Numerical methods in phase-change problems

Sergio R. Idelsohn
Mario A. Storti
Luis A. Crivelli
Grupo de Tecnología Mecánica del INTEC
Güemes 3450, 3000 - Santa Fe, Argentina.

Summary

This paper summarizes the state of the art of the numerical solution of phase-change problems. After describing the governing equations, a review of the existing methods is presented. The emphasis is put mainly on fixed domain techniques, but a brief description of the main front-tracking methods is included. A special section is devoted to the Newton-Raphson resolution with quadratic convergence of the non-linear system of equations.

INTRODUCTION

Heat conduction problems with phase change are receiving increasing attention, mainly because of the broad range of technological applications and applied research fields where this phenomenon has a prevailing role, in addition to the interesting mathematical aspects. In a wider context, they receive the name of Stefan problems, which comprehends general moving boundary phenomena. In civil engineering, for instance, the determination of the freezing time of certain ground regions or the degradation of frozen layers, follow this model. Casting of metals and alloys in metallurgy or solidification of crystals are other classic applications. In nuclear engineering, the applications range from theoretical reactor analysis to specific aspects of power plant safety such as the determination of the melting time of combustible rods under several types of accidents. Space-vehicle design requires numerical modeling of ablation of thermal protections under severe thermal loads during the reentry stage in the atmosphere. In the development of non-conventional energy resources, phase change problems occur in the design of heat storage devices based on high latent heat substances. The model applies also in other physical contexts such as phase change with diffusion or fluid flow in porous media.

Due to the inherent non-linearity of the energy balance at the interface governing the interface position, few analytical solutions of interest are available [1–4] giving place to the development of a great number of numerical algorithms [5–7] based on finite differences [8–26] finite elements [27–63] and, more recently, boundary elements [64–67]. All these methods can be classified into two main groups, front tracking methods and fixed domain methods. The main disadvantage of those pertaining to the first group is that they require some degree of regularity of the moving boundary as well as of its evolution. In contrast, fixed domain methods, which are based on weak formulations of the energy equation, can handle complex topological evolutions of the interface, but as a counterpart, the resolution of the temperature profile is somewhat degraded since they represent temperature gradients as a smooth function, even though it is discontinuous at the interface. However, the association of fixed domain techniques with finite elements is quite natural due to their common search for the modeling of complex geometries. Finally, among the fixed domain methods we can mention different strategies such as apparent heat capacity, enthalpy methods, change of the independent variable, fictitious heat flow method or the discontinuous integration method.
GOVERNING EQUATIONS

We will concentrate on the heat conduction equation with phase-change. Let us consider a bounded region $\Omega$, partially occupied by liquid $\Omega_L$ and solid $\Omega_s$, being $\Gamma$ the dividing interface where $T = T_m$ (see Figure 1). Mushy regions are not considered in this simplified model. Other simplifying hypotheses we will assume are continuity of density across the interface and negligible convection effects in the liquid phase. Sometimes, this last effect can be approximately taken into account by increasing the thermal conductivity in the fluid phase.

With these assumptions, the governing equations are

\[ \frac{\partial H}{\partial t} + \nabla \cdot (k \nabla T) + Q = 0, \quad \text{in } \Omega_L, \Omega_S \quad (\text{heat equation}) \]  

\[ T = T_m, \quad \text{at } \Gamma \quad (\text{isothermal p. change}) \]  

\[ \hat{n} \cdot (k \nabla T) = -\hat{n} \cdot (k \nabla T)_L = \rho L \nu, \quad \text{at } \Gamma \quad (\text{interface bal. eq.}) \]  

\[ T = T, \quad \text{at } \partial \Omega_T \quad (\text{Dirichlet b.c.}) \]  

\[ \hat{n} \cdot k \nabla T = \hat{q}, \quad \text{at } \partial \Omega_q \quad (\text{Neumann b.c.}) \]  

\[ T(x, t = 0) = T_0(x), \quad \text{in } \Omega \quad (\text{initial cond.}) \]  

where $\partial \Omega_T (\partial \Omega_q)$ is that portion of $\partial \Omega$ with Dirichlet (Neumann) boundary conditions, $k$ the thermal conductivity, $Q$ an internal heat source, $\rho$ the density, $H$ the specific enthalpy, $t$ the time, $\hat{n}$ the unit vector normal to the surface, $\nu$ the normal advance velocity of the moving boundary, $T$, $q$ the prescribed values of temperature and heat flow, $L$ the latent heat, and the subindices $s$ and $L$ denote the solid and liquid phases respectively.

Enthalpy is related to temperature by

\[ H(T) = \int_{T_{\text{ref}}}^{T} \left[ c(T) + L \delta(T' - T_m) \right] dT' \]

\[ = \int_{T_{\text{ref}}}^{T} c(T) \, dT' + L e(T - T_m) = H^c(T) + H^L(T) \]

where $\delta$ is Dirac's delta, $T_{\text{ref}}$ the reference temperature, $c$ the sensible specific heat, $e$ the Heaviside step function. $H^c, H^L$ denote that part of $H$ coming from the sensible and latent heat, respectively. Of course, the step in the enthalpy/temperature relationship stemming from $H^L$ is responsible for the highly non-linear character of the Stefan problem.
Two situations exist, essentially different, where the problem reduces to only one phase with a moving boundary. One of them is the case of a solid with a temperature infinitesimally below the melting temperature, this is the one phase Stefan problem, and is usually difficult to solve with fixed domain methods. The other case is the ablation of a thermal protection in which a heat load $q$ is added to the moving boundary while the material is removed as it changes phase, in this case, the heat load $q$ must be added to the right hand side of equation (2.2).

Even for pure substances with a definite melting temperature, it could happen that, for a particular geometry with an internal heat source, regions with finite volume at the melting temperature are produced. In these regions, called "mushy regions", both phases coexist and the enthalpy takes a value $H$ somewhere in between the interval $H(T_m^-) = H^c(T_m) < H < H(T_m^+) = H^c(T_m) + L$.

