KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

UNIWERSYTET GDAŃSKI

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY

Numerical Continuation Methods.

PWP Interdysciplinary Doctoral Studies in Mathematical Modeling
UDA-POKL.04.01.01-00-026/13-00
Projekt jest współfinansowany przez Unię Europejską w ramach Europejskiego Funduszu Społecznego

1 Introduction

The numerical continuation methods may be simply described as numerical approximation of the zero set of the continuous map $F : \mathbb{R}^{k+1} \to \mathbb{R}^k$, starting from some known zero $p_0 \in \mathbb{R}^{k+1}$. It is a natural problem in all applications related to the homotopy theory and nonlinear eigenvalue problems. For an extensive discussion on different attitudes towards this problem we can refer the reader to the monograph [1], and to review papers [2] and [11].

The numerical continuation methods (also known as *path following methods*) are focused on two generic ideas: predictor-corrector attitude and piecewise-linear methods, henceforth referred to as PL-continuation methods. Predictor-corrector methods may be applied in the case of a smooth map F (at least of class C^1) with a known derivative or its reasonable estimations. In contrast, PL-continuation methods do not require smoothness assumptions. In this paper we focus on the situation where a map F is of the C^1 class, but it is either not possible to find its derivative or it is very difficult to estimate. Moreover, the computational effort to calculate the value of the function F is relatively high, so it is important to limit considerably the number of points used to calculate the value of the function. That is why the predictor-corrector methods may not be applied and, instead, we are turning towards PL-continuation methods.

We are going to present both methods – but before we do let us introduce some general concepts that we will be referring to.

Let us start with the definition of the curve.

Definition 1. The curve in Euclidean space $\mathbb{R}^m$ is the map $\gamma : [a, b] \to \mathbb{R}^m$, of class C^1, satisfying the following conditions:

1. γ is an injective map;
2. $\gamma'(t) \neq 0$ for $t \in [a, b]$.

Now we will try to describe the set $f^{-1}(0)$ of the continuous map $f : \mathbb{R}^{k+1} \to \mathbb{R}^k$. Actually the only thing we can tell in this very general case is that the set $f^{-1}(0)$ is closed. But much more may be said in case of C^1 maps. Let us start with a few definitions.

Definition 2. The point $x_0 \in \mathbb{R}^{k+1}$ is called a regular point of the map $f : \mathbb{R}^{k+1} \to \mathbb{R}^k$ (class C^1) when the derivative $Df(x_0)$ is a linear map of the rank k (i.e. the maximum possible). If the point is not regular one we call it the critical point.

Definition 3. The point $y_0 \in \mathbb{R}^k$ is called a regular value of the map $f : \mathbb{R}^{k+1} \to \mathbb{R}^k$ (class C^1) if every $x \in f^{-1}(y_0)$ is a regular point. If y_0 is not a regular value we call it *the critical value.*

Let us now recall the implicit function theorem (presented in one of numerous versions):

Theorem 1 (cf. [8])**.** *Assume $f : \mathbb{R}^{k+1} \to \mathbb{R}^k$ is the map of class C^1, and $p \in \mathbb{R}^{k+1}$ is the regular zero of f. Then there exists such local class C^1 diffeomorphism $\varphi : V_0 \to U_p$ that*

1. *$0 \in V_0 \subset \mathbb{R}^{k+1}$ is the open set;*
2. *$p \in U_p \subset \mathbb{R}^{k+1}$ is the open set;*
3. *$\varphi(0) = p$;*
4. *$f(\varphi(x_1, ..., x_{k+1})) = (x_1, ..., x_k)$.*

This implies that

$$f^{-1}(0) \cap U_p = \varphi(0, ..., 0, x_{k+1}),$$

hence the set $f^{-1}(0)$ is locally the C^1 curve. Moreover we have this very general observation (cf. [10])

Corollary 1. *If* 0 *is the regular value of the map* f*, then the set* $f^{-1}(0)$ *is the sum of disjoint sets homeomorphic to the circle or to the open interval.*

Such regular curves look like very easy to follow – with no branching, no decisions to make – we just need to stick to our path. But this is only the case of regular zeroes. In case 0 is not a regular value different things may happen

- the dimension of $f^{-1}(0)$ may be (locally or globally) bigger than 1 – it is enough to think about the constant map $f(x) = 0$;
- the set $f^{-1}(0)$ may locally be something very different from the simple curve – it is enough to consider the map $f : \mathbb{R} \times \mathbb{R}^k \to \mathbb{R}^k$ given by $f(\lambda, x) = \lambda x - Ax$, where $A : \mathbb{R}^k \to \mathbb{R}^k$ is the linear map. In every point $(\lambda_i, 0)$, where λ_i is the eigenvalue of A the set of solutions is locally not a manifold. Actually it is the intersection of the real line with the n-dimensional subspace, where n is the dimension of $\mathrm{Ker}(\lambda I - A)$.

