International Journal of Pure and Applied Mathematics
Volume 78 No. 6 2012, 815-830
ISSN: 1311-8080 (printed version)
url: http://www.ijpam.eu

ijpam.eu

HIGH ORDER HALLEY TYPE DIRECTIONS

Bilel Kchouk[1 §], Jean-Pierre Dussault[2]

[1,2]Department of Mathematics
University of Sherbrooke
2500, Boul. de l'Université, Sherbrooke (Québec), J1K 2R1, CANADA

Abstract: The 1669-1670 Newton-Raphson's method, the 1694 Halley and the 1839 Chebyshev methods are well known algorithms used in non linear programming. By considering these three methods as displacement directions, we introduce in this paper new high order algorithms using these famous methods. We develop two families of such directions and study their convergence and their complexity. We show that these new methods are quite efficient in terms of convergence and costs. We finally introduce a global algorithm based on our analysis of high order methods.

AMS Subject Classification: 49M15
Key Words: Newton's method, Halley family method, automatic differentiation

1. Introduction

The Newton-Raphson's method [12], [15] is in particular used to solve unconstrained optimization problems min $f(x)$. Noting that this problem's local minimum is a stationnary point, we here restrict our study to solve equations $F(x) = 0$ where $F(x) = \nabla f(x)$, thus $\nabla F(x) = \nabla^2 f(x)$ is symmetric. The

Received: December 12, 2011

§Correspondence author

Newton's method is iterative: it creates a sequence $(x_k)_{k\geq 1}$ which converges quadratically, under suitable non-degeneracy assumptions, to the solution x^* verifying $F(x^*) = 0$. Observing that Newton, Halley [9], Chebyshev [2], and SuperHalley [8] share the same iterative scheme: $x_{k+1} = x + d_M$, where d_M is the method's direction, we here introduce new directions based on these algorithms.

First, we propose a brief reminder on these methods and J.-P. Dussault [3] "high order extrapolations". Then, we'll present the major idea of this paper: new high order directions, the methods, their convergence and their complexity. Finally, some research perspectives will be the conclusion of this article.

2. Preliminaries

2.1. Newton's Method

Let $F : D \subseteq X \to Y$ be an operator, where D is the definition set of F and X and Y Banach spaces. If F is Fréchet differentiable on an open convex set $D_0 \subseteq D$,then Newton's method to solve equation $F(x) = 0$ consists in developing the iterative sequence:

$$\boxed{x_{k+1} = x_k - \nabla F(x_k)^{-1} F(x_k)}$$

In 1818, Fourrier [4] proved the quadratic convergence of the method for the scalar case $X = \mathbb{R}$. In 1829, Cauchy [10] proved the quadratic convergence of the method under some assumptions on $F(x_0)$, whithout assuming the existence of a solution (nowadays refered to as semi-local convergence).

2.2. The Halley Methods Family

The following iterative scheme summarizes the Halley methods family :

$$\boxed{x_{k+1} = x_k - [I + \tfrac{1}{2}L(x_k)(I - \alpha L(x_k))^{-1}]\nabla F(x_k)^{-1} F(x_k)}$$

and

$$\boxed{L(x) = \nabla F(x)^{-1}\nabla^2 F(x)\nabla F(x)^{-1} F(x)}$$

The three best known cases are:

- $\alpha = 0$: Chebyshev's method;
- $\alpha = \frac{1}{2}$: Halley's method;

- $\alpha = 1$: SuperHalley's method.

Even if these three methods have a cubical convergence order, the SuperHalley's method is known to be the best on a number of iterations point of view. These three methods solve two linear systems. But only Chebyshev's method has to inverse a unique linear operator, whereas the two others have to inverse two distincts operators.

Remark 1. The order of a method or a direction refers to the highest differentiation order of the function F.

