PHYSICS

Discretization of continuum physics – a comparison of numerical methods from a physical point of view

Heiko Herrmann

Centre for Nonlinear Studies, Institute of Cybernetics at Tallinn University of Technology, Akadeemia tee 21, 12618 Tallinn, Estonia; hh@cens.ioc.ee

Abstract. For numerical calculations in continuum physics partial differential equations and the space-time are discretized. This can be done in different ways. Common approaches are finite difference methods and finite element methods, more rarely finite volume methods are used. Each method has different mathematical properties, which have been discussed in the literature, but they also imply a different physical meaning. This issue is discussed in this article and the connection of finite volume methods to thermodynamics of discrete systems is shown.

Key words: continuum physics, numerical methods, partial differential equations, finite elements, finite volumes, finite differences.

1. INTRODUCTION

There exist many different numerical methods for partial differential equations. Some of these methods can only be applied to special types of differential equations, e.g. elliptical partial differential equations, others are difficult to use in complicated geometries. Numerical methods for partial differential equations use different discretizations of space and time and have different mathematical properties, like convergence and accuracy. These properties have been discussed in the literature. Yet, one problem that is not well discussed is that the choice of the numerical method also implies a certain view about the physical system under consideration, i.e. the physical description is changed from a continuous description to a discrete one. In the following sections finite difference, finite element, and finite volume methods are applied to a model problem and the change in the point of view of physics is discussed. This includes the change of the former continuous space into a discrete one, and the continuity of the solution and also of the (numerical) fluxes.

1.1. A model problem

As numerical calculations for the systems of partial differential equations, i.e. the balance equations, are very time-consuming and difficult to handle, we will consider a simple model problem for exemplary calculations: the Laplace Problem.

A numerical solution U is calculated, which approximates the (exact) solution u to the two-dimensional Laplace Problem for mixed boundary conditions:

$$-\Delta u \;=\; f \quad \text{in } \Omega, \tag{1}$$

$$u \;=\; u_D \quad \text{on } \Gamma_D, \tag{2}$$

$$\frac{\partial u}{\partial n} \;=\; g \quad \text{on } \Gamma_N, \tag{3}$$

where $\Omega \subset \mathbb{R}^2$ is a bounded Lipschitz domain with polygonal boundary Γ. Dirichlet conditions are assumed on a closed subset Γ_D of Γ, while Neumann boundary conditions are assumed on $\Gamma_N := \Gamma \setminus \Gamma_D$.

Inhomogeneous Dirichlet conditions are incorporated by using the decomposition $v = u - u_D$ so that $v = 0$ on Γ_D, i.e.,

$$v \in H_D^1(\Omega) := \left\{ w \in H^1(\Omega) | w = 0 \text{ on } \Gamma_D \right\}. \tag{4}$$

2. FINITE DIFFERENCE METHODS

Finite difference methods (FDMs) [1] emphasize the grid/mesh points. Using FDMs one calculates a solution on a regular mesh; the approximated solution exists only pointwise. The construction of a continuous solution is a different matter, which can be tried afterwards either by fitting by a global function or interpolation (for example, by using splines).

2.1. The discretization of the model problem

The discretization with FDMs is based on the simplest approach and probably the first one that comes to mind: replacing the differential operators by difference quotients.

With the help of symmetric differences the discrete Laplace operator can be constructed in the following way. First write the Laplace operator as a sum of second order Cartesian derivatives:

$$\Delta u(x,y) \;:=\; \frac{\partial^2}{\partial x^2}u(x,y) + \frac{\partial^2}{\partial y^2}u(x,y), \tag{5}$$

then discretize each second derivative by using symmetric differences

$$\frac{\partial^2}{\partial x^2}u(x) \;=\; \frac{1}{h^2}\left(u(x)|_{x=x_{j+1}} - 2u(x)|_{x=x_j} + u(x)|_{x=x_{j-1}}\right) \tag{6}$$

with $x_{j-1} = x - h$ and $x_{j+1} = x + h$ and sum up:

$$-\frac{1}{h^2}\left(u(x_{j+1},y_j) - 2u(x_j,y_j) + u(x_{j-1},y_j)\right) - \frac{1}{h^2}\left(u(x_j,y_{j+1}) - 2u(x_j,y_j) + u(x_j,y_{j-1})\right) = f(x_j,y_j). \tag{7}$$