Fortunately we have strong theoretical (although not necessarily practical) reasons to assume that 0 is the regular value. Let us look at the well-known Sard's theorem (presented in the version adjusted to what we need):

Theorem 2 (Sard)**.** *Assume* $f : \mathbb{R}^{k+1} \to \mathbb{R}^k$ *is the map of class* C^2*. Then the set of critical values of* f *has* k*-dimensional Lebesgue measure* 0*.*

The immediate consequence of this is that in any neighbourhood of the critical value $0 \in \mathbb{R}^k$ there exist regular values $c \in \mathbb{R}^k$ of the map f. This means that even in case the set $f^{-1}(0)$ contains intersecting curves, the sets $f^{-1}(c)$ (for the right selection of $c \in \mathbb{R}^k$) consists only of mutually disjoint curves.

2 Piecewise Linear Methods

We are going to start with Piecewise-Linear continuation method (later denoted just as PL. We start with certain decomposition $\mathbb{T}$ of the domain into the simplices, and then define certain piecewise affine approximation $f_{\mathbb{T}} : \mathbb{R}^{k+1} \to \mathbb{R}^k$ of the map f. Then we are looking for the set $f_{\mathbb{T}}^{-1}(0)$, approximating the set $f^{-1}(0)$. This is just a short description – to be more precise we need to give some definitions.

Definition 4. We say that the points $v_0, v_1, ..., v_l \in \mathbb{R}^{k+1}$ are *affine independent* when vectors $v_1 - v_0, v_2 - v_0, ..., v_l - v_0$ are linearly independent.

We can see that the definition does not depend on the selection of the point v_0. We can also see that the affine independent set in $\mathbb{R}^{k+1}$ may not contain more than $k + 2$ points.

Definition 5. The set $\sigma \subset \mathbb{R}^{k+1}$ is called l-dimensional simplex when it is a convex hull of the set of $l + 1$ affine independent points $\{v_0, v_1, ..., v_l\} \subset \mathbb{R}^k$, i.e.

$$\sigma = \mathrm{conv}\{v_0, v_1, ..., v_l\} = \{t_0 v_0 + t_1 v_1 + ... + t_l v_l : t_i \in \mathbb{R}, t_i \geq 0, t_0 + t_1 + ... + t_l = 1\}.$$

Each point v_i $(i = 0, ..., l)$ is called a vertex of the simplex σ.

We should observe that each point $x \in \sigma$ may be uniquely represented as the convex combination of vertices of the simplex, i.e.

$$x = t_0 v_0 + t_1 v_1 + ... + t_l v_l,$$

where $t_0 + t_1 + ... + t_l = 1$ and $t_i \geq 0$ for $i = 0, 1, ..., l$. The sequence $(t_0, t_1, ..., t_l)$ is called *barycentric coordinates of x in the simplex σ*.

Definition 6. Let $\sigma \subset \mathbb{R}^{k+1}$ be the $(k+1)$-dimensional simplex spanned by the vertices $\{v_0, v_1, .., v_{k+1}\}$. Then each k-dimensional simplex spanned by the $(k+1)$-element subset of $\{v_0, v_1, .., v_{k+1}\}$ is called the k-dimensional face of the simplex σ.

We can see that each point belonging to the k-dimensional face V_i of the simplex σ, spanned by the vertices $\{v_0, .., v_{i-1}, v_{i+1}, ...v_{k+1}\}$ (meaning all vertices of σ different from v_i), have barycentric coordinates with $t_i = 0$.

When we divide the space $\mathbb{R}^{k+1}$ into simplices – in the right way – then we say that we have defined the *triangulation* of the space $\mathbb{R}^k$. It is described more precisely by the following definition.