**DISCRETIZATION SCHEMES**

We can see in Table I, the most representative methods classified in the two main groups as was discussed in the introduction.

**Table I. Classification of numerical algorithms**

- **Front tracking methods:**
  - Fixed mesh, finite differences
  - Variable mesh, finite differences
  - Moving mesh, finite elements
  - Moving boundary element

- **Fixed domain methods:**
  - Apparent heat capacity
  - Enthalpy based formulations
  - Apparent heat capacity with post-iteration correction
  - Fictitious heat flow
  - Freezing index
  - Discontinuous integration

**FRONT TRACKING METHODS**

**Fixed mesh finite differences**

Finite differences approximate derivatives by finite quotients on a structured mesh. The moving interface will lie, in general, in between two consecutive nodes and, then, the finite differences formulae can be modified explicitly for these neighboring nodes to take into account the phase change effect. Consider, for simplicity, the one-dimensional case and let $s = x_i + \alpha \Delta x$ the position of the interface (see Figure 2). Then finite difference approximations for second derivatives at the neighboring nodes, and the first derivative from one side of the interface are:

$$
\left( \frac{\partial^2 T}{\partial x^2} \right)_{i} \approx \frac{2}{\Delta x^2} \left[ \frac{T_{i-1} - 2 T_i + T_{i+1}}{\alpha + 1} \right]
$$

(6)

$$
\left( \frac{\partial T}{\partial x} \right)_{i+\alpha} \approx \frac{1}{\Delta x} \left[ \frac{\alpha + 1}{\alpha} T_{i-1} - \frac{\alpha + 1}{\alpha} T_{i+\alpha} \right]
$$

(7)

Replacing the discrete form of the laplacian (6) in (1), and discretizing the temporal derivative, the temperature field can be advanced in time. Next, using equation (7) the jump in heat flux across the interface can be computed and the position of the interface advanced in time. Obviously, a high degree of regularity is needed in the evolution of the interface, not allowing multiple interfaces or mushy zones.
Moving mesh finite differences

This method is based in keeping the interface always on one node, deforming the mesh as it moves. The number of nodes in between the domain boundary and the interface is kept constant. The derivative of the temperature on a moving node becomes

\[
\frac{DT}{Dt} = \mathbf{v} \cdot \nabla T + \frac{\partial T}{\partial t}
\]  

(8)

\(\mathbf{v}\) being the mesh velocity at the node under consideration. The first term in the right hand side of equation (8) is a convective term due to the movement of the mesh. Once the temperature gradients at the interface nodes are obtained, the interface velocity can be calculated and its position updated, followed by a relocation of the nodes according to some prescribed strategy. In a one dimensional problem, the simplest choice is a linear interpolation of both the nodes coordinates and their velocities

\[ v_i = \left( \frac{x_i}{s} \right) v_t \]  

(9)

resulting in an equation of the form

\[
\frac{\partial T}{\partial s} \frac{x_i}{s} + \frac{\partial T}{\partial x_i} = k \frac{\partial^2 T}{\partial x_i^2}
\]  

(10)

where \(s\) and \(x_i\) denote the position of the interface and the \(i\)-th node, respectively.

Both finite differences algorithms are efficient in solving one-dimensional problems, due to the inherent simplicity in the algorithm for mesh adaption. In contrast, the extension to multidimensional problems is considerably more complex. Notwithstanding that extensions to the multi-dimensional case have been proposed by Lazardis [12], their complexity turns them impractical.

Moving finite elements

Finite elements are based on a "weak form" of the governing equations [70]. This is obtained weighting both the governing partial differential equation (1) and the "natural boundary condition" (3.2) with an arbitrary weight function and applying Gauss divergence theorem. The resulting weak form has only first order derivatives on both the weight functions and the temperature field, allowing a low degree of continuity between adjacent elements. The appropriated expression for the interpolation of the temperature field in a transient problem is:

\[
T(x, t) = \sum_{i=1}^{N} N_i(x) u_i(t)
\]  

(11)

where the \(N_i(x)\)'s are "shape functions" with local support and \(u_i, i = 1, \ldots, N\) the \(N\) nodal temperatures. In its Galerkin form, which is appropriate for symmetric equations like in the
Numerical methods in phase-change problems

Stefan problem, the weights and the shape functions are the same, resulting in a discrete first order ODE's system

\[ \mathbf{K}\mathbf{u} + \mathbf{C}\mathbf{\dot{u}} = \mathbf{g}(t) \]  \hspace{1cm} (12)

where \( \mathbf{K} \) is the conductivity matrix, \( \mathbf{C} \) the capacity matrix, \( \mathbf{u} \) the vector of nodal temperatures, \( \mathbf{g} \) the vector of thermal loads and the dot indicates time derivative. In the instance of moving meshes, the shape functions also depend on time, resulting in an interpolation of the form\[36,60-61\]

\[ T(x, t) = \sum_{i=1}^{N} N_i(x, t) u_i(t) \]  \hspace{1cm} (13)

Taking the time derivative of this expression generates a convective term associated to the movement of the grid, as in equation (8). In this way, a convective non-symmetric matrix is added to the discrete system, of the form

\[ (\mathbf{K} + \mathbf{K}^*)\mathbf{u} + \mathbf{C}\mathbf{\dot{u}} = \mathbf{g}(t), \quad \text{with} \quad K_{ij}^* = \int_{\Omega} N_i \rho c (\mathbf{v} \cdot \nabla N_j) \, d\Omega \]  \hspace{1cm} (14)

The next step is to find difference formulas for both the time derivative of the vector of nodal temperatures in (14) and for the advance velocity of the moving interface. Once the first are introduced in (14), the temperature field can be advanced in time. With this temperature field, the discrete version of (2) is used to update the interface position. A variety of strategies can be defined, by choosing the form in which both stages are coupled, for instance, the temperature field can be obtained implicitly followed by an explicit determination of the position of the interface, or both steps can be performed simultaneously in a fully implicit scheme.