The result is a so-called five-point-star:

$$-\frac{1}{h^2}\left(u(x_{j+1},y_j) + u(x_j,y_{j+1}) - 4u(x_j,y_j) + u(x_{j-1},y_j) + u(x_j,y_{j-1})\right) \;=\; f(x_j,y_j). \tag{8}$$

The result of an FDM is a set of approximate values of the function on a discrete set of points. Often the result is displayed with interpolating lines, which is *wrong*! The interpolation is a different procedure for which *new* assumptions have to be made; one has to be careful that these assumptions are not in contradiction with assumptions made for applying the FDM.

3. FINITE ELEMENT METHODS

Finite element methods (FEMs) [1] also emphasize the grid-nodes, but in addition take care of the inter-polation. A remarkable point is that the system of partial differential equations is not solved directly, but instead its weak formulation is solved. There exist different types of FEMs according to the type of the mesh and the ansatz-functions used:

- P_1: triangular mesh, linear ansatz-functions;
- P_2: triangular mesh, quadratic ansatz-functions;
- Q_1: quadrilateral mesh, (bi-)linear ansatz-functions;
- Q_2: quadrilateral mesh, (bi-)quadratic ansatz-functions;
- ..., including mixed FEMs.

The weak formulation of this problem reads: Seek $v \in H_D^1(\Omega)$ such that

$$\int_\Omega \nabla v \cdot \nabla w \, dx \;=\; \int_\Omega f w \, dx + \int_{\Gamma_N} g w \, ds - \int_\Omega \nabla u_D \cdot \nabla w \, dx, \quad w \in H_D^1(\Omega). \tag{9}$$

3.1. The discretization of the model problem: Galerkin discretization for a P_1-Q_1-FEM

The model problem is discretized using a standard Galerkin method for implementation [2], where $H^1(\Omega)$ and $H_D^1(\Omega)$ are replaced by subspaces S and $S_D = S \cap H_D^1$, which are finite dimensional. Let $U_D \in S$ be a function that approximates u_D on Γ_D. Then the discretized problem reads: Find $V \in S_D$ such that

$$\int_\Omega \nabla V \cdot \nabla W \, dx \;=\; \int_\Omega f W \, dx + \int_{\Gamma_N} g W \, ds - \int_\Omega \nabla U_D \cdot \nabla W \, dx, \quad W \in S_D. \tag{10}$$

Let $(\eta_1, \ldots, \eta_N)$ be a basis of S and $(\eta_{i_1}, \ldots, \eta_{i_M})$ be a basis of S_D, where $I = \{i_1, \ldots, i_M\} \subseteq \{1, \ldots, N\}$ is an index set of cardinality $M \leq N - 2$. Then the discretized problem (10) is equivalent to

$$\int_\Omega \nabla V \cdot \nabla \eta_j \, dx \;=\; \int_\Omega f \eta_j \, dx + \int_{\Gamma_N} g \eta_j \, ds - \int_\Omega \nabla U_D \cdot \nabla \eta_j \, dx, \quad W \in S_D. \tag{11}$$

Furthermore, let

$$V \;=\; \sum_{k \in I} x_k \eta_k, \tag{12}$$

$$U_D \;=\; \sum_{k=1}^N U_k \eta_k, \tag{13}$$

which means we choose the same ansatz for the solution as for the test function. Then the discretized problem results in a linear system of equations $Ax = b$, where the coefficient matrix and the right-hand side are defined as

$$A_{jk} \;=\; \int_\Omega \nabla \eta_j \cdot \nabla \eta_k \, dx, \tag{14}$$

$$b_j \;=\; \int_\Omega f \eta_j \, dx + \int_{\Gamma_N} g \eta_j \, ds - \sum_{k=1}^N U_k \int_\Omega \nabla \eta_j \cdot \nabla \eta_k \, dx. \tag{15}$$