Definition 7. Triangulation $\mathbb{T}$ of the space $\mathbb{R}^{k+1}$ is such decomposition of this space into countable collection of $(k+1)$-dimensional simplices, that

1. if for $\sigma_1, \sigma_2 \in \mathbb{T}$, $\sigma_1 \neq \sigma_2$ there is $\sigma_1 \cap \sigma_2 \neq \emptyset$, then the intersection $\sigma_1 \cap \sigma_2$ is the common face of both simplices;

2. each bounded set $A \subset \mathbb{R}^{k+1}$ intersects the finite number of simplices $\sigma \in \mathbb{T}$.

We are not going to define any triangulation here, but there are some standard ways to do it (as an example see Coxeter-Freudenthal-Kuhn triangulation given in [1]). Actually it doesn't matter how the triangulation is defined – we will just be using it. We are going to need some way to move throughout the triangulation.

Definition 8. The following operation:

1. for given simplex $\sigma \in \mathbb{T}$ we choose one of its vertices v_i and the only k-dimensional face $V_i \subset \sigma$, not containing the vertex v_i;

2. then we choose exactly one vertex v_i' that $\sigma' = \text{conv}\{v_i', V_i\} \in \mathbb{T}$;

assigning $v_i \mapsto v_i'$is called the pivoting step.

Now, given the predefined triangulation $\mathbb{T}$ of the space $\mathbb{R}^{k+1}$, we are ready to define the algorithm PL. For the map f and triangulation $\mathbb{T}$ we define the map $f_{\mathbb{T}} : \mathbb{R}^{k+1} \to \mathbb{R}^k$, that is piecewise affine. There is a very natural way to do it for all vertices v belonging to the triangulation $\mathbb{T}$ – by the formula:

$$f_{\mathbb{T}}(v) = f(v).$$

Then the map $f_{\mathbb{T}}$ is extended the the affine map for each simplex $\sigma \in \mathbb{T}$ separately by:

$$f_{\mathbb{T}}(t_0 v_0 + t_1 v_1 + ... + t_{k+1} v_{k+1}) = t_0 f_{\mathbb{T}}(v_0) + t_1 f_{\mathbb{T}}(v_1) + ... + t_{k+1} f_{\mathbb{T}}(v_{k+1}).$$

Hence the map $f_{\mathbb{T}}$ is defined in the entire space $\mathbb{R}^{k+1}$ (we must check that the map is well defined in the common parts of the simplices – but we may leave it as an exercise). Now we

may consider the map $f_{\mathbb{T}}$ to be the approximation of the map f. This approximation seems to be very useful since it is easy to handle when the numerical computations are performed.

Now we may start to look for the set of solutions of the equation

$$f_{\mathbb{T}}(x) = 0.$$

Basically in each simplex it turns into the linear problem:

$$\begin{cases} f_{\mathbb{T}}(x) = 0 \\ x \in \sigma \end{cases} \Leftrightarrow \begin{cases} t_0 f_1(v_0) + t_1 f_1(v_1) + ... + t_{k+1} f_1(v_{k+1}) = 0 \\ t_0 f_2(v_0) + t_1 f_2(v_1) + ... + t_{k+1} f_2(v_{k+1}) = 0 \\ .... \\ t_0 f_k(v_0) + t_1 f_k(v_1) + ... + t_{k+1} f_k(v_{k+1}) = 0 \\ t_0 + t_1 + ... + t_{k+1} = 1 \\ t_i \geq 0, \qquad i = 0, 1, ..., k+1. \end{cases} \tag{1}$$

So it is enough to solve such linear problem – in every simplex – and we have the solution set. Generally we know how to do it – solving such linear problems (even with additional bounds on t_i) is a typical task (both theoretically and numerically). But we can see that now we start to touch some issues related to the regularity... Typically the solution set of the linear system given above is a 1-dimensional linear subspace of $\mathbb{R}^{k+1}$ translated by some vector. When we intersect it with a $(k+1)$-dimensional simplex we get the line segment (or an empty set). But in may happen that the solution of the linear system is the (translated) subspace having the dimension greater then 1. But this is not the case we are looking for...