However, as the rate of deformation of the mesh increases the stability of the scheme deteriorates\[39\]. A sufficient condition to assure the existence of a unique solution for the system of equations is the positive definite character of the jacobian matrix. This restriction limits the maximum nodal velocity, which is inversely proportional to the minimum element size. Notwithstanding, this kind of algorithm is appropriated for ablation problems, where the advance velocity of the interface can be specified not only by the energy balance equation, but also as a complicated non-local function due to chemical or mechanical effects in the aerodynamic boundary layer.

Moving boundary elements

If the thermal properties are assumed constant and homogeneous (not depending on position) in both phases, then the only source of non-linearity is associated to the moving interface. This allows a “condensation” of the problem in the interiors of the two phases by means of the appropriate Green functions, reducing the unknown fields to their projections on the boundaries, including the moving part. Due to the complexity of the Green function in space-time, O’Neill\[64\] has considered the homogeneous quasi-steady form of equation (1). This assumption is valid in problems with low Stefan number, i.e. in the limit of high latent heat, and in the absence of internal heat sources. The pDE in each phase becomes,

\[ \Delta T = 0, \quad \text{in } \Omega_{L,3} \]  \hspace{1cm} (15)

The appropriated Green functions are

\[ G_i = \begin{cases} \log r_i, & \text{in } 2D \\ 1/r_i, & \text{in } 3D \end{cases} \quad r_i = |\mathbf{r} - \mathbf{r}_i| \]  \hspace{1cm} (16)
where $\mathbf{r}_i$ is the position vector of an interior point relative to point $P_i$ (see Figure 3). The Green function satisfies

$$\Delta G_i = \lambda_i \delta(\mathbf{r} - \mathbf{r}_i)$$  \hspace{1cm} (17)

In each phase, the following boundary integral relationship can be obtained through the use of Green functions

$$\int_{\Omega_{LS}} G_i \Delta T \, d\Omega = \int_{\partial\Omega_{LS}} \left[ T \frac{\partial G_i}{\partial n} - T_i \frac{\partial G_i}{\partial n} \right] \, dS + \int_{\Omega_{LS}} T \Delta G_i \, d\Omega = 0$$ \hspace{1cm} (18)

with

$$\int_{\Omega_{LS}} T \Delta G_i \, d\Omega = \lambda_i T_i$$ \hspace{1cm} (19)

and

$$\lambda_i = \begin{cases} 2\pi & P_i \text{ interior to } \Omega_{LS} \\ \pi & P_i \text{ on a smooth part of } \partial\Omega_{LS} \\ \theta & P_i \text{ on a vertex of } \partial\Omega_{LS} \text{ with included angle } \theta \end{cases}$$ \hspace{1cm} (20)

Figure 3 Boundary elements formulation

Only boundary values enter expression (18) thus eliminating the determination of the unknown variable in the domain interior. The following step, is to discretize temperature and heat flux at the boundary

$$T = \sum_{i=1}^{N} T_i N_i(s) \hspace{1cm} \text{for } s \in \partial\Omega_{LS}$$ \hspace{1cm} (21)

$$\frac{\partial T}{\partial n} = \sum_{i=1}^{N} g_i N_i(s)$$
Replacing in (18) and with the help of (19) the following system is obtained:

\[ \mathbf{AT} = \mathbf{g} \]  

On boundaries where temperature is prescribed we obtain directly the value of the normal heat flux and the converse is true on boundaries with imposed flux. By application of the energy balance equation (2) the velocity of the interface is computed and its position can be updated. With regards to the operation count, as compared with FEM, the number of degrees of freedom is lower for the BEM for an equivalent mesh size, but as a counterpart the matrix of the system is full. However, a careful counting of the number of operations involved shows a lower order in the algorithmic complexity of the BEM. In addition, and what is may be more important in practice, mesh generation is much easier with BEM, simplifying considerably the task of grid movement associated with moving boundary problems. We can say that, among front tracking methods, BEM represents a great advance toward the modeling of more complex boundary evolutions. A complete version of the algorithm, taking into account unsteady effects in each phase is also proposed by O’Neill[64]. In this case, the temporal dependent Green function is used, and the algorithm becomes considerably more complex.

**General remarks about front tracking methods**

**Advantages of the method are:**
- The position of the interface is obtained explicitly at each time step
- The discontinuity in temperature gradient across the interface can be better taken into account.
- They result, generally speaking, in more efficient algorithms, as they are strongly oriented to a particular problem.

**Disadvantages are:**
- They need a starting solution.
- They cannot treat complex evolutions of the interface, appearing and disappearing phases, for instance.
- The interface must be smooth.

**FIXED DOMAIN METHODS**

This kind of methods consist in solving a PDE in the complete domain, capturing the interface as a by-product. First, note that equation (1) is valid also at the interface in the sense of distributions. It becomes singular at the interface with a term proportional to Dirac’s delta. To deal with the singularity, it is incorporated in the weak form of the PDE\[1,9\]. The intensity of the delta distribution is proportional to the energy balance at the interface (2), therefore, equation (1) if discretized correctly, enforces the PDE in both phases and the energy balance at the interface. Multiplying the PDE (1) in the whole domain, the energy balance (2) at the interface and the natural boundary condition (3.2) are multiplied by a regular enough weight function \( \phi(x) \), and integrated over the domain and boundaries to obtain the following weighted residual formulation:

\[
\int_{\Omega_L + \Omega_S - \Gamma} \phi \left[ \nabla \cdot (k\nabla T) + Q - \rho \frac{\partial H}{\partial t} \right] d\Omega - \\
\int_{\Gamma(t)} \phi \left[ \mathbf{n} \cdot (k\nabla T|_S - k\nabla T|_L) - \rho(H_L - H_S) \cdot \mathbf{v}_f \right] ds - \\
\int_{\partial \Omega_s} \phi (\mathbf{n} \cdot k\nabla T - \mathbf{q}) ds = 0 \quad (23)
\]
The temporal term can be rewritten as

$$\int_{\Omega_{t,\lambda,\nu}} \rho \frac{\partial H}{\partial t} d\Omega = \frac{\partial}{\partial t} \left\{ \int_{\Omega_{t,\lambda,\nu}} \rho \frac{\partial H}{\partial t} d\Omega \right\} - \int_{T=T_{\infty}} \rho H d\nu ds \quad (24)$$

