High-order processing of singular data

Yosef Yomdin and Gal Zahavi

Abstract.

This paper provides a survey of some recent results (mostly in [68, 71, 72, 73, 74]) concerning high-order representation and processing of singular data. We present these results from a certain general point of view which we call a "Model-Net" approach: this is a method of representation and processing of various types of mathematical data, based on the explicit recovery of the hierarchy of data singularities. As an example we use a description of singularities and normal forms of level surfaces of "product functions" recently obtained in [68, 34] and on this base describe in detail the structure of the Model-net representation of such surfaces.

Then we discuss a "Taylor-net" representation of smooth functions consisting of a net of Taylor polynomials of a prescribed degree k (or k-jets) of this function stored at a certain grid in its domain. Following [72, 74] we present results on the stability of Hermite fitting, which is the main tool in acquisition of Taylor-net data.

Next we present (following [71, 73, 74]) a method for numerical solving PDE's based on Taylor-net representation of the unknown function. We extend this method also to the case of the Burgers equation near a formed shock-wave.

Finally, we shortly discuss (following [28, 56]) the problem of a non-linear acquisition of Model-nets from measurements, as well as some additional implementations of the Model-net approach.

§1. Introduction

This paper provides a survey of some recent results (mostly in [68, 71, 72, 73, 74]) concerning high-order representation and processing of singular data. We try to present these results from a certain general point of view which we call a "Model-Net" approach.

Model-nets provide an approach to a representation and processing of various types of mathematical data, based on the explicit recovery

of the hierarchy of data singularities. Various specific cases have been investigated in [14, 17], [24]-[27], [34, 56, 66, 68], [71]-[74], including representation of high-resolution images, of curves and surfaces, high-order numerical methods for solving linear and non-linear PDE's, compact description of a free configuration space of a robot manipulator, etc.

The representation in our approach is achieved via a "net of geometric models". This is a system of interconnected local analytic models, subordinated to a certain grid in (or partition of) the data domain, and respecting the hierarchy of singularities of the data. The main feature of the Model-net approach is that we try to avoid a "blind" approximation of functions with singularities by standard methods, ignoring the presence of singularities. Instead we construct an hierarchy of smooth functions in different dimensions, reflecting the hierarchy of singularities of the original data, and finally approximate only these smooth components.

Certainly, in each specific problem where singularities appear, they are addressed in this or another way. So the Model-net approach strongly overlaps with many classical methods for treatment of singularities. We discuss some of these overlaps in more details below. Still, we've found Model-nets conceptually adequate and providing a useful insight in solving specific problems via applications of Singularity Theory and high-order methods.

In the present paper we shall give a detailed description of a couple of the main mathematical ingredients in the Model-nets (restricted to instructive examples), and a very short overview of some other related mathematical problems, as well as of some implementations of Model-nets approach.

In Section 2 we start with an example: a Model-net representation of level surfaces of smooth functions. We use a description of singularities and normal forms of level surfaces of "product functions" recently obtained in [68, 34] and on this base describe the structure of a Model-net representation of such surfaces.

The problem of a representation of smooth components in Model-nets brings us in Section 3 to one of the main mathematical topics in this paper - "Taylor-nets". The idea to use nets of local Taylor polynomials in analysis and representation of smooth functions goes back to at least Whitney ([61]-[65]) and Kolmogorov ([43, 44]). In the last two papers it was shown that Taylor-net representation is essentially optimal for smooth functions from the point of view of the amount of information to be stored. The role of Taylor nets was stressed once more in a recent solution of the Whitney extension problem ([30, 31]; see also [15, 19] and

references there). In numerical algorithms Taylor methods are tradition-
ally used in high accuracy computations of ODE's (see [9, 12, 36, 50]
and references there) and to a much smaller extent in solving of PDE's
(see, for example, [14, 66, 40, 41, 42]). Recently Taylor discretization
has been applied to solving elliptic PDE's and to many other problems
in differential equations and Dynamics by the group of the Michigan
State University ([7]-[12], [33, 36], [50]-[52]). The results of this group
clearly illustrate the power and efficiency of Taylor models.

However, there is a serious stability problem which makes diffi-
cult using Taylor-nets in general numerical algorithms. This problem
is shared by all high-order methods and it consists in an oversensitivity
of high order derivatives to the noise in the data. Even the most initial
step in acquisition of high-order data from the low order measurements
- the interpolation process - is well known to be highly unstable.

We believe that Taylor nets provide a framework where the stability
problem can be addressed with adequate mathematical tools: "Hermite
fitting" and "Multi-order strategy".

In Hermite fitting we approximate the point-wise data with a poly-
nomial of lower degree than the maximally possible. Accordingly, the
requirement of the precise interpolation is replaced by the requirement
of the least square deviation.

In Section 4, following [72, 74], we analyze the robustness of the
Hermite fitting operator. We give some experimental and theoretical
results from [72, 74] which support the following conclusion:

*If we replace Hermite interpolation by "Hermite fitting", signifi-
cantly reducing the degree of the fitting polynomial with respect to the
maximal possible, the stiffness of the Hermite fitting operator is reduced
in many orders of magnitude.*

"Multi-order strategy" consists in a successive processing of the
data: from the lowest order to higher and higher ones. At each or-
der the maximal possible for this order accuracy is achieved, and then
the next order data is included as a "correction" to the previous step.
This approach was implemented in a Taylor-net discretization of Laplace
equation in [66] and it provided a significant increase in stability. We
plan to present its further applications separately.

We believe that the Hermite fitting of an appropriate order, com-
bined with the multi-order strategy, will ultimately allow for a robust
acquisition and processing of Taylor-net data in a noisy environment.

In Section 5 we describe, following [71, 72, 73, 74] a method to
solve evolution PDE's using Taylor-net representation of the unknown

function. It combines an algebraic jet-extension of the original equation and the Hermite fitting. We also present in Section 6 some new results concerning the case of the Burgers equation where the already formed shock-waves can be treated as the evolving singularities. We shortly discuss also the Model-net approach to the processing of the birth of a shock-wave.

Finally, in Section 7 we provide a very short overview of the problem of acquisition of singular data from measurements, as well as a very short overview of some implementations of the Model-net approach, not covered by the present paper.

§2. Example: level surfaces

In ([68], see also [34]) we propose an approach to a representation of level surfaces of smooth functions in $\mathbb{R}^3$. Specifically, we consider level surfaces $Y(c) = \{F(x_1, x_2, x_3) = c\}$ of smooth functions $F(x)$ of three variables $(x_1, x_2, x_3) = x$, of a special "product" form. Our approach is motivated by the following consideration: surfaces usually appear as the boundaries of three-dimensional bodies $B \subset \mathbb{R}^3$. Let us assume that a connected body $B \subset \mathbb{R}^3$ is defined by the inequalities $F_1(x) \geq 0, \ldots, F_m(x) \geq 0$. For example, this is always the case for the surfaces produced by the Computer-Aided Design - Computer-Aided Manufacturing (CAD-CAM) systems, widely used in engineering. The interior $\hat{B}$ is exactly one of the connected components G_0^i of the set $G_0 = \{F(x) > 0\}$, where $F = F_1 F_2 \ldots F_m$. So our surface is the boundary of $\hat{B} = G_0^i$, and it is a part of the level surface $Y(0) = \{F(x) = 0\}$

If we want to smooth out sharp edges and corners of our surface, one of possibilities is to shift it slightly inside the body B by taking the appropriate component of the surface $Y(\epsilon) = \{F(x, y, z) = \epsilon\}$, where ϵ is a small non-zero number. See Figure 1.

In the present paper we shall not discuss in detail the specific notion of a "genericity" or of a "general position", used in [68]. Roughly, we say that the property P of $F = F_1 F_2 \ldots F_m$ is generic if it is satisfied with a probability 1 for a randomly and independently picked $F_1, \ldots, F_m$.

However, it is important to stress that in our setting the product $F = F_1 F_2 \ldots F_m$ generically has non-isolated singularities along the crossing curves C_{ij} of the surfaces $C_i = \{F_i = 0\}$ and $C_j = \{F_j = 0\}$, $i, j = 1, \ldots, m$. From the point of view of the *classification of singularities of smooth functions in the standard setting* (see [16, 18]) this is a very degenerate situation, appearing only in "codimension infinity".
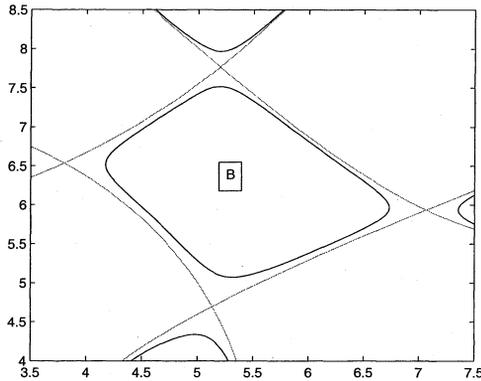
Fig. 1

So let $F = F_1 F_2 \ldots F_m$ be a generic product function. We denote by $C_i = \{F_i = 0\}$ the zero surfaces of the functions F_i, $i = 1, \ldots, m$, by $C_{ij} = \{F_i = 0\} \cap \{F_j = 0\}$, $i, j = 1, \ldots, m$, $i \neq j$, the zero curves of the couples of these functions, and by $w_{ijl} = \{F_i = 0\} \cap \{F_j = 0\} \cap \{F_l = 0\}$ - their "triple zeroes". The following proposition describes the singular structure of a generic function $F = F_1 F_2 \ldots F_m$:

Proposition 2.1. *For a generic function $F = F_1 F_2 \ldots F_m$ the critical set $\Sigma(F)$ of the function F consists of isolated non-degenerate (Morse) points w_i with $F(w_i) \neq 0$, of smooth curves C_{ij}, and of isolated triple points w_{ijl}, being the intersections of the zero surfaces C_i, C_j and C_l (and of the curves C_{ij}, C_{il} and C_{jl}). At the curves C_{ij} the zero surfaces C_i and C_j intersect transversally, and at the triple points w_{ijl} the corresponding triples of the zero surfaces C_i, C_j and C_l intersect transversally.*

The proof is given in [68].

The following theorem presents one of the main results of [68]. It claims that each "near singular" point of a generic surface can be represented by a standard model. We refer the reader to [68] for a proof, as well as for a discussion of some relations to the idea of R. Thom of an "organizing center" ([58]). For some "quantitative" results in Singularity Theory which play important role in our approach see ([67, 70]).

Let us describe our models. We use the following notations:

1. $Y_2^+(\epsilon) = \{y_1^2 + y_2^2 - y_3^2 = \epsilon\}$.

2. $Y_2^-(\epsilon) = \{y_1^2 + y_2^2 - y_3^2 = -\epsilon\}$.

