Study of iterative processes in computer graphic's problems

Jean-Marc Brun – Sebti Foufou – Abdelaziz Bouras
LISPI/LIGIA, bât. 710
Université Claude Bernard LYON 1
43, Bd du 11 Novembre 1918
69622 Villeurbanne Cedex

Philippe Arthaud
Coretech International
27-29, Bd du 11 Novembre 1918 B.P. 2131
69622 Villeurbanne Cedex

ABSTRACT. *The convergency of iterative processes can exhibit unexpected behaviours. In this paper an analysis of the convergency of the iterative processes is made through the application of the Newton method to the polynomial equations resolution. This leads to the definition of a dynamic relaxation scheme applied to computer graphics*

KEY WORDS : *Iterative processes, convergency, relaxation, equations, graphics*

1 Introduction

The behavior of a system evaluated at a finite order (generally 1) can be described although the corresponding equation have an unknown analytic solution; numerical solutions with iterative character are obtained. Theoretically these algorithms converge and their precision is only limited by number of iterations.
In practice these algorithms behavior varies a lot from a case of calculation to the other. If the Newton method [2] is used as an archetype of these algorithms, one observes generally an extremely rapid convergence from an initial value "reasonably" close to the solution. Unfortunately the method can converge very slowly [1], or refuse to converge in certain limit cases. A graphic representation for the method explains the reasons for this behavior (Figure 1).

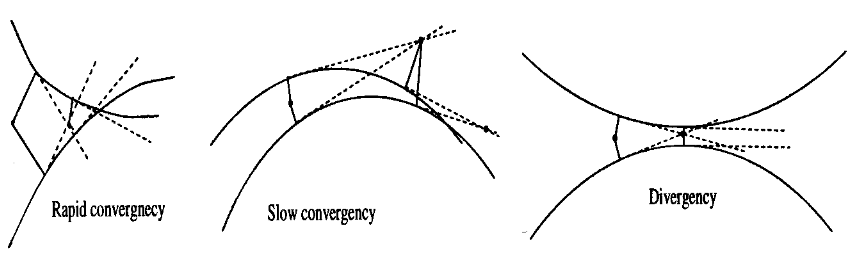

Figure 1: Behavior of the Newton algorithm

The closer the intersection is to a tangency case, the slower the convergence is and the bigger are the numerical errors. Up to a point where the numerical accuracy will be able to turn a case of tangency into a case of intersection or non intersection [4].
When the numerical precision transforms a case of tangency into a case of non intersection, the quasi convergent sequence 0,1,2,3 (Figure 2) goes too far and is followed by a a convergency in the opposite direction 4,5,6 etc. One might notice that this chaotic process

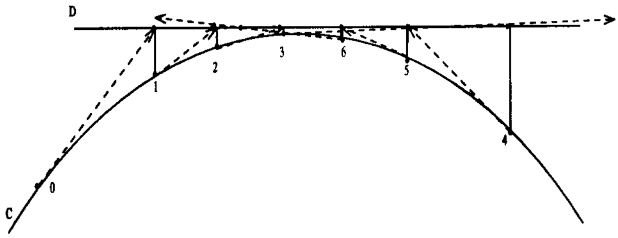

Figure 2: case of tangency

can result in a tangent to the curve C which would be so close to a parallel to D that an "overflow" has not to be excluded.

2 Fundamental problem underlying limit cases

A shaper study of limit cases shows that the tangency point accuracy precision is not commensurable to the numerical tangency. In such a study of the tangency point between a curve C and a line D [4], C is to be replaced by its osculator circle, and numerical errors turn to an error ε on the radius R of the circle (Figure 3).