2.3. Extrapolations

The high order extrapolations method have been proposed by J.-P. Dussault [3]. It can be summarized as follows : for a certain estimation x_k close to a regular root x^*, we can write $F(x_k) = r_k$. Thus, we have a system involving two vector variables x and r. Therefore, we can use the implicit functions theroem to define the vector function $x(r)$, as we know the value of $x(r_k)$. This theorem yields an expression of $\dot{x}(r) = \nabla_r x$, and thus allows us to define a linear extrapolation to guess $\hat{x}_{k+1} = x(r_k) + \dot{x}(r_k)(0 - r_k)$. We can easily verify that this linear extrapolation $\hat{x}_{k+1}$ is just a Newton step from x_k: using $F(x_k) - r_k = 0$, we obtain $\nabla_x F(x_k)\dot{x}(r_k) - I = 0$, then $\dot{x}(r_k) = \nabla_x F(x_k)^{-1}$ and as $r_k = F(x_k)$, $\hat{x}_{k+1}$ reduces to a Newton step. We also see that $x(0) = x^*$ is a regular solution of $F(x) = 0$.

2.3.1. Scalar Extrapolations

Another way to express Newton's method consists in writing:

$$\bar{r}_k = \frac{r_k}{||r_k||}$$

and

$$F(x_k) - \tau\bar{r}_k = 0$$

and devise extrapolation formulae from $\tau_k = ||r_k||$ to 0. Let us apply the implicit function theorem to obtain x as function of τ, with $x_k = x(\tau_k)$:

$$\nabla F(x_k)\dot{x}_k(\tau) - \bar{r}_k = 0.$$

It is now clear that $\dot{x}_k(\tau) = \nabla F(x_k)^{-1}\bar{r}_k$ and that the first-order extrapolation is nothing else than Newton's direction $d_N = \dot{x}_k(\tau)(0-\tau)$. Whenever F is many times continuously differentiable at x^*, we may generalize the extrapolation to a higer-order Taylor expansion of x:

$$\hat{x}_p(0) = x(\tau) + \dot{x}(\tau)(0-\tau) + ... + \frac{1}{p!}x^{(p)}(\tau)(0-\tau)^p$$

thus providing a super-linear convergence of order $p+1$. Even if high-order derivatives of F are required, this higer-order Taylor expansion only requires solving linear systems, all involving the same matrix $\nabla_x F(x_k)$. For example:

$$\nabla^2_{xx} F(x)\dot{x}(\tau)\dot{x}(\tau) + \nabla_x F(x)\ddot{x}(\tau) = 0$$

defines $\ddot{x}(\tau)$ as:

$$\ddot{x}(\tau) = \nabla_x F(x)^{-1}[\nabla^2_{xx} F(x)\dot{x}(\tau)\dot{x}(\tau)]$$

The extrapolation itself is:

$$\boxed{\hat{x}_2(0) = x(\tau) + \dot{x}(\tau)(0-\tau) + \tfrac{1}{2}\ddot{x}(\tau)(0-\tau)^2}$$

3. High order Halley Type Directions

3.1. Notations

In order to simplify, we denote by d_N, d_C, d_H respectively the Newton's, Chebyshev's and Halley's directions. Then, these directions are well defined by:

$$\boxed{F(x_k) + \nabla F(x_k) d_N = 0}$$

$$\boxed{F(x_k) + \nabla F(x_k) d_C + \tfrac{1}{2}\nabla^2 F(x_k) d_N d_N = 0}$$

$$\boxed{F(x_k) + \nabla F(x_k) d_H + \tfrac{1}{2}\nabla^2 F(x_k) d_H d_N = 0}$$

3.2. New Directions

From that remark we can now develop new high order directions.

Definition 1. For $p \geq 2$, a High order Halley ("HoH") type p-order direction (d) is a direction obeying to one of these relations:

$$\boxed{F(x_k) + \nabla F(x_k) d + \tfrac{1}{2}\nabla^2 F(x_k) d_1 d_2 + ... + \tfrac{1}{p!}\nabla^p F(x_k) d_l d_{l+1} ... d_{l+p} = 0} \quad (1)$$

or

$$\boxed{F(x_k) + \nabla F(x_k)d + \tfrac{1}{2}\nabla^2 F(x_k)dd_1 + ... + \tfrac{1}{p!}\nabla^p F(x_k)dd_{l-1}...d_{l+p-2} = 0} \quad (2)$$

where $d_i, i = 1, ..., l + p$ are linear combinations of directions defined by one of these two previous schemes (at an inferior or equal order), and $l = p(p+1)/2-1$.