3. $Y_3(\epsilon) = \{y_1 y_2 = \epsilon\}$.

4. $Y_4(\epsilon) = \{y_1 y_2 y_3 = \epsilon\}$.

Here $\epsilon > 0$ for near-singularities, and $\epsilon = 0$ for singular points (Figure 2).

Theorem 2.1. *Let $F = F_1 F_2 \ldots F_m$ be a generic product function. There exists a constant $K = K(F)$ such that for any $c \in \mathbb{R}$ and $Y(c) = \{F = c\}$ the following is true: at each regular point $x \in Y(c)$ of the surface $Y(c)$, where the sum of the absolute values of the main curvatures of $Y(c)$ at x exceeds K, this surface can be parametrized as follows:*

$$(2.1) \quad x_1 = \Psi_1(y_1, \ y_2, \ y_3), \ x_2 = \Psi_2(y_1, \ y_2, \ y_3), \ x_3 = \Psi_3(y_1, \ y_2, \ y_3),$$

with $(y_1, \ y_2, \ y_3) \in \tilde{Y}(\epsilon)$, where $\tilde{Y}(\epsilon) = Y_2^{\pm}(\epsilon)$, $Y_3(\epsilon)$, or $Y_4(\epsilon)$, and ϵ is a certain positive number.

At each singular point $x \in Y(c)$ of the surface $Y(c)$ the same parametrization holds, with $\epsilon = 0$ in the appropriate models.

This parametrization is valid in neighborhoods U_x of each point $x \in Y(c)$, and the size of U_x is uniformly in x bounded from below. The norm of the mapping $\Psi = (\Psi_1, \Psi_2, \Psi_3)$ is uniformly bounded in x from above.
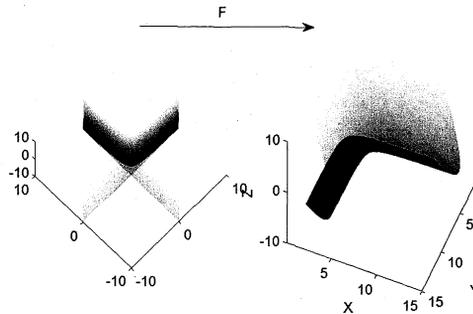

Fig. 2

In order to develop a Model-net representation of level surfaces $Y(c) = \{F = c\}$ we need a slightly extended version of Theorem 2.1.

As Proposition 2.1 above states, the singular set $\Sigma(F)$ of the product function F consists of isolated Morse critical points w_i with $F(w_i) \neq 0$, of smooth curves C_{ij}, and of isolated triple points w_{ijl}, being the intersections of the zero surfaces C_i, C_j and C_l (and of the curves C_{ij}, C_{il} and C_{jl}).

Theorem 2.2. *There exist neighborhoods W_i of the Morse points w_i of F, neighborhoods W_{ijl} of the triple points w_{ijl} and neighborhoods U_{ij} of the smooth curves $C_{ij} \setminus \bigcup_l W_{ijl}$ such that in each neighborhood W_i (respectively, W_{ijl}) there are smooth coordinates (y_1, y_2, y_3) in which F takes the form $F = y_1^2 \pm y_2^2 \pm y_3^2$ (respectively, $F = y_1 y_2 y_3$). In each neighborhood U_{ij} there are smooth functions (y_1, y_2) with the gradients of y_1, y_2 being linearly independent, such that F in this neighborhood takes the form $F = y_1 y_2$.*

Proof. This result is essentially proved in [68] (although it is not stated there in the above form). In particular, at the double curves C_{ij} we take $y_1 = F_i$, $y_2 = F_1 \ldots F_{i-1} F_{i+1} \ldots F_m$. Then $F = y_1 y_2$, and it is easy to check that outside of the neighborhoods $\bigcup_l W_{ijl}$ of the triple points w_{ijl} lying on the curve C_{ij} the gradients of y_1, y_2 are linearly independent. At the triple point w_{ijl} we take $y_1 = F_i$, $y_2 = F_j$, $y_3 = F_1 \ldots F_{i-1} F_{i+1} \ldots F_{j-1} F_{j+1} \ldots F_m$. Then $F = y_1 y_2 y_3$, and it is easy to check that the gradients of y_1, y_2, y_3 are linearly independent. Q.E.D.

Notice that Theorem 2.2 implies Theorem 2.1, with the prametrizing mappings $\Psi = (\Psi_1, \Psi_2, \Psi_3)$ being the inversions of the coordinate mappings $\{y_1(x_1, x_2, x_3), y_2(x_1, x_2, x_3), y_3(x_1, x_2, x_3)\}$ in the appropriate neighborhoods.

2.1. Model-net representation of F and of $Y(c)$

Now we are ready to describe the Model-net representation of the function F and of its level surfaces $Y(c)$. We distinguish a "representation" which preserves the full accuracy of the data, and "approximation" (treated in Section 2.2) where scalar parameters are truncated and smooth components of the data are approximated up to the allowed error.

2.1.1. Representation of F Let us introduce some convenient notations.

Definition 2.1. *The critical set $\Sigma(F)$ of F is called a "singular skeleton" of F. The union $\Sigma U(F)$ of all the neighborhoods W_i, W_{ijl}, and U_{ij} as defined in Theorem 2.2 is called a "near-singular support" of F.*

Now the the Model-net representation $\mathtt{MNR}(F)$ of the function F comprises:

1. The singular skeleton $\Sigma(F)$ of F.

2. The near-singular support $\Sigma U(F)$ of F.

3. The coordinate functions $(y_1,\ y_2,\ y_3)$ in each of the neighborhoods W_i, together with the model $F = y_1^2 \pm y_2^2 \pm y_3^2$.

4. The coordinate functions $(y_1,\ y_2,\ y_3)$ in each of the neighborhoods W_{ijl}, together with the model $F = y_1 y_2 y_3$.

5. The coordinate functions $(y_1,\ y_2)$ in each of the neighborhoods U_{ij}, together with the model $F = y_1 y_2$.

6. The function F itself outside of its near-singular support $\Sigma U(F)$.

From the point of view of Singularity Theory, what we keep at each singular point is the singular locus, the "normal form" of the singularity and the "normalizing transformation" (see [16, 18]). The normal forms are $F = y_1^2 \pm y_2^2 \pm y_3^2$ at the Morse points w_i, $F = y_1 y_2 y_3$ at the triple points w_{ijl} and $F = y_1 y_2$ along the curves C_{ij}. The normalizing coordinate transformation are given by the functions y_1, y_2, y_3 (respectively, y_1, y_2), as constructed in Theorem 2.2.

Let us stress that the $\mathtt{MNR}(F)$ represents F with a full accuracy: at this stage we just separate singular and non-singular data in the descriptions of the function F.

To clarify the structure of the data in $\mathtt{MNR}(F)$ let us assume that F is a real analytic function. In this case the segments of the singular curves C_{ij} are real analytic curves in $\mathbb{R}^3$ which can be parametrized by

$$(2.2) \qquad x_1 = \psi_{ij}^1(t),\ x_2 = \psi_{ij}^2(t),\ x_3 = \psi_{ij}^3(t),\ t \in [0,1],$$

with ψ_{ij}^1, ψ_{ij}^2, ψ_{ij}^3 real analytic functions. Together with the coordinates of the isolated singular points w_i, w_{ijl} these functions provide a complete description of the singular skeleton $\Sigma(F)$ of F. The precise geometry of the near-singular support of F is not important in our representation, so it can be described just as a δ-neighborhood of $\Sigma(F)$.

The functions (y_1, y_2, y_3) in each of the neighborhoods W_i, W_{ijl} are represented by their infinite Taylor series

$$(2.3) \qquad\qquad y_s = \sum_\alpha a_\alpha^s x^\alpha,\ s = 1, 2, 3,$$

with $\alpha = (\alpha_1,\ \alpha_2,\ \alpha_3)$ multi-indices.

To specify the representation of the functions $(y_1,\ y_2)$ in a neighborhood U_{ij} let us assume that the third coordinate x_3 is monotone on

the curve C_{ij}. Then according to the representation (2.2) of this curve, we can write

$$(2.4) \qquad y_s = \sum_\alpha b_\alpha^s(t)(\bar{x} - \psi(t))^\alpha, \ s = 1, 2,$$

with $\alpha = (\alpha_1, \ \alpha_2)$ multi-indices and

$$(\bar{x} - \psi(t))^\alpha = (x_1 - \psi_{ij}^1(t))^{\alpha_1}(x_2 - \psi_{ij}^2(t))^{\alpha_2}.$$

So what we have to store (besides the normal forms) are the coordinates of the isolated singularities and the functions $\psi_{ij}^s(t)$ in (2.2), all the Taylor coefficients a_α^s in (2.3) and all the Taylor coefficients functions $b_\alpha^s(t)$ in (2.4).

2.1.2. *Representation of $Y(c)$* Formally we can say that in order to represent a specific level surface $Y(c) = \{x, \ F(x) = c\}$ it is enough to store the representation of F and the value of the parameter c. However, this way we keep irrelevant information about F "far away" from $Y(c)$. To construct a Model-net representation $\text{MNR}(Y(c))$ of the level surface $Y(c)$ by itself we proceed as follows:

1. The regular part of $Y(c)$ outside of the near-singular support $\Sigma U(F)$ of F is stored as it is.

2. In each of the neighborhoods W_i, W_{ijl} the type of the model and the parametrizing transformations (2.1) of Theorem (2.1) above are stored.

3. In each of the neighborhoods U_{ij} the parametrizing transformations (2.1) of Theorem (2.1) above are stored. However, along the curves C_{ij} these transformations can be kept in a special form. Let us assume, as above, that the third coordinate x_3 is monotone on the curve C_{ij}. Then the transformations (2.1) can be written as

$$(2.5) \qquad x_1 = \Psi_1(y_1, \ y_2, \ x_3), \ x_2 = \Psi_2(y_1, \ y_2, \ x_3), \ x_3 = x_3.$$

Substituting here the parametrization (2.2) of the curve C_{ij} we get a one-parametric family of germs of two-dimensional coordinate transformations. We call this family a *profile of the surface $Y(c)$ along the curve C_{ij}*.

In the case of a real analytic F all these data can be kept in the form of convergent power series, as in Section 2.1.1 above.

2.2.   Model-net approximation of F and of $Y(c)$

In this step we perform a *finite-dimensional approximation* of F and Y with a prescribed accuracy. First of all, we truncate all the power series above to a finite number of terms, providing the required approximation accuracy. Secondly, we approximate with the prescribed accuracy each of the regular functions entering the $\mathrm{MBR}(F)$ and $\mathrm{MBR}(Y)$. These are: the function F itself outside of its near-singular support, the parametrizations (2.2) of the curves in the singular skeleton of F, and the the remaining Taylor coefficients functions $b^s_\alpha(t)$ in (2.4).