The coefficient matrix is symmetric and positive definite. So the linear system of equations has exactly one solution $x \in \mathbb{R}^M$, which gives the Galerkin solution

$$U \;=\; U_D + V \;=\; \sum_{j=1}^N U_j \eta_j + \sum_{k \in I} x_k \eta_k. \tag{16}$$
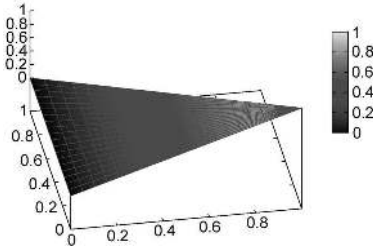
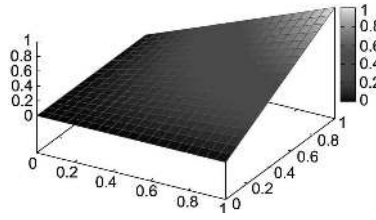**Fig. 1.** Linear ansatz-function on a triangle.

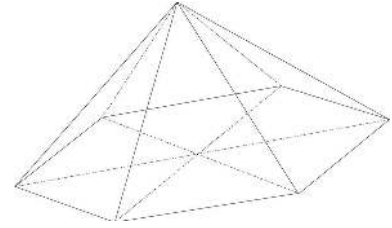**Fig. 2.** Bi-linear ansatz-function on a quadrilateral.

Fig. 3. A patch of ansatz-functions.

We assume a polygonal boundary Γ of the domain Ω; then $\bar{\Omega}$ can be covered by a regular triangulation $\mathcal{T}$ of triangles and quadrilaterals, i.e. $\bar{\Omega} = \cup_{T \in \mathcal{T}} T$ with each T being either a closed triangle or closed quadrilateral. The nodes $\mathcal{N}$ of the mesh lie on the vertices of the quadrilaterals or triangles, whereas no node lies on an edge of an element. The elements do not overlap and each edge $E \subset \Gamma$ of an element $T \in \mathcal{T}$ belongs either to $\bar{\Gamma}_N$ or to $\bar{\Gamma}_D$.

The coordinates of the vertices of an element determine the local stiffness matrix. For a triangular element T with the vertices (x_1, y_1), (x_2, y_2), and (x_3, y_3) and the corresponding basis functions η_1, η_2, and η_3,

$$\eta_j(x_k, y_k) = \delta_{jk}, \quad j, k = 1, 2, 3, \tag{17}$$

holds true. See Figs 1 and 2 for ansatz-functions over a triangle and quadrilateral and Fig. 3 for a patch of ansatz-functions.

4. TAKING ONE STEP BACK: HOW THE LOCAL BALANCES ARE DERIVED/ MOTIVATED?

The first two methods are mathematical approaches to the problem of finding an approximate solution to the system of partial differential equations. A more physical approach can be motivated by remembering that the system of partial differential equations under consideration is actually a set of balance equations and how the local balances have been derived/motivated [3].

The global balance equation for an arbitrary extensive quantity with specific density ψ, flux J^ψ, production density π^ψ, and supply density σ^ψ reads:

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{G(t)} \rho(x,t)\psi(x,t)\mathrm{d}V = -\oint_{\partial G(t)} J^\psi(x,t) \cdot \mathrm{d}a + \int_{G(t)} (\sigma^\psi(x,t) + \pi^\psi(x,t))\mathrm{d}V, \tag{18}$$

where $\mathrm{d}a = n\mathrm{d}a$, and n is the outward unit normal vector to the surface ∂G. As an example we consider the balance of mass. Then we have $\psi \equiv 1$ and $\sigma^1 \equiv 0$, $\pi^1 \equiv 0$, which expresses the conservation of mass:

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{G(t)} \rho(x,t)\mathrm{d}V = \frac{\mathrm{d}}{\mathrm{d}t} M^G(t) = -\oint_{\partial G(t)} J^1(x,t) \cdot \mathrm{d}a. \tag{19}$$

Applying the Reynolds transport theorem to the time derivative of the integral in equation (18) and the Gauss theorem

$$\oint_{\partial G(t)} J^\psi \cdot \mathrm{d}a = \int_{G(t)} \nabla \cdot J^\psi \mathrm{d}V \tag{20}$$

to the flux over the boundary in equation (18), we obtain

$$\int_{G(t)} \left(\frac{\partial}{\partial t}(\rho(x,t)\psi(x,t)) + \nabla \cdot (w(x,t)\rho(x,t)\psi(x,t)) \right) dV = - \int_{G(t)} \nabla \cdot J^{\psi}(x,t) dV$$
$$+ \int_{G(t)} (\sigma^{\psi}(x,t) + \pi^{\psi}(x,t)) dV, \qquad (21)$$

where w is the mapping velocity of the control volume.

