Abstract

We introduce a constructive method that provides the local solution of general implicit systems in arbitrary dimension via Hamiltonian type equations. A variant of this approach constructs parametrizations of the manifold, extending the usual implicit functions solution. We also investigate the unsolved critical case of the implicit functions theorem, define the notion of generalized solution and prove existence and basic properties. Relevant examples and counterexamples are also indicated. The applications concern new necessary conditions (with less Lagrange multipliers), perturbations and algorithms in non-convex optimization problems.

MSC: 26B10 34A12 49K21 49M37

keywords: local parametrizations, uniqueness, critical case, generalized solutions, nonlinear programming.

1 Introduction

In the Euclidean space $\mathbb{R}^d$, $d \in \mathbb{N}$, we consider the general implicit functions system:
where $l \in \mathbb{N}$, $1 \leq l \leq d - 1$ and $F_j \in C^1(\Omega)$, $F_j(x^0) = 0$, $j = 1, l$, $x^0 \in \Omega \subset R^d$ bounded domain, given.

The problem (1) has a long and well known history and we quote the monographs of Krantz and Parks [15], Dontchev and Rockafellar [12] for a comprehensive presentation, including important applications and recent research developments. We also mention the book by Thorpe [33], where related ideas are discussed from the point of view of differential geometry. In particular, it is known that one can associate to (1) a system of nonlinear (partial) differential equations (basically derived from the differentiation formula, under usual assumptions), see [15], Ch. 4.1.

In the recent paper [34], in dimension two and three, it was shown that one can associate to (1) other (essentially simpler) systems of ordinary differential equations, under the mere assumption that $F_j \in C^1(\Omega)$, $j = 1, l$ and in the absence of any independence-type condition. These new systems provide a constructive (local) parametrization of the solution of (1) around $x^0$ under certain conditions.

Moreover, it is possible to define a local generalized solution of (1) even in the critical case, in arbitrary dimension. Our approach to this old question is novel. A variant (in dimension three) was discussed in [23] as well, where it was proved that it is enough to use ordinary differential systems (especially of Hamiltonian type) in order to solve locally (1) via appropriate parametrizations of the unknowns. Several relevant numerical examples are also indicated.

In this paper, we discuss the solution of the general implicit system (1) in arbitrary dimension, by using a new iterated system of ordinary differential equations. The approach has a constructive character and we indicate two variants that give a parametrization of the unknowns in (1) or construct exactly the classical implicit solution (in function form, see Theorem 6). Obtaining parametrizations is advantageous since they may provide a better
description of the manifold. This is done in Section 2, under the usual nondegeneracy condition from the implicit function theorem. We underline that the existence question is well known (via the classical implicit functions theorem or the inverse function theorem, etc.), but a general and effective construction seems not to be available, to the best of our knowledge. The systems of ordinary differential equations that we use here are derived from a first order partial differential system of equations. An interesting fact is that, although just continuity is valid for the right-hand side, we also prove the uniqueness of the solution. This may be compared with the results from [3], [10] (see Remarks [1], [2]). As a first application, in the final part of Section 2, perturbations of the system (1) are investigated, both for implicit functions and implicit parametrizations.

We also recall that, in algebraic geometry, implicitization and parametrization (via rational functions) are important subjects, Gao [13], Wang [37], Schicho [30]. General parametrization methods are not known, Gao [13] and recent papers study approximate parametrization approaches, Dobiasova [11], Yang, Jüttler and Gonzales-Vega [38].

In Section 3, we show how to solve the critical case as well, under the assumption $F_j \in C^1(\Omega)$, $j = 1, 7$, in the absence of (2). We introduce the notion of (local) generalized solution, prove its existence and basic properties. We also indicate some relevant examples. The generalized solution obtained by our method covers all possible cases and is an extension of the notion of local solution from the implicit function theorem, in the classical case. Singular situations in the implicit functions theorem were discussed by different methods in [1], [8], [19] and a comprehensive account can be found in [15], Ch.5.4, where it is specified that a complete solution of the critical case is not known.

The implicit parametrizations constructed in this work are also useful in computations of integrals on implicitly defined manifolds, which seems to be yet unsolved. There is a recent interest in this direction due to important questions in the well known level set method for evolving surfaces, see [16] and its references, or in shape optimization, [35].

The last section is devoted to applications in nonlinear programming. We use reduced gradients to obtain optimality conditions in the simpler Fermat form, involving no or fewer multipliers, i.e. the constraints can be eliminated, at least partially. We also introduce two algorithms, including
one that works in the critical case as well. Some numerical examples are also provided together with a comparison with the relaxation approach \[32\] or the \texttt{fmincon} routines in MatLab.

2 Implicit parametrizations

In this section, we discuss the system (1) under the classical independence assumption. To fix ideas, we assume

$$D(F_1, F_2, \ldots, F_l)$$

$$\neq 0 \quad \text{in} \quad x^0 = (x_1^0, x_2^0, \ldots, x_l^0).$$  \hspace{1cm} (2)

The hypothesis (2) will be dropped in the next section.

Clearly, condition (2) remains valid on a neighbourhood $V \in \mathcal{V}(x^0)$, $V \subset \Omega$, under the $C^1(\Omega)$ assumption on $F_j(\cdot)$, $j = 1, l$ and we denote by $A(x)$, $x \in V$, the corresponding nonsingular $l \times l$ matrix from (2).

We introduce on $V$ the undetermined linear systems of equations with unknowns $v(x) \in \mathbb{R}^d$, $x \in V$:

$$v(x) \cdot \nabla F_j(x) = 0, \quad j = 1, l.$$  \hspace{1cm} (3)

We shall use $d - l$ solutions of (3) obtained by fixing successively the last $d - l$ components of the vector $v(x) \in \mathbb{R}^d$ to be the rows of the identity matrix in $\mathbb{R}^{d-l}$ multiplied by $\Delta(x) = \det A(x)$. Then, the first $l$ components are uniquely determined, by inverting $A(x)$, due to (2).

In this way, the obtained $d - l$ solutions of (3), denoted by $v_1(x), \ldots, v_{d-l}(x) \in \mathbb{R}^d$, are linear independent, for any $x \in V$.

Moreover, these vector fields are continuous in $V$ as $\nabla F_j(\cdot)$ are continuous in $V$ and the Cramer’s rule ensures the continuity of the solution for linear systems with respect to the coefficients. Other choices of solutions for (3), useful in this section, are possible (see Theorem \[6\]).

We introduce now $d-l$ nonlinear systems of first order partial differential equations associated to the vector fields $(v_j(x))_{j=1,d-l}$, $x \in V \subset \Omega$. Furthermore, we denote the sequence of independent variables by $t_1, t_2, \ldots, t_{d-l}$.

These systems have a nonstandard (iterated) character in the sense that the solution of one of them is used as initial condition in the next one. Consequently, the independent variables in the ”previous” systems enter as parameters in the next system via the initial conditions. Due to their simple
A Hamiltonian approach

structure (just one derivative in each equation), we stress that each system (4), (5), ..., (6), may be interpreted as an ordinary differential system in $V \subset \mathbb{R}^d$, with parameters, although partial differential notations are used:

$$\frac{\partial y_1(t_1)}{\partial t_1} = v_1(y_1(t_1)), \quad t_1 \in I_1 \subset \mathbb{R}, \quad y_1(0) = x^0;$$

$$\frac{\partial y_2(t_1, t_2)}{\partial t_2} = v_2(y_2(t_1, t_2)), \quad t_2 \in I_2(t_1) \subset \mathbb{R}, \quad y_2(t_1, 0) = y_1(t_1);$$

$$\ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots$$

$$\frac{\partial y_{d-l}(t_1, t_2, \ldots, t_{d-l})}{\partial t_{d-l}} = v_{d-l}(y_{d-l}(t_1, t_2, \ldots, t_{d-l})), \quad t_{d-l} \in I_{d-l}(t_1, \ldots, t_{d-l-1}), \quad y_{d-l}(t_1, \ldots, t_{d-l-1}, 0) = y_{d-l-1}(t_1, t_2, \ldots, t_{d-l-1}).$$

Here, the notations $I_1, I_2(t_1), \ldots, I_{d-l}(t_1, \ldots, t_{d-l-1})$ are $d-l$ real intervals, containing 0 in interior and depending, in principle, on the "previous" parameters. The existence of the solutions $y_1, y_2, \ldots, y_{d-l}$ follows by the Peano theorem due to the continuity of the vector fields $(v_j)_{j=1}^{d-l}$ on $V$. We show in the next two results that all these subsystems have the uniqueness property as well.