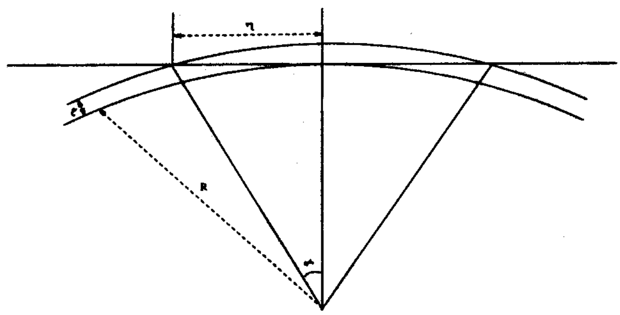

Figure 3: Utilization of the osculator circle in cases of tangency

When the error ε is negative there is neither tangency nor intersection, but when ε is positive the tangency gives place to two distinct points of intersection on D on either side of the tangency point with an error η. One clearly notice that errors ε and η are connected to the angle α by relationships :

$$\alpha = \frac{\eta}{R} \quad , \quad \varepsilon = R\frac{\alpha^2}{2} + O(\alpha^4) \tag{1}$$

hence:

$$\frac{\varepsilon}{R} = \frac{\eta^2}{2R^2} \quad , \quad \frac{\eta}{\varepsilon} = 2\frac{R}{\eta} \tag{2}$$

Therefore if $\eta \longrightarrow 0$, $\frac{\eta}{\varepsilon} \longrightarrow \infty$. both magnitudes are not comparable.

3 iterative Processes and equations

A natural approach of the behavior of iterative processes consists in studying the processes of equation resolution and especially the Newton method when the formal solution of an equation is unknown or that the equations setting cannot be achieved. Although what we can learn from this study account is of great importance, one can not limit the the

behavior of iterative systems to the Newton method, moreover the Newton method is not to be limited to the resolution of polynomial equations [2].
Indeed one might notice that an equation $P(x) = 0$ focuses only on behaviors of oscillation around an extremity, jump from a solution to another and of a more or less rapid convergence, but no regular or/and explosive divergence.
Other behaviors can be found in the resolution of the system :

$$\begin{cases} y = f(x) \\ x = g(y) \end{cases} \tag{3}$$

An iterative process which alternately evaluates a function then the other is generally used :

$$y_1 = f(x_0),\ x_1 = g(y_1),\ y_2 = f(x_1),\ x_2 = g(y_2),\ \cdots \tag{4}$$

According to the relative position of the curves $y = f(x)$ and $x = g(y)$ this process converges or diverges steadily, some solutions can not be reached and, at first the process can seem to be erratic.

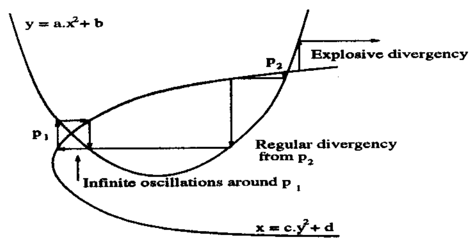

Figure 4: An iterative processes

One can, however, invert divergent processes when following the curves :

$$x = f^{-1}(y), y = g^{-1}(x) \tag{5}$$

4 Study of iterative processes as series

An iterative process can be seen as series of values or points and the process is convergent when the associated series is convergent too.
Indeed iterative processes make up the series p_n by means of the calculation of dp_n :

$$dp_n = f(p_n) \quad , \quad p_{n+1} = p_n + dp_n \tag{6}$$

$$p_{n+1} = p_0 + dp_0 + dp_1 + dp_2 + \cdots + dp_n \tag{7}$$

So one might notice that any convergent continuation/series (7) is bounded by a geometric series :

$$G_n = G_0(1 + q + q^2 + \cdots + q^n) \tag{8}$$

of the reason $-1 < q < 1$. Such as $q_n = \frac{dp_{n+1}}{dp_n}$ and $| q_n | < | q |$.
A very rapid convergence will correspond to q_n tending towards zero when n increases. It is typically the case of the Newton method in favorable cases.
On the other hand, in unfavorable cases, especially in the cases of tangency, the reason q_n

does not tend towards zero but towards a limit $q \neq 0$ and the convergence is all the more slow since $\mid q \mid$ is close to 1. When the convergence is very slow, if the reason q_n rightly tends towards a limit, one can safely predict the limit of the series :

$$p_\infty = p_n + \frac{dp_n}{1 - q_n} \tag{9}$$

This method of convergence study is general and can be applied to any iterative process. Therefore one can use it for the resolution of an equation through the Newton method as well as for the resolution of the coupled system (3) or for geometric problems like projection on a curve or parametric surface.