Applying Gauss theorem and using (24) the weak form is obtained

$$\frac{\partial}{\partial t} \int_{\Omega} \rho H \phi d\Omega + \int_{\Omega} \nabla \phi \cdot \nabla T \, d\Omega = \int_{\Omega} \phi Q \, d\Omega + \int_{\partial\Omega} \phi \mathbf{q} \cdot d\mathbf{s} \quad (25)$$

Next, the time interval is divided in $M$ sub-intervals of length $\Delta t$, being $t_n = n\Delta t$ the $n$-th time stage. By standard backward Euler discretization, the following implicit scheme is obtained, where the superscript $n$ indicates approximated values at time stage $t_n$. It can be shown that, under certain regularity hypotheses, the solution of (26) converges in the weak sense to (25) when the time step goes to zero. Several methods have been proposed depending on the primary variable chosen, and we will now address the review the most representative among them.

**Apparent heat capacity**

Here, the main variable is temperature and the strategy is to recover some kind of modified heat transfer equation, like (12). To take into account the latent heat effect, a high specific heat is added in a narrow interval of temperatures, as shown in Figure 4. Bonacina et.al.[10] proposes the following equivalent physical constants

$$k_{\text{eq}} = k_S + \frac{k_m - k_S}{2} \left[ |T - (T_m - \nu)| \right] \quad \text{for } |T - T_m| \leq \nu \quad (27)$$

In numerical experiments it has been found that the result are relatively independent of the regularization parameter $\nu$ when a one phase problem is solved. In contrast, numerical results are strongly affected for general two phase Stefan problems.

Comini et.al.[31] adapted this technique to FEM, obtaining an equivalent heat capacity as the derivative of enthalpy with respect to temperature. Enthalpy is interpolated within the element using the same shape functions as temperature. For a two-dimensional problem, the equivalent specific heat is obtained as

$$c_{\text{eq}} = \frac{1}{2} \left[ \frac{\partial H}{\partial x} \frac{\partial H}{\partial y} \right] \quad (28)$$

Later, Morgan et.al.[32] corrected this expression to avoid convergence problems. Lemmnon[33] also recommends a technique of this type for the heat capacity, as well as for an equivalent thermal conductivity.

The main disadvantage of this kind of schemes is that the time step is restricted to be small enough in order that the interface does not advance more than one element per time step, otherwise the latent heat from those elements which the interface has traversed completely in one step is lost. Usually, codes based on this algorithm perform a check on time step size to avoid this loss of heat.
To solve this problem, Bathe and Kosgoftaar [27] proposed a method to divide the region that has been traversed by the interface in two regions. In the first one the latent heat effect is replaced by a fictitious heat source, equivalent to the absorbed or generated heat. In the second region, they use an equivalent heat capacity. In this way, a better energy balance is achieved at the cost of a loss in algorithm simplicity.

Another alternative to the definition of the apparent heat capacity was proposed by Storti et al. [50]. The method starts with the following exact expression for the Dirac’s δ contribution of the latent heat to the apparent heat capacity

$$\int_{\Omega} \phi \delta(T - T_m) \, d\Omega = \int_{T=T_m} \phi \frac{1}{|\nabla T|} \, d\Omega$$

for any smooth function \( \phi \). This expression is valid in the sense of distributions for regular enough temperature profiles. The proposed method consists in considering a constant \( c_m \) by averaging the previous contribution to the latent heat over the element volume \( \Omega_e \)

$$c_e = c + c_m, \quad c_m = \frac{1}{\Omega_e} \int_{T=T_m} \frac{L}{|\nabla T|} \, ds$$

where \( ds \) stands for the differential element of surface at the phase-change front. A subtle point with respect to this algorithm is that, in the continuum case, the temperature gradient is discontinuous across the interface, and expression (29) is not well defined. However, the temperature gradient is continuous inside the element at the discrete level, and the use of (29) is completely justified and unambiguous. The main advantage of this choice for the apparent heat capacity over previous ones is that it gives an exact average of the heat capacity.

Runnels and Carey [56] proposed the so called “accelerated heat capacity method”, which in fact uses the same formula (29) proposed firstly in [50]. They use a standard Galerkin
discretization leading to a non-linear system of ODE's which is integrated in time by the Crank-Nicholson scheme. In fact, expression (29) has also been used, before them, to derive an exact expression for the jacobian of the system of equations (25) in the context of Newton-Raphson methods by Storti et.al. [50], and for steady-state phase-change problems with advection by Pardo et.al. [63]. These non-linear solution strategies will be described in section 4.

