I. INTRODUCCIÓN AL MÉTODO DE MULLER.

Originalmente en [2] se planteó un método para encontrar raíces de funciones algebraicas a través de un procedimiento iterativo que consiste en utilizar un polinomio interpolante de segundo grado para encontrar su raíz más cercana al valor de la ecuación que se quiere resolver. Partiendo de dicha premisa es que a continuación se presenta una deducción y análisis del método, sin embargo, primero es necesario definir algunos preliminares matemáticos que serán útiles para justificar argumentos hechos posteriormente y en consecuencia presentar el método en su totalidad.

II. PRELIMINARES MATEMÁTICOS.

II-A. Polinomio interpolante.

Si se conoce el valor de una función $f(x)$ en n+1 puntos defínase el polinomio interpolante como un polinomio de grado n que pasa por los n+1 puntos previamente establecidos y se utiliza para aproximar los valores de $f(x)$ en las cercanías de cada punto. La deducción más detallada de cómo se puede encontrar dicho polinomio y toda la teoría detrás se salen del ámbito de esta investigación, sin embargo, una explicación puede encontrarse en cualquier texto sobre análisis y métodos numéricos (véase por ejemplo [1]).

II-B. Polinomio interpolante de Newton.

Definición II.1. *Dado un conjunto distinto de puntos $(x_n, f(x_n)), \ldots, (x_0, f(x_0))$ se define el polinomio interpolante de Newton como un polinomio de grado n de la forma*

$$\begin{aligned} P(x) &= f[x_n] + f[x_n, x_{n-1}](x - x_n) \\ &\quad + f[x_n, x_{n-1}, x_{n-2}](x - x_n)(x - x_{n-1}) \\ &\quad + \cdots + f[x_n, \ldots, x_0](x - x_n)(x - x_{n-1}) \cdots (x - x_1) \end{aligned} \tag{II.1}$$

donde $f[x_n, \ldots, x_i]$ es la operación conocida como diferencia dividida de Newton, la cual se definirá más adelante. Del anterior planteamiento es posible observar cómo los n+1 puntos acaban definiendo un polinomio de grado n que sirve para interpolar la función en los valores conocidos.

II-C. Cota de error de un polinomio interpolante

Considerando entonces que un polinomio interpolante aproxima una función en las cercanías de los n+1 puntos que lo definen, es posible estimar cuál es la magnitud máxima del error que se comente al realizar tal aproximación. Sea $P(x)$ el polinomio interpolante y $f(x)$ una función con n+1 derivadas continuas definidas en el menor intervalo I que contenga los números reales $[x_n, x_{n-1}, \ldots, x_0]$, se define el error cometido como la expresión

$$f(x) - P(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_n) \cdots (x - x_0) \tag{II.2}$$

con $\xi \in I$.
De nuevo la demostración de la anterior expresión va más allá del ámbito de este trabajo y puede encontrarse en cualquier libro de análisis numérico.

II-D. Diferencias divididas.

Dada una función $f(x)$ se definen sus diferencias divididas como una operación recursiva de la forma

$$\begin{aligned} f[x_n] &= f(x_n) \\ f[x_n, x_{n-1}] &= \frac{f[x_{n-1}] - f[x_n]}{x_{n-1} - x_n} \\ &\vdots \\ f[x_n, \ldots, x_0] &= \frac{f[x_{n-1}, \ldots, x_0] - f[x_n, \ldots, x_1]}{x_0 - x_n} \end{aligned} \tag{II.3}$$

Donde $f[x_n, x_{n-1}]$ es la primera diferencia dividida, $f[x_n, x_{n-1}, x_{n-2}]$ es la segunda diferencia dividida y así sucesivamente.

Lemma II.1. $f[x_n, \ldots, x_0] = \frac{f^n(\xi)}{n!}$ *para algún $\xi \in I$, donde I es el intervalo más pequeño que contiene los números reales $[x_n, \ldots, x_0]$.*

Demostración. Primero se plantea una función $f(x) \in C^n I$ aproximada por un polinomio interpolante $P(x)$ de grado n partiendo de n+1 puntos $[x_n, \ldots, x_0]$ establecido anteriormente en la definición (II.1). Dado que $P(x)$ es una interpolación de $f(x)$ se plantea la función $g(x) = f(x) - P(x)$ la cual tiene n+1 raíces, así que por el teorema de Rolle generalizado se tiene que:

$$\begin{aligned} g^n(\xi) &= f^n(\xi) - P^n(\xi) = 0 \\ g^n(\xi) &= f^n(\xi) - f[x_n, \ldots, x_0]n! = 0 \\ f[x_n, \ldots, x_0] &= \frac{f^n(\xi)}{n!} \end{aligned} \tag{II.4}$$