5 Equation resolution by the Newton method

The Newton method can be used to solve any equation $P(x) = 0$ if one knows (how to calculate) $P'(x)$ and if one can avoid to leave the $P(x)$ domain of definition. The behavior of the process is generally excellent, i.e. that it converges with an increasing acceleration of the convergence ($q_n \longrightarrow 0$) and the process can be safely stopped at the iteration n by checking the correction dx_n because the final error on the value of x is bounded by :

$$x - (x_n + dx_n) < dx_n\left(1 - \frac{1}{1-q}\right) \simeq q dx_n \tag{10}$$

5.1 Study of a polynomial equation

To study the behavior of the Newton method we are going to apply it to the resolution of polynomial equation (12), starting from a value x_0.

$$P(x) = x^6 + 6x^5 + 6x^4 - 18x^3 - 31x^2 - 24x - 36 = 0 \tag{11}$$

$$P(x) = (x+3)(x+3)(x+2)(x-2)(x^2+1) = 0 \tag{12}$$

on the $P(x)$ graph (figure 5) one can see that all the solutions are included in the area $a < x < b$ (area where $P(x) \simeq 0$) depending on the drawing scale. The study of the convergence follows the following schema :

$$\begin{array}{lllll}
x_0 \rightarrow & dx_0 \searrow & & & \\
& & q_0 & \rightarrow & \varepsilon_0 = (1 - \frac{1}{1-q_0})dx_1 \\
x_1 \rightarrow & dx_1 \nearrow\searrow & & & \\
& & q_1 & \rightarrow & \varepsilon_1 \\
x_2 \rightarrow & dx_2 \nearrow\searrow & & & \\
& & q_2 & \rightarrow & \varepsilon_2 \\
x_3 \rightarrow & dx_3 \nearrow & & &
\end{array}$$

and ends up with the results presented on the table below.

Iteration	x_i	dx_i	q_{i-1}	ε_{i-1}
0	$x_0 = 10$	$dx_0 = -2$		
1	$x_1 = 8$	$dx_1 = -1.6$	$q_0 = 0.8$	$\varepsilon_0 = 6.4$
2	$x_2 = 6.4$	$dx_2 = -1.273$	$q_1 = 0.795$	$\varepsilon_1 = 4.94$
3	$x_3 = 5.127$	$dx_3 = -1.006$	$q_2 = 0.79$	$\varepsilon_2 = 3.8$
4	$x_4 = 4.121$	$dx_4 = -0.784$	$q_3 = 0.78$	$\varepsilon_3 = 2.78$

One can see that the process converges very slowly, with an important limit error ε_i, but that the convergence accelerates since $q_{i+1} < q_i$.
One will notice that this will be a typical behavior of polynomial equations because $P(x)$

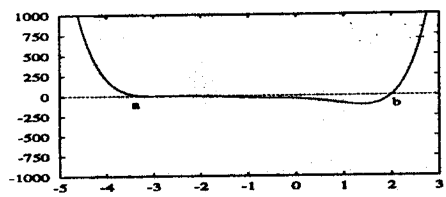

Figure 5: Approximat Graph of $P(x)$

behaves as its highest term when $x \longrightarrow \infty$ [7]. on the other hand in the general case there is no reason to observe this kind of behavior, the equation $sin(x) = 0$ for instance will give solutions around any initial value x_0 (sufficiently close to $k\pi$).

