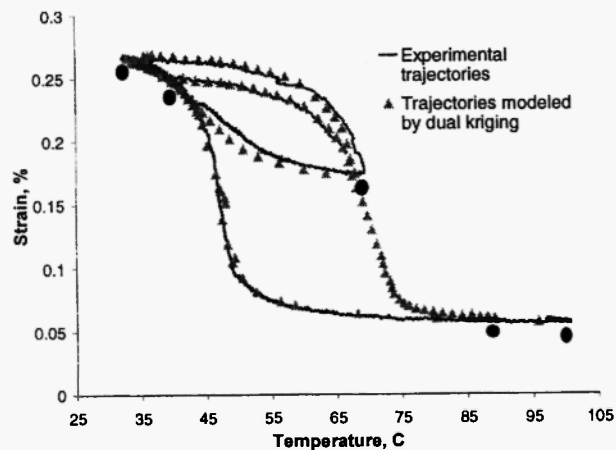

Fig. 6: Experimental and calculated trajectories of partial thermal cycle obtained for a load of 53 MPa.

4.4. Experimental validation of the main cycle interpolation

Validation consists in reconstructing strain-stress isothermal behavior curves from strain-temperature isostress ones. Initial data both for strain-temperature and strain-stress curves are taken from /7/. The kriged volume is created from two strain-temperature curves presented in Figure 7. The default stress interval is between 128 and 174 MPa, but it can be extended within reasonable limits.

Parametrical representation of the hysteretic volume obtained in this way makes it possible to determine the isothermal behavior in the main cycle corresponding to T = 18°C and T = -34°C. The comparison between calculated and experimental isothermal curves is presented in Figure 8.
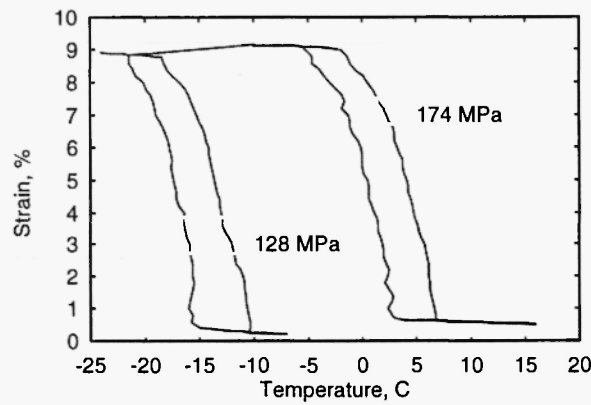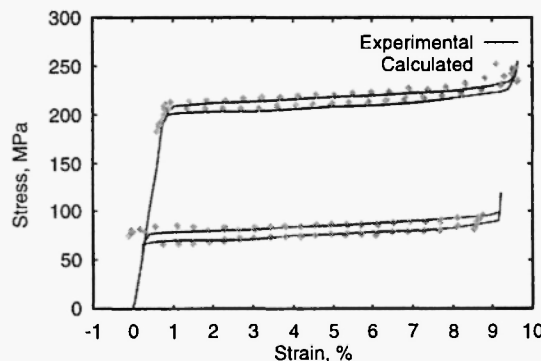
Fig. 7: Experimental isostress curves /7/.

Fig. 8: Comparison between calculated and experimental isothermal curves /7/.

5. CONCLUSION

The model presented in this article shows the benefits of three-dimensional kriging to simulate the mechanical behavior of SMA. The most interesting feature of this model is the description of the main

and partial cycles by a single analytical formulation. Initial data are provided experimentally in order to obtain a material law defined in a stress-strain-temperature space that enables simulation of the complex thermo-mechanical path of loading.

Several ideas could be tested to improve this methodology, from both the experimental and numerical points of view:

1) The test bench could be improved in order to provide more reliable experimental results, in particular concerning partial cycles data. Indeed, partial cycles are constructed as a function of the average strain of the specimen. Thus, in order to control more precisely the release of partial cycles, a displacement control seems to be more appropriate than the currently programmed temperature control.

2) The precision of the method strongly depends on the degree of specimen stabilization. Even the samples that are definitely settled in the main cycle can change their behavior in partial cycles. Moreover, any partial cycle modifies the profile of the following ones.

3) Further validation should be performed for new thermal cycles (obtained for a stress different from the ones used to construct the model) and for isothermal strain-stress experimental curves.

REFERENCES

1. F. Trochu, *Engineering with Computers*, 9, 160-177 (1993).

2. A. Limaien, Computer-Aided CMM Inspection Planning and Verification, Ph.D. Thesis, McMaster University, September 1996.

3. P. Terriault, M.-A. Meunier and F. Trochu, *Journal of Intelligent Material Systems and Structures*, 8(7), 553-630 (1997).