Basically, any conventional approximation method can be used to approximate the regular data. The most important point is that *all the singularities and near-singularities have been sorted out, so an exponential decrease of the error (in the number of the approximation's degrees of freedom) can be expected.*

In a similar way we approximate also the smooth components in the Model-net representation $\mathrm{MBR}(Y)$ of the level surface Y.

In the next section we shortly describe a method of "Taylor-nets" which is especially convenient for a representation of the regular components in the Model-net approximation. At this moment let us stress one important general feature of our approach: *we do not try to fit our local models with one another exactly. Just an agreement between the neighboring models within the prescribed accuracy is required.* This approach is usually called a "non-conforming representation", and it makes our representation very flexible. We do not discuss this feature in detail in the present paper (see [24, 66]).

§3.   Taylor-net representation of smooth functions

Taylor-net representation of a smooth function consists of a net of Taylor polynomials of a prescribed degree k (or k-jets) of this function stored at a certain grid in its domain. So we keep explicitly high order derivatives of our function at each grid point. This is a highly non-orthodox decision from the point of view of the traditional numerical analysis, in particular, because of the sensitivity of the high-order derivatives to the noise. However, explicit storage of high-order derivatives brings very serious advantages in accuracy and in processing efficiency, so we believe that an effort to overcome the stability problem is well justified. Fortunately, Taylor-net representation is strongly supported by the results of the Whitney extension theory (see [61]-[65], [15, 19, 30, 31] and references there) and by the Kolmogorov theory of optimal representation of smooth functions (see [43, 44] and many other publications).

The question of an optimal representation of smooth functions has been investigated by Kolmogorov in his work on ϵ-entropy of functional classes. *The ϵ-entropy is the logarithm of the minimal number of ϵ-balls covering a certain relatively compact set.* The problem of an optimal representation of smooth functions in Kolmogorov's setting can be shortly described as follows: *How many bits do we need to memorize a C^k-function of n variables up to a prescribed accuracy $\epsilon > 0$?* Mathematically, this is exactly the question of computing the ϵ-entropy of the subset of C^k-functions with uniformly bounded derivatives, with respect to the C^0-norm.

It was shown in [43, 44] that asymptotically, the best way to memorize a C^k-function up to the accuracy $\epsilon > 0$ is to store the coefficients of its k-th order Taylor polynomials (k-jets) at each point of some grid with the step $h = O(\epsilon^{\frac{1}{k+1}})$, taking into account that the neighboring jets are strongly correlated (this correlation is also a central tool in Whitney's extension theory).

The corresponding estimate for the ϵ-entropy of the class of C^k-functions in the space C^0 is an expression of order $\left(\frac{1}{\epsilon}\right)^{\frac{n}{k}}$.

One of the main trade-offs in any numerical approach, based on a grid representation of the data, is between the density of the grid versus the processing complexity at each grid-point. Kolmogorov's representation tends to increase as far as possible the analytic power and flexibility of the local data representation at each grid-point, strongly expanding in this way this grid-point's "covering area".

As a result, a density of the grid can be strongly reduced, while preserving the required approximation accuracy. This reduction may lead to a major efficiency improvement, especially in the problems with the large number of unknowns and parameters.

Let us give a simple (and purely illustrative) example. Assume we have to approximate a function f of 10 variables on the unit cube Q, with the accuracy of $\epsilon = 0.01$. We use a uniform grid in Q with the step-size h and a Tailor polynomial approximation at each grid-point. Assuming that the derivatives of f up to the third order, are bounded by 100, we get according to the Tailor remainder formula, that the accuracy of the first order Tailor polynomial approximation within the distance h from the grid-point is $100h^2$, while the accuracy of the third order approximation is $10h^4$. Hence to get a required overall approximation accuracy of $\epsilon = 0.01$, we must take $h = 0.01$ in the first case and $h = 0.16$ in the second case. The size of the covering area of each point increases sixteen times. Hence the number of the required grid-points in the 10-dimensional cube Q drops 16^{10} times. On the other side, the complexity

of the local representation and processing at each grid-point is roughly the number of the stored coefficients of the Taylor polynomial. For the third degree Taylor polynomial it is of order 200, while for the first degree Taylor polynomial it is 11. The 1:20 jump in local complexity is more than compensated by the 16^{10} reduction in the number of grid-points.

Also in dimensions 1 and 2 and with the representation of degree 2 a very strong accuracy improvement can be achieved. In particular, the Taylor-net discretization of the Laplace equation in the plane presented in [66] uses jets of degree two and a five-points neighbor stencil, where the Hermite fitting is applied. This scheme provides a discretization error of order h^{10}.

It is important to stress, that Taylor-net representation allows for a "point-wise processing" based on the "Jet-calculus", or "Differential Algebra" (which is just the translation of the basic analytic operations to the jet language. See [16, 18, 24, 27] for some examples). A very serious progress in this direction has been achieved by the group of the Michigan State University ([7]-[12], [33, 36], [50]-[52]) which has developed, in particular, a computer system "COSY INFINITY" implementing in a highly efficient way a good part of the "Jet-calculus library". We mention also [60] where Differential Algebra is combined with Hermite fitting (in a somewhat different sense from what is presented below) in a simulation of electron optical systems.

§4.   Hermite fitting

We return to one of our main mathematical problems: high-order numerical methods are usually very sensitive to the noise in the data. In this section we address one of the manifestations of this difficulty, namely the well-known instability of Hermite interpolation. We analyze (mostly following [72, 74]) the robustness and accuracy of the interpolation and "fitting" operators in the process of acquisition of high order data.

The most straightforward solution to the problem of reconstructing high-order derivatives of a function from its values (or from its jets of a lower order) on a certain grid is to interpolate these point-wise data with polynomials of the required degree. However, the interpolation process is well known to be highly unstable. Instead we use a "polynomial fitting" where the degree of the fitted polynomials is significantly smaller than the number of the input parameters. Accordingly, the requirement of the precise interpolation is replaced by the requirement of the least square fitting. Extending this procedure to the point-wise input data containing not only the values of the function, but also the values of some of its

derivatives (as our PDE discretization in Section 5 below requires) we get the "Hermite fitting".

While it is classically well known that replacing interpolation by fitting improves stability (see [3, 5, 6, 32, 35, 49, 54]), we are not aware of any systematic study of the sensitivity of various fitting schemes, especially in a setting where the input scheme *of a relatively high order* is fixed while the degree of an approximating polynomial varies.

In [72, 74] we've performed an initial numerical study of the "stiffness" of Hermite interpolation and fitting operators in some typical situations. Before we summarize these results, let us stress that our observations and conjectures are certainly very preliminary. Besides a need for additional experiments, there are important theoretical questions that have to be answered in order to properly interpret experimental data. First of all, this concerns the dependence of the SVD decomposition of the Hermite fitting on the choice of the norms in the polynomial and jet spaces, as well as the role of geometric scaling of the data. Some of these questions have been initially addressed in [74, 72], Section 3 (especially, Subsection 3.5.3). Below we provide some considerations, showing that in Hermit fitting of jet-data of a relatively high order a certain "rigidity" appears which reduces the dependence of the SVD decomposition on the possible choices.

We restrict ourselves to the following fitting scheme: the points $x_1, \ldots, x_s$ in $[-1.5, 1.5]$ are at equal distance $\frac{3}{s}$ from one another, symmetrically around 0. At each x_j the input Taylor polynomial (jet) of order N is given, forming the input "jet-vector" $\bar{f}_s^N$. For each degree $K \leq s(N+1) - 1$ we find a polynomial $\bar{f}^K$ of degree K whose derivatives up to order N at each grid-point provide the least square deviation from the input data $\bar{f}_s^N$. So $\bar{f}^K = G^{(K,N,s)}(\bar{f}_s^N)$, and $G^{(K,N,s)}$ is our Hermite fitting operator. In particular, for $K = s(N+1) - 1$ we get the Hermite interpolation operator $G^{(N,s)}$.

The following estimate of the accuracy of the Hermite fitting is straightforward: let f be a $K+1$ times continuously differentiable function in a neighborhood of the origin, with the $K+1$-st order derivative uniformly bounded by M. Take the input "jet-vector" $\bar{f}_s^N$ formed by the Taylor polynomials of degree N of f at the points $x_1, \ldots, x_s$. Finally, let $\bar{f}^K = G^{(K,N,s)}(\bar{f}_s^N)$ be the Hermite fitting polynomial of $\bar{f}_s^N$.

Proposition 4.1. *At each point x_i and for each $j = 0, 1, \ldots, N$ we have*

$$|f^{(j)}(x_i) - (\bar{f}^K)^{(j)}(x_i)| \leq CMh^{K+1-j}.$$

Proof. We consider the Taylor polynomial $T_K(x)$ of degree K of f at the origin. By Taylor remainder formula we find that a quadratic

deviation of T_K from $\bar{f}_s^N$ is bounded by $C_1 s M h^{K+1}$. But the deviation of $\bar{f}^K$ from $\bar{f}_s^N$ is even smaller since $\bar{f}^K$ is the least square deviation polynomial. Therefore, the same bound is valid for the deviation of $\bar{f}^K$ from $\bar{f}_s^N$ at each of the points x_i. Taking into account the natural weighting of the norm by the powers of h we get the required result.

Q.E.D.

Exactly in the same way we can prove that a similar bound holds not only at the points x_i but anywhere in the considered interval.

In order to analyze the Hermite interpolation operator $G^{(N,s)}$ we decompose the input space into the subspaces of *"increasing smoothness"*. Our goal in this decomposition is to describe the stiffness of the Hermite fitting, as applied to the jets of different *"smoothness"*. To achieve this goal we study a Singular Value Decomposition of the operator $G^{(N,s)}$, and we find the corresponding bases of the singular vectors V_i in the input space and U_i in the output space. We present below the singular values of the Hermite Interpolation operator in some specific examples.

We fix the Euclidean coefficient-wise norm in the space of polynomials of any degree. The norm in the input space of the jets is also fixed to be Euclidean with respect to all the jets coefficients. Here we have a certain choice: the norm in the space of polynomials can be chosen in a different way. However, what is essential for our applications to PDE's discretization in Section 5 below is that *the norm of the input jets at each point of the grid is the same as the norm of the output polynomial*. For input jets of a relatively high degree this requirement significantly reduces the influence of the choice of the norm on the singular value decomposition of the Hermite fitting operator.

For the degree of the output polynomial comparable with the degrees of the input jets (see tables below) we can expect this restriction to make the singular values similar to "eigenvalues" (i.e. being "almost invariant"). We have experimentally tested this conclusion in some examples. In particular, in [72, 74] we've compared the results in the tables below with the results for another norm: instead of the polynomial coefficients we use the derivatives (rescaling by the factorial of the index).