5.2 Behavior in the solutions' zone

The studied equation shows however, a certain number of typical behavior in the zone where solutions (- 3, -2,2) are found [6]; these behaviors are directly linked to the departure point of the process.
The study of the convergence by means of values dx_i and reasons q_i of the geometric series which estimates the limit x_∞ shows that :

- $q_i \longrightarrow 0$ when the convergence is rapid ($x_0 = 2.15$)
- $q_i \longrightarrow q$ constant when the convergence is slow in the vicinity of a double solution ($x_0 = -3.5$). The limit value of the geometric series gives then the solution with a great accuracy.
- The explosion of the process characterizes itself by an important correction dx_i, where $x_0 = -2.409$.
- The jump from a solution to another characterizes itself by strong and irregular variations of the coefficients q_i. (with $x_0 = 1.325$ one jumps over the solution $x = -2$ to find $x = 2$. With $x_0 = 1.325$ one forgets $x = 2$ and jumps over $x = -2$ to find $x = -3$ as for $x_0 = -3.5$).

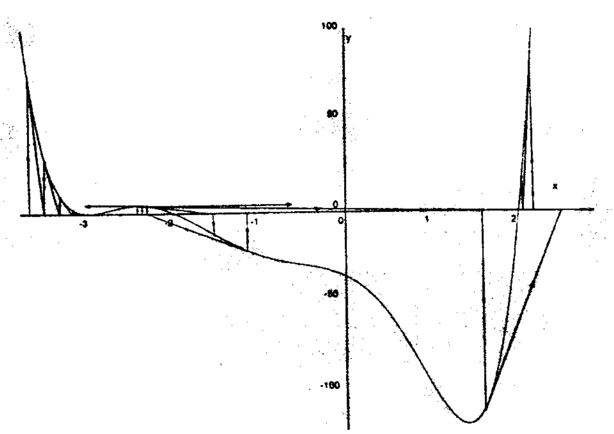

Figure 6: Behavior in the zone where solutions are found

- Alternate convergence around a point of inflexion of the curve $P(x)$ where $(q_i < 0)$, $(x_0 = -1.135)$.

So one observes that regular convergence zones (with q_i constant or tending steadily towards 0) are zones where there is no other solution to look for than the limit, whereas the zones where q_i varies irregularly are zones to study in a closer way.

5.3 control of the process behavior through relaxation

With a pure Newton process, this close study can only be considered by multiplying the points of departure so as to split the area into zones :

- Convergent in regular manner
- Without possible solution (zones with extremity)

By and large behaviors can be more difficult to control, and phenomena of aliasing can appear with periodic functions. To look for the solution of $cos(x) = 0$, starting from $x_0 = Atan(\frac{1}{2\pi}) = 0.1578311$ make jump 2π at each iteration, if one is very close to this value one might believe to have found a convergent process with a slow convergency ! Therefore a robust method to control the convergency of iterative processes consists in limiting the process steps when they are too important. As a result, one has to replace the original series by a series :

$$x = x_0 + dx_0^1 + dx_1^1 + \cdots + dx_n^1 + \cdots \tag{13}$$

The terms $dx_i^1 = r.dx_i$ are calculated by relaxation [11] on the calculation process of the corrections dx_i. The convergence will be slower if $r < 1$ and more rapid if $r > 1$, provided that $r < r_{lim}$. The term relaxation or on-relaxation will be used then.

6 Iterative processes in parametric form

Roughly, an iterative process is characterized by the search of a value of parameter t for which a given condition $C(t)$ is achieved. This condition can appear as $C(t) = 0$ (0 scalar or vectorial) without possibility to solve the equation $C(t) = 0$. Then one can define $g(t)$ and the series t_i :

$$\begin{aligned} t &= t_0 + g(t_0) + \cdots + g(t_n) + \cdots \\ t_{n+1} &= t_n + g(t_n) \end{aligned} \tag{14}$$

The function $g(t)$ was built in an attempt to near to the solution of the equation $C(t) = 0$, so it has to check $g(t) = 0$ when $C(t) = 0$. The functions $g(t)$ will be (by construction)