**Theorem 1** Under assumption (2), if $l = d - 1$, then the system (4)-(6) consists just of one subsystem of dimension $d$ with the uniqueness property.

**Proof.** By the implicit functions theorem, under hypothesis (2), around $x^0 = (x_1^0, x_2^0, \ldots, x_d^0)$, there are some open sets (expressed as products of open intervals) such that $x_1 = \tilde{x}_1(x_d), \ldots, x_{d-1} = \tilde{x}_{d-1}(x_d)$, uniquely defined. Replacing these expressions in the last equation of the differential system, we get an ordinary differential equation in $x_d$ with continuous right-hand side, that has a unique solution (it can be integrated in fact) on some
interval. The essential point here is that the right-hand side is non zero around the origin, due to (2) and one can divide by it, etc. Next, the above equalities give all the components, uniquely determined on some interval around the origin.

Remark 1
The cases \( d = 1 \) (one ODE), \( d = 2, l = 1 \) (one Hamiltonian system) are discussed in [3], using similar ideas.

Theorem 2
Under assumption (2), if \( 1 \leq l \leq d - 2, d \geq 3 \), every subsystem in (4)-(6) has a unique solution.

Proof. We shall proceed by induction on \( d \), while \( l \leq d - 2 \) is arbitrarily fixed. In Theorem 1 and Remark 1 we have clarified the small dimension cases, including \( d = 3, l = 2 \). Here, we start the induction with \( d = 3, l = 1 \), which also gives a hint on the general argument.

Denoting by \((x_0, y_0, z_0)\) the initial condition in this case, using (3) and (4)-(6) and fixing that \( F_x(x_0, y_0, z_0) \neq 0 \) in (2), we obtain two iterated Hamiltonian systems (see [23], [34]). We write just the first one (the second one is similar):

\[
\begin{align*}
    x' &= -F_y(x, y, z), & t \in I_1, \\
    y' &= F_x(x, y, z), & t \in I_1, \\
    z' &= 0, & t \in I_1, \\
    x(0) &= x_0, & y(0) = y_0, & z(0) = z_0; \\
\end{align*}
\]

(7)

It is clear that \( z(t) = z_0 \) is the constant solution and what remains is in fact a simple Hamiltonian system with \( d = 2, l = 1 \) that has unique solution by Theorem 1.

In the general case, we assume that the statement is valid for \( d - 1 \) and any arbitrarily given \( l \leq d - 3 \). One can reduce any subsystem in (4)-(6) from dimension \( d \) to \( d - 1 \) as above in (7) since the construction of the right-hand side \( v_j(x) \) involves \( d - 1 - l \geq 1 \) null components. If we restrict the implicit system (1) to \( R^{d-1} \) by fixing the existing constant independent variable, then the hypothesis (2) remains valid for the reduced implicit system and \( A(x) \) is again the nonsingular matrix giving the maximal rank.
A Hamiltonian approach

We notice that the modified $v_j(x)$ satisfies as well the reduced algebraic system \((3)\) since we have just removed one null component of the original $v_j(x)$. Therefore, the reduced differential subsystem comes as well from a reduced implicit system satisfying \((2)\).

We have the following two variants in dimension $d$. The first one is $l = d - 2$.

Then, in dimension $d - 1$, the uniqueness of the reduced differential system follows by Theorem \[1\] If $l \leq d - 3$, then we can apply the hypothesis of the induction and again we get uniqueness. This ends the proof.

**Remark 2**

Note that for systems associated to divergence free fields, the uniqueness results of \[10\] are valid under certain Sobolev type regularity conditions. However, under our hypotheses, we have just continuity in the right-hand side of the differential system \((4)-(6)\) and \[10\] cannot be applied.

The next theorems prove more properties and clarify the use of the above setting.

**Theorem 3** a) There are closed intervals $I_j \subset R$, $0 \in \text{int}I_j$, independent of the parameters, such that $I_j \subset I_{j'}(t_1, t_2, \ldots, t_{j-1})$, $j = 1, d - l$.

b) The solutions of the systems \((4) - (6)\) are of class $C^1$ in any existence point and we have:

$$
\frac{\partial y_{d-l}}{\partial t_k}(t_1, \ldots, t_{d-l}) = v_k(y_{d-l}(t_1, \ldots, t_{d-l})), \quad k = 1, d - l.
$$

**Proof.** Each systems \((4) - (6)\) is solved locally in $V$ and any point from the obtained trajectories may serve as an initial condition for the "next" system to be locally solved in $V$ as well. The existence of local solutions is ensured by Peano theorem, which also gives an estimate of the existence intervals.

We denote by $M = \max \{|v_j|_{C^1(V)}, j = 1, d - l\}$.

Take $V_j, j = 1, d - l - 1$ such that $x_0 \in V_j \subset V_{j+1} \subset V$, open subsets. Let $b_1 = \text{dist}(x_0, \partial V_1) > 0$, then we may choose $I_1 = \left[-\frac{b_1 M}{M}, \frac{b_1 M}{M}\right]$ and the local solution of \((4)\) is obtained in $V_1$. Fix $b_2 = \min \{\text{dist}(x, y); x \in \partial V_1, y \in \partial V_2\}$. Then the solution of \((5)\) exists in $I_2 = \left[-\frac{b_2 M}{M}, \frac{b_2 M}{M}\right]$ for any initial data from $V_1$ and with the trajectory contained in $V_2$. 


This argument can be iterated up to the system (6) and the number of iteration steps is finite. This proves the first point.

Clearly, \( y_1 \) satisfies the statement on differentiability, in the origin. Then, \( y_2 \) satisfies it as well since \( y_2(t_1,0) = y_1(t_1) \). And so on, this extends step by step up to \( y_{d-l} \) which is continuously differentiable in all its arguments in the origin. The formula, in the origin, follow from (4) - (6).

For \( k = d-l \), we use (6) and 
\[
\frac{\partial y_{d-l}}{\partial t_{d-l-1}}(0, \ldots, 0) = \frac{\partial y_{d-l-1}}{\partial t_{d-l-1}}(0, \ldots, 0) = v_{d-l-1}(y_{d-l-1}(0, \ldots, 0)) = v_{d-l-1}(x^0).
\]

This proceeds iteratively up to \( k = 1 \).

Due to the uniqueness property, we denote by \( x^1 = y_{d-l}(t_1, \ldots, t_{d-l}) \), for some \((t_1, \ldots, t_{d-l})\) in the existence set and by \( \tilde{y}_{d-l} \), the solution of (4) - (6) with initial condition \( x^1 \). Since we are in the autonomous case with respect to all the independent variables, we have \( \tilde{y}_{d-l}(s_1, \ldots, s_{d-l}) = y_{d-l}((s_1, \ldots, s_{d-l}) + (t_1, \ldots, t_{d-l})) \). The proof is finished by applying the above relation in the new origin (in \( x^1 \)).

**Remark 3**

In [23], for \( d = 3 \), two iterated Hamiltonian systems are used. A related analysis via specific ODE’s arguments together with relevant numerical examples are indicated. The system (4) - (6) is a generalization of this situation and we underline that, as in [23], one can approximate easily its solution, for instance with MatLab. The system (4) is an usual ordinary differential system and we get its approximate solution in the discretization points of \( I_1 \); then the system (5) is solved for each initial condition defined for the values of the parameter \( t_1 \) given by these discretization points in \( I_1 \) and so on. Finally, one obtains the approximate values of \( y_{d-l}(t_1, t_2, \ldots, t_{d-l}) \) on a discretization grid of \( I_1 \times I_2 \times \ldots \times I_{d-l} \). In fact, all the solutions may be computed on their maximal existence interval. This can be achieved very simply and very quickly, by standard numerical routines for ODE’s.

**Theorem 4** For every \( k = 1, l, j = 1, d-l \), we have

\[
F_k(y_j(t_1, t_2, \ldots, t_j)) = 0, \quad (t_1, t_2, \ldots, t_j) \in I_1 \times I_2 \times \ldots \times I_j.
\]
Proof. We notice first that, for any $k = 1, l$, we have:

$$\frac{\partial}{\partial t_1} F_k(y_1(t_1)) = \nabla F_k(y_1(t_1)) \cdot v_1(y_1(t_1)) = 0, \forall t_1 \in I_1,$$

since $v_1$ is orthogonal to $\nabla F_k$, $k = 1, l$, by (3).

Moreover, $F_k(y_1(0)) = F_k(x^0) = 0$, $k = 1, l$, by (1.1). This gives (2.6) for $j = 1$. The argument follows by induction after $j$.