The next table presents the singular values of the Hermite fitting operators on the three-point second order input (N=2) for the maximal possible degree $K = 8$ (interpolation), and for $K = 7, 6, 5, 4, 3, 2$.

Degree	$K = 8$	$K = 7$	$K = 6$	$K = 5$	$K = 4$	$K = 3$	$K = 2$
λ_1	0.40265	0.40265	0.40265	0.40271	0.40271	0.41342	0.44721
λ_2	0.42268	0.42268	0.42268	0.42268	0.42309	0.42309	0.44721
λ_3	0.60835	0.60835	0.60835	0.60839	0.60839	0.61439	0.63246
λ_4	0.70422	0.70422	0.70424	0.70424	0.71265	0.71265	0
λ_5	2.2788	2.2788	2.2788	2.2874	2.2874	0	0
λ_6	30.055	30.055	30.122	30.122	0	0	0
λ_7	180.12	180.32	180.32	0	0	0	0
λ_8	8963.3	8963.3	0	0	0	0	0
λ_9	1.53E+05	0	0	0	0	0	0

(4.1)

Notice that the singular values in this table for different K satisfy $\lambda_i^{K''} \approx \lambda_i^{K'}$. Moreover, for $K'' > K'$ the data above exhibit an "interlacing"

$$(4.2) \qquad \lambda_{i+1}^{K'} \geq \lambda_i^{K'} \geq \lambda_i^{K''}$$

of the singular values. In [72, 74] we prove this fact using the Cauchy interlacing theorem.

The following table presents the singular values of the Hermite fitting to a zero order jets ($N = 0$) sampled on 15 points, equally spaced on the interval $[-1.5, 1.5]$, for $K \leq 15$.

(4.3)

Degree	$K = 15$	$K = 14$	$K = 12$
λ_1	0.240112293	0.240112293	0.240112293
λ_2	0.29155981	0.29155981	0.29155981
λ_3	0.826716928	0.826716928	0.826716928
λ_4	3.188015031	3.188015031	3.188015031
λ_5	17.33801918	17.33801918	17.33801918
λ_6	121.5379498	121.5379498	121.5379498
λ_7	1055.429813	1055.429813	1055.429814
λ_8	11099.79737	11099.79737	11099.79737
λ_9	139622.9286	139622.9286	139623.0265
λ_{10}	2090470.162	2090471.036	2090471.036
λ_{11}	37341260.67	37341270.12	37370046.12
λ_{12}	804613192.4	805092644.3	805092644.3
λ_{13}	21435670947	21445551410	0
λ_{14}	7.44422E+11	0	0
λ_{15}	3.86761E+13	0	0
Degree	$K = 11$	$K = 10$	$K = 9$
λ_1	0.240112293	0.240112293	0.240112293
λ_2	0.29155981	0.29155981	0.29155981
λ_3	0.826716928	0.826716928	0.826716928
λ_4	3.188015031	3.188015031	3.188015074
λ_5	17.33801918	17.33801925	17.33801925
λ_6	121.53795	121.53795	121.5384342
λ_7	1055.429814	1055.432069	1055.432069
λ_8	11099.81073	11099.81073	11119.1032
λ_9	139623.0265	139805.963	139805.963
λ_{10}	2092562.983	2092562.983	0
λ_{11}	37370046.12	0	0
λ_{12}	0	0	0
λ_{13}	0	0	0
λ_{14}	0	0	0
λ_{15}	0	0	0

In the next table we take three sampling points at equal distances. However, in each column in this table the input is different: the order N

of the input jets at the three sampling points is shown over each column of the table. This order is chosen in such a way that in each column the degree of the fitting polynomial plus one is equal to the number of the degrees of freedom of the input. So here we always have an exact Hermite interpolation.

(4.4)

Degree	$K = 21, N = 6$	$K = 18, N = 5$	$K = 15, N = 4$
λ_1	0.360437645	0.36472918	0.372406663
λ_2	0.363562265	0.370581455	0.379292565
λ_3	0.415647949	0.431788249	0.465462266
λ_4	0.427276416	0.458292554	0.487276202
λ_5	0.504989836	0.526836232	0.619940883
λ_6	0.521312988	0.616384875	0.705307924
λ_7	0.622538724	0.706851955	**2.272098427**
λ_8	0.705455456	**2.271932616**	29.72196737
λ_9	**2.271904922**	29.71456651	179.6045381
λ_{10}	29.71383269	179.5979873	8811.484996
λ_{11}	179.5969907	8806.90793	142455.3704
λ_{12}	8806.142731	142244.144	17859407.86
λ_{13}	142193.1041	17763671.94	254365662.9
λ_{14}	17619862.71	253344356.1	53631680658
λ_{15}	253274918.6	53145607750	1.65943E+12
λ_{16}	52860538116	1.57545E+12	0
λ_{17}	1.55338E+12	5.59641E+14	0
λ_{18}	5.39069E+14	1.21428E+16	0
λ_{19}	1.20845E+16	0	0
λ_{20}	5.80461E+18	0	0
λ_{21}	2.59599E+20	0	0
Degree	$K = 12, N = 3$	$K = 9, N = 2$	$K = 6, N = 1$
λ_1	0.38203034	0.402650826	0.4315
λ_2	0.397468701	0.422678113	0.5127
λ_3	0.495501188	0.608347055	0.7552
λ_4	0.594626284	0.704219496	**2.3085**
λ_5	0.714442477	**2.278831052**	30.9946
λ_6	**2.273008659**	30.0548157	180.7111
λ_7	29.74871692	180.1176427	0
λ_8	179.6340419	8963.263908	0
λ_9	8856.993934	153126.6911	0
λ_{10}	144423.3667	0	0
λ_{11}	18498528.84	0	0
λ_{12}	255469801.1	0	0
λ_{13}	0	0	0
λ_{14}	0	0	0
λ_{15}	0	0	0
λ_{16}	0	0	0
λ_{17}	0	0	0
λ_{18}	0	0	0
λ_{19}	0	0	0
λ_{20}	0	0	0
λ_{21}	0	0	0

These examples suggest that roughly till the index $N + 2$ the singular values of the Hermite interpolation operator $G^{(N,s)}$ are of "order of one". Then there is a relatively narrow interval of indices i where the growth of λ_i looks exponential, but with a moderate ratio. Finally the ratio of the exponential growth of λ_i jumps up.

As the tables (4.1) and (4.3) show, the same pattern holds also for the largest singular value (the norm) of the Hermite fitting operator, as the input is fixed while the degree of the fitting polynomial varies.

In [72, 74] we provide more examples of various Hermite interpolation schemes. In all these examples the above observation holds.

It is relatively easy to explain the first part of the above pattern in singular values: the polynomial of an order $K \leq N$ can be reconstructed from neighboring jets of the same or higher order just ba an "averaging". Let us stress once more that this "normalization" is provided by our choice of the same norm for the input jets and for the output polynomial.

However, it is exactly the "intermediate range" where we can hope to get a high-order fitting, still preserving a low noise sensitivity. So it would be important to find a mathematical explanation of the behavior of λ_i in the intermediate and the final ranges. We hope that the following simple results (which overlap with some classical and modern results in the same direction - see [4, 5, 6, 21, 54]) provide at least a partial explanation:

Theorem 4.1. *Let $x_1, \ldots, x_s$ be a uniform grid in $[-1, 1]$. Let $Y = (y_1, \ldots, y_s)$, with $\sum_{j=1}^{s} y_j^2 = 1$ be a zero order interpolating data. Let $P_Y^d(x)$ be the polynomial of degree $d \leq s$ providing the least square deviation from Y. Then $M = \max_{[-1,1]} |P_Y^d(x)|$ does not exceed $2T_d(\frac{2+\frac{d}{s}}{2-\frac{d}{s}})$ where $T_d(x) = \cos(d \arccos(x))$ is the d-th Chebyshev polynomial. In particular, $M \leq 2 \exp(5d\sqrt{\frac{d}{s}})$.*

Proof. Since $\sum_{j=1}^{s} y_j^2 = 1$ we have $|y_j| \leq 1$ for each j. The square deviation of Y from the identical zero polynomial is equal to $\sum_{j=1}^{s} y_j^2 = 1$. The polynomial $P_Y^d(x)$ provides the least square deviation from Y and hence $\sum_{j=1}^{s} (y_j - P_Y^d(x_j))^2 \leq 1$. We conclude that $|y_j - P_Y^d(x_j)| \leq 1$ for each j, and hence $|P_Y^d(x_j)| \leq 2$ for each j.   Q.E.D.

Now we use the following simple lemma (compare [4, 5, 21, 29, 55]):

Lemma 4.1. *Let $P(x)$ be a polynomial of degree d. If $|P(x_j)| \leq 1$, $j = 1, \ldots, s$ then $M = \max_{[-1,1]} |P(x)|$ does not exceed $T_d(\frac{2+\frac{d}{s}}{2-\frac{d}{s}})$ where $T_d(x) = \cos(d \arccos(x))$ is the d-th Chebyshev polynomial. In particular, $M \leq \exp(5d\sqrt{\frac{d}{s}})$.*

Proof. Let for a subset A of $[-1, 1]$ $M(\epsilon, A)$ denote the minimal number of ϵ-intervals covering A. Let $\Delta \subset [-1, 1]$ be the set on which $|P(x)| \leq 1$. Δ consists of at most $\frac{d}{2}$ intervals. Hence we have the following bound for $M(\epsilon, \Delta)$:

$$(4.5) \qquad M(\epsilon, \Delta) \leq \frac{1}{\epsilon}\mu(\Delta) + \frac{d}{2},$$

where $\mu(\Delta)$ is the Lebesgue measure of Δ. Now by assumption the set Δ contains the uniform grid $\{x_j\}$, $j = 1, \ldots, s$ and hence $M(\frac{2}{s}, \Delta) \geq s$. We conclude that

$$(4.6) \qquad\qquad \mu = \mu(\Delta) \geq 2 - \frac{d}{s} \ .$$

Now we use the following Remez inequality (see [55, 29]):

Let $P(x)$ be a polynomial of degree d. Then

$$(4.7) \qquad\qquad \max{}_{[-1,1]}|P(x)| \leq T_d\big(\frac{4 - \mu(\Delta)}{\mu(\Delta)}\big),$$

where $\Delta \subset [-1, 1]$ is the set in $[-1, 1]$ on which $|P(x)| \leq 1$. $T_d(x) = \cos(d\,arccos(x))$ is the d-th Chebyshev polynomial.