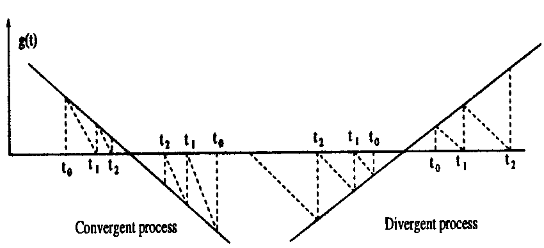

Figure 7: Linear iterative processes

symmetrical around the solution $g(t) = 0$. If one merges $g(t)$ and its tangent around

$g(t) = 0$ one obtains the cases in figure 7. If the slope of the tangent is negative the process is convergent, otherwise it is divergent, the series that calculates t can be written :

$$t = t_0 + g(t_0) + g(t_0 + g(t_0)) + \cdots \tag{15}$$

with $g(t_i) = a(t_i - t)$ one finds :

$$\begin{aligned} t_{i+1} &= t_i + g(t_i) = t_i + a(t_i - t) = (1+a)t_i - at \qquad (16) \\ g(t_{i+1}) &= a(((1+a)t_i - at) - t) = a(1+a)(t_i - t) = (1+a)g(t) \\ t &= t_0 + g(t_0)(1 + (1+a) + (1+a)^2 + \cdots) \\ t &= t_0 + g(t_0)\frac{1}{1-(1+a)} = t_0 + rg(t_0) \qquad (17) \end{aligned}$$

limit of the geometric series with reason $(1+a)$ that converges when $1+a < 1$, i.e. $a < 0$. One might notice that when $a > 0$ the series diverges but one can always find the intersection between the line and $g(t) = 0$. In the general case the curve $g(t)$ is not a line and to assimilate the convergence to that of a geometric series consists in replacing the curve $g(t)$ by the cord. One obtains then :

$$t_{i+1} = t_i + r_i.g(t_i) \quad , \quad r_i = \frac{t_i - t_{i-1}}{g(t_{i-1}) - g(t_i)} \tag{18}$$

that can be applied as soon as $t1 = t0 + g(t0)$ and $g(t1)$ are calculated.
The initial process has been replaced by a dynamic relaxation process with a coefficient r_i which converges more rapidly than the initial process and replaces a divergent process by the convergent symmetrical process. The only condition needed to obtain a convergent

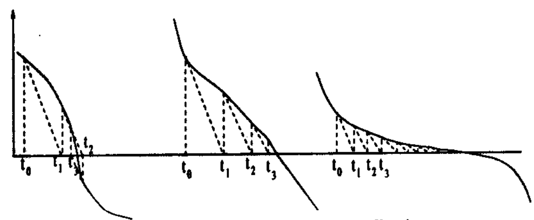

Figure 8: Non linear iterative processes

process is to have a function $g(t)$ continuous and symmetrical along the solution (where $g(t)$=0).
However, an unreasoned application can provoke explosions of the process, especially when r_i is important and when it varies a lot, which is the case in zones "parallel" to the solution where $g(t) = g(t_{i-1})$ and in the vicinity of limit cases (quasi-tangency for example).
Therefore the reasoned application of the dynamic relaxation principle consists in slowing down the process when $g(t_i)$ is too important, and conversely in accelerating it, when r_i is stable, even if it leads to important corrections.

