

TRABAJO FINAL DE GRADO

COMPLEJIDAD EN EL
TEOREMA FUNDAMENTAL DEL
ALGEBRA

AUTOR:

FEDERICO CORREA

MARZO 2024

ORIENTADOR:

DIEGO ARMENTANO

FACULTAD DE CIENCIAS ECONÓMICAS Y DE ADMINISTRACIÓN,
UDELAR

LICENCIATURA EN MATEMÁTICA
UNIVERSIDAD DE LA REPÚBLICA
MONTEVIDEO, URUGUAY

¿Qué más quisiera que pasar la vida entera
como estudiante el día de la primavera?

Contents

1. Breve introducción al Análisis Numérico	3
1.1. Complejidad algorítmica en Análisis Numérico	3
1.2. Enfoque Geométrico	4
1.3. Métodos de Homotopía	5
1.4. El problema de encontrar raíces de polinomios	7
2. Teoría de Kantorovich	9
2.1. Puntos Fijos	9
2.2. Aproximaciones sucesivas y cálculos aproximados	11
2.3. Convergencia Cuadrática	12
3. Teoría α de Smale	15
3.1. Método de Newton	15
3.2. Teoría γ de Smale	20
3.3. Ceros aproximados	24
3.4. Teoría α de Smale	26
4. Complejidad del Teorema Fundamental del Álgebra	31
4.1. Método de Homotopía	34
5. Homogeneización y Numero de Condición	37
5.1. Cota Geodésica para la complejidad del método de Homotopía	43
5.2. Complejidad promedio $\mathbb{E}(K)$	47
5.3. Cuencas de atracción	50
5.4. Simulaciones	53
5.5. Código	55

Capítulo 1

Breve introducción al Análisis Numérico

1.1. Complejidad algorítmica en Análisis Numérico

Abel y Galois, dos valores del tango allá arriba los tenemos: escindieron la matemática con sus resultados, dejando zanjado el asunto de si es posible dar soluciones a ecuaciones polinomiales con exactitud entre otras cosas. A partir de los años 80 con el trabajo de Smale [23] se da impulso a una nueva corriente en el análisis numérico iniciada por Kantorovich, que hasta el momento se centraba únicamente en pruebas de convergencia y velocidad asintótica de convergencia en el entorno de un cero.

Una buena medida de la *complejidad* es tomar el numero de operaciones aritméticas requeridas a partir de ciertos datos de entrada que llamaremos *Input* para llegar a los resultados, o soluciones deseadas a las que llamaremos *Output*.

En computación científica se dice que un algoritmo definido por una “maquina” M es *tratable*, o de *tiempo polinomial*, si el tiempo de ejecución asociado al input x , que denotamos por $K(x)$ satisface:

$$K(x) \leq C \cdot (\text{size}(x))^q$$

siendo C, q constantes que solo dependen de la maquina M

Pero aquí nos encontramos con un problema: la mayoría de los problemas en análisis numérico no podemos resolverlos de manera exacta sino aproximada, es por esto que se introduce un parámetro $\varepsilon > 0$. Ahora diremos que un problema es tratable en contexto de análisis numérico, asumiendo que M es una maquina que manipula números reales si dado $\varepsilon \in (0, 1)$ y x como input, se tiene que:

$$K(x, \varepsilon) \leq C \cdot (\text{size}(X) + |\log(\varepsilon)|)^q.$$

Hay que tener presente que existen input para los cuales el algoritmo o bien no termina o falla. A este tipo de input los llamaremos *ill-posed* o mal puesto.

El *numero de condición* μ trata de capturar esta propiedad, un numero de condición grande tiene asociado un tiempo de calculo grande. En general nos van a interesar algoritmos tales que:

$$K(x, \varepsilon) \leq C(\text{size}(x) + |\log(\varepsilon)| + \mu(x))^q \quad (1.1)$$

Con la idea de hacer desaparecer la constante de aproximación, haciendo que el output final sea una solución “aproximada” en un sentido mas estricto que definiremos mas adelante, Smale propone:

$$K(x) \leq C(\text{size}(x) + |\log(\mu(x))|)^q \quad (1.2)$$

Una vez que llegamos a probar que un algoritmo cumple con estas cotas sobre su complejidad, la pregunta que vamos a intentar responder es la siguiente:

¿Cual es la complejidad media esperada?

es decir, dada una medida de probabilidad sobre el Input, quisiéramos probar que

$$\mathbb{E}(K) \leq C(\text{size})^q$$

es decir, que en promedio la complejidad está acotada por una constante C multiplicada por el tamaño del espacio de input a la q .

Algunas preguntas que vamos a poder responder con este enfoque serán:

¿que algoritmos tienen tiempo polinomial en media?

¿cuales algoritmos son mas eficientes para un problema dado?

Y así podemos hacer un camino formal para el desarrollo de la fundación de la complejidad algorítmica en el análisis numérico. De esta área hay mucho por aprender, problemas no resueltos y cada vez mas lineas de investigación, puede decirse que lo que se sabe actualmente sobre el tema es mas bien limitado.

1.2. Enfoque Geométrico

Vamos a llamar \mathcal{X} e \mathcal{Y} a los espacios de input y output asociados a un problema computacional que supondremos además son variedades diferenciables.