**Enthalpy formulation**

Here, enthalpy is the main variable to be discretized. Temperature is obtained afterwards inverting the function $H(T)$, that is,

$$T = \begin{cases} H/c & H < cT_m \\ T_m & cT_m \leq H \leq cT_m + L \\ (H - L)/c & cT_m + L < H \end{cases}$$

(31)

This method was first proposed by Atthey[8] in a finite difference context for space discretization and explicit integration in time of the form

$$H_{m+1}^n = H_m^n + \frac{k\Delta t}{\rho \Delta x^2} (T_m^N - 2T_m^m + T_{m-1}^N) + Q(T_m^m)\frac{\Delta t}{\rho}$$

(32)

where $(\cdot)^m_n$ indicates the corresponding quantity at node $m$, temporal stage $n$. Of course, time step is restricted by a stability condition of the form

$$\frac{\Delta t}{\Delta x^2} \leq \frac{\rho c_p}{2k}$$

(33)

This algorithm can be used to treat problems with mushy zones, but has the disadvantage that spurious plateaux in temperature histories are observed, in places where a sharp change in slope at the interface should be found instead. This effect was first noted by Voller et.al.[71], and it is caused by the fact that all the heat needed to melt the element must be supplied to the associated node to advance the interface to the next node. White[21, 22], presented an implicit variant of this algorithm.

Meyer[16], proposed the first implicit method, regularizing the enthalpy/temperature relationship in an interval of width $2\nu$ as

$$H(t) = \begin{cases} c_s T & T < T_m - \nu \\ H(T_m - \nu) + (L/2\nu)(T - T_m + \nu) & |T - T_m| \leq \nu \\ H(T_m + \nu) + c_L(T - T_m - \nu) & T > T_m + \nu \end{cases}$$

(34)

Together with a standard difference scheme for the spatial second derivative, the algebraic system

$$A(u) + i(u) = 0$$

(35)

is obtained, where $A$ is the discrete laplacian, $i$ is the vector of nodal enthalpies and $u$ the nodal temperatures. Meyer affirms that the numerical results are relatively independent of $\nu$, but the numerical experiments of Voller et.al.[71] do not confirm this assertion, especially for low $\nu$.

To improve the behavior of explicit methods, with regards to the artificial temperature plateaux, Voller and Cross[20] proposed to compute the time when the interface traverses the actual node. This is supposed to happen when the node attains the “phase change enthalpy” $H_m = cT_m + L/2$. Then, if it is verified that

$$H_i^n > H_m > H_i^{n+1}$$

(36)
the interface is at the phase change enthalpy at \( t_i = t_n + \alpha \Delta t \), where

\[
\alpha = \frac{H_m - H_i^n}{H_i^{n+1} - H_i^n}
\]

and recomputing the nodal temperature as

\[
T_i^{n+\alpha} = T_i^n + \alpha(T_i^{n+1} - T_i^n)
\]

Bell[72] showed that this algorithm is appropriate for nodes lying near the part of the boundary where temperature is prescribed, because there, the rate of change of enthalpy is approximately constant, but it is not so for nodes that are far from the boundary.

**Apparent heat capacity with post-iteration correction**

Another possibility is a combination of the apparent heat capacity and full enthalpy methods which was first proposed by Pham[25,57] and extended later by Comini et.al. [58]. In this version, Comini uses a three level algorithm combined with the standard apparent heat capacity method to obtain an implicit scheme, but avoiding inner iteration within each time step. This scheme is used as a predictor to be followed by a corrector, the nodal enthalpy increments are calculated from the predictor temperatures and the corrected temperatures are obtained through inversion of the \( H(T) \) relationship. This kind of schemes have received the name of post-iterative and do not require the less reliable checks on time step needed in conventional apparent heat capacity codes, that we mentioned in §3.2.1.

In a different approach, Yao and Chait[79] replaced the usual “differential” heat capacity by a “nominal” heat capacity defined as

\[
H(T) = c^N(T)T
\]

The method has the simplicity of the apparent heat capacity formulation but allows for the use of larger time steps as it includes a post-iterative correction. the non-linear system is solved by the Gauss-Seidel SOR iterative scheme.

**Fictitious heat flow**

In this method, the latent heat effect is taken into account by a fictitious heat source on those elements that change phase. This scheme was initially proposed in a FEM context by Rolph and Bathe[41] and later by Roose and Storrer[45]. The first order system of ODE’s (12) is modified as

\[
C\ddot{u} + k\dot{u} - g + g_k = 0
\]

where \( g_k \) is the internal heat source term. This heat source has a duration given by the amount of heat needed to melt the volume associated to each node

\[
\Omega_i = \int_\Omega N_i \, d\Omega = \sum_e \int_{\Omega_e} N_i \, d\Omega
\]

where \( e \) is an element index running over all the elements that meet at node \( i \). The fictitious heat source is defined by the following strategy

\[
g_{kj} = \begin{cases} 
0 & \text{if } \begin{cases} T^{n-1} < T_m \text{ and } T_j^n < T_m \\
T^{n-1} > T_m \text{ and } T_j^n > T_m 
\end{cases} \\
\frac{T_m - T_j^n}{\Omega_i} & \text{if } \begin{cases} T^{n-1} < T_m \text{ and } T_j^n \geq T_m \\
T^{n-1} > T_m \text{ and } T_j^n \leq T_m 
\end{cases}
\end{cases}
\]

(42)
where \( T^n_j \) stands for the \( j \)-th iteration at time stage \( t^n \), all quantities are referred to node \( i \). This amount is added to equation (40) until

\[
\sum_l g_{kl} = L \Omega_i
\]  

(43)

**Freezing index**

This kind of algorithm is based on a transformation of the main variable in such a way that the new unknown is continuous and differentiable over the domain. It was first introduced by Duvaut[73] and applied later by Fremond[74], Fremond and Blanchard[46] and Kikuchi and Ichikawa[43, 44], among others. The new variable is called the *freezing index* and is defined as

\[
M(x, t) = \int_0^t k_i T(x, t) \, dt
\]  

(44)

where \( k_i \) is the thermal conductivity of the phase occupying point \( x \). In the case of constant conductivity, temperature can be obtained through the inverse relation

\[
T = \frac{1}{k} \frac{dM}{dt}
\]  

(45)

Fremond and Blanchard[46] regularized the enthalpy/temperature relationship in the phase-change zone by means of an homographic transformation, to allow the inclusion of mushy zones.