4. A. Amengual, A.A. Likhachev and E. Cesari, *Scripta Materialia*, 34(10), 1549-1554 [[year]].

5. J. Ortin, *Journal de Physique*, **IV**, 4, November (1991).

6. J. Ortin and A. Planes, *Journal de Physique*, **IV**, 1, November (1991).

7. E. Patoor, E. Eberhardt and M. Berveiller, *Acta Metall.*, 35(11), 2779-2789 (1987).

8. F. Trochu, N. Sacépé, O. Volkov and S. Turenne, accepted in: *Proceedings of the International Conference of Martensitic Transformations ICOMAT'98*, Bariloche, Argentina, 1998.

APPENDIX: PARAMETRICAL KRIGING

In this appendix, the principle of parametrical kriging that is a particular case of the more general theory of dual kriging is given. Briefly speaking, dual kriging consists of approximating a set of discrete data by a continuous analytical function. As shown in /1/, dual kriging is, in fact, a generalization of well-known approximation methods like spline or Bezier interpolation.

As a preliminary, we briefly review certain notions from the analysis of functions used in the sequel. An *isomorphism*, or *isomorphic* mapping, from A onto A' is a one-to-one mapping from the set A onto the set A' which determines an equivalence relation splitting any set on which it is defined into disjoint equivalence classes – the classes of pairwise isomorphic systems (cf. *Encyclopaedia of Mathematics*, Kluwer Academic Publishers, 1995). For example, the function that associates the hysteretic volume in the strain-stress-temperature space with a three-dimensional unitary cube in the parametrical space is an isomorphism, since it should take into account pairwise isomorphic systems of points, lines and surfaces representing partial hysteresis curves.

Consider two sets of points in the n-dimensional Euclidean space: the first one contains parameter values denoted *[x]*, the second function values denoted *[f]*. Kriging interpolation consists of constructing an isomorphic mapping from *[x]* to *[f]*. The only restriction imposed here on the approximating function is that it should be continuous.

The approximating function is divided into two parts: average term $a(x)$ called *drift* and fluctuation term $b(x)$ called *covariance*, so that $f(x)$ can be written as:

$$f(x) = a(x) + b(x) \tag{3}$$

where $a(x)$ and $b(x)$ are interpolated in two subsets of continuous functions $f_{drift}(x)$ and $f_{cover}(x)$, respectively the drift and the covariance, which define together a *profile* for kriging, i.e., the set of parameters that determine the structure of the kriging interpolator. Expression (3) may then be presented in the form:

$$f(x) = \sum_{i=1}^{m} a_i f_{drift}^i(x) + \sum_{i=1}^{n} b_i f_{cov\,ar}(\| x - x_i \|) \tag{4}$$

or in vector notation:

$$f(x) = [f_{cov\,ar}(h), f_{drift}(x)] \cdot \begin{bmatrix} b \\ a \end{bmatrix} \tag{5}$$

with the distance $h = \| x - x_i \|$ where i corresponds to the index of b.

The problem of approximation reduces to the determination of a set of coefficients a_i, b_i, that will be denoted hereafter as vector $[b, a] = (b_1, b_2 \dots b_N, a_0, \dots a_m)$. the required coefficients can be found by solving the following linear system of equations:

$$\begin{bmatrix} [f] \\ 0 \end{bmatrix} = \begin{bmatrix} f(x_1) \\ f(x_2) \\ . \\ f(x_n) \\ 0 \end{bmatrix} = \begin{bmatrix} f_{cov\,ar}(0) & f_{cov\,ar}(h_{21}) & . & f_{cov\,ar}(h_{n1}) & f_{drift}(x_1) \\ f_{cov\,ar}(h_{12}) & f_{cov\,ar}(0) & . & f_{cov\,ar}(h_{n2}) & f_{drift}(x_2) \\ . & . & . & . & . \\ f_{cov\,ar}(h_{1n}) & f_{cov\,ar}(h_{2n}) & . & f_{cov\,ar}(0) & f_{drift}(x_n) \\ f_{drift}(x_1) & f_{drift}(x_2) & . & f_{drift}(x_n) & 0 \end{bmatrix} \cdot \begin{bmatrix} b_1 \\ b_2 \\ . \\ b_n \\ a \end{bmatrix} = [K] \cdot \begin{bmatrix} b \\ a \end{bmatrix} \tag{6}$$
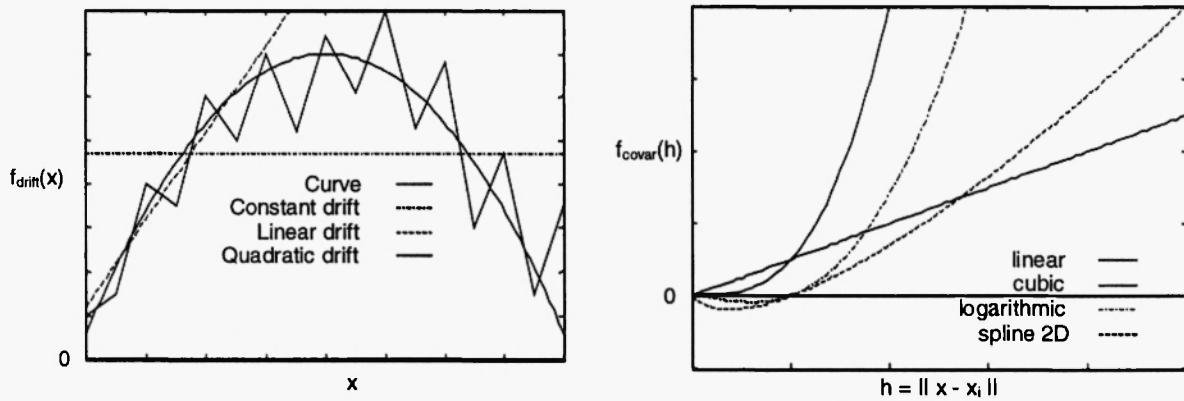
Fig. 9: Drift and covariance functions most frequently used in kriging approximation.