Remark 2. The previous definition is available for any $p \geq 2$; in this paper, we concentrate our study on the case $p = 3$. Therefore, it is obviously more interesting to study the first family, as, even if it has to solve multiple linear systems, all these systems involve the same linear operator to be inversed. Whereas the second family need to inverse several (up to p) matrices.

3.3. Known Directions are HoH Type Directions

More than Newton's, Halley's and Chebyshev's, the other directions seen in this paper belong also to the HoH definition.

Proposition 1. *SuperHalley's direction is a HoH of second order direction.*

Proof. SuperHalley's method is given by:

$$x_{k+1} = x_k - [I + \frac{1}{2}L(x_k)(I - L(x_k))^{-1}]\nabla F(x_k)^{-1}F(x_k), k = 0, 1, 2, ...$$

and

$$L(x) = \nabla F(x)^{-1}\nabla^2 F(x)\nabla F(x)^{-1}F(x)$$

Thus, SuperHalley's direction can be reformulated as:

$$\boxed{F(x_k) + \nabla F(x_k)d_{SH} + \tfrac{1}{2}\nabla^2 F(x_k)d_V d_N = 0}$$

and

$$\boxed{F(x_k) + \nabla F(x_k)d_V + \tfrac{1}{2}\nabla^2 F(x_k)d_V(2d_N) = 0}$$

Therefore, SuperHalley's direction obeys to the relation (2). □

By reconsidering the extrapolations we first saw that the first order extrapolation is exactly Newton's method.

Proposition 2. *The second order extrapolation is exactly Chebyshev's method. Consequently, it is a second order HoH method.*

Proof. For the second order extrapolation method, the direction is given by:

$$\hat{x}_2 = x + \dot{x}(0-\tau) + \frac{1}{2}\ddot{x}(0-\tau)^2$$

Then,

$$\hat{d}_2 = \hat{x}_2 - x$$

And,

$$\dot{x}(0-\tau) = -\nabla F(x)^{-1}F(x) = d_N$$

Moreover,

$$\ddot{x} = -\nabla F(x)^{-1}\nabla^2 F(x)\dot{x}\dot{x}.$$

Therefore,

$$\hat{d}_2 = \dot{x}(0-\tau) + \frac{1}{2}\ddot{x}(0-\tau)^2$$

$$\hat{d}_2 = -\nabla F(x)^{-1}F(x) - \frac{1}{2}\nabla^{-1}F(x)\nabla^2 F(x)\dot{x}(0-\tau)\dot{x}(0-\tau)$$

$$\boxed{\hat{d}_2 = -\nabla F(x)^{-1}F(x) - \tfrac{1}{2}\nabla F(x)^{-1}\nabla^2 F(x)[\nabla F(x)^{-1}F(x)]^2]}$$

Finally, $\hat{d}_2 = d_C$ which is Chebyshev's direction. Consequently, the second order extrapolation method is exactly the Chebyshev's method □

Proposition 3. *The third order extrapolation direction is a third order Halley type direction.*

Proof. The third order extrapolation direction is given by:

$$\hat{x}_3 = x + \dot{x}(0-\tau) + \frac{1}{2}\ddot{x}(0-\tau)^2 + \frac{1}{6}\dddot{x}(0-\tau)^3$$

As we have

$$\dot{x}(0-\tau) = -\nabla F(x)^{-1}F(x) = d_N$$

And

$$\ddot{x} = -\nabla F(x)^{-1}\nabla^2 F(x)\dot{x}\dot{x}$$

Thus

$$\dddot{x} = -\nabla F(x)^{-1}[\nabla^3 F(x)\dot{x}^3 + 2\nabla^2 F(x)\dot{x}\ddot{x} + \nabla^2 F(x)\ddot{x}\dot{x}]$$