We assume that for $j = 1, r, r \in \mathbb{N}, r \leq d - l - 1$, we have (8) for any $k = 1, l$ and for any $(t_1, t_2, \ldots, t_r) \in I_1 \times I_2 \times \ldots \times I_r$.

We show that this is also valid for $j = r + 1$. First we remark that

$$F_k(y_{r+1}(t_1, t_2, \ldots, t_r, 0)) = F_k(y_r(t_1, t_2, \ldots, t_r)) = 0, \forall k = 1, l, \quad (9)$$

due to the induction hypothesis.

We also notice that

$$\frac{\partial}{\partial t_{r+1}} F_k(y_{r+1}(t_1, t_2, \ldots, t_{r+1})) =$$

$$= \nabla F_k(y_{r+1}(t_1, t_2, \ldots, t_{r+1})) \cdot v_{r+1}(y_{r+1}(t_1, t_2, \ldots, t_{r+1})) = 0, \quad (10)$$

$$\forall (t_1, t_2, \ldots, t_{r+1}) \in I_1 \times I_2 \times \ldots \times I_{r+1}, \forall k = 1, l,$$

due to the differential equation satisfied by $y_{r+1}$ on $I_{r+1}$ and to the orthogonality relation (3) satisfied by the construction of $v_{r+1}(\cdot)$. By (9), (10), we get (8) for $j = r + 1$ and this ends the proof.

Under hypothesis (2), the local solution of (1) is a $d - l$ dimensional manifold around $x^0$. We expect that $y_{d-l}(t_1, t_2, \ldots, t_{d-l})$ is a local parametrization of this manifold on $I_1 \times I_2 \times \ldots \times I_{d-l}$.

**Theorem 5** If $F_k \in C^1(\Omega)$, $k = 1, l$, with the independence property (2), and the $I_j$ are sufficiently small, $j = 1, d - l$, then the mapping

$$y_{d-l} : I_1 \times I_2 \times \ldots \times I_{d-l} \rightarrow \mathbb{R}^d$$

is regular and one-to-one on its image.
Proof. We get that \( y_j \in C^1(I_1 \times I_2 \times \ldots \times I_j) \) by Theorem 3. The matrix \( B \) of partial derivatives of \( y_{d-l} = (y_{d-l}^1, y_{d-l}^2, \ldots, y_{d-l}^d) \), where the superscripts denote the components of the vector \( y_{d-l} \), is:

\[
B = \begin{pmatrix}
\frac{\partial y_{d-l}^1}{\partial t_1} & \frac{\partial y_{d-l}^1}{\partial t_2} & \cdots & \frac{\partial y_{d-l}^1}{\partial t_{d-l}} \\
\frac{\partial y_{d-l}^2}{\partial t_1} & \frac{\partial y_{d-l}^2}{\partial t_2} & \cdots & \frac{\partial y_{d-l}^2}{\partial t_{d-l}} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial y_{d-l}^d}{\partial t_1} & \frac{\partial y_{d-l}^d}{\partial t_2} & \cdots & \frac{\partial y_{d-l}^d}{\partial t_{d-l}}
\end{pmatrix}
\] (11)

We denote by \( M_{d-l} \) the \((d-l) \times (d-l)\) matrix of the last \( d-l \) rows in \( B \) and we compute its determinant. Notice that the last column in \( M_{d-l} \) is given by the last \( d-l \) components of the vector \( v_{d-l} \), that is \((0,0,\ldots,0,\Delta(x))^T\) due to the way we have constructed \( v_{d-l} \) in (3), \( x \) being here the appropriate point in \( V \) obtained as the value of the solution \( y_{d-l}(t_1, t_2, \ldots, t_{d-l}) \), for some \((t_1, t_2, \ldots, t_{d-l}) \in I_1 \times I_2 \times \ldots \times I_{d-l}\). We write shortly \( \Delta(y_{d-l}) \) for \( \Delta(x) \) due to the way we have constructed \( v_{d-l} \) in (3), being here the appropriate point in \( V \) obtained as the value of the solution \( y_{d-l}(t_1, t_2, \ldots, t_{d-l}) \), for some \((t_1, t_2, \ldots, t_{d-l}) \in I_1 \times I_2 \times \ldots \times I_{d-l}\). We cut the last row and the last column in \( M_{d-l} \), we denote the obtained matrix by \( M_{d-l-1} \) and we have:

\[
\det M_{d-l} = \Delta(y_{d-l}) \det M_{d-l-1}.
\] (12)

Taking into account the equation of \( y_{d-l} \) (see (6)) and the fact that the components of \( v_{d-l} \), from order \( l+1 \) to order \( d-1 \) are 0 (as mentioned above), the initial condition in (6) gives by integration:

\[
y_{d-l}^{l+1} = y_{d-l-1}^{l+1}; \ldots; y_{d-l}^{d-1} = y_{d-l-1}^{d-1}
\]

and they are independent of \( t_{d-l} \). Therefore, we can write

\[
M_{d-l-1} = \begin{pmatrix}
\frac{\partial y_{d-l-1}^{l+1}}{\partial t_1} & \frac{\partial y_{d-l-1}^{l+1}}{\partial t_2} & \cdots & \frac{\partial y_{d-l-1}^{l+1}}{\partial t_{d-l-1}} \\
\frac{\partial y_{d-l-1}^{d+1}}{\partial t_1} & \frac{\partial y_{d-l-1}^{d+1}}{\partial t_2} & \cdots & \frac{\partial y_{d-l-1}^{d+1}}{\partial t_{d-l-1}} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial y_{d-l-1}^{d-1}}{\partial t_1} & \frac{\partial y_{d-l-1}^{d-1}}{\partial t_2} & \cdots & \frac{\partial y_{d-l-1}^{d-1}}{\partial t_{d-l-1}}
\end{pmatrix}
\] (13)

Relation (13) shows that in fact \( M_{d-l-1} \) has a similar structure as \( M_{d-l} \), associated to \( y_{d-l-1} \). Using the differential system satisfied by \( y_{d-l-1} \) and
the structure of $v_{d-l-1}$, we see again that the last column in $M_{d-l-1}$ is of the form $(0, 0, \ldots, 0, \Delta(y_{d-l-1}))^T$ (and of length $d-l-1$). Here, the determinant $\Delta(y_{d-l-1})$ is $\Delta(x)$ computed in the point $x = y_{d-l-1}(t_1, t_2, \ldots, t_{d-l-1})$.

One can iterate the above arguments to obtain

$$
\det M_{d-l} = \Delta(y_{d-l})\det M_{d-l-1} = \Delta(y_{d-l})\Delta(y_{d-l-1})\det M_{d-l-2} = \ldots = \Delta(y_{d-l})\Delta(y_{d-l-1})\ldots\Delta(y_1) \neq 0,
$$

where the notations $M_{d-l-2}$, etc., are obvious. Relations (11) - (14) end the proof.

We consider now another solution choice in (3). We shall use $d-l$ solutions of (3) obtained by fixing the last $d-l$ components of the vector $v(x) \in \mathbb{R}^d$ to be the rows of the identity matrix in $\mathbb{R}^{d-l}$. The next result shows that we construct exactly the solution of the classical implicit functions theorem, which follows as a special case of our approach.

**Theorem 6** If $F_k \in C^1(\Omega)$, the last $d-l$ components of $y_{d-l}$ have the form $(t_1 + x^0_{l+1}, t_2 + x^0_{l+2}, \ldots, t_{d-l} + x^0_d)$, that is the first $l$ components of $y_{d-l}$ give the unique solution of the implicit system (1) on $(x^0_{l+1}, x^0_{l+2}, \ldots, x^0_d) + (I_1 \times I_2 \times \ldots \times I_{d-l})$.

**Proof.** By inspection and induction, one can see that the last $d-l$ components of $y_j(t_1, t_2, \ldots, t_j)$ are $(t_1 + x^0_{l+1}, t_2 + x^0_{l+2}, \ldots, t_j + x^0_{l+j}, t_{j+1} + x^0_{l+j+1}, \ldots, t_{d-l} + x^0_d)$.

This is due to the special choice of the last components of the vectors $v_k$ in (3), as rows of the identity matrix, allowing explicit integration. Then, we have just to remark that by redenoting the last $d-l$ components of $y_{d-l}$ as $(s_{l+1}, s_{l+2}, \ldots, s_d)$, then the first $l$ components of $y_{d-l}$ are functions of $(s_{l+1}, s_{l+2}, \ldots, s_d)$, defined on $(x^0_{l+1}, x^0_{l+2}, \ldots, x^0_d) + (I_1 \times I_2 \times \ldots \times I_{d-l})$, solving (1) due to Theorem 4.