**Discontinuous integration**

Once equation (11) is replaced in (26) a system of equations is obtained, which can be written as

\[
K^n u^n + i(u^n) = g^n
\]  

(46)

where \( i \) is the vector of nodal enthalpies given by

\[
i_j = \frac{1}{\Delta t} \int_\Omega \rho H(T(x)) N_j(x) \, d\Omega
\]  

(47)

The relationship \( i(u) \) inherits some good and bad properties from the constitutive relation \( H(T) \), for instance it can be multiple valued, but the inverse relationship \( u(i) \) is not. Moreover \( i \) is the gradient of some convex, but possibly non-differentiable, functional \( \Phi(u) \) [69]. At some points in the temperature vector space, this functional does not have a single gradient vector, but a *convex set* of gradients. These points are those which have "mushy elements", that is, elements where all the nodes are at the melting temperature. The existence of this functional guarantees that a unique solution to (46) exists, in the sense that there exists a pair \((i, u)\) such that it is a solution of (46) and \( i \) is an enthalpy vector "admissible" to the vector of nodal temperatures \( u \), i.e., belonging to his set of gradients.

The difficult points of this method are first, to compute the vector of nodal enthalpies (47) and, second, to solve the nonlinear system (46) for the temperature vector. With respect to the first, we note that, as the enthalpy is a discontinuous function in those elements intersected by the interface, a numerical integration by Gaussis method is not appropriate at all, since it is designed to integrate accurately smooth functions. But what is more important, it would produce "jumps" in the residual, i.e., the residual would be non continuous as a function of the temperature vector even in the case where no mushy zones exist. To see this, it suffices to consider a temperature field for which the interface is arbitrarily close to a Gauss point, then an
Numerical methods in phase-change problems

infinitesimal change in a nodal temperatures, may advance the interface to beyond the Gauss point, and a finite amount of heat would be released, producing a jump in the residual vector. This jumps have catastrophic consequences on the solution of the system of equations with any method. Increasing the number of Gauss points is a rather poor solution, it has the effect of splitting the jump into smaller jumps and too many Gauss points are required to achieve a satisfactory improvement.

To overcome this difficulty, Crivelli et al.[48] proposed to integrate the enthalpy vector in such a way as to obtain a variation as smooth as possible. First, the nodal enthalpies are separated into two contributions coming from the sensible heat and the phase change, in the same way as in (5). The first contribution is smooth and can be integrated with the standard Gauss routine. The second contribution can be integrated exactly for “simplex” elements (linear one-dimensional elements[48], triangular in 2D and tetrahedra in 3D). For other elements, an approximated integration can be done by sub-dividing the element in “simplex” elements, and applying the exact integration to them. For instance, a quadrilateral element in 2D can be subdivided in two or four triangular subelements. For quadrilateral elements, an auxiliary isoparametric transformation from the master element coordinates \((\xi, \eta)\) can be devised in such a way that the interface becomes straight in the new coordinates \((\xi', \eta')\).[49] the integration is then performed over each subelement via the standard. The method has been later extended to non-isothermal problems[52,53]. A similar strategy is used by Tacke[81] in a FVM context to eliminate spurious oscillations in temperatures histories.

It’s worth to note, that this is the only method that corresponds to a straightforward weighted residual formulation of the original PDE’s. As we will see in the next section. This leads to the development of a consistent heat capacity matrix for the phase-change front[50] which is important in the solution of the non-linear system. Furthermore, a consistent explicit algorithm can be devised with the aide of this capacity matrix for the phase change front[51].

NON-LINEAR SYSTEM SOLUTION STRATEGIES

The non-linearity arising from the movement of the interface is, in most problems, much stringer than other non-linearities due to variations in the thermal coefficients. This results in great difficulties in the solution of the non-linear discrete system. Some authors, mainly in the finite difference community[16], employ successive relaxation procedures[75]. Blanchard[46] adopts a conjugate gradient method with preconditioning. However, in FEM codes it is recommended the use of incremental techniques[41, 76] of the type

\[ S(u^j_n)\Delta u^j_n = -r(u^j_n) \]

\[ u^{j+1}_n = u^j_n + \Delta u^j_n \quad (48) \]

where \(u^j_n\) stands for the \(j\)-th iteration at time stage \(t^n\), \(S\) is the iteration matrix and \(r(u)\) the residual of the non-linear system of equations.

Secant methods

It is common in Stephan problems to observe a lack of convergence in the iterative scheme, in the form of oscillating temperatures at nodes connected to elements which are changing phase. This can be easily understood considering a one-dimensional projection of the residual along an arbitrary direction \(\Delta u\) (as in the line-search method to be described next in this section, see equation (49)). The projected residual would look as shown in Figure 5. Here, \(x\) is the scalar parameter along the direction considered, and \(f(x)\) the projection of the residual, so that we have a one dimensional non-linear equation of the form \(f(x) = 0\). At some points, a given amount of latent heat is released producing an “S” shaped curve. This behavior, when combined with secant methods, for instance, can results in a lack of convergence as shown in Figures 5-6. This occurs, for instance, when the interface goes through a Gauss point in the standard finite
element method, the slope of the curve is small outside a narrow interval associated with the phase change, if the slope of the secant method is based on the regular part, outside of the phase change portion, the above mentioned oscillatory behavior is obtained and the method may not converge at all (see Figure 5). On the other hand, if the slope is based on the phase change portion, a monotone, but very slow convergence is obtained (see Figure 6). Referring to the discontinuous element method if Gauss integration is used to compute the nodal enthalpies, then the “S” shaped bends degenerate in jumps, and the convergence problems become worse. This is the reason why discontinuous integration is needed as proposed in [48] and [49]. These problems can be partially overcome, using under-relaxation for the initial iterations.

![Figure 5](image)

**Figure 5** Convergence process for a secant method with a secant slope based on the regular part. No convergence is obtained. Note the oscillating behavior.