$\square$

III. DEDUCCIÓN DEL MÉTODO.

Para encontrar alguna raíz de la función algebraica $f(x)$ primero se parte de tres puntos $\{(x_2, f(x_2)), (x_1, f(x_1)), (x_0, f(x_0))\}$ que sirven como base para encontrar un polinomio interpolante de segundo grado que aproxime la función $f(x)$. Sin perder generalidad se asume que $x_2 < x_1 < x_0$ y se define $w(x) = x - x_2$, por lo que el polinomio interpolante tiene la forma

$$P(w) = aw^2 + bw + c \tag{III.1}$$

El motivo de esto es que se quiere encontrar el valor más pequeño $x - x_2$ que sea al mismo tiempo una raíz del polinomio así que lo planteado en III.1 simplifica ligeramente el álgebra a desarrollar para encontrar los valores buscados en la ecuación.

De lo anterior entonces es necesario encontrar los valores de a, b, c para saber cuál es el polinomio interpolante en los puntos establecido, para esto se plantea el siguiente sistema de ecuaciones

$$P(w(x_2)) = c = f(x_2) \tag{III.2}$$

$$P(w(x_1)) = a(x_1 - x_2)^2 + b(x_1 - x_2) + f(x_2) = f(x_1) \tag{III.3}$$

$$P(w(x_0)) = a(x_0 - x_2)^2 + b(x_0 - x_2) + f(x_2) = f(x_0) \tag{III.4}$$

Sustrayendo III.3 de III.4

$$\begin{aligned} a\Big((x_0 - x_2)^2 - (x_1 - x_2)^2\Big) + b\Big((x_0 - x_2) - (x_1 - x_2)\Big) \\ = f(x_0) - f(x_1) \end{aligned}$$

$$\begin{aligned} a\Big[\Big((x_0 - x_2) - (x_1 - x_2)\Big)\Big((x_0 - x_2) + (x_1 - x_2)\Big)\Big] \\ + b\Big((x_0 - x_2) - (x_1 - x_2)\Big) = f(x_0) - f(x_1) \end{aligned}$$

$$a\Big[(x_0 - x_2) + (x_1 - x_2)\Big] + b = \frac{f(x_0) - f(x_1)}{(x_0 - x_2) - (x_1 - x_2)}$$
$$a\Big[(x_0 - x_2) + (x_1 - x_2)\Big] + b = \frac{f(x_0) - f(x_1)}{x_0 - x_1}$$
$$a\Big[(x_0 - x_2) + (x_1 - x_2)\Big] + b = f[x_0, x_1]$$
$$b = f[x_0, x_1] - a\Big[(x_0 - x_2) + (x_1 - x_2)\Big] \qquad \text{(III.5)}$$

Ahora sustituyendo III.5 en III.3

$$a(x_1 - x_2)^2 + \Big(f[x_0, x_1] - a(x_0 + x_1 + 2x_2)\Big)(x_1 - x_2) + f(x_2) = f(x_1)$$

$$a(x_1 - x_2)^2 - a(x_0 + x_1 - 2x_2)(x_1 - x_2) + f[x_0, x_1](x_1 - x_2) + f(x_2) = f(x_1)$$

$$a\Big((x_1 - x_2)(x_1 - x_2 - x_0 - x_1 + 2x_2)\Big) + f[x_0, x_1](x_1 - x_2) + f(x_2) = f(x_1)$$
$$a\Big((x_1 - x_2)(x_2 - x_0)\Big) + f[x_0, x_1](x_1, x_2) + f(x_2) = f(x_1)$$
$$a(x_2 - x_0) + f[x_0, x_1] = \frac{f(x_1) - f(x_2)}{x_1 - x_2}$$
$$a(x_2 - x_0) = f[x_1, x_2] - f[x_0, x_2]$$
$$a = \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0} = f[x_2, x_1, x_0] \qquad \text{(III.6)}$$