Setting $\hat{d}_3 = \hat{x}_3 - x$

$$\hat{d}_3 = \dot{x}(0-\tau) + \frac{1}{2}\ddot{x}(0-\tau)^2 + \frac{1}{6}\dddot{x}(0-\tau)^3.$$

As we had:

$$d_N = -\nabla F(x)^{-1} F(x),$$

$$d_C = d_N - \frac{1}{2}\nabla F(x)^{-1}\nabla^2 F(x) d_N^2,$$

we can conclude:

$$\boxed{\hat{d}_3 = d_N - \tfrac{1}{2}\nabla F(x)^{-1}\nabla^2 F(x)[d_N(2d_C - d_N)] - \tfrac{1}{6}\nabla F(x)^{-1}\nabla^3 F(x) d_N^3},$$

which is a direction obeying to the third order Halley type relation (1).

$$\boxed{F(x_k) + \nabla F(x_k)\hat{d}_3 + \tfrac{1}{2}\nabla^2 F(x_k) d_N(2d_C - d_N) + \tfrac{1}{6}\nabla^3 F(x_k) d_N^3 = 0}$$

□

3.4. Convergence

In this section, we study the HoH type algorithms convergence. We assume that the function F obeys to the classical hypotheses for convergence, in particular it is p-Fréchet differentiable.

3.4.1. Quartic Convergence

From relations (1) and (2), we are able to define directions which algorithms will have a quartic convergence. But, not all the third order Halley type directions have a fourth order convergence. Thus, it's necessary to choose carefully the directions d_i to inject in the definition scheme. As a counter-example, a direction using only Newton's direction d_N will only have a cubical convergence.

Theorem 1. *In the scalar case ($F : \mathbb{R} \to \mathbb{R}$), some third order Halley type directions have a fourth order convergence. In particular, directions defined by:*

$$F(x_k) + \nabla F(x_k) d_1 + \frac{1}{2}\nabla^2 F(x_k) d_C^2 + \frac{1}{6}\nabla^3 F(x_k) d_C^3 = 0 \tag{3}$$

$$F(x_k) + \nabla F(x_k) d_2 + \frac{1}{2}\nabla^2 F(x_k) d_C^2 + \frac{1}{6}\nabla^3 F(x_k) d_N^3 = 0 \tag{4}$$

$$F(x_k) + \nabla F(x_k)\hat{d}_3 + \frac{1}{2}\nabla^2 F(x_k) d_N(2d_C - d_N) + \frac{1}{6}\nabla^3 F(x_k) d_N^3 = 0 \tag{5}$$

$$F(x_k) + \nabla F(x_k) d_{SH3} + \frac{1}{2}\nabla^2 F(x_k) d_{SH3}\, d_{SH} + \frac{1}{6}\nabla^3 F(x_k) d_{SH3}\, d_{SH}^2 = 0 \tag{6}$$

have a quartic convergence order.

Proof. The main idea of the proof consists in defining the directions as polynomials of the step error. We here provide details only for the equation (6) direction. The proof for other directions is similar. In order to simplify, we denote by f the function F and by f' its first derivative, f'' its second derivative, and so on.

$$f(x^*) = 0.$$

First, we define several quantities: the error at the k step:

$$e_k = x_k - x^*;$$

the error at the $k+1$ step:

$$e_{k+1} = x_{k+1} - x^*;$$

a usefull coefficient:

$$c_k = \frac{f^{(k)}(x^*)}{k!f'(x^*)};$$

the Newton's direction at step k:

$$d_N = -\frac{f(x_k)}{f'(x_k)};$$

the SuperHalley's direction at step k:

$$d_{SH} = -\frac{f + \frac{1}{2}f'' d_N d_V}{f'}, \tag{7}$$

with

$$d_V = -\frac{f}{f' + f'' d_N}.$$

Using theses notations,functions and derivatives expressed as polynomials of e_k are written:

$$f(x_k) = f'(x^*)[e_k + c_2 e_k^2 + c_3 e_k^3 + \mathcal{O}(e_k^4)], \tag{8}$$

$$f'(x_k) = f'(x^*)[1 + 2c_2 e_k + 3c_3 e_k^2 + \mathcal{O}(e_k^3)], \tag{9}$$

$$f''(x_k) = f'(x^*)[2c_2 + 6c_3 e_k + \mathcal{O}(e_k^2)], \tag{10}$$

$$f'''(x_k) = f'(x^*)[6c_3 + \mathcal{O}(e_k)], \tag{11}$$

$$\frac{1}{1+x} = \sum_{n=0}^{\infty}(-1)^n x^n. \tag{12}$$

From equations(12) and (9) we have:

$$\frac{1}{f'(x_k)} = \frac{1}{f'(x^*)}[1 - 2c_2 e_k + (4c_2^2 - 3c_3)e_k^2 + \mathcal{O}(e_k^3)]. \tag{13}$$

Next, we deduce a formula for the Newton's direction ($d_N = -\frac{f(x_k)}{f'(x_k)}$) in terms of e_k^p.

$$d_N = -e_k + c_2 e_k^2 + (2c_3 - 2c_2^2)e_k^3 + \mathcal{O}(e_k^4). \tag{14}$$

Thus, we have all the quantities to express:

$$d_V = -e_k - c_2 ek^2 + (2c_2^2 - 4c_3)ek^3 + \mathcal{O}(ek^4). \tag{15}$$

Then, from equations (7), (8), (10), (13), (14) and (15) we obtain:

$$d_{SH} = -e_k - (4c_2^2 + c_3)ek^3 + \mathcal{O}(ek^4). \tag{16}$$

We can now express d_{SH3} as:

$$d_{SH3} = -\frac{f}{f' + \frac{1}{2}f'' d_{SH} + \frac{1}{6}f''' d_{SH}^2}.$$

Then, from a similar proof, we can deduce:

$$d_{SH3} = -e_k + \mathcal{O}(ek^4). \tag{17}$$

Finally:

$$e_{k+1} = d_{SH3} + e_k,$$

$$\boxed{e_{k+1} = \mathcal{O}(e_k^4)}.$$

Which is exactly a quartic convergence. □

3.4.2. Illustration

We consider three functions from the Moré, Garbow and Hillstrom [11] collection. Four algorithms will be tested on these functions; Newton, Chebyshev, SuperHalley and a third order Halley type algorithm called d_{SH3} and defined by:

$$\boxed{F(x_k) + \nabla F(x_k) d_{SH3} + \frac{1}{2}\nabla^2 F(x_k) d_{SH3}\, d_{SH} + \frac{1}{6}\nabla^3 F(x_k) d_{SH3}\, d_{SH}^2 = 0}$$

The tables (1, 2, 3) represent the error at each step.

Iteration	1	2	3	4
Newton	10^{-2}	10^{-4}	10^{-7}	10^{-14}
Chebyshev	10^{-2}	10^{-6}	10^{-15}	0
SuperHalley	10^{-3}	10^{-10}	10^{-16}	0
SH3	10^{-4}	10^{-15}	0	0

Table 1: Iteration error on Beale's function with a 10^{-2} beginning error

Iteration	1	2	3	4
Newton	10^{-2}	10^{-4}	10^{-9}	10^{-16}
Chebyshev	10^{-3}	10^{-7}	10^{-16}	0
SuperHalley	10^{-5}	10^{-12}	0	0
SH3	10^{-5}	0	0	0

Table 2: Iteration error on Wood's function with a 10^{-2} beginning error

Iteration	1	2	3	4
Newton	10^{-2}	10^{-5}	10^{-7}	10^{-15}
Chebyshev	10^{-2}	10^{-8}	0	0
SuperHalley	10^{-4}	10^{-13}	0	0
SH3	10^{-5}	0	0	0

Table 3: Iteration error on Rosenbrock's function with a 10^{-2} beginning error

We can easily verify that this direction d_{SH3} has a quite good convergence behavior, better than SuperHalley or Chebyshev methods, which have a cubic convergence order.