As these equations hold true for *arbitrary* control volumes, one can conclude that the integrand must vanish. This results in local balance equations of the form

$$\frac{\partial}{\partial t}(\rho(x,t)\psi(x,t)) + \nabla \cdot (w(x,t)\rho(x,t)\psi(x,t)) + \nabla \cdot J^{\psi}(x,t) = (\sigma^{\psi}(x,t) + \pi^{\psi}(x,t)). \qquad (22)$$

This is exactly what finite volume methods (FVMs) are based on.

5. FINITE VOLUME METHODS

Finite volume methods [1] in one sense reverse the derivation of the local balance equation and are therefore the most physical approach of the mentioned methods. There are many different types of FVMs, some focusing on the mesh(-nodes) and interpolation and some focusing on the control volumes.

Finite volume methods are used if differential equations have a divergence form, i.e. contain parts like

$$Lu \quad := \quad -\nabla \cdot (K\nabla u + cu) + ru \;=\; f, \qquad (23)$$

with $K : \Omega \to \mathbb{R}^{d,d}$, $c : \Omega \to \mathbb{R}^d$, and $r, f : \Omega \to \mathbb{R}$, a part which is also included in parabolic equations of the form

$$\frac{\partial u}{\partial t} \quad = \quad Lu, \qquad (24)$$

which is a generalized form of balance equations.

5.1. The discretization of the model problem

Our stationary model problem reads

$$-\nabla \cdot (k\nabla u) \;=\; f \quad \text{in } \Omega, \qquad (25)$$
$$u \;=\; 0 \quad \text{at } \partial\Omega, \qquad (26)$$

with $k \equiv 1$.

The starting point is to divide the area into control volumes, which have a polygonal boundary, and integrate the differential equation over the control volumes:

$$-\int_{\Omega_i} \nabla \cdot (k\nabla u) \, dx \;=\; \int_{\Omega_i} f \, dx, \quad i \in \Lambda. \qquad (27)$$

Here $\Lambda := \{i \in \overline{\Lambda} | a_i \in \Omega\}$, a_i is an inner node of the associated Delaunay triangulation of Ω, and Ω_i is the Voronoi polygon (Wigner–Seitz cell) of a_i (see Figs 4, 5).
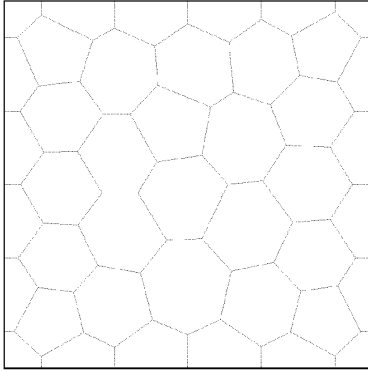
Fig. 4. The control volumes of an FVM mesh (Voronoi diagram).
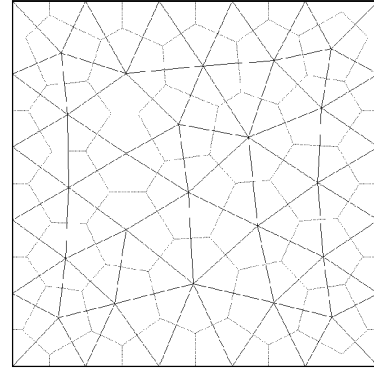

Fig. 5. Voronoi diagram and Delaunay mesh.