The uniqueness comes from the implicit function theorem.

**Remark 4**
We underline that, although Theorem 6 provides the classical solution of the implicit functions theorem, a parametrization may be more advantageous in applications since it offers a more complete description of the corresponding manifold by removing the condition to obtain just functions. One can use maximal solutions of (4) - (6) and, in many examples, the (local) maximal solution from Theorem 5 may give even a global description of the manifold, [23], [34]. In applications, the choice of other solutions of (3) is also possible and of interest [24], in order to improve the description of the manifold.

**Remark 5**

Beside the existence statement, Theorem 6 gives a construction recipe for the implicit functions solution and an evaluation of its existence neighborhood (via Theorem 3), in the system (1).

For instance, if in the proof of Theorem 3 we take $V = B(x_0, R)$ and $V_j = B(x_0, jR/(d-l)^{-1})$, then $I_j = [-R/(d-l)M, R/(d-l)M]$, for $j = 1, 2, ..., d-l$. This may be compared with [5], [28] where other types of arguments are used.

We consider now general perturbations of (1) having the form

$$F^\lambda_k(x_1, ..., x_d) = 0, \quad k = \overline{1,l}, \quad \lambda \in (-1, 1), \quad (15)$$

where $F^\lambda_k \in C^2(\Omega \times (-1, 1))$, $F^0_k = F_k$ and $F^\lambda_k(x^0) = 0, \quad k = \overline{1,l}$. Hypothesis (2) remains clearly valid for the perturbation as well, for $\lambda$ small.

We denote by $(S^\lambda)$ the differential system similar to (4) - (6), associated to the perturbed implicit system (15) and by $v^\lambda_j$ the corresponding solutions of (3), appearing in the right-hand side of $(S^\lambda)$. Then, $v^\lambda_j$ are in $C^1(V_1 \times (-\lambda_0, \lambda_0))$, under our hypotheses, for some $V_1 \subset V$ independent of $\lambda \in (-\lambda_0, \lambda_0)$, for $\lambda_0$ small. The same ideas as in Thm. 3 or Rem. 5 and the obvious property

$$M^\lambda = \max \left\{ |v^\lambda_j|_{C(V)}, \quad j = \overline{1,d-l} \right\} \rightarrow M = \max \left\{ |v_j|_{C(V)}, \quad j = \overline{1,d-l} \right\}$$

give the existence of the closed intervals with the origin in their interior $I_j, j = \overline{1,d-l}$, independent of $\lambda$, such that the solution of $S^\lambda$ is defined on $I_1 \times I_2 \times \ldots \times I_{d-l}$.

We denote by $y^\lambda_1(t_1), ..., y^\lambda_{d-l}(t_1, t_2, \ldots, t_{d-l})$, the unique solution of $(S^\lambda)$, defined in $I_1 \times \ldots \times I_{d-l}$. By making translations with respect to
A Hamiltonian approach

the initial conditions in each subsystem of \((S_\lambda)\), the initial conditions become 0 and the differentiability properties of the solution, with respect to \(\lambda\), are a consequence of standard results on the differentiability with respect to the parameters in ODE’s (since \(v_j^\lambda \in C^1(V_1 \times (-\lambda_0, \lambda_0))\)) and of an inductive argument as before. Denoting by \(z_1^\lambda(t_1), \ldots, z_{d-l}^\lambda(t_1, t_2, \ldots, t_{d-l})\) the derivative of the above solution with respect to \(\lambda \in (-\lambda_0, \lambda_0)\), their system in variations associated to (15) and (1) can be obtained by differentiation in \((S_\lambda)\) with respect to \(\lambda \in (-\lambda_0, \lambda_0)\), etc.

**Proposition 1** We have:

a) the last \(d - l\) components of \(z_1^\lambda(t_1), \ldots, z_{d-l}^\lambda(t_1, t_2, \ldots, t_{d-l})\) are null.

b) for any \(j = 1, \ldots, d - l\) and \((t_1, t_2, \ldots, t_{d-l}) \in I_1 \times \ldots \times I_{d-l}\), \(z_j^\lambda(t_1, t_2, \ldots, t_j)\) is the unique solution of:

\[
\nabla y F_k^\lambda(y_j^\lambda) z_j^\lambda + \partial_\lambda F_k^\lambda(y_j^\lambda) = 0, \quad k = 1, \ldots, l.
\]

**Proof.** The first statement is a clear consequence of Theorem 6 and of the above discussion. Since we have already established above the differentiability properties of \(y_j^\lambda\) with respect to \(\lambda\) on some given open set, one can differentiate with respect to \(\lambda\) in (15) with \(x_j\) replaced by \(y_j^\lambda\), to obtain (16). Notice that the solution of the linear system (16) is unique due to (2) and to a).

**Remark 6**

One can obtain for \(z_1^\lambda(t_1), \ldots, z_{d-l}^\lambda(t_1, t_2, \ldots, t_{d-l})\) the relation (16) even for implicit parametrizations as in Theorem 5, but point a) is not valid and (16) is not uniquely determining (without supplementary information) \(z_1^\lambda(t_1), \ldots, z_{d-l}^\lambda(t_1, t_2, \ldots, t_{d-l})\). To obtain the necessary supplementary information, one has to use directly the differential systems (4)-(6) and to compute the corresponding system in variations.

Consider now, as an example, the special case of perturbations of the form

\[
F_j(x_1, \ldots, x_d) + \lambda h_j(x_1, \ldots, x_d) = 0, \quad j = 1, \ldots, l, \quad \lambda \in (-1, 1),
\]

(17)

where \(h_j \in C^2(\Omega)\), \(h_j(x^0) = 0\).
If, moreover, \( l = 1 \) and the equation \( F(x_1, \ldots, x_d) = 0, F \in C^2(\Omega) \), together with the associated initial condition, represents the boundary of a subdomain in \( \Omega \) (where \( F < 0 \), for instance) then the geometric perturbation defined by (17) may be very complex, including topological and boundary perturbations [21], [14], [31]. Computing the equation in variations as in Proposition 1, the perturbations (17) generate a directional derivative in the implicit system (1). Consequently, by the above geometric interpretation, we may define, for \( l = 1 \), a new type of geometric directional derivative of domains. This is more general than the speed method or the topological derivatives [21] and has applications in shape optimization, fixed domain methods, see [22], [35].

3 Generalized solutions

In this section, we discuss the problem (1) for \( F_j \in C^1(\Omega), j = 1, \ldots, l \), in the absence of the hypothesis (2) i.e. all determinants of maximal order \( l \) may be null in \( x^0 \). We remark that there is \( \{x^n\} \subset \Omega \), such that:

\[
x^n \to x^0, \quad \text{rank}J(x^n) = l, \quad n \in \mathbb{N},
\]

where \( J(x^n) \) denotes the Jacobian matrix of \( F_1, F_2, \ldots, F_l \in C^1(\Omega) \), in \( x^n \).

Notice that in case (18) is not fulfilled, it means that rank \( J(x) < l \) in \( x \in W \), where \( W \) is a neighborhood of \( x^0 \). Then \( F_1, F_2, \ldots, F_l \) are not functionally independent in \( W \) and the problem (1) can be reformulated by using less functionals [26], [29]. That is (18) is in fact always valid, except for not well formulated problems, including redundant equations. One may classify the systems of type (1), from this point of view, in well-posed and ill-posed systems. Notice as well that (18) is fulfilled if (2) holds, i.e. (18) is the generalization of (2), valid for all well-posed implicit systems.

Due to (18), in each \( x^n \), one can use the results of the previous section for the system

\[
F_j(x) - F_j(x^n) = 0, \quad j = 1, \ldots, l, \quad x \in \Omega,
\]

where we can find locally the solution of (19) around \( x^n \), in a neighborhood depending on \( n \).
From (18), we also have $F_j(x^n) \to F_j(x^0) = 0$, for $n \to \infty$, $j = 1, \ldots, l$, since $F_j \in C^1(\Omega)$.

We denote by $T_n$ the closure in $\overline{\Omega}$ of the manifold defined by (19). It is compact and connected. We also have that $\{T_n\}$ are uniformly bounded since $\Omega$ is bounded and, on a subsequence denoted by $\alpha$, we get

$$T_n \to T_\alpha, \quad n \to \infty,$$

in the Hausdorff-Pompeiu metric [21], [17], where $T_\alpha$ is some compact connected subset in $\mathbb{R}^d$.