3.5. Per Iteration Cost

3.5.1. Automatic Differentiation

Automatic differentiation (AD) is a technique which goes back as far as the 50's. This method can be very usefull nowadays thanks to the computer developpement. One can see more about AD in [6]. Using AD, one can rely on the availability of a gradient ∇f of some real valued (differentiable) function at a computational cost as worst 5 times the cost of computing f itself and the factor 5 is very conservative. This is competitive with carefully hand coded gradients. We use this complexity bound to assess the overall complexity of our HoH algorithm.

3.5.2. Per Iteration Cost: Example

In this part, we will consider the per iteration cost of a particular direction. The linear algebra complexity expressions are based on [5]. Calculations for other directions are similar. This direction (d_1) is defined by:

$$\boxed{F(x_k) + \nabla F(x_k)d_1 + \frac{1}{2}\nabla^2 F(x_k)d_C^2 + \frac{1}{6}\nabla^3 F(x_k)d_C^3 = 0}$$

Or:

$$\boxed{d_1 = d_N - \frac{1}{2}\nabla F^{-1}(x_k)\nabla^2 F(x_k)d_C^2 - \frac{1}{6}\nabla F^{-1}(x_k)\nabla^3 F(x_k)d_C^3}$$

- For the Newton's direction : $d_N = -\nabla F(x_k)^{-1}F(x_k)$, costs are given by:

Operation	Cost
Cholesky factorization	$\frac{1}{3}n^3 + \frac{1}{2}n^2 - \frac{5}{6}n$
Solving linear system	$2n^2 - n$
Obtaining derivative ∇F with AD	nc
Cost of F	c
Vector subtraction	n
Total	$\frac{1}{3}n^3 + \frac{5}{2}n^2 - \frac{5}{6}n + (n+1)c$

- For the Chebyshev's direction (having already Newton's direction): $d_C = d_N - \frac{1}{2}\nabla F(x_k)^{-1}\nabla^2 F(x_k)d_N^2$, costs are given by:

Operation	Cost
$\nabla^2 F(x_k)d_N d_N = \nabla(\nabla(F d_N)d_N)$	$25c + 30n$
Solving the linear system	$2n^2 - n$
Vector sum and dot product	$2n$
Total	$2n^2 + 31n + 25c$

- For the third order Halley type direction d_1 (having Chebyshev's and Newton's direction) costs are:

Operation	Cost
$\nabla^2 F(x_k)d_C d_C = \nabla(\nabla(F d_C)d_C)$	$25c + 30n$
$\nabla^2 F(x_k)d_C d_C = \nabla(\nabla(F d_C)d_C)$	$125c + 155n$
Solving linear system	$2n^2 - n$
2 vector sums and 2 dot products	$4n$
Total	$2n^2 + 188n + 150c$

Finally this direction d_1 will cost: $\frac{1}{3}n^3 + \frac{13}{2}n^2 + \frac{1309}{6}n + (n + 176)c$. If we approximate c by $\mathcal{O}(n^2)$ (which is quite realistic, one can observe for example functions from [11]), then this direction will have a per iteraton cost $\mathcal{O}(n^3)$. In [7], Gundersen and Steihaug study sparse matrices (which can constitute a new perspective), but their calculations dont consider the gain obtained from technique explained in Section 2.5.1 (for example to consider $\nabla^2 F(x_k)d_1 d_2$ as $\nabla(\nabla(F d_1)d_2))$. Therefore Gundersen and Steihaug obtain a coeffcient for n^3 bigger than 1/3 for Newton or Chebyshev.

4. Research Perspectives

4.1. Mutli-Dimensional Convergence

Convergence of these new directions has been proved previously in the real case. Then, it will be interesting to demonstrate this convergence in Banach spaces by extending Kantorovich theorem for Newton's method [14]. This theorem simplify the convergence proof by majorizing Newton's error ($||x_k - x^*||$) by a suitable real valued sequence. The same scheme can be used to demonstrate Chebyshev's or SuperHalley's method.