Then, by using the Gauss law and $\partial\Omega = \cup_{j \in A_i}\Gamma_{ij}$, we obtain

$$-\int_{\Omega_i} \nabla \cdot (k\nabla u)\, dx \;=\; \int_{\partial\Omega_i} \nu \cdot (k\nabla u)\, d\sigma \;=\; \sum_{j \in A_i} \int_{\Gamma_{ij}} \nu_{ij} \cdot (k\nabla u)\, d\sigma. \tag{28}$$

The approximation of the normal derivative which represents the flux is the crucial point. As an example difference quotients are used here, which leads to a finite difference-like scheme

$$\int_{\Omega_i} \nabla \cdot (k\nabla u)\, dx \approx \sum_{j \in A_i} k\frac{u(a_j)-u(a_i)}{d_{ij}} m_{ij}, \tag{29}$$

where $\Gamma_{ij} = \partial\Omega_i \cap \partial\Omega_j$ $(j \in \Lambda_i)$, m_{ij} is the length of Γ_{ij}, and $d_{ij} = |a_i - a_j|$. Using

$$\int_{\Omega_i} f\, dx \approx f(a_i)m_i, \quad m_i := |\Omega_i| \tag{30}$$

as an approximation for the integral over f, we finally arrive at

$$\sum_{j \in A_i} k\frac{u(a_j)-u(a_i)}{d_{ij}} m_{ij} = f(a_i)m_i, \quad i \in \Lambda. \tag{31}$$

6. INTERMEDIATE COMPARISON OF THE METHODS

In previous sections the discretization of the continuum for numerical methods was discussed. Figure 6 shows an illustrative comparison of how different methods "approximate" the solution.

As Figure 6 shows, the results of different numerical methods reveal significant differences. By using a finite difference method, the solution is approximated only on a discrete set of points. Finite element methods calculate the approximate solution at discrete points *and* the interpolation between the points. Finite volume methods result in an approximate (averaged) value, which belongs to the computational cell as a whole. To summarize:

FDM: solution and flux are non-continuous, exist only on a discrete set of points;
FEM: solution exists on a continuous space, fluxes show jumps;
FVM: solution exists on a continuous space, but has jumps, fluxes are continuous.

The behaviour of FVMs to assign the averaged value to the computational cell and the calculation method by use of fluxes over the boundary of the cell shows a significant similarity to thermodynamics of discrete systems.
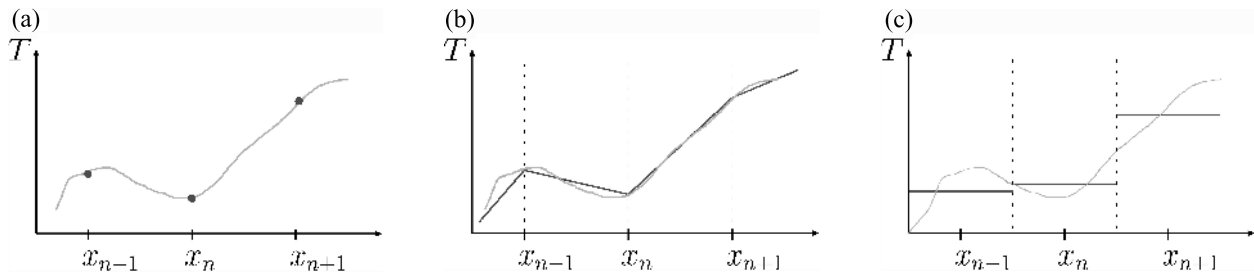
Fig. 6. Results of different methods. (a) An FDM result: Approximative solution at points. The continuum has been eliminated. (b) An FEM result: Approximative solution at points with interpolation. The continuum has been preserved, no jumps are allowed. (c) An FVM result: Averaged value over the control volume. The continuum has been preserved, jumps at interfaces between control volumes are possible.

7. THERMODYNAMICS OF DISCRETE SYSTEMS

In thermodynamics of discrete systems the system is described as a whole; no internal structure or distribution is taken into account. Exchange with the environment is due to fluxes over the boundary (see Figs 7 and 8).

Different situations in thermodynamics of discrete systems:

One equilibrium system in contact with a reservoir: This situation is the standard case discussed in thermodynamics books. It has been well known for a long time. See Fig. 7a.