**Definition 1** $T = \bigcup T_\alpha$ is the (local) generalized solution of (1) in $x^0$. The union is taken for all the sequences and subsequences satisfying (18), (20).

This notion was introduced in [31] and further discussed in [23], in dimension two and three, by exploiting continuity properties with respect to data in Hamiltonian systems. The present treatment in arbitrary dimension is based on general convergence properties and allows a relaxation of the regularity conditions.

The above definition covers all critical or non critical cases. See Remark 7 as well. For instance, if in (1) we have just one equation and $x^0$ is an isolated extremum for the respective function, then the generalized solution is just $\{x^0\}$. If the respective function is identically zero in the open set $O \subset \Omega$ and $x^0$ is on the boundary of $O$, then (18) is satisfied and the generalized solution is the boundary of $O$ or some subset of it - see Proposition 2 and Example 1 below. A complete description of the level sets (even of positive Lebesgue measure) may be obtained in this way via the generalized solutions. The generalized solution is not a manifold and may be not a compact subset (for instance, if $\Omega$ is unbounded), but it is connected, [21], Appendix 3. The approximating generalized solution, i.e. $\bigcup T_{n_0}$ (for some "big" $n_0$ in (20) and for several choices of the approximating sequences of $x_0$ in (18)), may be not connected. One can easily approximate the generalized solutions, by the techniques from Section 2 applied to the corresponding terms from the sequence $\{x_n\}$ close enough to $x^0$. Due to the properties of the Hausdorff-Pompeiu distance, the approximation is uniform in the space variables. If not enough sequences are taken into account, it is possible to
obtain (locally) just a subset of $T$. For instance, in the equation $x^2 - y^2 = 0$, around the origin, with one approximating sequence $(x_n, y_n) \to (0, 0)$, such that $|y_n| < x_n$, just some part of the solution is generated at the limit.

Taking into account a supplementary sequence such that $x_n < -|y_n|$ the whole solution is obtained (locally) by Definition 1. An algorithm for the approximation of the generalized solution is discussed in [24], including many relevant examples.

Let $M \subset \overline{\Omega}$ denote the connected component of the solution of (1), containing the critical point $x^0$. If $\text{int}M$ is nonvoid, then it does not contain $x^0$, due to (18), that is $x^0 \in \partial M$.

**Proposition 2** We have: $x^0 \in T_\alpha \subset T \subset \partial M_{x_0}, \forall \alpha$, where $\partial M_{x_0}$ is the connected component of $\partial M$ containing $x_0$. In particular

$$F_j(x) = 0, \ j = 1, l, \ \forall \ x \in T. \quad (21)$$

**Proof.** By (18), we have $x^n \in T_n, \ \forall \ n$ and we get $x^0 \in T_\alpha$ by the definition of the Hausdorff-Pompeiu convergence. The next inclusion follows by Definition 1.

The same argument gives that, for any $x \in T$, then $x \in T_\beta$ for some subsequence $\beta$, and there are $\lambda_n \in T_n$ (here $T_n$ is the subsequence convergent to $T_\beta$) such that $\lambda_n \to x$ for $n \to \infty$. By (19), we see that $F_j(\lambda_n) = F_j(x^n) \to F_j(x^0) = 0, \ j = 1, l$, on a subsequence. Then, by continuity, $F_j(\lambda_n) \to F_j(x) = 0$ as claimed and (21) is proved.

Consequently, $T_\alpha \subset M, \forall \alpha$. If $\text{int}M$ is nonvoid, then it is formed just of points not satisfying (2) since $\nabla F_j$ are null. Then $\{x^n\}$ are disjoint from $\overline{\text{int}M}$ ($M$ is not necessarily a Caratheodory set and may be distinct from $\overline{\text{int}M}$) and, consequently, $T_\alpha \subset \partial M$. By Prop.A3.2 in [21], each $T_\alpha$ is connected and contains $x_0$, by the above argument. If $\partial M$ has more connected components, then it yields $T_\alpha \subset \partial M_{x_0}, \forall \alpha$. Definition 1 ends the proof.

**Remark 7**
If $x^0$ is a regular point, i.e. (2) is satisfied, then we denote by $S$ the manifold giving the (local) solution of (1) around $x^0$. Then $S$ coincides with the generalized solution around $x^0$. In Definition 1, we may choose $x^n \to x^0$, $x^n \in S$ and the uniqueness property from the implicit functions theorem gives (for this choice) that $T_n = S$ locally, for $n$ big enough. This choice of $\{x^n\}$ satisfies (18) since $J(x^n) \to J(x^0)$, so $x^n$ satisfies (2) for $n$ big enough. We see that in the classical case, one obtains $T = S$ (locally), that is Definition 1 gives indeed a generalization of the classical local solution of the implicit functions theorem.

**Example 1** In $\mathbb{R}^2$, take $d = 2, l = 1$ and

$$f(x_1, x_2) = \begin{cases} x_1^2(x_2^2 - x_1^2)^2 & \text{if } x_1 < 0, \ |x_2| \leq |x_1| \\ 0 & \text{otherwise}. \end{cases}$$

(22)

Clearly $f$ is in $C^1(\mathbb{R}^2)$ and $\nabla f(x_1, x_2) = 0$, on the second line of (22).

Take $x^0 = (0, 0)$ and $x^n \to x^0$, $x^n = (x^n_1, x^n_2)$, $x^n_1 < 0$, $|x^n_2| < |x^n_1|$.

In such points $x^n$, one can use Theorem 5 and (3), together with the relations (4) - (6), give the Hamiltonian system (in dimension two, iterated systems are not necessary):

$$\begin{align*}
x'_1(t) &= -4x_1^2x_2(x_2^2 - x_1^2), \\
x'_2(t) &= 2x_1(x_2^2 - x_1^2)(x_2^2 - 3x_1^2), \\
(x_1(0), x_2(0)) &= x^n.
\end{align*}$$

(23)

Here, we have chosen $(-f_{x_2}, f_{x_1})$ as the solution of (3).

In Figure 1, we represent the solution $T_n$ of (23) obtained with MatLab, for $x^n = (-\frac{1}{n}, 0)$, $n = 2, 5$. The generalized solution of the implicit function problem (1) corresponding to (22) is given by $T = \{(x_1, x_2) \in \mathbb{R}^2; x_1 = \pm x_2, x_1 \leq 0\}$, the boundary of the critical set of $f(\cdot, \cdot)$, to which $x^0$ belongs.

The generalized solution contains the essential information about the solution set of (1), since it gives its boundary (and in Proposition 2 the inclusion becomes equality, in this example).

If we define

$$f_1(x_1, x_2) = x_1^2[(x_1^2 + x_2^2 - 1)_+]^2$$

A Hamiltonian approach

17
Figure 1:
A Hamiltonian approach

and \( x^0 = (0,1) \), then \( \partial M \) is connected and the corresponding generalized solution is \( \partial M \) without the lower half of the unit circle. The inclusion in Proposition 2 is strict and \( M \) is not Caratheodory, in this case. This is also related to the local character of the construction from Section 2. See Ex. 2 in [23] as well.

We continue now with a partial converse of Proposition 2 that shows that the notion of generalized solution is a strict extension of the classical notion of solution.

**Proposition 3** Let \( x^0 \) be the unique critical point of (1) in the interior of the closed ball \( B(x^0) \). Then, \( T = M \subset B(x^0) \).

**Proof.** Due to Proposition 2, we have just to prove \( M \subset T \).

Let \( A \) be a connected component of \( M - \{x^0\} \). It is open in the relative topology of \( M \cap B(x^0) \) since all the points except \( x^0 \) are regular and the implicit functions theorem can be applied. It is also maximal in the sense that it cannot be strictly extended in \( B(x^0) \). Notice that in the relative topology of \( M \), we have \( \partial A \subset \partial B(x^0) \cup \{x^0\} \), by the implicit function theorem. Consequently, \( A \subset A \cup \{x^0\} \) since \( A \) is maximal and the part of \( \partial A \) contained in \( \partial B(x^0) \) is also contained in \( A \).

We have \( x^0 \in \overline{A} \). Otherwise, by the above relation, it yields \( A = \overline{A} \), that is \( A \) is both closed and open in \( M \) and this contradicts \( M \) connected.