That is why we must now take care of the regularity oz zeroes of the map $f_{\mathbb{T}}$ somehow, to avoid the situation that was described above. Fortunately it appears that – similarly as in the case of C^1 maps – also in this case we may define the regular values of the map $f_{\mathbb{T}}$:

Definition 9. We say that the point $x \in \mathbb{R}^{k+1}$is the regular value of the map $f_{\mathbb{T}}$ when the following two conditions are met:

1. x does not belong to any face of any simplex $\sigma \in \mathbb{T}$ with the dimension less than k (especially x is not a vertex of any simplex from $\mathbb{T}$);

2. if $x \in \sigma \in \mathbb{T}$,then the derivative $Df_{\mathbb{T}}|_{\sigma}(x)$ has maximal rank (i.e. equal to k).

If for $y \in \mathbb{R}^k$ the set $f_{\mathbb{T}}^{-1}(y)$ contains only regular points of $f_{\mathbb{T}}$, then we say that y is the regular value of $f_{\mathbb{T}}$.

We are not going to give the technical details here – we just mention that the theorem of Sard's type works here as well: even if 0 is not a regular value of $f_{\mathbb{T}}$ we can disturb it with arbitrarily small constant (but appropriately selected) so the disturbed map has 0 as a regular value.

Let us return to our algorithm. It is obvious that there is no need to solve the equation (1) in every simplex. Assuming that 0 is the regular value of $f_{\mathbb{T}}$, we can see that solution of the linear equation (if exists) is the one dimensional linear subspace translated with some vector. If it is intersected with the simplex σ then we have a line segment. So the set $f_{\mathbb{T}}^{-1}(0)$ is just the polygonal chain in $\mathbb{R}^{k+1}$. And to trace such polygonal line we have to know where to start, and taking the simplex σ intersection the polygonal chain, identify two k-dimensional faces of the simplex, that intersects. Having identified the face we need to perform the pivoting step and move to the next simplex. The presentation of this idea is given in Figure 1.

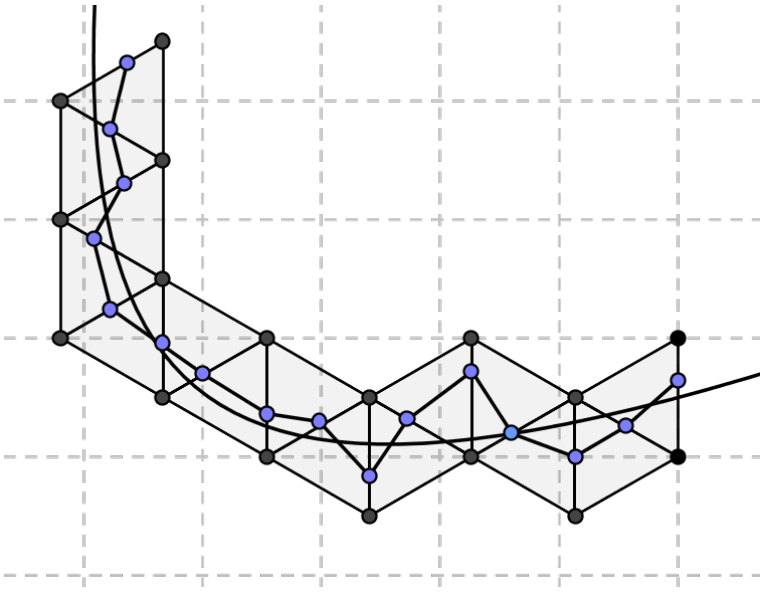

Figure 1: PL algorithm path following

Fortunately it appears that it is even not necessary to solve the linear systems (1) with the appropriate restrictions. Our goal is to identify the two faces of σ containing solutions of (1) – and not necessarily to find the entire line segment. When we change our attitude this way it appears that it is good idea to define – for any k-dimensional face $V = \text{conv}\{v_0, v_1, ..., v_k\}$ – the so-called *labeling matrix*, given by:

$$\mathcal{L}(V) = \begin{bmatrix} 1 & ... & 1 \\ f(v_0) & ... & f(v_k) \end{bmatrix},$$

where vectors $f(v_i)$ form columns of the matrix.

Definition 10. We say that the k-simplex V is completely labeled if

1. the matrix $\mathcal{L}(V)$ is nonsingular;
2. the matrix $\mathcal{L}(V)^{-1}$ is lexicographically positive, i.e. the first nonzero value in each row is positive.