One non-equilibrium system in contact with a reservoir: This situation is also well understood, but usually not discussed in books, an exception is [4]. See Fig. 7b.

Two or more non-equilibrium systems in contact with each other and a reservoir: This situation has only been discussed recently in [5,6]. It is not mentioned in books at all. A generalization from two systems to more should be possible by partial contact quantities. See Fig. 8.
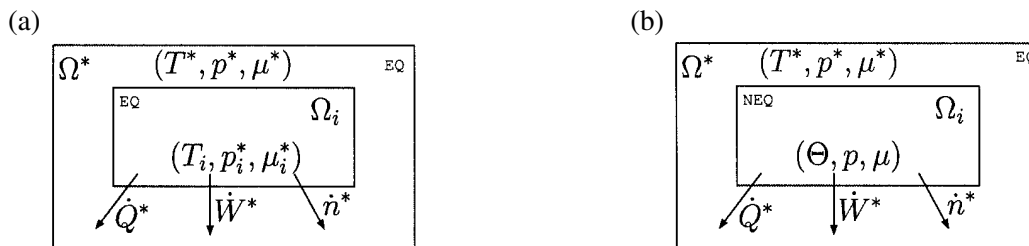

Fig. 7. Two discrete systems in contact with each other and with the environment. (a) Equilibrium. (b) Non-equilibrium.
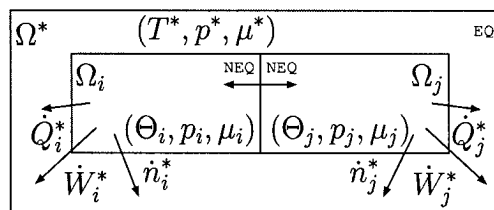

Fig. 8. Two discrete systems in contact with each other and with the environment.

7.1. Contact quantities and replacement quantities

The familiar quantities from thermostatics like the (thermostatic) temperature, pressure, and chemical potential do not exist in non-equilibrium situations. However, non-equilibrium counterparts can be defined[1] (see [4]): the contact temperature Θ, dynamic pressure p, and dynamic chemical potential μ.

These non-equilibrium contact quantities are defined by the exchange of the system with its equilibrium environment by defining the inequalities

$$\left(\frac{1}{\Theta} - \frac{1}{T^*}\right)\dot{Q} \geq 0, \tag{32}$$

$$(p - p^*)\dot{V} \geq 0, \tag{33}$$

$$(\mu^* - \mu)\dot{n}^e \geq 0. \tag{34}$$

For a discussion of these inequalities see [4].

The situation changes if a contact between two non-equilibrium systems is considered. The first idea that one might have is to use the contact temperature of the non-equilibrium environment, but this does not work. Instead, the non-equilibrium environment is replaced by an equilibrium environment which causes the same fluxes. The thermostatic temperature of the replaced environment is called the replacement temperature θ^*, and analogue for the pressure π^* and chemical potential ν^*. The defining inequalities read:

$$\left(\frac{1}{\Theta} - \frac{1}{\theta^*}\right)\dot{Q} \geq 0, \tag{35}$$

$$(p - \pi^*)\dot{V} \geq 0, \tag{36}$$

$$(\nu^* - \mu)\dot{n}^e \geq 0. \tag{37}$$

The difference is that in this case the replacement quantities (θ^*, π^*, and ν^*) and not the contact quantities are defined by the inequalities.

7.2. Compound deficiency and excess quantities

A system can be composed of subsystems. Of course, the accuracy depends on how the system is described – as a whole, without taking the composition into account or as a compound system. This difference is the compound deficiency; it can be taken into account by introducing excess quantities [5,6].

At the boundary of a subsystem the actual value of the quantity is the averaged quantity plus the excess quantity (Fig. 9).

In this way excess quantities can be used to formulate fluxes and jump conditions between computational cells.
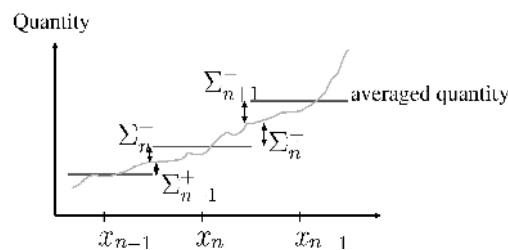

Fig. 9. Excess quantity.