One can consider a sequence \( x^n \in A, x^n \to x^0 \) and the associated manifolds \( T_n \). Notice that \( A = T_n \) by the implicit functions theorem. It follows that \( A \cup \{x^0\} = \lim T_n \), \( A \cup \{x^0\} \subset T \). As \( A \) is an arbitrary component of \( M - \{x_0\} \), we get the conclusion and finish the proof.

**Proposition 4** Let \( F_j \in C^1(\Omega), \ j = 1, \ldots, l \) and \( x^n \to x^0, \ x^n, x^0 \in \Omega \). Denote by \( \tilde{T}_n, \tilde{T}_0 \) the generalized solutions of (1) contained in the bounded domain \( \Omega \), corresponding to the initial conditions \( x^n \), respectively \( x^0 \). Then

\[
\limsup_{n \to \infty} \tilde{T}_n \subset \tilde{T}_0. \tag{24}
\]

**Proof.** Let \( \tilde{x}_{nk} \in \tilde{T}_{nk}, \tilde{x}_{nk} \to \tilde{x} \), where \( n_k \to \infty \) is some subsequence. We show that \( \tilde{x} \in \tilde{T}_0 \).

By Definition 1, there is \( \tilde{x}_{nk} \in \Omega \), such that (2) is satisfied in \( \tilde{x}_{nk} \) and \( |\tilde{x}_{nk} - x^{nk}| < \frac{1}{n_k} \) (here, we also use the characterization of the Hausdorff-Pompeiu limit) and there are \( y_{nk} \in T_{\tilde{x}_{nk}} \) such that \( |y_{nk} - \tilde{x}_{nk}| < \frac{1}{n_k} \).
Consequently, \( y_{nk} \rightarrow \hat{x} \) for \( n_k \rightarrow \infty \). Here, \( T_{\hat{x}_{nk}} \) is the solution of (19) corresponding to \( \hat{x}_{nk} \). By using the sequences \( \hat{x}_{nk} \rightarrow x^0 \) and \( y_{nk} \in T_{\hat{x}_{nk}} \), \( y_{nk} \rightarrow \hat{x} \), we see that \( \hat{x} \in T_0 \) due to Definition 18 and the proof is finished.

**Example 2** Let \( x_n \rightarrow 0 \), \( x_n < 0 \) be some strictly increasing sequence and \( g_n : \mathbb{R}^2 \rightarrow \mathbb{R} \) be given by

\[
g_n(x, y) = \begin{cases} 
    c_n[(x - x_n)^2 + y^2 - \frac{1}{4} \min\{ |x_{n+1} - x_n|^2; |x_n - x_{n-1}|^2 \}]^2, & \text{if} \\
    |x - x_n|^2 + |y|^2 \leq \frac{1}{4} \min\{ |x_{n+1} - x_n|^2; |x_n - x_{n-1}|^2 \} \\
    0 & \text{otherwise},
\end{cases}
\]

where \( c_n > 0 \) is some "big" constant. We consider the function \( F : \mathbb{R}^2 \rightarrow \mathbb{R} \) by

\[
F(x, y) = f(x, y) + \sum_{n=1}^{\infty} g_n(x, y),
\]

where \( f \) is given in (22). Clearly \( F \) is in \( C^1(\mathbb{R}^2) \) and \((x_n, 0)\) are local maximum points of \( F \) if \( c_n \) are "big". The sum in (25) has always just maximum two non zero terms due to the form of the supp \( g_n \).

Take the sequence \( x^n = (x_n, 0) \rightarrow (0, 0) = x^0 \). Then, the implicit equations \( F(x, y) = F(x_n, 0) \) have the unique solution \( x^n = (x_n, 0) \) in a neighbourhood of \( x^n \) and \( T_{x^n} = (x_n, 0) \). In the point \((0, 0)\), we have \( T_0 \) as in Example 1. We see in this example that the inclusion in (24) may be strict.

### 4 Reduced gradients in nonlinear programming

In constrained optimization, projected gradient methods are a classical tool, but their application may be hindered by the difficulty to effectively compute projections on the admissible set, Ciarlet [6]. Based on the results from the previous sections, we use here the reduction approach to eliminate, totally or partially, the constraints (and the Lagrange multipliers), that allows optimality conditions in a more effective way, decreasing the dimension. Local and global algorithms and numerical examples are also discussed, under weak assumptions. The elimination of certain unknowns has advantages at computational level.
A Hamiltonian approach

In the recent papers [32], [20], dimensional reduction is obtained via new relaxation procedures associated to implicit functions. Our approach is certainly different and ensures good numerical results. In the case of polynomial and semi-algebraic optimization, Thm.6.5, Thm.7.5, in the setting of global optimization, a stronger constraint qualification is used.

We consider now the classical minimization problem with equality constraints:

\[(P) \quad \text{Min} \{g(x_1, \ldots, x_d)\} \]

subject to \((1)\). It is known that by Theorem 6 we can replace it (around \(x^0\)) by the unconstrained problem for \((t_1, t_2, \ldots, t_{d-l}) \in (I_1 \times I_2 \times \ldots \times I_{d-l})\):

\[(P_1) \quad \text{Min} \{g(y^0_{d-l}, y^0_{d-l}, \ldots, y^0_{d-l}, t_1 + x^0_{l+1}, t_2 + x^0_{l+2}, \ldots, t_{d-l} + x^0_d)\}, \]

where \((y^0_{d-l}, y^0_{d-l}, \ldots, y^0_{d-l}, t_1 + x^0_{l+1}, t_2 + x^0_{l+2}, \ldots, t_{d-l} + x^0_d)\) are the components of \(y_{d-l}\), the solution of \((4)-(6)\), corresponding to this case. This methodology can be extended to the case of implicit parametrizations.

By Theorem 6, Theorem 3 and the chain rule, one easily obtains the (known) first order optimality conditions in the Fermat form, involving the tangential gradient to the constraints manifold:

**Proposition 5** If \(x^0\) is a local solution of \((P)\) satisfying that \(g\) and \(F_i, i = 1, l\), are in \(C^1(R^d)\) and \((2)\) holds, then we have:

\[\nabla g(x^0).v_j(x^0) = 0 \quad j = 1, d - l. \quad (26)\]

In fact, this is equivalent with the classical Lagrange multipliers rule, since under \((26)\), \(\nabla g(x^0)\) is in the normal space, which has the basis given by \(\nabla F_i(x^0), i = 1, l\).

In this non convex setting, we introduce the following algorithm of projected gradient type, based on the use of the tangential gradient:

**Algorithm 1** 1) choose \(n = 0\), \(\delta > 0\) (a tolerance parameter) and denote by \(t^n = (t^n_1, \ldots, t^n_{d-l})\) such that \(y_{d-l}(t^n_1, \ldots, t^n_{d-l}) = x^n\) in \((4)-(6)\).

2) compute \(\rho^{n+1} \in [0, \alpha_n]\) via the line search:

\[\text{Min} g[y_{d-l}(t^n - \rho[\nabla g(x^n).v_j(x^n)]_{j=1,d-l})].\]

3) set:
\[x^{n+1} = y_{d-l}(t^n - \rho^{n+1}[\nabla g(x^n).v_j(x^n)]_{j=1,d-l}),\]
\[t^{n+1} = t^n - \rho^{n+1}[\nabla g(x^n).v_j(x^n)]_{j=1,d-l},\]
4) If \(|g(x^n) - g(x^{n+1})| < \delta\), then STOP! Otherwise \(n:=n+1\) and GO TO Step 2).

**Remark 8**

The algorithm works practically in \(V\), where the system (4)-(6) is defined and the parameter \(\alpha_n\) in the line search with limited minimization rule has to be chosen "small", such that we remain in \(V\) and the system (4)-(6) can be solved around \(t^n\), in Step 2). In Step 3) we perform the "projection" on the constraints manifold \(M \subset \Omega\). The points \(x^n\) generated by this algorithm are always admissible for \((P)\). No convexity properties are assumed. The definition of \((P_1)\) uses the implicit function Theorem 6 which is appropriate for optimality conditions, while for the Algorithm 1 the general implicit parametrization method has to be taken into account. The same is valid for the subsequent problem \((Q_1)\) and the related results.