As a corollary we get (see [29], Corollary 3.2):

$$(4.8) \qquad\qquad \max{}_{[-1,1]}|P(x)| \leq \exp(5d\sqrt{2 - \mu(\Delta)}).$$

Now combining (4.12) and (4.13) with (4.11) we obtain

$$(4.9) \qquad\qquad M = \max{}_{[-1,1]}|P(x)| \leq T_d\big(\frac{2 + \frac{d}{s}}{2 - \frac{d}{s}}\big).$$

In particular,

$$(4.10) \qquad\qquad M \leq \exp\big(5d\sqrt{\frac{d}{s}}\big).$$

This completes the proof of the lemma and of Theorem 4.1.   Q.E.D.

Corollary 1. *The maximum norm of the Hermite fitting of order d on the uniform grid of s points in $[-1, 1]$ does not exceed $2T_d(\frac{2+\frac{d}{s}}{2-\frac{d}{s}}) \leq 2\exp(5d\sqrt{\frac{d}{s}})$.*

Notice that for d fixed and s growing the both bounds in Theorem 4.1 and Corollary 1 above tend to 2. This justifies to some extent our expectation that the robustness of the Hermite fitting of order d on the uniform grid of s points in $[-1, 1]$ increases with s.

Because of the relation between the norm of the Hermite fitting of different orders, and the corresponding singular values of the Hermite interpolation, this corollary provides also a partial explanation to the accelerating growth of the singular values observed above.

In a similar way we can prove the following extension of Theorem 4.1:

Theorem 4.2. *Let* $x_1, \ldots, x_s$ *be a uniform grid in* $[-1, 1]$. *Let* $J = (j_1, \ldots, j_s)$, *with* $\sum_{l=1}^{s} ||j_l||^2 = 1$ *be a k-th order interpolating data (the scalar product on the jet space being coordinate-wise). Let* $P_J^d(x)$ *be the polynomial of degree* d, $k \le d \le (k+1)s - 1$, *providing the least square deviation from* J. *Then* $M = \max_{[-1, 1]} |P_J^d(x)|$ *does not exceed* $3 + 2(\frac{1}{s})^k T_d(\frac{2 + \frac{d}{s}}{2 - \frac{d}{s}})$. *In particular,* $M \le 3 + 2(\frac{1}{s})^k \exp(5d\sqrt{\frac{d}{s}})$.

Proof. Since $\sum_{l=1}^{s} ||j_l||^2 = 1$ we have $||j_l||^2 \le 1$ for each j. The square deviation of J from the identical zero polynomial is equal to 1 and hence the square deviation of $P_J^d(x)$ from J does not exceed 1. Exactly as in the proof of Theorem 3.8.1 we conclude that the absolute value of each of the derivatives up to order k of $P_J^d(x_l)$, $l = 1, \ldots, s$ does not exceed 2.

Now we apply Lemma 4.1 to the k-th derivative $(P_J^d)^{(k)}(x)$ of $P_J^d(x)$. We conclude that the maximum of its absolute value on $[-1, 1]$ does not exceed $A = 2T_d(\frac{2 + \frac{d}{s}}{2 - \frac{d}{s}})$. Next, the $k - 1$-st derivative of $P_J^d(x)$ is bounded by 2 at each grid-point $x_1, \ldots, x_s$, and therefore its maximum on $[-1, 1]$ is at most $2 + \frac{1}{s}A$. For the maximum of the $k - 2$-d derivative we get a bound $2 + \frac{1}{s}(2 + \frac{1}{s}A) = 2 + 2\frac{1}{s} + (\frac{1}{s})^2 A$. Continuing in the same way (and assuming $d, s \ge 3$) we finally get $M = \max_{[-1, 1]} |P_J^d(x)| \le 3 + (\frac{1}{s})^k A$. This completes the proof of Theorem 4.2. Q.E.D.

Assuming that our bounds are reasonably sharp, the immediate conclusion would be that passing to the k-th order jet data improves the fitting stability by a factor $(\frac{1}{s})^k$, independently of d.

Remark. The result of Lemma 4.1 above has been extended to arbitrary discrete subsets in higher dimension in [69], where a general "discrete Remez inequality" is proved. We plan to present the corresponding generalizations of the above results on Hermite fitting to higher dimensions and to arbitrary discrete "sampling subsets" in [72].

The results and the numerical observations in [71, 74] open a number of seemingly interesting mathematical questions:

1. Do the bounds of Theorem 4.1 (Theorem 4.2) accurately reflect the true behavior of the singular values? Our experimental data provide a certain initial support for the approach of Theorems 4.1 and 4.2.

2. In the process of a Singular Value Decomposition we produce a sequence of orthogonal piecewise-polynomials V_i and a sequence of orthogonal polynomials U_i (orthogonality - with respect to the coordinate-wise scalar product). What are their properties and their relation to the classical orthogonal polynomials?

3. Is it true that the "smoothness" of the piecewise-polynomials V_i (in the Whitney sense - see [61]-[65]) is measured by the size of the singular value λ_i?

4. One can hope that the estimates above provide an approach to the following general problem:

Assume that we can measure the values of a C^k-function f (with the C^k-norm of f explicitly bounded) at any required point with a known accuracy. What is the optimal strategy to reconstruct a Taylor-net representation of f and what accuracy of reconstruction can be achieved?

Assume in addition that we can interpret the measurement errors as a random noise. What is the answer in this case?

Notice that the approach of [69] provides a partial answer to this question in terms of the geometry of the sampling set.

Already at this initial stage the numerical and theoretical results above show that if we replace Hermite interpolation by Hermite fitting, significantly reducing the degree of the fitting polynomial with respect to the maximal possible, the stiffness of the Hermite fitting operator is reduced in many orders of magnitude.

§5. Jet discretization of PDE's

In this section we explain, following [71, 73, 74] how to write down a system of ordinary differential equations which describes the evolution of the Taylor-net data for the solution of an evolution PDE. For other discretization methods for solving different types of PDE's, based on Taylor-nets, and for a general background see [14, 40, 41, 42, 66], [7]-[12], [33, 36], [50]-[52].

Assuming that the equation and its solutions are sufficiently smooth, we reduce the initial PDE to a relatively compact system of ordinary differential equations. The unknowns of this system are the (time depending) Taylor coefficients of the solution with respect to the space variable, up to a certain fixed order N, computed at a certain fixed space-grid $x_1, \ldots, x_s$. More accurately, we represent the solution as:

$$(5.1) \qquad u(x,t)|_{x=x_i} \approx P_i(x,t) = \sum_{j=0}^{N} \frac{u_i^j(t)}{j!}(x - x_i)^j.$$

The derivatives $u_i^j(t), \quad i = 1, \ldots, s, \quad j = 0, \ldots, N$, of the Taylor polynomials $P_i(x,t)$, form the unknowns of the constructed system of the ordinary differential equations. The initial values $u_i^j(0), \quad i =$

$1, \ldots, s$, $j = 0, \ldots, N$ at the moment $t = 0$ are provided by the Taylor coefficients at x_i of the initial data $u(x, 0) = \psi(x)$ for the original PDE.

The following general example illustrates the construction: let us assume that the original evolution PDE has the form $u_t = F(u, u_x)$. Then we get

$$(5.2) \quad \frac{d}{dt} u_i^0 = F(u, u_x) = F(u_i^0, u_i^1),$$

$$\frac{d}{dt} u_i^1 = \frac{d}{dt} \frac{\partial}{\partial x} u = \frac{\partial}{\partial x} \frac{\partial}{\partial t} u$$

$$(5.3) \qquad = \frac{\partial}{\partial x} F(u, u_x) = F_u(u_i^0, u_i^1) u_i^1 + F_{u_x}(u_i^0, u_i^1) u_i^2,$$

etc. Continuing this way, we express the time derivatives of each of the unknown Taylor coefficients $u_i^j(t)$, $i = 1, \ldots, s$, $j = 0, \ldots, N$ through the algebraic expressions of these coefficients.

The problem is that in the expression for the time derivative of the last Taylor coefficients $u_i^N(t)$, $i = 1, \ldots, s$, the next derivative of u of order $N + 1$ with respect to x appears. Such derivatives are not the part of our unknowns, so to "close up" the ODE system under construction, we have to express them through the derivatives up to the order N at the neighboring points. We do this via the Hermite fitting. Let us consider two examples.

5.1. First order wave equation $u_t = u_x$

We start with a toy model, of the first order wave equation

$$(5.4) \qquad\qquad u_t = u_x.$$

We assume that the solution $u \in C^p([0, 2\pi] \times [0, T])$ is p times smooth, $p = N + l$, and we consider periodic boundary conditions.

The derivatives of u satisfy the equation

$$(5.5) \quad \frac{d}{dt} \frac{\partial u}{\partial x} = \frac{\partial^2 u}{\partial t \partial x} = \frac{\partial^2 u}{\partial x \partial t} = \frac{\partial}{\partial x} \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

Similarly, in general we obtain

$$(5.6) \quad \frac{d}{dt} \frac{\partial^j u}{\partial x^j} = \frac{\partial^{j+1} u}{\partial x^{j+1}}, \quad j = 1, 2, \cdots, N$$

As it was explained above, the derivatives $u_i^j(t)$, $i = 1, \ldots, s$, $j = 0, \ldots, N$, of the Taylor polynomials $P_i(x)$ of the solution $u(x,t)$ with respect to x at the grid-points $x_1, \ldots, x_s$ form the unknowns of the constructed system of the ordinary differential equations. According to the above calculation, this system takes a very simple form:

$$(5.7) \qquad \frac{d}{dt} u_i^j = u_i^{j+1}, \quad i = 1, \ldots, s, \quad j = 0, \ldots, N.$$

Any solution u of (5.13) satisfies these equations exactly, but the resulting system is not closed: the time derivative of u_i^N is u_i^{N+1}, which is not a part of our unknowns. Therefore we replace u^{N+1} with an approximation obtained as a result of an interpolation of the neighboring derivatives up to order N. The new system obtained in this way is closed and thus solvable in the usual sense.

The interpolation scheme can be chosen in various ways. In [73, 74, 71] numerical experiments have been conducted for a three point Hermite interpolation, with only two nearest neighbor grid-points used (here h denotes the step-size of the grid):

$$u^{(N+1)}(x,t) \approx \frac{-105}{8h^3} u^{(N-2)}(x-h,t) + \frac{-33}{8h^2} u^{(N-1)}(x-h,t)$$

$$+ \frac{-3}{8h} u^{(N)}(x-h,t) + \frac{-18}{h^2} u^{(N-1)}(x,t) + \frac{105}{8h^3} u^{(N-2)}(x+h,t)$$

$$(5.8) \qquad\qquad + \frac{-33}{8h^2} u^{(N-1)}(x+h,t) + \frac{3}{8h} u^{(N)}(x+h,t).$$

In matrix language this is expressed through a multiplication of the coordinates of the three neighboring jets by the matrix D defined as follows:

$$(5.9) \quad D = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ -\frac{105}{8h^3} & -\frac{33}{8h^2} & -\frac{3}{8h} & 0 & -\frac{18}{h^2} & 0 & \frac{105}{8h^3} & -\frac{33}{8h^2} & \frac{3}{8h} \end{pmatrix}$$

The right-hand side matrix of the resulting system of ODE's (see [73, 74, 71]) is block-diagonal, with the main blocks corresponding to the grid-points x_i. Each block has ones over the main diagonal, and the last 1 is replaced by the matrix D.