6.1 Example of dynamic relaxation : projection of a point on a curve

To project a point (x_1, y_1) on a parametric curve $(x(t), y(t))$ turns to find the value of t for which :

$$\begin{cases} x(t) = x_1 \\ y(t) = y_1 \end{cases} \tag{19}$$

Let P a known point, one looks for $p(t)$ on the curve so that the projection of P on the tangent in t to the curve be merged with $p(t)$. We have applied the study of the convergence (dynamic relaxation) on the process of the projection of a point on a Bézier curve. The results are presented in the table below :

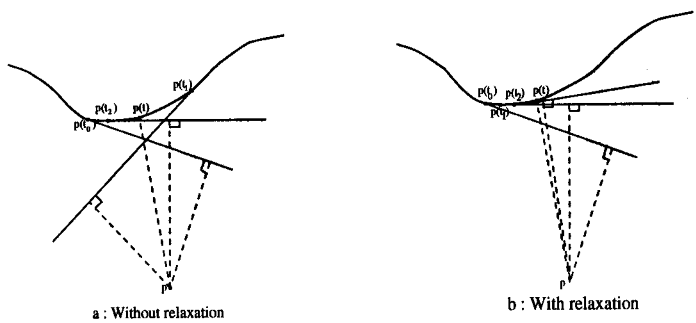

Figure 9: Projection of a point on a curve (Three iterations)

Iteration	*Without relaxation*		*with relaxation*	
	t_i	dt_i	t_i	dt_i
1	0	0.3993	0	0.3993
2	0.3993	-0.2534	0.0999	0.2959
3	0.1459	0.2238	0.3861	-0.2346
4	0.3697	-0.2091	0.2596	0.0073
5	0.1606	0.1989	0.2634	-7.6812e-04
6	0.3595	-0.1922	0.2630	8.3075e-07
7	0.1673	0.1872	0.2630	8.5722e-11
⋮				
1000	0.3589	-0.1993		

One might notice that without the dynamic relaxation, the algorithm keeps oscillating whereas with relaxation it converges after 7 iterations. This behavior is shown in figure 10.

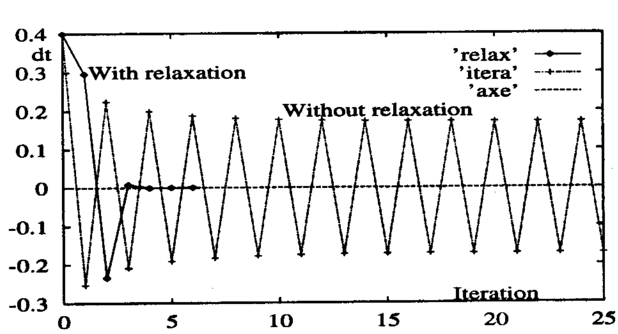

Figure 10: dt plotted against iteration number

7 Acceleration of convergence for series

In some series the variations dt_i cannot always be calculated according to t_i. In such series one cannot use the dynamic relaxation method. However it is possible to study the convergence of the series and to replace the series t_n by an other one corresponding to the successive limits of the geometric series which evaluate, by extrapolation, the sum of non calculated terms. This process can be applied recursively and compared to other classic processes such as the extrapolation of Richardson, the process Δ^2 of Aitken [9] or the ε algorithm of Wynn [8]. The diagram is represented by the figure (11), from the series $t_i^0 = t_i$ one calculates $dt_i^0 = t_{i+1}^0 - t_i^0$ then a series of limits $t_i^1 = t_i^0 + r_i^0 dt_i^0$ with

$$r_i^0 = \frac{t_{i+1}^0 - t_i^0}{dt_i^0 dt_{i+1}^0} \tag{20}$$

and so on and so forth :

$$t_i^{j+1} = t_i^j + r_i^j dt_i^j \quad , \quad dt_i^j = t_{i+1}^j - t_i^j \quad , \quad r_i^j = \frac{t_{i+1}^j - t_i^j}{dt_i^j - dt_{i+1}^j} \tag{21}$$

One might notice that values t_0^j which improve the accuracy require the calculation of $(2j+1)$ terms. By applying this method to the calculation of a root of the equation (12)

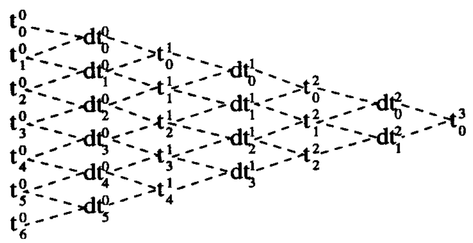

Figure 11: Acceleration of convergence for series

by starting from $t_0 = -3.5$ one finds for instance :