Sustituyendo III.6 en III.5

$$b = f[x_0, x_1] - f[x_2, x_1, x_0]\Big[(x_0 - x_2) + (x_1 - x_2)\Big]$$
$$b = f[x_0, x_1] - f[x_2, x_1, x_0](x_0 - x_2) - f[x_2, x_1, x_0](x_1 - x_2)$$
$$b = f[x_0, x_1] - f[x_0, x_1] + f[x_2, x_1] - f[x_2, x_1, x_0](x_1 - x_2)$$
$$b = f[x_2, x_1] + f[x_2, x_1, x_0](x_2 - x_1)$$
$$b = f[x_2, x_1] + \frac{(x_1 - x_2)\frac{\left(f(x_0) - f(x_1)\right)}{x_0 - x_1} + (x_1 - x_2)\left(\frac{f(x_1) - f(x_2)}{x_1 - x_2}\right)}{x_0 - x_2}$$

$$b = f[x_2, x_1] + \frac{\frac{x_2 - x_1}{x_0 - x_1}\big(f(x_0) - f(x_1)\big) + f(x_1) - f(x_2)}{x_0 - x_2} - \frac{f(x_0) - f(x_2)}{x_0 - x_2} + f[x_2, x_0]$$

$$b = f[x_2, x_1] + f[x_2, x_0] + \frac{\frac{x_2 - x_1}{x_0 - x_1}\big(f(x_0) - f(x_1)\big) + f(x_1) - f(x_2) - f(x_0) + f(x_2)}{x_0 - x_2}$$

$$b = f[x_2, x_1] + f[x_2, x_0] + \frac{\frac{x_2 - x_1}{x_0 - x_1}\big(f(x_0) - f(x_1)\big) - f(x_0 - f(x_1))}{x_0 - x_2}$$
$$b = f[x_2, x_1] + f[x_2, x_0] + \frac{\big(f(x_0) - f(x_1)\big)\left(\frac{x_2 - x_1}{x_0 - x_1} - 1\right)}{x_0 - x_2}$$
$$b = f[x_2, x_1] + f[x_2, x_0] + \frac{\big(f(x_0) - f(x_1)\big)\left(\frac{x_2 - x_1 - x_0 + x_1}{x_0 - x_1}\right)}{x_0 - x_2}$$
$$b = f[x_2, x_1] + f[x_2, x_0] + \frac{\big(f(x_0) - f(x_1)\big)\left(\frac{x_2 - x_0}{x_0 - x_1}\right)}{x_0 - x_2}$$
$$b = f[x_2, x_1] + f[x_{2,x} 0] - \frac{f(x_0 - f(x_1))}{x_0 - x_1}\frac{x_0 - x_2}{x_0 - x_2}$$
$$b = f[x_2, x_1] + f[x_2, x_0] - f[x_1, x_0] \qquad \text{(III.7)}$$

Con los planteamientos anteriores entonces sabemos que

$$a = f[x_2, x_1, x_0]$$
$$b = f[x_2, x_1] + f[x_2, x_0] - f[x_1, x_0]$$
$$c = f(x_2)$$

lo cual significa que se tienen los datos suficientes para despejar (III.1) y encontrar w, así que utilizando la fórmula cuadrática se tiene que

$$w = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = -\frac{2c}{b \pm \sqrt{b^2 - 4ac}}$$

y haciendo las sustituciones correspondientes se puede concluir entonces que

$$x = x_2 - \frac{2f(x_2)}{b \pm \sqrt{b^2 - 4f(x_2)f[x_2, x_1, x_0]}}$$

aplicando de forma sucesiva la ecuación anterior se tiene

$$x_{n+1} = x_n - \frac{2f(x_n)}{b \pm \sqrt{b^2 - 4f(x_n)f[x_n, x_{n-1}, x_{n-2}]}} \qquad \text{(III.8)}$$

lo cual concluye la deducción del método de Muller para encontrar raíces de funciones algebraicas.

IV. Análisis de convergencia.

IV-A. *Demostración de convergencia.*

Sea p la raíz de la función $f(x)$ que se busca encontrar con el algoritmo, se asume que si $x_n \to p$ entonces por el lemma II.1 $b \to f'(p)$ y $a \to f''(p)$ entonces sustituyendo en (III.8) se tiene que

$$p = p - \frac{2f(p)}{f'(p) \pm \sqrt{\left(f'(p)\right)^2 - 4f(p)f''(p)}} \qquad \text{(IV.1)}$$

y ya que por definición $f(p) = 0$ se puede concluir que la expresión (IV.1) es igual a 0 siempre que $f'(p) \neq 0$ así que el método converge a una raíz de la ecuación.