where $h_{IJ} = \| x_I - x_J \|$ represents the distance between two points x_I and x_J. *[K]* defines a *kriging matrix*. Note that the last equation results from the *unbiased estimation condition* (cf. /1/).

Since the matrix *[K]* is assumed to be non-degenerate, one obtains a vector of the unknown multipliers *[b,a]* by multiplying the system (6) by $[K]^{-1}$. Substituting the vector *[b,a]* into equation (5), the final kriged function becomes:

$$f(x) = [f_{covar}(h), f_{drift}(x)] \cdot [K]^{-1} \cdot \begin{bmatrix} [f] \\ 0 \end{bmatrix} \tag{7}$$

Parametric kriging differs from simple kriging by the fact that each component of the vector *[x]* is associated with its proper unique profile *[f_{covar}, f_{drift}]*. This property enables us to define a particular data structure with a vector of direction corresponding to the ith component of vector *[x]*:

$$k_{x_i} = [f_{covar}(h), f_{drift}(x)] \cdot [K]^{-1} \tag{8}$$

This structure is built by knowing only the samples *[x]* and the type of kriging profiles *[f_{covar}, f_{drift}]*. In simple kriging, all k_x are identical, since only one profile for all the variables is considered.

For example, let us consider the case of parametric surface kriging. We denote by *[x,y]* a set of points coordinates and by *[f_x, f_y]*, the set of corresponding function values. Following the procedure explained in the previous paragraph, the two-dimensional set *[x,y]* generates two directions of kriging, k_x, k_y, and the final kriged function takes the form:

$$f(x, y) = k_x \cdot \begin{bmatrix} [f] & 0 \\ 0 & 0 \end{bmatrix} \cdot k_y \tag{9}$$

where *[f]* stands for a quadratic matrix containing *[f_x, f_y]* values.

225

linear profile for y
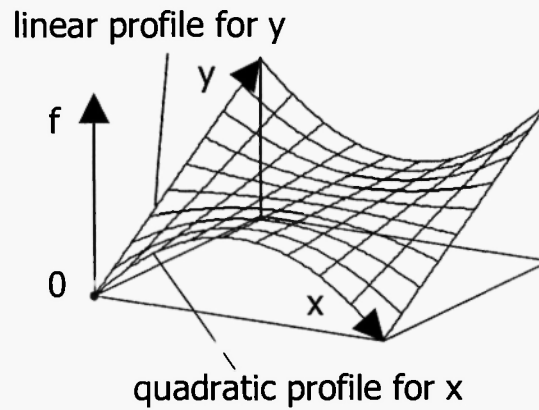

quadratic profile for x

Fig. 10: Curvilinear surface described by a set of parametric co-ordinates (x, y).

A similar procedure can be used to build parametric volume kriging interpolation on three-dimensional data sets $[x,y,z]$ and $[f_x, f_y, f_z]$. In this case, $[f]$ represents a cubic matrix which can be defined along three directions k_x, k_y, k_z. the procedure to obtain the value of a kriged function $f(x,y,z)$ stands as follows: multiplication of the cubic matrix $[f]$ by k_z in order to decrease its dimension by 1. Thus, the final solutions takes the form similar to (9):

$$f(x, y, z) = k_x \cdot \begin{bmatrix} [f]k_z & 0 \\ 0 & 0 \end{bmatrix} \cdot k_y \qquad (10)$$

Similarly, n-dimensional volume kriging may be derived using the algorithm exhibited above.