Newton	Successive Limits	ε Algorithm
$t_2^0 = -3.186974$	$t_0^1 = -2.91316$	$t_2 = -2.91316$
$t_4^0 = -3.057165$	$t_0^2 = -3.00886$	$t_4 = -3.02154$
$t_6^0 = -3.01535167$	$t_0^3 = -2.99988$	$t_6 = -2.99802$

One can see that the improvements brought about in relation to the Newton method are even more spectacular than the ε algorithm of Wynn, without touching to the efficiency of the dynamic relaxation which gives improvements to each term t_i and not only for even terms :

$$t_3 = -3.0220 \quad , \quad t_4 = -3.0018 \quad , \quad t_5 = -2.999965 \quad , \quad t_6 = -3.000003$$

8 Conclusion

The analysis that has been made of iterative processes highlighted the fact that the prominently rapid convergence of the Newton method could produce in turn chaotic behaviors or continued divergence. These behaviors could be characterized by the variations of a dynamic relaxation coefficient which stemmed from the definition of a limit of the process

by a geometric series. The use of this dynamic relaxation coefficient makes it possible to replace the initial process by a more rapidly converging one in limit cases and provides a more realistic limit of the error made. However, this method remains applicable to divergent processes.
Finally, when there is no solution, the dynamic relaxation provides an oscillating process. When it is not possible to use the dynamic relaxation schema another one is proposed to accelerate the convergency of very slowly convergent series. The results obtained by this method of series (series (series (...))) make a promising alternative to classic methods.

Références

[1] W.H. Press & S.A. Tenkolsky & W.T. Wetterling & B.P. Flannery, *Numerical Recipes in C, the art of scientific computing*, Cambridge université presse, 1992.

[2] J. Stoer & R. Bulirsch, *Introduction to Numerical Analysis*, Springer, New York.

[3] M.E. Rosen, *A Review of Quasi-Newton Methods in Nonlinear Equation Solving and Unconstrained Optimization*, Proc. ACM 21st Nat. Conf., ACM Pub. P-66, Thompson Book Co., Washingthon, D. C., pp. 37-41, 1966.

[4] G. Mullenheim, *Convergence of a Surface/Surface intersection algorithm*, Computer Aided Geometric Design, v7, p415-423, 1990.

[5] A.E.Midleditch & T.W. Stay & S.B. Tor, *Intersection algorithms for lines and circles*, acm transaction on graphics, v7, n1, p25-40 Jan 1989.

[6] M. Mignotte & M. Payafar, *Distance entre les racines d'un polynôme*, Analyse numérique/Numerical analysis, v13, n2, p181-192 1979.

[7] J. Dumontet, *A propos de l'optimalité des algorithmes némuriques : étude du calcul de x^n*, Informatique théorique/Theoretical computer science, v12, n2, p109-123, 1978.

[8] C. Brezinski, *algorithmes d'accélération de la convergence étude numérique*, Edition Technip, 1978.

[9] J.P. Delahaye, *Optimalité du procédé Δ^2 d'aitken pour l'accélération de la convergence linéaire*, Analyse numérique/Numerical analysis, v15, n4, p321-330 1981.

[10] M. Cosnard, *Caractérisation de la convergence locale de la méthode des approximations successives*, Analyse numérique/Numerical analysis, v11, n3, p225-240, 1977.

[11] J. Cea R. Glowinski, *Sur des méthode d'optimisation par relaxation*, Analyse numérique/Numerical analysis.

[12] L. Gonzalez-vega & H. Lombardi & T. Recio & M.F. Roy, *Spécialisation de la suite de Strum et sous-résultats (I)*, Informatique théorique/Theoretical computer science, v24, n6, p661-588, 1990.