---

[1]　At least as long as one non-equilibrium system in contact with a reservoir is considered.

8. CONNECTION BETWEEN FINITE VOLUME METHODS AND THERMODYNAMICS OF DISCRETE SYSTEMS

Each of the control volumes of the FVM forms a discrete system (Schottky system) in the sense of thermodynamics. The value of the variable belongs to the system as a whole and the exchange of quantities is described by fluxes over the boundaries of the system.

One point where a good knowledge of thermodynamics of discrete systems can be useful is the approximation of the flux over the boundary

$$\oint_{\partial G(t)} J^\psi(x,t) \cdot \mathrm{d}a. \tag{38}$$

An example of the usefulness of thermodynamical considerations in computational algorithms is the derivation of a second order accurate algorithm for hyperbolic first order partial differential equations, which is stable for the Courant number $\Delta t/c\Delta x \leq 1$ in [7] – it makes use of excess quantities to define jump relations at the boundary between cells. This algorithm is equivalent to the wave-propagation algorithm proposed by Bale et al. in [8], who actually proved the stability and accuracy.

9. CONCLUSIONS

Different numerical methods for partial differential equations were discussed. A correspondence between FVMs and thermodynamics of discrete systems was shown by introducing interacting non-equilibrium subsystems. The value of variables belongs to the discrete system (control volume) as a whole. This leads to jumps at the interfaces between control volumes and a stepwise approximation of the solution using an FVM. Compared to the shape functions used for an FEM, this approximation seems rude, but the advantage is that one can handle the fluxes between the control volumes. Finite element methods on the other hand give a steady approximation of the solution, but the value of the variables varies across the cells; therefore a correspondence to discrete systems is not given.

ACKNOWLEDGEMENTS

REFERENCES

1. Knabner, P. and Angermann, L. *Numerik partieller Differentialgleichungen*. Springer, 2000.
2. Alberty, J., Carstensen, C. and Funken, S. A. Remarks around 50 lines of matlab: short finite element implementation. *Num. Alg.*, 1999, **20**, 117–137.
3. Muschik, W., Papenfuß, C. and Ehrentraut, H. *Concepts of Continuum Thermodynamics*. Technische Universität Berlin und Kielce University of Technology, 1996.
4. Muschik, W. *Aspects of Non-Equilibrium Thermodynamics*. World Scientific, Singapore, 1990.
5. Muschik, W. and Berezovski, A. Thermodynamic interaction between two discrete systems in non-equilibrium. *J. Non-equilib. Thermodyn.*, 2004, **29**, 237–255.
6. Muschik, W. and Berezovski, A. Non-equilibrium contact quantities and compound deficiency at interfaces between discrete systems. *Proc. Estonian Acad. Sci. Phys. Math.*, 2007, **56**, 133–145.
7. Berezovski, A. and Maugin, G. A. Simulation of thermoelastic wave propagation by means of a composite wave-propagation algorithm. *J. Comput. Phys.*, 2001, **168**, 249–264.
8. Bale, D. S., LeVeque, R. J., Mitran, S. and Rossmanith, J. A. A wave propagation method for conservation laws and balance laws with spatially varying flux functions. *SIAM J. Sci. Comput.*, 2003, **24**, 955–978.

Pideva keskkonna füüsika probleemide diskretiseerimine – numbriliste meetodite võrdlus füüsika seisukohalt

Heiko Herrmann

Numbriliste meetodite kasutamisel pideva keskkonna füüsikas diskretiseeritakse diferentsiaalvõrrandid ja pidev keskkond. Seda võib teha erinevalt. Tavaliselt lähtutakse lõplike vahede või lõplike elementide meetoditest, harvem lõplike mahtude meetodist. Meetoditel on eri matemaatilised omadused, mida on käsitletud kirjanduses, ja erinev füüsikaline sisu. Viimast küsimust on käsitletud antud artiklis, kus on esitatud seosed lõplike mahtude meetodite ja diskreetsete süsteemide termodünaamika vahel.