According to Theorem 5.1 below, the discretization error for $N = 2$ is of order h^6. Numerical experiments in [74, 71] support this expected accuracy. Notice that in each specific case the eigenvalues of the right-hand side matrix were purely imaginary. However, the general question of stability remains open.

5.2. Heat equation $u_t = u_{xx}$

We consider this equation on the segment $x \in [0, 2\pi]$ with periodic boundary conditions. As above, we have

$$(5.10) \qquad \frac{d}{dt}\frac{\partial u}{\partial x} = \frac{\partial^2 u}{\partial t \partial x} = \frac{\partial^2 u}{\partial x \partial t} = \frac{\partial}{\partial x}\frac{\partial u}{\partial t} = \frac{\partial^3 u}{\partial x^3}$$

Similarly, in general we obtain

$$(5.11) \qquad \frac{d}{dt}\frac{\partial^j u}{\partial x^j} = \frac{\partial^{j+2} u}{\partial x^{j+2}}, \quad j = 1, 2, \cdots, N$$

So at each grid-point we have to express through the neighbors two derivatives of u - of orders $N + 1$ and $N + 2$. The corresponding matrix of the Hermite interpolation takes the form

$$(5.12)$$
$$D = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ -\frac{105}{8h^3} & -\frac{33}{8h^2} & -\frac{3}{8h} & 0 & -\frac{18}{h^2} & 0 & \frac{105}{8h^3} & -\frac{33}{8h^2} & \frac{3}{8h} \\ \frac{72}{h^4} & \frac{39}{2h^3} & \frac{3}{2h^2} & -\frac{144}{h^4} & 0 & -\frac{36}{h^2} & \frac{72}{h^4} & -\frac{39}{2h^3} & \frac{3}{2h^2} \end{pmatrix}$$

Also here the right-hand side matrix of the resulting system of ODE's is block-diagonal, with the main blocks corresponding to the grid-points x_i. Each block has ones shifted twice over the main diagonal, and the last two of them are replaced by the matrix D.

According to Theorem 5.1 below, the discretization error for $N = 2$ is of order h^5. Numerical experiments in [74, 72] support this expected accuracy. Notice that in each specific case the eigenvalues of the right-hand side matrix had negative real part. However, the general question of stability also here remains open.

In [73, 74, 72] some additional cases have been analyzed, including some integro-differential equations, the wave equation, and the (non-linear) Burgers equation. The last case we present below.

5.3. Burgers equation $u_t = uu_x$

The suggested approach works also for non-linear equations. In this section we consider, following [74, 72], an example of the Burgers equation *in a smooth region, away of the shock-waves*. A much more challenging case of a shock-wave neighborhood will be described in Section 6 below (where we also touch shortly the most difficult part: a birth of the shock-wave).

As above, we have

$$(5.13) \qquad \frac{d}{dt}\frac{\partial u}{\partial x} = \frac{\partial}{\partial x}\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}(uu_x) = (u_x)^2 + uu_{xx} \ ,$$

and similarly

$$(5.14)$$
$$\frac{d}{dt}\frac{\partial^2 u}{\partial x^2} = \frac{\partial^2}{\partial x^2}\frac{\partial u}{\partial t} = \frac{\partial^2}{\partial x^2}(uu_x) = \frac{\partial}{\partial x}[(u_x)^2 + uu_{xx}] = 3u_x u_{xx} + uu_{xxx} \ .$$

Therefore, our system (for the jet order $N = 2$) takes the form

$$(5.15)$$
$$\frac{d}{dt}u_i^0 = u_i^0 u_i^1 \ , \frac{d}{dt}u_i^1 = (u_i^1)^2 + u_i^0 u_i^2 \ , \frac{d}{dt}u_i^2 = 3u_i^1 u_i^2 + u_i^0 u_i^3 \ , \ i = 1,\ldots,s \ .$$

To get our system in closed form it remains to replace the term u_i^3 by a combination of the derivatives up to the order 2 at the neighboring points. In particular, if we use, as above, the three point Hermite interpolation, we have

$$u_i^3 = \frac{-105}{8h^3}u_{i-1}^0 + \frac{-33}{8h^2}u_{i-1}^1 + \frac{-3}{8h}u_{i-1}^2 + \frac{-18}{h^2}u_i^1 +$$
$$(5.16) \qquad\qquad \frac{105}{8h^3}u_{i+1}^0 + \frac{-33}{8h^2}u_{i+1}^1 + \frac{3}{8h}u_{i+1}^2 \ .$$

As above, the discretization error (until the solution remains smooth!) is of order h^6. Numerical experiments in [74, 72] confirm this estimate and a general stability of the scheme. However, in [74, 72] we did not discuss how to adapt our method to the formation and tracking of singularities. An initial step in this direction is presented in Section 6 below.

5.4. Discretization error

Let us return to a Taylor-net discretization of a general PDE. In general, the evolution equations for the Taylor coefficients in Taylor-net representation of the unknown function u are obtained, as above, algebraically by the "Jet extension" of the original equation, combined with the Hermite fitting. The last is used to express the derivatives of the orders higher that N, which naturally arise in the Jet extension, through the Taylor coefficients up to order N at the neighboring grid-points.

One of the main features of the proposed discretization scheme is that its order of accuracy (expressed as the power of the grid size h) may be significantly higher than the order N of the Taylor polynomials

explicitly used. Such a high order of accuracy is achieved via the use of the high order Hermite fitting with the neighboring grid-points. (This is true also for other discretization schemes based on Taylor-nets - see [14, 66]. The reason is that having at each grid-point a jet of a relatively high degree, we possess even over a small neighbor stencil enough degrees of freedom to cancel low order terms in the discretization error).

In [74, 71] we study the discretization error of the scheme discussed in the present paper, and prove the following result:

Theorem 5.1. *Let in the equation* $u_t = F(x, u, u_x, \ldots, \frac{\partial^n u}{\partial x^n})$ *the function F be q times continuously differentiable with respect to all its variables, $q = N + n + K + 1$. Then the N-th order Taylor discretization scheme as above, with the Hermite fitting of order $K \le l(N + 1) - 1$ at the stencil of l neighbors, has a discretization error of order h^r, where h is the step-size of the scheme and $r = K - N - n + 1$. In particular, for the equations $u_t = u_x$ and $u_t = uu_x$ and for $N = 2$, $l = 3$, $K = 8$, we get $r = 6$ while for the equation $u_t = u_{xx}$ the order r is equal to 5.*

In many situations we would like to sacrifice some of the accuracy of the scheme in order to make it more robust. To achieve this goal, we choose, using the results and the tables of Section 4 above, the neighbor stencil and the fitting order in such a way that the norm of our Hermite fitting operator is relatively small. Notice that only this operator brings into the constructed system of ODE's the negative powers of the grid-step h, as well as the "large singular values" of the interpolation.

In particular, in most of our computations in [74, 71] we use jets of order 2, so $N = 2$, and a full order (K=8) three points Hermite interpolation. Consider now, for example, the equation $u_t = u_x$. According to Theorem 5.1 the discretization error is of order h^6. If we replace the Hermite interpolation by a fitting of degree, say, $K = 6$, we still get the discretization error of order h^4, which is better than h^3 provided by second-order jets in direct computations. On the other hand, the table (4.6) in Section 4 shows that we have reduced the norm of the Fitting matrix (and hence of the right hand side matrix in the resulting ODE) in roughly two orders of magnitude.

As far as the stability of our discretization scheme is concerned, in all the specific examples of its numerical implementation for linear PDE's considered in [74, 71] the following fact was confirmed numerically: the eigenvalues of the right hand side matrix have negative real parts. We could not prove this property in general, so the stability of our scheme remains an important open question.

We hope that a proper development of the results of Sections 4 on stiffness of the Hermite fitting, and of Section 5 on jet-discretization of

evolution PDE's will allow one to construct robust and noise-insensitive
high accuracy numerical schemes.

§6. Burgers equation: near a shock-wave

In this section we prove some initial results concerning a Model-net
discretization of the Burgers equation near a shock-wave. Our purpose
is just to demonstrate how a Model-net approach leads to a natural
high-order numerical scheme of a high theoretical accuracy. We do not
try here to compare this scheme with known methods, so we give only
a few references from a huge literature on the subject.

For the structure of the shock-waves and a general overview of com-
putational approaches and problems see [46, 47, 48]. For a survey of ef-
ficient modern computational methods see [57], and for a more detailed
presentation see [53, 45] and references therein. A detailed treatment of
shock-waves from the point of view of Singularity Theory can be found in
[20] (see also [13]). The approach presented below can be also considered
as a form of the "tracking method".

Let us start with a well-developed shock-wave: the most challenging
case of a formation of singularity we shall shortly discuss later.

In this case the solution $u(x,t)$ is a piecewise-smooth function. It
has jumps at the points $z_j(t)$, $j = 1, \ldots, q$, which are just the positions of
each of the shock-waves at the moment t, and the solution preserves the
smoothness of the initial data between the jumps ([46]). The evolution
of $u(x,t)$ in time is governed by the usual equation $u_t = uu_x$ at smooth
points, while each jump point $z(t)$ satisfies

$$(6.1) \qquad \frac{d}{dt}z(t) = \frac{1}{2}(u_-(z,t) + u_+(z,t)).$$

Here $u_-(z,t)$ and $u_+(z,t)$ denote the limit values of u from the left and
from the right at z, respectively.