IV-B. *Análisis del error y orden de convergencia.*

Para establecer una expresión que defina el error del método anteriormente planteado se define $P(p)$ como el polinomio interpolante de segundo grado evaluado en la raíz p y $f(p)$ como la función aproximada evaluada en su raíz. Además por (III.8) se sabe que x_{n+1} es una raíz del polinomio interpolante, entonces por (II.2) se sabe que

$$P(p) - f(p) = -\frac{f'''(\xi)}{6}(p - x_n)(p - x_{n-1})(p - x_{n-2}) \qquad \text{(IV.2)}$$

luego $f(p) = 0$ y utilizando el teorema del valor medio se sabe que

$$\frac{P(p) - P(x_{n+1})}{p - x_{n+1}} = f'(p),\ P(p) = f'(p)(p - x_{n+1})$$

ya que $P(x_{n+1}) = 0$, entonces sustituyendo en (IV.2) se tiene

$$p - x_{n+1} = -\frac{f'''(p)}{6f'(p)}(p - x_n)(p - x_{n-1})(p - x_{n-2}) \qquad \text{(IV.3)}$$

lo cual es un punto de partida suficiente para establecer el error del método y posteriormente su orden de convergencia. Utilizando la notación de error $|e_{n+1}| = |p - x_{n+1}|$ y definiendo $M = |\frac{f'''(p)}{6f'(p)}|$ se puede reescribir IV.3 como

$$|e_{n+1}| = M|e_n||e_{n-1}||e_{n-2}| \qquad \text{(IV.4)}$$

y utilizando la definición de orden de convergencia se tiene que

$$|e_{n+1}| = C|e_n|^{\alpha} \tag{IV.5}$$

lo cual implica también que

$$|e_n| = C|e_{n-1}|^{\alpha} \tag{IV.6}$$

y despejando $|e_{n-1}|$ se tiene

$$|e_{n-1}| = \left(\frac{1}{C}|e_n|\right)^{\frac{1}{\alpha}} \tag{IV.7}$$

y de forma similar

$$|e_{n-1}| = C|e_{n-2}|^{\alpha}$$

así que sustituyendo la ecuación anterior en (IV.7) y despejando $|e_{n-2}|$ se tiene

$$|e_{n-2}| = \frac{1}{C^{\frac{1}{\alpha}+\frac{1}{\alpha^2}}}|e_n|^{\frac{1}{\alpha^2}} \tag{IV.8}$$

por lo que reemplazando (IV.5), (IV.6), (IV.7) y (IV.8) en (IV.4) se tiene que

$$C|e_n|^{\alpha} = M|e_n|\frac{1}{C^{\frac{1}{\alpha}}}|e_n|^{\frac{1}{\alpha}}\left(\frac{1}{C^{\frac{1}{\alpha}+\frac{1}{\alpha^2}}}\right)|e_n|^{\frac{1}{\alpha^2}}$$

simplificando entonces se tiene que

$$C|e_n|^{\alpha} = \frac{M}{C^{\frac{2}{\alpha}+\frac{1}{\alpha^2}}}|e_n|^{1+\frac{1}{\alpha}+\frac{1}{\alpha^2}} \tag{IV.9}$$

lo cual implica que $C = \frac{M}{C^{\frac{2}{\alpha}+\frac{1}{\alpha^2}}}$ y $\alpha = 1 + \frac{1}{\alpha} + \frac{1}{\alpha^2}$ por lo que despejando α se llega a la ecuación $\alpha^3 - \alpha^2 - \alpha - 1 = 0$ de la que resolviendo se sabe entonces que $\alpha \approx 1.84$, lo cual indica que el orden de convergencia del método de Muller es aproximadamente 1.84: menor que el método de Newton pero mejor que bisección por ejemplo. Sabiendo entonces su orden de convergencia se tiene que

$$|p - x_1| = C|p - x_0|^{1.84}$$

$$C^{0.84}|p - x_1| = (C|p - x_0|)^{1.84}$$

y de forma inductiva se tiene que

$$C^{0.84}|p - x_{n+1}| = (C|p - x_0|)^{1.84^n} \tag{IV.10}$$

entonces se debe cumplir que $C|p - x_0| < 1$ para que posteriormente cuando $n \to \infty$, $x_n \to p$. En otras palabras, la convergencia del método está garantizada únicamente si el punto inicial escogido está lo suficientemente cerca de la raíz a encontrar.