One may show (see [1], def. 12.3.1 together with property 12.3.2), that the k-dimensional face V of one of the simplices belonging to the triangulation $\mathbb{T}$, is completely labeled if and only if there exists such neighbourhood of 0 in $\mathbb{R}^k$, that for any c belonging to this neighbourhood, we have $f_{\mathbb{T}}^{-1}(c) \cap V \neq \emptyset$. So all completely labeled faces not only contain zeroes of $f_{\mathbb{T}}$, but also these zeroes are stable with respect to small perturbations. When we take these results into the consideration we can see that instead of looking for a solution of the problem (1) in the simplex σ it is enough to identify two of its k-dimensional, completely labeled faces. It may appear that this is not easier then solving the system of linear equations (with restrictions) but one may give methods identifying the completely labeled faces without any need to invert the matrix (for details refer to [1] chapter 12.4).

And we can state another very important property:

Theorem 3 ([1], 12.3.8, door-in-door-out principle)**.** *In case the simplex $\sigma \in \mathbb{T}$ contains at least one k-dimensional face, that is completely labeled, then it contains exactly two such faces.*

Hence starting from any initial simplex $\sigma_0 \in \mathbb{T}$, containing zero of the map $f_{\mathbb{T}}$, it is enough to identify two of its completely labeled faces, make a pivoting step with respect to one of them, jump into the next simplex, find its second completely labeled face, make the pivoting step, etc... Now we may describe how the algorithm works:

1. We start from the simplex $\sigma_0 \in \mathbb{T}$ containing k-dimensional face V_0, which is completely labeled.

2. Having the simplex $\sigma_n \in \mathbb{T}$ and its completely labeled face V_n we have a pivoting step with respect to this face and find the next simplex $\sigma_{n+1} \in \mathbb{T}$.

3. We find the completely labeled face $V_{n+1} \neq V_n$ of σ_{n+1} (which was not processed before) and return to the step 2.

And we may prove that the method really works. This is stated in the next theorem:

Theorem 4 ([1], 15.5.2)**.** *Assume that $f : \mathbb{R}^{k+1} \to \mathbb{R}^k$ is of class C^2 and $x \mapsto Df(x)$ satisfies the Lipschitz condition with a constant Lon certain convex set $U \subset \mathbb{R}^{k+1}$. Then*

$$|f(x) - f_{\mathbb{T}}(x)| \leq \frac{1}{2} L \delta^2,$$

for $x \in U$ and δ equal to the largest diameter of the simplex belonging to the triangulation $\mathbb{T}$.

3 Prediction-Correction method

The next method we will discuss (i.e. *Prediction-Correction method* or shortly PC method) – starts from the completely different attitude. Now we will be assuming that $f : \mathbb{R}^{k+1} \to \mathbb{R}^k$ is C^1 map and 0 is its regular value. Moreover, we know one of the zeroes of the map f, i.e. $x_0 \in \mathbb{R}^{k+1}$ satisfying $f(x_0) = 0$. Hence (as we mentioned at the beginning of this review), there exists such curve $\gamma : (a, b) \to \mathbb{R}^k$, of class C^1, $0 \in (a, b)$, that

- $\gamma(0) = x_0$;
- $f(\gamma(t)) = 0$;
- the rank of the derivative $Df(\gamma(t))$ equals k;
- $\gamma'(t) \neq 0$;
- $|\gamma'(t)| = 1$.

for $t \in (a, b)$. The last condition is known as *the arc-length parametrization* and it should be considered to be of purely technical nature.

We should note that since $f(\gamma(t)) = 0$, then also

$$Df(\gamma(t))(\gamma'(t)) = 0,$$

implying that $\gamma'(t) \in \operatorname{Ker} Df(\gamma(t))$. But we know that the kernel of the derivative $Df(\gamma(t))$ is one-dimensional – hence the vector $\gamma'(t)$ spans the kernel of the map $Df(\gamma(t))$. That is why the square matrix (of dimension $k + 1$), given by

$$\begin{bmatrix} Df(\gamma(t)) \\ \gamma'(t)^T \end{bmatrix}$$

is always nonsingular (in the last formula v^T denotes the transpose of the vector v). This means that the determinant

$$\det \begin{bmatrix} Df(\gamma(t)) \\ \gamma'(t)^T \end{bmatrix}$$

has constant sign. Thus, we may have the following definition:

Definition 11. The curve γ is positively oriented if

$$\det \begin{bmatrix} Df(\gamma(t)) \\ \gamma'(t)^T \end{bmatrix} > 0,$$

for $t \in (a, b)$.