Accordingly, we use a Model-net representation $\mathtt{MNR}(u(x,t))$ of the
solution $u(x,t)$ (along the lines presented in Section 2). In this special
case the $\mathtt{MNR}(u(x,t))$ comprises the coordinates $z_j(t)$, $j = 1, \ldots, q$ of
the jump-points in $[0,1]$, a fixed grid x_i, $i = 1, \ldots, s$ in $[0,1]$, and the
collection of Taylor polynomials

$$(6.2) \qquad u(x,t) \approx P_i(x,t) = \sum_{j=0}^{N} \frac{u_i^j(t)}{j!}(x - x_i)^j$$

representing the solution $u(x,t)$ at each of the *regular* grid-points x_i which are not the nearest neighbors of the jump-pints. At each *singular* grid-point x_l (which is the nearest neighbors of the jump-pint $z_m(t)$) we keep a somewhat more complicated *singular local model*. Each of these singular models comprises the coordinate $z_m(t)$ of its jump-point and *two Taylor polynomials*

$$u(x,t)_- \approx P_{l-}(x,t) = \sum_{j=0}^{N} \frac{u_{l-}^j(t)}{j!}(x - x_l)^j \ ,$$

$$(6.3) \qquad u(x,t)_+ \approx P_{l+}(x,t) = \sum_{j=0}^{N} \frac{u_{l+}^j(t)}{j!}(x - x_l)^j \ .$$

representing the values of $u(x,t)$ to the left and to the right of the jump-point z_m.

Now the unknowns of the constructed system of the ordinary differential equations are:

1. The derivatives $u_i^j(t)$, $j = 0, \ldots, N$, of the Taylor polynomials $P_i(x,t)$ in (6.2) at the regular grid-points x_i.

2. The derivatives $u_{l\pm}^j(t)$, $j = 0, \ldots, N$, of the Taylor polynomials $P_{l\pm}(x,t)$ in (6.3) at the singular grid-points x_l.

3. The coordinates $z_j(t)$, $j = 1, \ldots, q$ of the jump-points in $[0,1]$.

The system is constructed as follows:

1. For regular points x_i it is the same as in Section 5 above. In particular, for N=2 and the three points central Hermite interpolation we get

$$(6.4) \quad \frac{d}{dt}u_i^0 = u_i^0 u_i^1 \ , \frac{d}{dt}u_i^1 = (u_i^1)^2 + u_i^0 u_i^2 \ , \frac{d}{dt}u_i^2 = 3u_i^1 u_i^2 + u_i^0 u_i^3 \ ,$$

with the term u_i^3 replaced by

$$u_i^3 = \frac{-105}{8h^3}u_{i-1}^0 + \frac{-33}{8h^2}u_{i-1}^1 + \frac{-3}{8h}u_{i-1}^2 + \frac{-18}{h^2}u_i^1 +$$
$$(6.5) \qquad \frac{105}{8h^3}u_{i+1}^0 + \frac{-33}{8h^2}u_{i+1}^1 + \frac{3}{8h}u_{i+1}^2 \ .$$

Notice that for the regular point x_i both the points x_{i-1} and x_{i+1} belong to the same continuity interval of u.

2. For singular points x_l the equations (6.4) remains the same for the derivatives $u_{l\pm}^j(t)$. However, in a singular case the term u_{l-}^3 is replaced by the interpolated value $HF_-^3(x_l)$. Here HF_-^3 is a result of the Hermite interpolation of the third derivative, taken over a certain stencil of the grid-points *to the left of x_l*. At the point x_l itself the jet $P_{l-}(x,t)$ is used in this interpolation.

Accordingly, the term u_{l+}^3 is replaced by the interpolated value $HF_+^3(x_l)$, where HF_+^3 is a result of the Hermite interpolation of the third derivative, taken over a certain stencil of the grid-points *to the right of x_l*, with the jet $P_{l+}(x,t)$ used at x_l itself.

3. Finally, to get the right-hand side of the differential equation for the coordinate of the jump-point, we have to express the limit values $u_-(z_m,t)$ and $u_+(z_m,t)$ of $u(x,t)$ from the left and from the right at z_m through the values of the variables u_i^j and z_m. To achieve this goal, we construct the k-degree Hermite interpolation (fitting) polynomials HF_-^0 and HF_+^0 of the function u to the left and to the right of z_m. For HF_-^0 the interpolation (fitting) is taken over a certain stencil of the grid-points *to the left of x_l*. At the point x_l itself the jet $P_{l-}(x,t)$ is used in this interpolation. For HF_+^0 the interpolation (fitting) is taken over a symmetric stencil of the grid-points *to the right of x_l*. At the point x_l itself the jet $P_{l+}(x,t)$ is used.

Thus the equation for z_m takes the form

$$(6.6) \qquad \frac{d}{dt} z_m = \frac{1}{2}(HF_-^0(z_m) + HF_+^0(z_m)).$$

Notice that the right hand side of this equation is a polynomial both in u_i^j and in z_m.

6.1. Discretization accuracy

In this section we prove the following result:

Theorem 6.1. *Consider the Burgers equation $u_t = uu_x$ near a formed shock-wave of its solution. Then the N-th order Taylor-net discretization scheme as above, with the Hermite fitting of order $K \le l(N+1) - 1$ at the stencil of l neighbors for the $N+1$-st derivative and with the Hermite fitting of order $k = K - N - 1 \le p(N+1) - 1$ at the stencil of p neighbors (to the left or to the right) for the limit values u_- and u_+, has a discretization error of order h^r, where h is the step-size of the scheme and $r = K - N$. In particular, for $N = 2$, $l = 3$, $K = 8$, $k = 6$, $p = 3$ we get $r = 6$.*

Proof. The equations (6.4) and (6.6) would remain precise, if we substituted to the right-hand side the precise values of of the $N + 1$-st derivative of u (of u_- and u_+, respectively). However, we replace these values with their expression through the neighboring points via Hermite Fitting. According to Proposition 4.1 the maximal error in this computation for the $N+1$-st derivative of u is of order h^r. Now the value of $k = r - 1$ in Theorem 6.1 was chosen to provide the same accuracy for the values of u_- and u_+.

The right-hand sides of the equations (6.4) and (6.6) are Lipschitzian in their arguments. Hence the overall error in the right hand side of the system is of order h^r. We've shown that if u is a true solution of the Burgers equation then the right hand side of our discretized sistem of ODE's differs from the true values of the time derivatives not more than to h^r. This completes the proof of the theorem. Q.E.D.

6.2. Birth of a singularity

Shock-wave appears in a solution of the Burgers equation when the characteristics cross one another for the first time. Generically this happens at cuspidal points of the envelop of the characteristics (see [46, 20]).

The standard method of characteristic lines for a quasi-linear Burgers equation produces a regular system of ODE's along the characteristics in the "phase-space". Then the original solution can be reconstructed via the projection onto the geometric space and a certain discontinuous selection of the inverse projection branches, governed by the differential equation (6.4). Notice that generically at the birth-point of a shock-wave the projection has a Whitney cusp singularity ([20]).

Our Model-net discretization method can be extended to this situation as follows: we add a special model describing the evolution at time of the solution $u(x, t)$ as we approach the birth-point of a shockwave. The parameters of this model are essentially the Taylor coefficients of the solution in the phase-space. Since the corresponding equations in the phase-space are regular, a simple polynomial system of ODE's for these Taylor coefficients can be produced in the same way as above. The solution $u(x, t)$ itself is given then by an explicit analytic expression (involving solutions of third degree polynomial equations). Finally, the evolution equation for the shock position is obtained from equation (6.6) by substituting the appropriate expressions for u_- and u_+. We plan to provide a more detailed description as well as simulation results separately.

§7. More on Model-nets

Let us mention very shortly some other mathematical problems related to applications of Model-nets in computations.

7.1. Model-net data acquisition from measurements

The Model-net representation contains "geometric parameters" of the local models (their singular skeleton - see Section 2 above). These parameters enter the data in a *non-linear way*. The simplest example is provided by piecewise smooth functions, like the solutions of Burgers equations considered above. Nonlinear parameters here are the coordinates of the jump points. This fact leads to a challenging mathematical problem of a non-linear reconstruction of the data from measurements, which turns out to be closely related to Complex Analysis, Moment Theory and Semi-algebraic Geometry. See [57, 22, 59, 28, 56] and references therein for discussion of various approaches to this problem. Some new connections of this problem with Fuchsian differential equations, as well as with Padé approximations and holonomic combinatorial systems were recently found in [1, 2, 38, 39].

7.2. Some implementations of Model-nets

One of the most important practical problems related to application of Model-nets concerns representation and processing of digital images. Such a representation has been suggested in [17, 26]. See also [23] and references therein, where a general analysis of the performance of edges-based methods in images representation is given, as well as [37]. However, in general the "geometric" methods, as for today, suffer from an inability to achieve a full visual quality for high resolution photo-realistic images of the real worlds. *In fact, the mere possibility of a faithful capturing such images with geometric models presents one of important open problems in Image Processing, sometimes called "the vectorization problem".*

Let us express our strong belief that a full visual quality Model-net representation for high resolution photo-realistic images of the real worlds is possible. As achieved, it promises to bring a major advance in image compression and capturing, in particular, via the approach of [2, 28, 38, 39, 22, 56, 59] and of the present paper.

Another implemented application of Model-nets concerns the motion planning problem in Robotics. It is presented in [24, 25, 27]. The Model-nets are used here to compactly represent and process the free configuration space of the moving system. We expect that the approach of the present paper can be combined with the approach of [24, 25, 27]

providing, in particular, a framework for an efficient dynamical motion planning.

References

[1] D. Batenkov, Reconstruction of picewise D-finite functions, preprint, Weizmann Inst., 2008.

[2] D. Batenkov, N. Sarig and Y. Yomdin, Robust non-linear reconstruction of piecewise-smooth functions from their Fourier data, in preparation.

[3] I. S. Berezin and N. P. Zhidkov, Computing Methods, Addison-Wesley Publising Company, 1965.

[4] S. Bernstein, Sur une propriété des polynomes, Proc. Kharkov Math. Society, Serie 2, **14** (1913), 1–6.

[5] S. Bernstein, Sur une formule d'interpolation, Coptes Rendus AS, **191** (1930), 635–637.

[6] S. Bernstein, Sur la limitation des valeurs d'un polynome $P_n(x)$ de degré n sur tout un segment par ses valeurs en $n+1$ points du segment, Isvestiya AN SSSR, (1931), 1025–1050.

[7] M. Berz, Algorithms for higher derivatives in many variables with applications to beam physics In: Automatic Differentiation of Algorithms, Breckenridge, CO, 1991, SIAM, Philadelphia, PA, 1991, pp. 147–156.

[8] M. Berz and G. Hoffstätter, Computation and application of Taylor polynomials with interval remainder bounds, Reliab. Comput., **4** (1998), 83–97.

[9] M. Berz and K. Makino, Verified integration of ODEs and flows using differential algebraic methods on high-order Taylor models, Reliab. Comput., **4** (1998), 361–369.

[10] M. Berz and K. Makino, New methods for high-dimensional verified quadrature, In: Proceedings of the International Conference on Interval Methods and their Applications on Global Optimization, Interval 1998, Nanjing, Reliab. Comput., **5**, 1999, pp. 13–22.

[11] M. Berz and J. Hoefkens, Verified high-order inversion of functional dependencies and interval Newton methods, Reliab. Comput., **7** (2001), 379–398.