In this algorithm, \(\Omega\) is a bounded domain, \(g \in C^1(\Omega)\) is bounded from below and the constraints are as in (1) with hypothesis (2) satisfied in \(x^0\). We denote by \(G(t) = g(y_{d-l}(t))\), defined in a neighborhood of the origin in \(R^{d-l}\) and of class \(C^1\) due to (4)-(6) and Theorem 3. The sequence \(\{g(x^n) = G(t^n)\}\) is non increasing and convergent in this general setting, ensuring the convergence of the algorithm. The sequence \(\{x^n\}\) is bounded. Moreover, we have \(\nabla G(t^n) = [\nabla g(x^n).v_j(x^n)]_{j=1,d-l}\) by Theorem 3 and the Algorithm 1 is in fact a transcription of the classical gradient method for the unconstrained problem \((P_1)\). One can discuss other (very rich) variants of such local algorithms with their convergence (to stationary points, in general), under supplementary hypotheses if necessary, Bertsekas [1], Patriksson [27]. The new point in Algorithm 1 is that one can effectively compute the "projection" \(y_{d-l}\).

We discuss now the general case of both equality and inequality constraints:

\[\begin{array}{c}
(Q) \quad \text{Min}\{g(x_1, \ldots, x_d)\} \\
\text{subject to } (1) \text{ and to } \quad G_j(x) \leq 0 \quad j = 1, m, \quad (27)
\end{array}\]
A Hamiltonian approach

where $g, F_i, G_j$ are in $C^1(R^d)$. The Mangasarian-Fromovitz condition in this case consists of \cite{2} and there is $d \in R^d$ such that

$$\nabla F_i(x^0) d = 0, \quad i = 1, l, \quad \nabla G_j(x^0) d < 0, \quad j \in I(x^0),$$

(28)

with $I(x^0)$ being the set of indices of active inequality constraints in $x^0$. See \cite{2}, § 2.3.4 or \cite{7}, § 6 for excellent presentations. The necessary and sufficient metric regularity condition from \cite{36} cannot be used here due to the lack of convexity.

The reduced problem is again obtained via Theorem 6:

\begin{equation}
\begin{aligned}
(Q_1) \quad \text{Min} \{g(y^1_{d-l}, y^2_{d-l}, \ldots, y^l_{d-l}, t_1 + x^0_{l+1}, t_2 + x^0_{l+2}, \ldots, t_{d-l} + x^0_d)\},
\end{aligned}
\end{equation}

subject to the constraints (27), in the "reduced" form:

$$G_j(y^1_{d-l}, y^2_{d-l}, \ldots, y^l_{d-l}, t_1 + x^0_{l+1}, \ldots, t_{d-l} + x^0_d) \leq 0 \quad j = 1, m,$$

(29)

Lemma 1 The minimization problem $(Q_1)$ satisfies the Mangasarian-Fromovitz condition in the origin of $R^{d-l}$.

Proof. By the first part in (28), we see that $d$ is in the tangent space to the manifold (1) since $\nabla F_i(x^0), i = 1, l$ is a basis in the normal space to the manifold given (1), under hypothesis (2). Then $d = \sum_{s=1}^{d-l} \alpha_s v_s$ with $\alpha_s$ some scalars, since $v_s, s = 1, d-l$, gives a base in the tangent space. By the second part in (28) we get $\sum_{s=1}^{d-l} \alpha_s \nabla G_j(x^0) v_s < 0$. Using the derivation formula from Theorem 3, this may be rewritten as $\sum_{s=1}^{d-l} \alpha_s \frac{\partial}{\partial t_s} g_j(0, 0, \ldots, 0) < 0$, where

$$g_j(t_1, \ldots, t_{d-l}) = G_j(y^1_{d-l}, y^2_{d-l}, \ldots, y^l_{d-l}, t_1 + x^0_{l+1}, \ldots, t_{d-l} + x^0_d).$$

is the composed mapping. This shows that the Mangasarian-Fromovitz hypothesis is satisfied in the origin of $R^{d-l}$ with the vector $(\alpha_1, \ldots, \alpha_{d-l})$.

If $x^0$ is a local solution of $(Q)$, by Lemma 1 one can apply the classical KKT theorem, \cite{6}, to the problem $(Q_1)$ in the origin of $R^{d-l}$ that becomes a local solution for $(Q_1)$. Using again the derivation formula, we get:

Theorem 7 Let $x^0$ be a local minimum for $(Q)$. Then, there are $\beta_j \geq 0, j = 1, m$ such that
\[ 0 = \nabla g(x^0).v_s(x^0) + \sum_{j=1}^{m} \beta_j \nabla G_j(x^0).v_s(x^0), \quad s = 1,d-l, \]

\[ 0 = \beta_j G_j(x^0), \quad j = 1,m. \]

**Remark 9**

This is a simplified version of the KKT conditions since it eliminates the Lagrange multipliers for the equality constraints. It is possible to eliminate completely the Lagrange multipliers: if \( x^0 \) is a local solution of problem \((Q)\), then one can remove the inactive inequality constraints at \( x^0 \). This is a consequence of the remark that the inequality constraints that are not active at \( x^0 \) define a neighborhood of \( x^0 \). The minimum property of \( x^0 \) is preserved in this neighborhood, just under the equality constraints supplemented by the active constraints rewritten as equalities. Under the independence condition for all these constraints, one can write optimality conditions as in the Proposition 5.

We relax now the hypotheses in the problem \((Q)\) and we describe a direct minimization algorithm of global type. It looks for the solution in a maximal neighborhood of \( x^0 \), corresponding to the maximal solutions of the subsystems in \((4) - (6)\) (the maximal existence intervals may depend on the respective initial conditions). See Remark 4 and \([23], [34]\).

We assume in the sequel that \( g \) and \( G_j, j = 1,m \), are just in \( C(R^d) \) and \( F_i, i = 1,l \), are in \( C^1(R^d) \) and satisfy condition (2) in \( x^0 \). This last condition can be removed in fact, working with generalized solutions, according to the subsequent Remark 11. Notice that \( x^0 \) is here just an admissible point for \((Q)\) and not a local minimum as in Theorem 7. We can also add the abstract constraint \( x \in D \), some given subset in \( R^d \), such that \( x^0 \in D \).

The main observation is that in solving numerically \((4) - (6)\), now using the variant corresponding to Theorem 5, we obtain automatically a discretization of the manifold defined by \((1)\), in a maximal neighborhood of \( x^0 \), as explained above. Let us denote by \( n \) the discretization parameter. For instance, \( 1/n \) can characterize the size of the discretization for the parameters \( t_1, \ldots, t_{d-l} \), \( n \) or may be linked to the length of the intervals where the maximal solution is computed, etc. We denote by \( C_n \) the set of all these discretized points that, moreover, satisfy all the constraints (the inequality and the other restrictions have to be just checked). They give the approximating admissible set and we formulate the algorithm:
Algorithm 2 1) choose $n = 1$, the discretization step $1/n$ and the solution intervals $I^n_1, \ldots, I^n_{d-1}$, the tolerance parameter $\delta$.

2) compute the discrete set of admissible points $C_n$, starting from $x^0$, via (4) - (6) and by testing the validity of (27) and D.

3) find in $C_n$ the approximating minimum of $(Q)$, denoted by $x^n$.

4) test if the solution is satisfactory by $|g(x^n) - g(x^{n-1})| \leq \delta$.

5) If YES, then STOP. If NO, then $n := n + 1$ and GO TO step 1).

In step 4) other tests (on the solutions, on the gradients, etc.) may be used. The approximating minimum $x^n \in C_n$ may be not unique and the Algorithm 2 finds all all of them. One can adapt the convergence test to such situations.

Theorem 8 The algorithm is convergent as $n \to \infty$.

This is a consequence of the density of $\bigcup C_n$ in the admissible set, according to Theorem 5.

Remark 10

The set defined by the equality constraints may have several connected components. See Example 4. Starting from $x^0$, Algorithm 2 will minimize just on the component that contains $x^0$. Initial guesses from all the admissible components are necessary if we want to minimize on all of them.

Remark 11

If condition (2) is not fulfilled, then one can use the generalized solution of (1) as explained in Section 3 (see Proposition 3), since the Hausdorff-Pompeiu distance ensures the uniform convergence of approximating points. The computed minimum may satisfy (1) or the minimum property with some small error tolerance and the convergence property with respect to the discretization parameters is ensured. An algorithm for the computation of the generalized solution, with relevant examples is studied in [24].

Finally, we indicate some illustrative numerical examples and compare our results with other approximation methods, from MatLab or [32].