Following [1] (def. 2.1.7) we may repeat the definition:

Definition 12. Let A be the matrix of the dimension $k \times (k+1)$ and having the rank k. Then the unique vector $\mathcal{T}(A) \in \mathbb{R}^{k+1}$ satisfying

1. $A(\mathcal{T}(A)) = 0$;
2. $|\mathcal{T}(A)| = 1$;
3. $\det \begin{bmatrix} A \\ \mathcal{T}(A)^T \end{bmatrix} > 0$;

is called the tangent vector induced by the matrix A.

Let us now consider the following initial value problem

$$\begin{cases} \gamma'(t) = \mathcal{T}(Df(\gamma(t))) \\ \gamma(0) = x_0. \end{cases} \tag{2}$$

And we can see that the solution of the problem (2) is closely related to the solution of our path following problem $f(\gamma(t)) = 0$.

Lemma 1 ([1], 2.1.12)**.** *If γ is the solution of the initial value problem* (2)*, then $f(\gamma(t)) = 0$.*

Proof. First of all let us observe that

$$\frac{d}{dt} f(\gamma(t)) = Df(\gamma(t))(\gamma'(t)).$$

And if $\gamma(t)$ satisfies

$$\gamma'(t) = \mathcal{T}(Df(\gamma(t))),$$

then (by the definition of $\mathcal{T}$) there must be $Df(\gamma(t))(\gamma'(t)) = 0$, implying that $f(\gamma(t))$ is constant. Because of the initial condition $\gamma(0) = x_0$ it must be constantly equal to 0. □

Hence instead of searching the curve γ in the set $f^{-1}(0)$ we may look for the solution of the initial value problem (2). Now we may use different methods developed to numerically solve initial value problems of this kind, but here we may actually combine two attitudes: solving differential equation, with additional use of the fact that the solution is such curve γ that $f(\gamma(t)) = 0$. This leads to the so-called *Prediction-Correction method*. The algorithm consists of two repeatedly performed steps: prediction and correction.

Assume now we are performing the n-step of our procedure, so we have the n-th point x_n belonging (approximately) to the curve γ. The prediction step assumes we are moving on the integral curve for the problem (2) – so we find the approximate solution of the differential equation. We can use something simple here – like Euler method. Hence, for the fixed step length $h > 0$, we have:

$$y_{n+1} = x_i + h\mathcal{T}(Df(x_n)). \tag{3}$$

This leads to the point that – as we suppose – belongs to the curve $\gamma(t)$. But this may happen (and usually it does) that $f(y_n) \neq 0$. Staring the next Euler method step from this point could place us on some other integral curve, actually it may happen that we will depart the set $f^{-1}(0)$ very quickly... So, instead of performing the next Euler step, we should perform the correction step. We should now look for the zero of f that is the closest one to the point y_n (the one we have just found by Euler formula (3)). Now we look for $x_{n+1} \in \mathbb{R}^{k+1}$, satisfying

$$|x_{n+1} - y_{n+1}| = \min_{x \in f^{-1}(0)} |x - y_{n+1}|. \tag{4}$$

How the zero of f being close to y_{n+1} may be found? We may use Newton method – but not the classical one. The classical Newton method may be applied to certain maps $f : \mathbb{R}^m \to \mathbb{R}^m$ of class C^1. In our case the dimensions of the domain and codomain differ, so we need to refer to certain modification given in [3]. The iterations of this version of Newton method are given by:

$$w_{l+1} = w_l - Df(w_l)^+(f(w_l)), \tag{5}$$

where A^+ is the generalized (Moore-Penrose) inverse of the matrix A (see [4]).

So, the correction step is to use the equation (5) as long as we arrive to the assumed accuracy. The last value w_{l+1} may be considered to be the next approximation x_{n+1}.

Following this procedure we get the sequence of points $x_n \in \mathbb{R}^{k+1}$, that (joined with line segments) define the polygonal chain approximating the curve $\gamma(t)$. The idea is presented in the Figure 2.