### Quasi-Newton methods

A practical way to automatically adjust the relaxation parameter consists in applying a line-search scheme, where, once a “search direction” $\Delta u^j_n$ is found by a secant or Newton-Raphson method, the length of the step is chosen in such a way to cancel the projection of the residual in that search direction

\[
\begin{align*}
   f(\sigma) &= \Delta u^j_n \cdot r(u^j_n + \sigma \Delta u^j_n) = 0 \\
   u^{j+1}_n &= u^j_n + \sigma \Delta u^j_n \\
   \end{align*}
\]  

(49)

It has been also suggested[14] to modify the iteration matrix in order to control this lack of convergence. This can be done feeding the information contained in the successive
residual vectors back in the iteration matrix. Quasi-Newton schemes\cite{77, 78} perform rank-one corrections to the iteration matrix in such a way as to satisfy the relation

$$Gd = r(u) - r(\bar{u}) = y$$

where $\bar{u}$ is a point lying close to $u$. The rank-one modification can be written as

$$G = S + \frac{y - Sd}{e^Td} e^T, \quad d = u - \bar{u}$$

where $e$ is an arbitrary vector such that $e^Td \neq 0$. In a FEM context, a straightforward application of this strategy will destroy the band structure of the iteration matrix, to preserve its active profile it is preferable to perform only a correction to the diagonal terms, i.e.,

$$G = S + \text{diag}\{\alpha_i\}, \quad \alpha_i = \frac{y_i - \sum_{j=1}^{N} S_{ij}d_j}{d_i}$$

where $\text{diag}\{\alpha_i\}$ stands for a diagonal matrix with $\alpha_i$ in the $i$-th diagonal entry.

A standard form to check convergence is controlling the variation in the nodal temperatures vector in such way that convergence is achieved when:

$$\frac{\|\Delta u\|}{\|u\|} < \epsilon_u$$
where $\epsilon_u$ is the prescribed tolerance. However it is more convenient to evaluate convergence in terms of the equilibrium equations, that is

$$\frac{\|r(u)\|}{\|Ku\|} < \epsilon_r.$$  \hspace{1cm} (54)

Firstly, this last equation is invariant under a shift in the reference temperature (e.g. changing from Celsius to Kelvin) in contrast with (53), and furthermore, a very slow rate of convergence might trigger the first stop criterion inadequately.

**Newton-Raphson methods**

The use of implicit schemes allows in principle, the use of arbitrarily large time steps, from the stability point of view, but requires efficient techniques to solve the associated non-linear system of equations. However, in complex multidimensional geometries it is highly desirable to have the possibility to choose the time step independently of the mesh size, since sometimes the FEM mesh has small elements due to restrictions of the mesh generation process rather than due to a desired local refinement.

Prior to the discussion of the choice of the algorithm to solve this non-linear system, some observations can be made regarding the system of equations itself. Considering constant conductivities ($k \neq k(T)$), it can be shown that the non-linear discrete system associated to a standard Galerkin discretization for a problem with an arbitrarily smooth $H(T)$ relationship, is symmetric and has a positive definite Jacobian of the system of equations, defined as:

$$J_{jk} = \frac{\partial r_j}{\partial u_k}$$  \hspace{1cm} (55)

It is quite natural and desirable to preserve this characteristic for phase-change problems. However, it must be remarked that rather natural definitions of the discrete enthalpies do not lead to symmetric contributions to the Jacobian. This is the case if, for instance, a term of the form $C(u)u$ is introduced in the system, where $u$ is the unknown temperature vector to be found and $C_{ij} = \int_{\Omega} N_i \rho \tilde{c} N_j$ an approximation to the capacity matrix based on an apparent heat capacity $\tilde{c}$. The contribution to the Jacobian is

$$\frac{\partial i_l}{\partial u_j} = C_{lj} + \frac{\partial C_{lk}}{\partial u_j} u_k,$$

and is not symmetric in general. As can be seen, the first term is always symmetric but the second one, deriving from the non-linearities, may not be symmetric.

In contrast, a weak formulation together with nodal enthalpies defined as in (47), always leads to symmetric positive-definite Jacobians, since

$$\frac{\partial i_l}{\partial u_j} = \int_{\Omega} \rho N_i \frac{dH}{dT} N_j \, d\Omega$$  \hspace{1cm} (57)

For phase change problems, Storti et al.[50] proposed the use of (57) with the contribution of the singularities (represented by the Dirac $\delta$ function) at the phase change front evaluated by means of (29). This contribution is

$$C_{ij}^L = \int_{T=T_m} N_i N_j \frac{|\nabla T|}{|\nabla T|} \, ds$$  \hspace{1cm} (58)

The following numerical examples, show that the combination of discontinuous integration as proposed in [48] and [49] with the Newton-Raphson method, based on the previous expression [50], leads to quadratic convergence.
NUMERICAL EXAMPLES

In a paper whose main objective is to summarize the state of the art of the numerical solutions of phase-change problems, it would be desirable to present numerical examples that offer a comparison of the different methods described. However, due to the great number of algorithms proposed in the literature, it is not possible to achieve this goal. Therefore, we will show some examples using algorithms in which the authors have been involved, in which fixed mesh methods with Newton–Raphson techniques for the non-linear problem have been used.

The first two examples are very simple. They have been selected to show the importance of using a consistent Jacobian matrix, in order to obtain quadratic convergence of the residual, including the terms described in equation (58).

The third example is an application to a practical problem which is used to show the power of the fixed mesh methods to deal with irregular phase-change surfaces, in particular, with surfaces that may appear and disappear.