4.2. Shamanskii-Halley Method

In [1], Brent studies how many times one can use the same matrix in some algorithms. Thus, we can consider a method inspired by Shamanskii's. As a matter of fact, Shamanskii's method can be seen, in the case of two iterations, as:

$$x_{k+\frac{1}{2}} = x_k - \nabla F(x_k)^{-1} F(x_k)$$

$$x_{k+1} = x_{k+\frac{1}{2}} - \nabla F(x_k)^{-1} F(x_{k+\frac{1}{2}})$$

Therefore, this method uses the same jacobian for multiple iterations. More generally, we can state the next theorem.

Theorem 2. (see [13]) *If x^* is a regular solution of system $F(x) = 0$ (if $\nabla F(x^*)$ is invertible), then the iterative process:*

$$x_{k+\frac{i}{d}} = x_{k+\frac{i-1}{d}} - \nabla F(x_k)^{-1} F(x_{k+\frac{i-1}{d}})$$

converges locally to x^* and the sequence $(x_k)_{k\geq 1}$ for integer values of k converges with an order $d + 1$.

From this method, we deduce a new algorithm based on Shamanskii's and Chebyshev's method. This method can be written as:

$$x_{k+\frac{1}{2}} = x_k - \nabla F(x_k)^{-1} F(x_k) - \frac{1}{2} \nabla F(x_k)^{-1} \nabla^2 (x_k) [\nabla F(x_k)^{-1} F(x_k)]^2$$

$$x_{k+1} = x_{k+\frac{1}{2}} - \nabla F(x_k)^{-1} F(x_{k+\frac{1}{2}}) - \frac{1}{2} \nabla F(x_k)^{-1} \nabla^2 (x_{k+\frac{1}{2}}) [\nabla F(x_k)^{-1} F(x_{k+\frac{1}{2}})]^2$$

From the same idea, we can generalize this method for multiple iterations instead of two:

$$\begin{aligned} x_{k+\frac{i}{d}} &= x_{k+\frac{i-1}{d}} - \nabla F(x_k)^{-1} F(x_{k+\frac{i-1}{d}}) \\ &\quad - \frac{1}{2} \nabla F(x_k)^{-1} \nabla^2 (x_{k+\frac{i-1}{d}}) [\nabla F(x_k)^{-1} F(x_{k+\frac{i-1}{d}})]^2 \end{aligned}$$

The main goal of such a method is to get how many times we use the same jacobian. Finally, we can also use one of the HoH type direction instead of Chebyshev's.

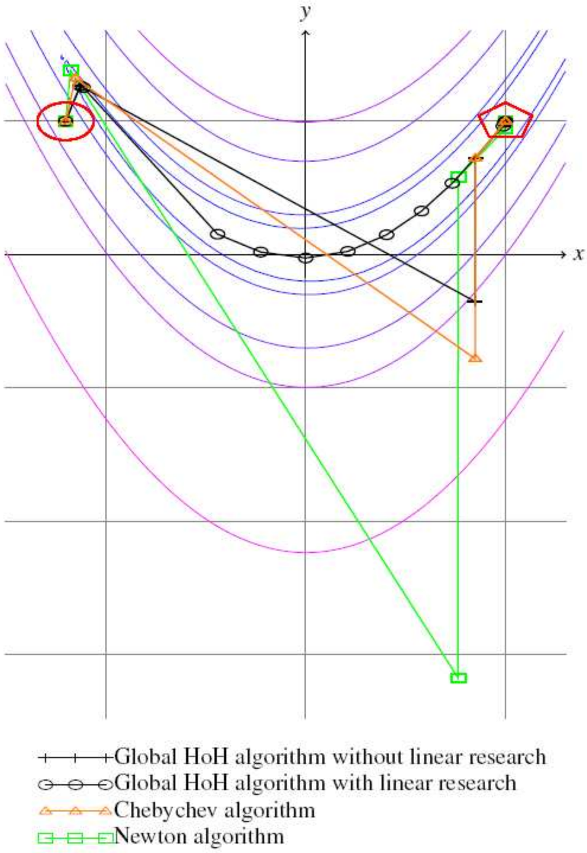

Figure 1: Newton's, Chebyshev's and Global HoH's direction tested on Rosenbrock's function

4.3. A Global HoH Algorithm

Another important aspect to develop in Banach spaces is a global method: this method should use a first, second or a third order Halley type direction depending on the current point x_k of the algorihm. For example, this method can start with Newton's direction, and therefore, using a general criterion, continues to use the same direction or choose another one. This criterion could be the ratio cost/convergence speed. Brent [1] developed such an approach based on "efficiency" of an algorithm.