Example 3
We consider first a minimization problem on the torus in $\mathbb{R}^3$, with radii 2 respectively 1, defined implicitly by $F = 0$, and with initial point $(x_0, y_0, z_0) = (\sqrt{5}, 2, 0)$:

$$
\begin{align*}
\min \{xyz\} \\
F(x, y, z) &= (x^2 + y^2 + z^2 + 3)^2 - 16(x^2 + y^2)
\end{align*}
$$

The obtained results are given below, compared with the application of the fmincon routine of MatLab:

$$
\begin{align*}
\min &= -2,7154 \\
x_{\min} &= 1,7841; y_{\min} = 1,8199; z_{\min} = -0,8363 \\
f\text{mincon} : \min &= -2,7153; x_{\min} = 1,802; y_{\min} = 1,802; z_{\min} = -0,836
\end{align*}
$$

Using other starting points like (1,0,0) or (3,0,0) is not allowed by MatLab that finds no other admissible solutions in these cases, while our approach works.

**Example 4**

Now, we consider two equality restrictions, given by $F$ and $P$, that represent a torus intersected with a paraboloid, see Fig.2 and Fig.3. Two initial points are taken into account since the intersection has two components.

$$
\begin{align*}
\min \{x^3 + 5y - 7\sin z\} \\
P(x, y, z) &= \frac{2\sqrt{3}}{3} x - y^2 - z^2 \\
(x_0, y_0, z_0) &= (\sqrt{3}, 1, 1); (x_0, y_0, z_0) = (\sqrt{3}, -1, 1)
\end{align*}
$$

The numerical results and a comparison with MatLab routine fmincon is indicated below:

$$
\begin{align*}
(\sqrt{3}, -1, 1) : \text{minimal value} &= 0.498975897823261 \\
\text{solution} : &\ (1.0668890550184, -0.814925789648031, 0.7536319333331335) \\
(\sqrt{3}, 1, 1) : \text{minimal value} &= -7.65929313197537 \\
\text{solution} : &\ (1.10697710321061, 0.817093948780941, 0.781479124977557)
\end{align*}
$$
In the second case, fmincon stops after 42 iterations with the message that constraints are not satisfied within the tolerance. In the first case, fmincon finds basically the same solution.

**Remark 12**

In [32], an example in $\mathbb{R}^6$, with three equality constraints, is discussed. Reworking it via Algorithm 2 starting from the two points indicated there on p.451, we obtain the new points

$$(0.5631, -3.2581, 0.51593, 0.4692, 1.4635, 3.589),$$

$$(0.56166, -3.3154, 0.50897, 0.5047, 1.4365, 3.6777)$$

with the cost values 343,7695, respectively 383,7265. This improves the quoted experiment and can be directly checked. It does not contradict [32] since our algorithm needs no bounds on the independent variables and extends the search domain, which is an advantage from the point of view of global optimization. The necessary working time, on a medium performance laptop, is several minutes. More details on the experiment and some high dimensional numerical examples are indicated in [25].

Figure 2: The admissible set
References

[1] D. P. Bertsekas, Nonlinear Programming (3rd Edition), Athena Scientific, Nashua, MASS (2016).

[2] J.F. Bonnans and A. Shapiro, Perturbation analysis of optimization problems, Springer Verlag, New York (2000).

[3] F. Bouchut, L. Desvilletes, On two-dimensional Hamiltonian transport equations with continuous coefficients, Diff. Int. Eqns., vol.14, no.8, pp. 1015-1024 (2001).

[4] G.J. Butler and H.I. Freedman, Further critical cases of the scalar implicit function theorem, Aequationes Math. 8, pp. 203211 (1972).

[5] H.C. Chang, W. He and N. Prabhu, The analytic domain in the implicit function theorem, JIPAM, Vol. 4, Iss. 1, Article 12 (2003).

[6] Ph. Ciarlet, Introduction to numerical linear algebra and optimization, Cambridge Univ. Press, New York (1989).

[7] F.H. Clarke, Optimization and nonsmooth analysis, John Wiley & Sons, New York (1983).
A Hamiltonian approach

[8] E. Coddington and N. Levinson, Theory of ordinary differential equations, McGraw-Hill, New York (1955).

[9] M. Delfour and J.-P. Zolesio, Shapes and Geometry, SIAM, Philadelphia (2001).

[10] R.J. DiPerna, P.L. Lions, Ordinary differential equations, transport theory and Sobolev spaces, Invent. Math. 98, pp.511-547 (1989).

[11] K. Dobiasova, Parametrizing implicit curves, WDS’08 Proceedings of Contributed Papers, MATHFYZPRESS, pp. 19-22, Prague (2008).

[12] A.L. Dontchev and R.T. Rockafellar, Implicit functions and solution mappings, Springer, New York (2009).

[13] Xiao-Shan Gao, Conversion between implicit and parametric representations of algebraic varieties, Mathematical mechanization and applications, Academic Press, pp. 253-271, San Diego (2000).

[14] A. Henrot, M. Pierre, Variation et optimisation de formes: une analyse géométrique, Springer Verlag, Berlin (2005).

[15] S.G. Krantz and H.R. Parks, The implicit function theorem, Birkhäuser, Boston (2002).

[16] C. Kublik and R. Tsai, Integration over curves and surfaces defined by the closest point mapping, preprint ArXiv 1504.05478v4 (2015).

[17] C. Kuratowski, Introduction to set theory and topology, Pergamon Press, Oxford (1962).

[18] J.B. Lasserre, An introduction to polynomial and semi-algebraic optimization, Cambridge University Press, Cambridge (2015).

[19] S. Lefschetz, Differential equations: geometric theory, Interscience, New York (1957).

[20] A. Mitsos, B. Chachuat, P.I. Barton, McCormick-based relaxations of algorithms, SIAM J. Optim. 20(2), pp.573-601 (2009).

[21] P. Neittaanmäki, J. Sprekels, D. Tiba, Optimization of elliptic systems. Theory and applications, Springer, New York (2006).
[22] P. Neittaanmaki, D. Tiba, Fixed domain approaches in shape optimization problems, Inverse Problems, vol.28, p.1-35, (2012) doi:10.1088/0266-5611/28/9/093001

[23] M.R. Nicolai and D. Tiba, Implicit functions and parametrizations in dimension three: generalized solutions, DCDS - A vol. 35, no.6, pp.2701 - 2710 (2015). doi:10.3934/dcds.2015.35.2701

[24] M.R. Nicolai, An algorithm for solving implicit systems in the critical case, Ann. Acad. Rom. Sci. Ser. Math. Appl. Vol. 7, no. 2, pp.310 - 322, (2015).

[25] M.R. Nicolai, High dimensional applications of implicit parametrizations in nonlinear programming, Ann. Acad. Rom. Sci. Ser. Math. Appl., Vol. 8, no.1, pp.44 -55, (2016).

[26] M. Nicolescu, N. Dinculeanu and S. Marcus, Analiză Matematică, vol. I, Ed. 4, Ed. Didactică și Pedagogică, București (1971).

[27] M. Patriksson, Nonlinear Programming and Variational Inequality Problems: A Unified Approach, Springer (2013).

[28] Phan Phien, Some quantitative results on Lipschitz inverse and implicit functions theorems, arXiv: 1204.4916v2 (2012).

[29] W. Rudin, Principles of mathematical analysis, Second Edition, McGraw-Hill, New York (1964).

[30] J. Schicho, Rational parametrizations of algebraic surfaces, Thesis, J. Kepler Univ. Linz (1995).

[31] J. Sokolowski, J.-P. Zolesio, Introduction to shape optimization. Shape sensitivity analysis, Springer Verlag, Berlin (1992).

[32] M.D. Stuber, J.K. Scott, P.I. Barton, Convex and concave relaxations of implicit functions, Optimization methods and software, 30(3), pp.424-460 (2015).

[33] J.A. Thorpe, Elementary topics in differential geometry, Springer Verlag, New York (1979).
[34] D. Tiba, The implicit functions theorem and implicit parametrizations, Ann. Acad. Rom. Sci. Ser. Math. Appl. 5, no. 1-2, pp. 193-208, (2013). [35] D.Tiba, Boundary Observation in Shape Optimization, in "'New trends in differential equations, control theory, and optimization’", V. Barbu, C. Lefter, I. Vrabie (Eds.), World Scientific Publishing, Singapore, pp.301 - 314 (2016).

[36] D. Tiba, C. Zalinescu, On the necessity of some constraint qualification conditions in convex programming, J.Convex Anal. vol.11, no.1, pp.95-110 (2004).

[37] D. Wang, Irreducible decomposition of algebraic varieties via characteristic set method and Gröbner basis method, CAGD 9, pp. 471-484, (1992).

[38] H. Yang, B. Jüttler, L. Gonzales-Vega, An evolution-based approach for approximate parametrization of implicitly defined curves by polynomial parametric spline curves, Math. Comp. Sci. 4, no. 4, pp. 463-479 (2010)