[12] M. Berz, K. Makino and J. Hoefkens, Verified integration of dynamics in the solar system, In: Proceedings of the Third World Congress of Nonlinear Analysts, Part 1, Catania, 2000, Nonlinear Anal., **47**, 2001, pp. 179–190.

[13] D. Bessis and J. D. Fournier, Pole condensation and the Rieman surface assicuated with a shock in Burgers' equation, J. Physique Lett., **45** (1984), 833–841.

[14] E. Bichuch and Y. Yomdin, High order Taylor discretization for numerical solution of parabolic PDE's, preprint, 1995.

[15] E. Bierstone, P. Milman and W. Pawlucki, Differentiable functions defined in closed sets. A problem of Whitney, Invent. Math., **151** (2003), 329–352.

[16] J. M. Boardman, Singularities of differential maps, Publ. Math. Inst. Hautes Études Sci., **33** (1967), 21–57.

[17] M. Briskin, Y. Elichai and Y. Yomdin, How can Singularity theory help in Image processing, In: Pattern Formation in Biology, Vision and Dynamix, (eds. M. Gromov and A. Carbone), World Scientific, 2000, pp. 392–423.

[18] Th. Bröcker and L. Lander, Differentiable Germs and Catastrophes, London Math. Soc. Lecture Notes Ser., **17**, Cambridge Univ. Press, 1975.

[19] Y. Brudnyi and P. Shvartsman, Whitney's extension problem for multivariate $C^{1,\omega}$-functions, Trans. Amer. Math. Soc., **353** (2001), 2487–2512.

[20] R. Caflisch, N. Ercolani, T. Hou and Y. Landis, Multi-valued solutions and branch point singularities for non-linear hyperbolic or elliptic systems, Comm. Pure Appl. Math., **XLVI** (1993), 453–499.

[21] D. Coppersmith and T. J. Rivlin, The growth of polynomials bounded at equally spaced points, SIAM J. Math. Anal., **23** (1992), 970–983.

[22] M. Elad, P. Milanfar and G. H. Golub, Shape from moments—an estimation theory perspective, IEEE Trans. Signal Process., **52** (2004), 1814–1829.

[23] J. Elder, Are Edges Incomplete?, Int. J. of Comp. Vision, **34**, 97–122.

[24] Y. Elichai and Y. Yomdin, Flexible high-order discretization, preprint, 1989.

[25] Y. Elichai and Y. Yomdin, Global motion planning algorithm based on a high order discretization and on hierarchies of singularities, In: Proc. 28th CDC, Tampa, Florida, 1989, pp. 1173–1174.

[26] Y. Elichai and Y. Yomdin, Normal forms representation: a technology for image compression, SPIE, **1903**, Image and Video Processing, 1993, pp. 204–214.

[27] Y. Elichai and Y. Yomdin, Flexible high order discretization of geometric data for global motion planning, In: Algorithmic Complexity of Algebraic and Geometric Models, Theoret. Comput. Sci., **157**, 1996, pp. 53–77.

[28] B. Ettinger, N. Sarig and Y. Yomdin, Linear versus non-linear acquisition of step-functions, J. Geom. Anal., **18** (2008), 369–399.

[29] T. Erdelyi, Remez-type inequalities and their applications, J. Comp. Appl. Math., **47** (1993), 167–209.

[30] Ch. Fefferman, A generalized sharp Whitney theorem for jets, Rev. Mat. Iberoamericana, **21** (2005), 577–688.

[31] Ch. Fefferman, Whitney's extension problem for C^m, Ann. of Math. (2), **164** (2006), 313–359.

[32] T. Fort, Finite differences and difference relations in the real domain, Clarendon Press, 1948.

[33] J. Grote, M. Berz and K. Makino, High-order representation of Poincaré maps, In: Automatic Differentiation: Applications, Theory, and Implementations, Lect. Notes Comput. Sci. Eng., **50**, Springer, Berlin, 2006, pp. 59–66.

[34] D. Haviv and Y. Yomdin, Uniform approximation of near-singular surfaces, Theoret. Comput. Sci., **392** (2008), 92–100.

[35] F. B. Hildebrand, Introduction to Numerical Analysis, Second Ed., Dover Publ, Inc., New-York, 1987.

[36] J. Hoefkens, M. Berz and K. Makino, Computing validated solutions of implicit differential equations, In: Challenges in Computational Mathematics, Pohang, 2001, Adv. Comput. Math., **19**, 2003, pp. 231–253.

[37] R. Kazinnik, S. Dekel and N. Dyn, Low bit-rate image coding using adaptive geometric piecewise polynomial approximation, IEEE Trans. Image Process., **16** (2007), 2225–2233.

[38] V. Kisun'ko, Cauchy type integrals and a *D*-moment problem, C. R. Math. Acad. Sci. Soc. R. Can., **29** (2007), 115–122.

[39] V. Kisun'ko, *D*-moment problem and applications, in preparation.

[40] E. Kochavi, R. Segev and Y. Yomdin, Numerical solution of field problems by nonconforming Taylor discretization, Appl. Math. Modelling, **15** (1991), 152–157.

[41] E. Kochavi, R. Segev and Y. Yomdin, Modified algorithms for nonconforming Taylor discretization, Computers & Structures, **49** (1993), 969–979.

[42] A. Kolesnikov and A. J. Baker, An efficient high-order Taylor weak statement formulation for the Navier–Stokes equations, J. Comput. Phys., **173** (2001), 549–574.

[43] A. N. Kolmogorov, Asymptotic characteristics of some completely bounded metric spaces, Dokl. Akad. Nauk SSSR, **108** (1956), 585–589.

[44] A. N. Kolmogorov and V. M. Tihomirov, ε-entropy and ε-capacity of sets in functional space, Amer. Math. Soc. Transl., **17** (1961), 277–364.

[45] A. Kurganov and E. Tadmor, New high-resolution central schemes for nonlinear conservation laws and convection-diffusion equations, J. Comput. Phys., **160** (2000), 241–282.

[46] P. Lax, The formation and decay of shock waves, Amer. Math. Monthly, **79** (1972), 227–241.

[47] P. Lax, Gibbs phenomena, J. Sci. Comput., **28** (2006), 445–449.

[48] P. Lax, Computational fluid dynamics, J. Sci. Comput., **31** (2007), 185–193.

[49] R. Lorentz, Multivariate Birkhoff Interpolation, Lecture Notes in Math., **1516**, Springer-Verlag, 1992.

[50] S. Manikonda and M. Berz, An accurate high-order method to solve the Helmholtz boundary value problem for the 3D Laplace equation, Int. J. Pure Appl. Math., **23** (2005), 365–378.

[51] S. Manikonda and M. Berz, Multipole expansion solution of the Laplace equation using surface data, Nuclear Insstruments and Methods in Physical Reseaerch A, **558** (2006), 175–183.

[52] S. Manikonda, M. Berz and K. Makino, High-order verified solutions of the 3D Laplace equation, Transactions on Computers, Issue 11, **Vol. 4** (2005), 1604–1610.

[53] H. Nessyahu and E. Tadmor, Non-oscillatory central differencing for hyperbolic conservation laws, J. Comput. Phys., **87** (1990), 408–463.

[54] E. A. Rakhmanov, Bounds for polynomials with a unit discrete norm, Ann. of Math. (2), **165** (2007), 55–88.

[55] E. J. Remez, Sur une propriete des polynomes de Tchebycheff, Comm. Inst. Sci. Kharkov, **13** (1936), 93–95.

[56] N. Sarig and Y. Yomdin, Signal acquisition from measurements via non-linear models, C. R. Math. Acad. Sci. R. Can., **29** (2007), 97–114.

[57] E. Tadmor, High resolution methods for time dependent problems with piecewise smooth solutions, In: Proceedings of the International Congress of Mathematicians, Vol. III, Beijing, 2002, Higher Ed. Press, Beijing, 2002, pp. 747–757.

[58] R. Thom, Stabilité Structurelle et Morphogénèse, W. A. Benjamin, Inc., 1972; English edition, Structural Stability and Morphogenesis, W. A. Benjamin, Inc., 1976.

[59] M. Vetterli, P. Marziliano and T. Blu, Sampling signals with finite rate of innovation, IEEE Trans. Signal Process., **50** (2002), 1417–1428.

[60] L. Wang, J. Rouse, H. Liu, E. Munro and X. Zhu, Simulation of electron optical systems by differential algebraic methods combined with Hermite fitting for practical lens fields, Microelectronic Engineering, **73-74** (2004), 90–96.

[61] H. Whitney, Analytic extension of differentiable functions defined in closed sets, Trans. Amer. Math. Soc., **36** (1934), 63–89.

[62] H. Whitney, Differentiable functions defined in closed sets I, Trans. Amer. Math. Soc., **36** (1934), 369–387.

[63] H. Whitney, Functions differentiable on the boundaries of regions, Ann. of Math. (2), **35** (1934), 482–485.

[64] H. Whitney, Differentiable functions defined in arbitrary subsets of Euclidean space, Trans. Amer. Math. Soc., **40** (1936), 309–317.

[65] H. Whitney, On the extension of differentiable functions, Bull. Amer. Math. Soc., **50** (1944), 76–81.

[66] Z. Wiener and Y. Yomdin, From formal numerical solutions of elliptic PDE's to the true ones, Mathematics of Computations, **69** (2000), 197–235.

[67] Y. Yomdin, Some quantitative results in singularity theory, Ann. Polon. Math., **87** (2005), 277–299.

[68] Y. Yomdin, Generic singularities of surfaces, In: Singularity Theory, Dedicated to Jean-Paul Brasselet on His 60th Birthday, Proceedings of the 2005 Marseille Singularity School and Conference, (eds. D. Cheniot, N. Dutertre, C. Murolo, D. Trotman and A. Pichon), World Scientific, 2007, pp. 357–376.

[69] Y. Yomdin, Discrete Remez inequality, preprint, Weizmann Inst., 2008.

[70] Y. Yomdin and G. Comte, Tame Geometry with Applications in Smooth Analysis, Lecture Notes in Math., **1834**, Springer-Verlag, 2004.

[71] Y. Yomdin and G. Zahavi, Taylor discretization of evolution PDE's, preprint, Weizmann Inst., 2007.

[72] Y. Yomdin and G. Zahavi, Robustness of Hermite fitting, in preparation.

[73] G. Zahavi, MA thesis, The Weizmann Inst., Rehovot, 2005.

[74] G. Zahavi, Ph. D. thesis, Weizmann Inst., 2007.

Department of Mathematics
The Weizmann Institute of Science
Rehovot 76100
Israel

E-mail address: `yosef.yomdin@weizmann.ac.il`
`gal.zahavi@weizmann.ac.il`