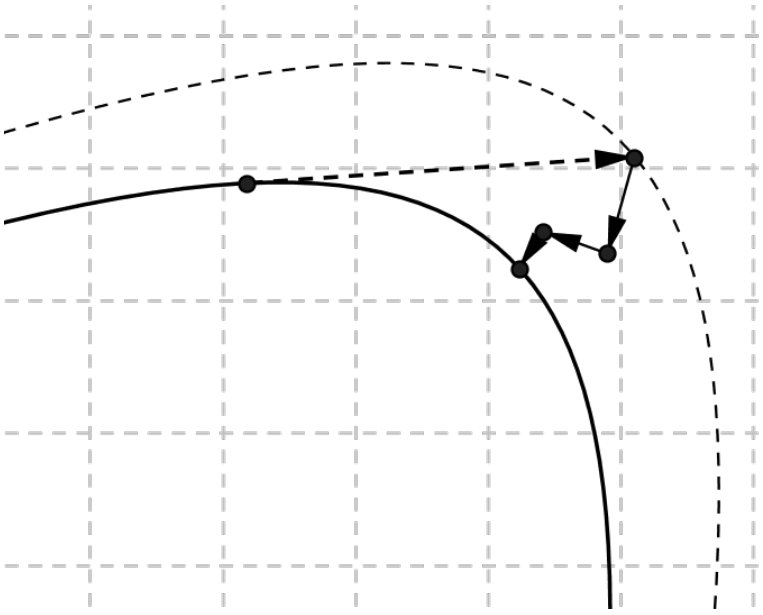

Figure 2: PC path following algorithm. The solid line corresponds to the traced curve. The dashed arrow show the prediction step transferring us to different integral curve (marked with the dashed line). Solid arrows show three correction steps.

The description presented above is very far from being complete. We have not described how to – effectively – perform the following steps:

- Find the approximation of $Df(x_n)$.
- Find the right value of Euler step h.
- Find $\mathcal{T}(Df(x_n))$.
- Find $Df(y_n)^+$

The basic ideas helping to deal with these issues may be found in [1]. And by the end of this review we will present the theorem that actually shows that the method works – and the sequence of points generated by the method may be considered to be the approximation of the curve γ.

Theorem 5 ([2], theorem. 3.2)**.** *Let $f : \mathbb{R}^{k+1} \to \mathbb{R}^k$ be such class C^1 map, that has 0 as a regular value. Let $\gamma_h(t)$ be the polygonal chain generated by the PC algorithm with Euler step equal to $h > 0$. Let $\gamma(t)$ be such curve in $f^{-1}(0)$ that $\gamma(0) = \gamma_h(0) = x_0$ and both curves have the arc-length parametrization. If $A(x)$ is the approximation of $Df(x)$, satisfying $\|A(x) - Df(x)\| = O(h)$ uniformly in the set of solutions, then*

$$|f(\gamma_h(t))| \leq O(h^2)$$

and

$$|\gamma_h(t) - \gamma(t)| \leq O(h^2).$$

References

[1] E. L. Allgower, K. Georg, *Introduction to Numerical Continuation Methods*, Society for Industrial and Applied Mathematics, 2003.

[2] E. L. Allgower, K. Georg, Continuation and Path Following, *Acta Numerica*, Vol. 2, pp. 1 - 64, 1993.

[3] A. Ben-Israel, A Newton-Raphson method for the solution of systems of equations, J. Math. Anal. Appl., 15 (1966), pp. 243-252

[4] A. Ben-Israel, T. Greville, Generalized Inverses. Springer-Verlag, 2003.

[5] J.C. Butcher, Numerical Methods for Ordinary Differential Equations, New York: John Wiley & Sons, 2003

[6] Y. J. Cho, Y-Q. Chen, Topological Degree Theory and Applications, 2006, Chapman and Hall

[7] X. C., Tetsuro Yamamoto, Newton-like methods for solving underdetermined nonlinear equations with nondifferentiable terms, Journal of Computational and Applied Mathematics, Volume 55, Issue 3, 30 November 1994, Pages 311-324

[8] M. W. Hirsch, Differential Topology, Springer-Verlag New York, 1976

[9] T.Y. Li, Numerical solution of multivariate polynomial systems by homotopy continuation methods, *Acta Numerica (1997)*, pp. 399-436

[10] J. Milnor, Topology from differentiable viewpoint, Princeton University Press, 1997

[11] Werner C. Rheinboldt, Numerical continuation methods: a perspective, Journal of Computational and Applied Mathematics, Volume 124, Issues 1-2, 1 December 2000, Pages 229-244