V. VENTAJAS Y DESVENTAJAS.

V-A. Ventajas.

- Su convergencia es bastante rápida, siendo considerablemente mejor que métodos como bisección.
- Permite encontrar no solo raíces reales sino también complejas.
- No requiere el cálculo de ninguna derivada para su implementación.

V-B. Desventajas.

- Se necesita conocer un valor cercano a la raíz que se desea encontrar.
- Puede resultar en raíces complejas incluso si solo se buscan raíces reales.
- No es posible garantizar convergencia si la raíz está en un punto crítico (véase IV.1).

VI. PSEUDOCÓDIGO.

El siguiente algoritmo asume la utilización de aritmética de números complejos dentro del mismo, además la operación de diferencias divididas calculada de forma externa con un procedimiento similar al mostrado en (II.3). La condición de paro es que el valor absoluto de la función evaluada en la última aproximación encontrada sea menor a una tolerancia ϵ o que la cantidad de iteraciones supere un valor límite previamente establecido.

Data: x_2, x_1, x_0 valores distintos para interpolar $f(x)$, un valor ϵ de tolerancia y un valor N que define la cantidad máxima de iteraciones.
Result: La raíz de $f(x)$ más cercana a los puntos iniciales dados o un mensaje de error en caso de que no se cumpla la tolerancia mínima

```
begin
    x_{n+1} ← 0
    x_n ← x_2
    x_{n-1} ← x_1
    x_{n-2} ← x_0
    x_{n+1} ← 3
    while |f(x_{n+1})| < ϵ or i ≥ N do
        a ← f[x_n, x_{n-1}, x_{n-2}]
        b ← f[x_n, x_{n-1}] + f[x_n, x_{n-2}] − f[x_{n-1}, x_{n-2}]
        c ← f(x_n)
        D ← √(b² − 4ac)
        /* Escoge el valor de mayor magnitud
           para el denominador                  */
        if |b − D| ≤ |b + D| then
            E ← b + D
        else
            E ← b − d
        end
        x_{n+1} ← x_n − 2c/E
        x_{n-2} ← x_{n-1}
        x_{n-1} ← x_n
        x_n ← x_{n+1}
        i ← i + 1
    end
    if |f(x_{n+1})| < ϵ then
        return x_{n+1}
    else
        Output: No se pudo encontrar un valor para la raíz.
    end
end
```

Algorithm 1: Método de Muller

VII. EJEMPLOS.

VII-A. Visualizando algunas iteraciones

Como primer ejemplo para visualizar el algoritmo se puede utilizar el polinomio deducido en (IV.9) y encontrar todas sus raíces haciendo uso del algoritmo. Primero se presenta una gráfica del polinomio como función compleja, representando en los ejes x, y de la gráfica el plano complejo y en el eje z de la gráfica el módulo de cada valor de la función, tal representación resulta útil para visualizar dónde se encuentran las raíces en el plano complejo. Las tres raíces del polinomio con seis decimales correctos son:

$$(+1.839286 + 0.000000i)$$
$$(-0.419643 - 0.606290i)$$
$$(-0.419643 + 0.606290i)$$