For the time being, we tested a basic global HoH algorithm. The criterion to decide which direction should be taken could be summarized as "the higher extrapolation order as far as it's a desent direction". Therefore, this global HoH direction evaluates at each current point x_k the sign of $\nabla F(x_k)\hat{d}_3$. If it's negative, then the third order extrapolation direction is choosen; otherwise, algorithm checks sign of $\nabla F(x_k)d_C$, and takes Chebyshev's direction if it's negative. If not, Newton's direction is used by default. Moreover, we used an Armijo type simple line search to enforse global convergence. Thus, this global HoH algorithm can be written as: $x_{k+1} = x_k + \theta\hat{d}_i$, where $i = max(1;2;3)/sign(\nabla F(x_k)\hat{d}_i) \leq 0$.

4.3.1. Graphical Illustration

The next figure illustrates Newton's, Chebyshev's and our global HoH directions tested on Rosenbrock's function. One can see efficiency of HoH direction compared to classical methods. This figure also underlines purpose of a linear research: on one hand it slows down the algorithm, but in the other hand it ensure convergence of the method.

In conclusion, we presented a new approach for high order methods for non linear and differentiable optimization. We can see these methods as high order directions, and by combining subtly such directions, we can obtain new higher order methods. We developed two families in this paper: the first corresponds to classical Halley's method and allows us to expand the definition of this algorithm; while the second one develops Chebyshev's method. From both of these families, we can obtain new methods having an interesting behavior in terms of convergence. And more precisely, the second family (Chebyshev) has quite good complexities. This should constitute a forthcoming research.

References

[1] R.P. Brent, Some efficient algorithms for solving systems of nonlinear equations, *SIAM JNA*, **10**, No. 2 (1973), 327-344.

[2] P.L. Chebyshev, *Collected Works*, Number 5, Moscou-Leningrad (1951).

[3] J.-P. Dussault, High-order Newton-penalty algorithms, *J. Comput. Appl. Math.*, **182**, No. 1 (2005), 117-133.

[4] J.B.J. Fourrier, *Question d'Analyse Algébrique*, Oeuvres complétes (II), Gauthiers-Villars, Paris (1890), 243-253.

[5] L. Fox, *An Introduction to Numerical Linear Algebra*, Oxford (1964), 328p.

[6] A. Griewank, A. Walther, *Evaluating Derivatives, Principles and Techniques of Algorithmic Differentiation*, Philadelphia (2008).

[7] G. Gundersen, T. Steihaug, On large scale unconstrained optimization problems and higher order methods, *Optimization Methods and Software*, **25**, No. 3 (2010), 337-358.

[8] J.M. Gutiérrez, M.A. Hernández, An acceleration of Newton's method: Super-Halley method, *Appl. Math. Comput.*, **117**, No-s: 2-3 (2001), 223-239.

[9] E. Halley, A new, exact and easy method of finding roots of any equations generally, and that without any previous reduction, *Philos. Trans. Roy. Soc. London*, **18** (1694), 136-145.

[10] A.L. Cauchy, *Note sur la détermination approximative des racines d'une équation algébrique ou transcendante, Leons sur le calcul différentiel*, Buré frères (1823), 573-609.

[11] B.S. Moré, J.J. Garbow, K.E. Hillstrom, Testing unconstrained optimization software, *ACM Transactions on Mathematical Software*, **7**, No. 1 (1981), 17-41.

[12] I. Newton, *Methodus Fluxionium et Serierum Infinitarum* (1664-1671).

[13] J.T. Ortega, W.C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, New York (1970), 572p.

[14] L.B. Rall, R.A. Tapia, *The Kantorovich Theorem and Error Estimates for Newton's Method*, Mathematics Research Center, 1970.

[15] J. Raphson, *Analysis Aequationum Universalis*, London (1690).