Freezing of a rod

A rod of length \( l = 1 \) is initially at \( T(x,0) = 0 \), at \( t = 0 \) the temperature on one end is lowered to \( T_w = -2 \) and an adiabatic condition is imposed on the other end. The freezing temperature is \( T_m = -1 \), the physical constants are set to one, i.e., \( k = \rho c = 1 \). The latent heat is \( L = 10 \) corresponding to a Stefan number of \( \text{Ste} = 0.1 \). In Figure 7 we can see the convergence histories for the first and second time steps, with \( \Delta t = 1 \). In each plot in full line, the convergence with a consistent Jacobian is shown, i.e., including the term (58), results obtained with a secant method, that does not include that term, are shown in dashed line. In both cases, a line-search method has been used to stabilize the iterative process. The convergence criterion (54) is used in all cases. Quadratic convergence is observed when the consistent Jacobian is used, whereas the secant method attains the prescribed tolerance \( (\epsilon_r = 10^{-6}) \) in three times more iterations in the first step, and seems not to converge at all in the second step.

![Figure 7](image_url)
Freezing of a square region

This example consists of a square region $ABCD$ (see Figure 8) of side 1m long and with constant physical properties. The conductivity is $k = 1 \text{W/m} \circ \text{C}$; the heat capacity is $\rho c = 1 \text{J/m}^3\circ \text{C}$; the latent heat is $L = 10 \text{J/m}^3$, and the melting temperature is $T_m = 0 \circ \text{C}$. For $t < 0$ the temperature is uniform and equal to $0.3 \circ \text{C}$. For $t > 0$ the temperature on sides $AD$ and $CD$ are lowered to $T_w = -1 \circ \text{C}$, while $AB$ and $BC$ are assumed to be adiabatic. The Stefan number for this problem is $\text{Ste} = \rho c (T_m - T_w) / L = 0.1$. The mesh has 900 elements and is refined near the Dirichlet boundaries. Figure 9 compares convergence histories for the secant and Newton-Raphson method for the fourth and eighth time steps. The convergence criterion is chosen as $\epsilon_r = 10^{-4}$. Quadratic convergence is observed for the Newton-Raphson method, whereas linear convergence rates that need 6 to 20 iterations to reduce the error by one order of magnitude are obtained with the secant method.

![Figure 8 Finite element mesh and problem description for numerical example of freezing of a square section](image)

Freezing of a buried pipe

This example examines the freezing of the soil surrounding a buried pipe with a constant wall temperature kept at $1 \circ \text{C}$. The geometry is shown in Figure 10, and the dimensions are: $AB = 0.2\text{m}$, $BC = 2R = 0.1\text{m}$ and $AF = FE = 1\text{m}$. Adiabatic conditions were imposed on sides $AB$, $CD$, $DE$ and $EF$. The temperature on the pipe wall was assumed to be equal to that of the circulating fluid on $BC$ and convective heat transfer is assumed to occur on the surface $AF$, that is,

$$q = h (T - T_{air})^{4/3}, \quad \text{on } AF$$

(59)
Figure 9 Convergence histories for the two-dimensional problem. a,b) Secant, c, d) Newton-Raphson, a,c) fourth time step, b, d) eight time step.

Figure 10 Finite element idealization of a buried pipe.
The physical properties are:

\[
\begin{align*}
\begin{align*}
& k = 1.9 \text{W/m}^\circ \text{K} \\
& \rho c = 1.59 \times 10^6 \text{J/m}^3\text{K} \\
& k = 1.35 \text{W/m}^\circ \text{K} \\
& \rho c = 1.64 \times 10^6 \text{J/m}^3\text{K}
\end{align*}
\end{align*}
\begin{align*}
& T < T_m \text{ (frozen soil)} \\
& T > T_m \text{ (unfrozen soil)}
\end{align*}
\]

\[ T_m = 0 \, ^\circ \text{C} \]
\[ L = 7.24 \times 10^7 \text{J/m}^3 \]
\[ h = 1.5764 \text{W/m}^2\text{K}^{4/3} \]

The region is initially at \( T(\mathbf{x}, 0) = 5 \, ^\circ \text{C} \) and at time \( t = 0 \) the temperature of the air is suddenly lowered to \( T_{\text{air}} = -15 \, ^\circ \text{C} \).

Results are depicted in Figures 11-13 corresponding to isotherms at \( t = 30, 40 \) and 50 days, respectively.

![Figure 11 Isoterms after 30 days for the buried pipe problem](image)

Note that the interface splits in two somewhere between 40 and 50 days, so that for \( t > 50 \) days two interfaces coexist. Fixed grid schemes are well suited to treat this anomalous behavior, in contrast with front tracking algorithms that exhibit serious difficulties or break down in this situation.
Figure 12 Isotherms after 40 days for the buried pipe problem

Figure 13 Isotherms after 50 days for the buried pipe problem
CONCLUSIONS

In this work, we have reviewed the most representative numerical algorithms for modeling phase-change problems. Implicit methods are preferred over explicit ones for general purpose codes, because for small Stefan number problems, the critical time step is controlled by the comparatively small time scale of each phase, whereas the time scale of the solution is controlled by the high apparent capacity concentrated at the interphase. With respect to the relative advantage of front-tracking and fixed-domain methods, the first group is more efficient and more accurate, but has problems to model complex geometries. The second one is a more robust alternative which may deal with both, simple and very complicated geometries, in particular with surfaces that may appear and disappear.

Finally, regarding non-linear discrete system solution strategies, the Newton-Raphson method which exhibits quadratic convergence is more efficient and robust than secant methods.

ACKNOWLEDGEMENTS

Thanks are due to Prof. J. C. Heinrich for revising the manuscript. Also, the authors wish to express their gratitude to Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET, Argentina) for its financial support.

References


40 M. Hogge and P. Gerrekens, “Two-dimensional deforming finite element methods for surface ablation”, AIAA 18th Thermophysics Conference, (Montreal, Canada, June 1-3, 1983)


Numerical methods in phase-change problems


60 Y. Ruan and N. Zabaras, “Moving and deforming finite element simulation of two-dimensional Stefan problems”, *Communications in Applied Numerical Methods*, 16, pp. 495-506 (1990)