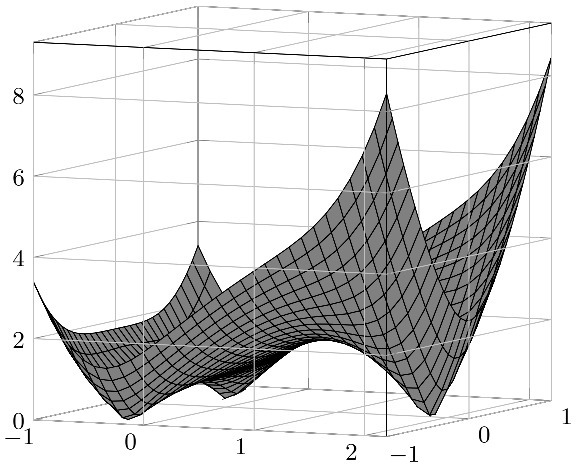

Figura VII.1. Polinomio $z^3 - z^2 - z - 1$

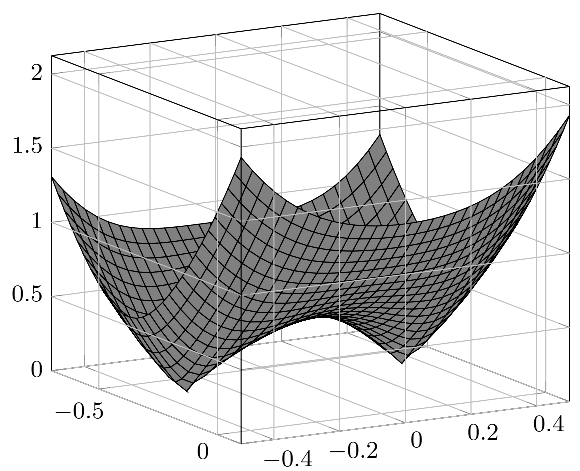

Figura VII.2. Primera iteración

Una gráfica de dicho polinomio puede verse en la figura VII.1 (con z siendo un número complejo). Se consideran algunas iteraciones del método de Muller con su respectiva forma y sus raíces de modo que se pueda visualizar su comportamiento y sucesiva aproximación a una de las raíces. Eligiendo los puntos más cercanos a cada raíz es posible utilizar el método para encontrarlas todas.

Cuadro I
ITERACIONES DEL MÉTODO.

Iteración	**Raíz 1**	**Raíz 2**
1	$+1.625000 + 0.330718i$	$+1.625000 - 0.330718i$
2	$+0.219445 - 1.020150i$	$+0.326009 - 0.148102i$
3	$-0.381386 - 0.049646i$	$-0.380227 + 0.412363i$

El cuadro I muestra las sucesivas aproximaciones de las raíces de los polinomios interpolantes hacia la tercera raíz de la función original $(-0.419643 + 0.606290i)$, siendo que con únicamente tres iteraciones la magnitud de la distancia de la raíz más cercana del tercer polinomio interpolante y la raíz de la función original es de tan solo 0.197892. Dos iteraciones más garantizan que $|f(x_n) < \epsilon|$ si se asume $\epsilon = 10^{-6}$.

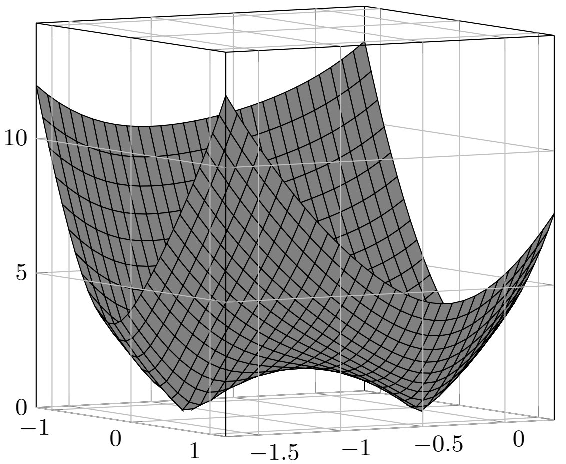

Figura VII.3. Segunda iteración

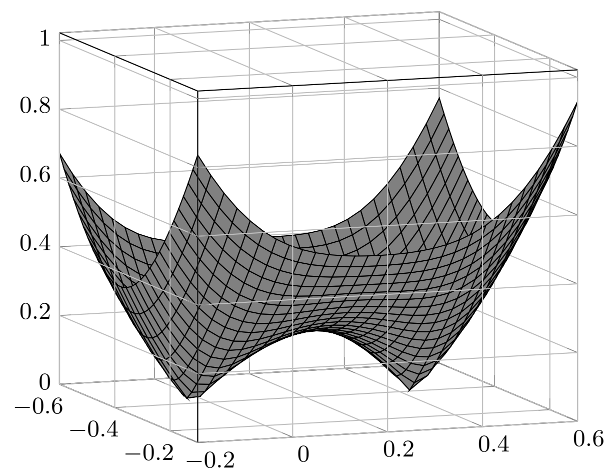

Figura VII.4. Tercera iteración

REFERENCIAS

[1] Kendall E. Atkinson. *An introduction to numerical analysis*. Wiley, New York u.a, 2. ed. edition, 1989.

[2] David E. Muller. A method for solving algebraic equations using an automatic computer. *Mathematics of computation*, 10(56):208–215, Jan 1, 1956.