La variedad solución $\mathcal{V} \subset \mathcal{X} \times \mathcal{Y}$ es el conjunto de pares (x, y) tales que y es el output que se obtiene del algoritmo a partir del input x . Esta variedad viene equipada por las proyecciones π_1 y π_2 que resultan de restringir las proyecciones canónicas del espacio producto a \mathcal{V}

$$\begin{array}{ccc} & \mathcal{V} \subset \mathcal{X} \times \mathcal{Y} & \\ \swarrow \pi_1 & & \searrow \pi_2 \\ \mathcal{X} & & \mathcal{Y} \end{array}$$

Los algoritmos intentar “invertir” el mapa π_1 y es por esto que los puntos críticos de este mapa jugaran un papel muy importante en la complejidad del algoritmo.

Sea $D\pi_1(x, y) : T_{(x,y)}\mathcal{V} \rightarrow T_x\mathcal{X}$ la derivada de π_1 , vamos a definir la variedad crítica Σ a partir de este mapa como:

$$\Sigma' := \{(x, y) \in \mathcal{V} : \text{rank}(D\pi_1(x, y)) < \dim \mathcal{X}\}$$

y a su proyección, sobre \mathcal{X}

$$\Sigma := \pi_1(\Sigma') \subset \mathcal{X}$$

la llamaremos variedad discriminante.

En el caso en que $\dim \mathcal{X} = \dim \mathcal{V}$ tenemos que la inversa de π_1 esta bien definida para $(x, y) \in \mathcal{V} - \Sigma'$, por el teorema de función implícita, existe una función diferenciable localmente definida entre ciertos entornos abiertos U_x y U_y en \mathcal{V} de x e y respectivamente :

$$\mathcal{S}(x, y) := \pi_2 \circ \pi_1^{-1} \Big|_{U_x} : U_x \rightarrow U_y.$$

Su derivada

$$D\mathcal{S}(x, y) : T_x\mathcal{X} \rightarrow T_y\mathcal{Y},$$

es el operador condición.

$T_x\mathcal{X}$ es el espacio tangente en el punto x de \mathcal{X} , una manera de definirlo:

$$T_x\mathcal{X} = \{\gamma'(0) : \gamma : (-\varepsilon, \varepsilon) \rightarrow \mathcal{X}, \gamma(0) = x\}.$$

Definición. 1.2.1 (Numero de condición).

$$\mu(x, y) := \max_{\dot{x} \in T_x\mathcal{X}, \|\dot{x}\|_x^2=1} \|D\mathcal{S}(x, y)\dot{x}\|_y$$

Este número es una cota superior en primer orden del peor caso de sensibilidad del error del output con respecto a pequeñas perturbaciones en el input.

1.3. Métodos de Homotopía

Sea $F \in \mathcal{X}$ en el caso de sistemas de ecuaciones complejos homogeneos que definiremos en detalle mas adelante ($\mathcal{X} = \mathcal{H}_{(d)} = \mathcal{H}_{d_1} \times \dots \times \mathcal{H}_{d_k}$, con $(d) = (d_1, \dots, d_k) \in \mathbb{N}^k$ y \mathcal{H}_{d_i} es el conjunto de los polinomios homogeneos de grado d_i). Los métodos de homotopia consisten en considerar un sistema F_0 cuya solución z_0 es conocida, $(F_0, z_0) \in \mathcal{V}$ junto con un camino $F_t \in \mathcal{X}$ con $0 \leq t \leq 1$ que conecte F_0 y $F_1 := F$ para luego aproximar el levantado del camino (F_t, z_t) en \mathcal{V} . Este procedimiento da como resultado z_1 que debería ser el output para F . Este procedimiento funciona siempre que F_t no interseque la variedad discriminante, es decir: $F_t \in \mathcal{X} \setminus \Sigma$ para todo $t \in [0, 1]$.

La manera algorítmica de hacer esto es aproximar este camino por una colección finita de puntos:

$$(F_{t_k}, \tilde{z}_{t_k}), \quad 0 = t_0 \leq t_k \leq t_K = 1$$

de manera que cada \tilde{z}_{t_k} sea una buena aproximación de z_{t_k} . Con este algoritmo la complejidad esta determinada por el numero de elementos de esta colección, es decir K . En este trabajo estudiaremos en detalle el método de aproximación:

$$\tilde{z}_{t_{k+1}} := N_{F_{t_{k+1}}}(\tilde{z}_{t_k})$$

donde:

$$N_F(x) = x - (DF(x)|_{x^\perp})^{-1}F(x)$$

es el mapa de Newton asociado al sistema F , un método clásico para encontrar soluciones a sistemas no lineales. Para definir la aproximación numérica de la curva de la homotopia, la estrategia es lograr que en cada paso el punto \tilde{z}_i sea una *solución aproximada* de F_i es decir:

$$\|N_{F_{t_i}}^n(\tilde{z}_i) - z_i\| \leq \left(\frac{1}{2}\right)^{2^n - 1} \|\tilde{z}_i - z_i\|, \quad \forall n \geq 1$$

con $(F_i, z_i) \in \mathcal{V}$.

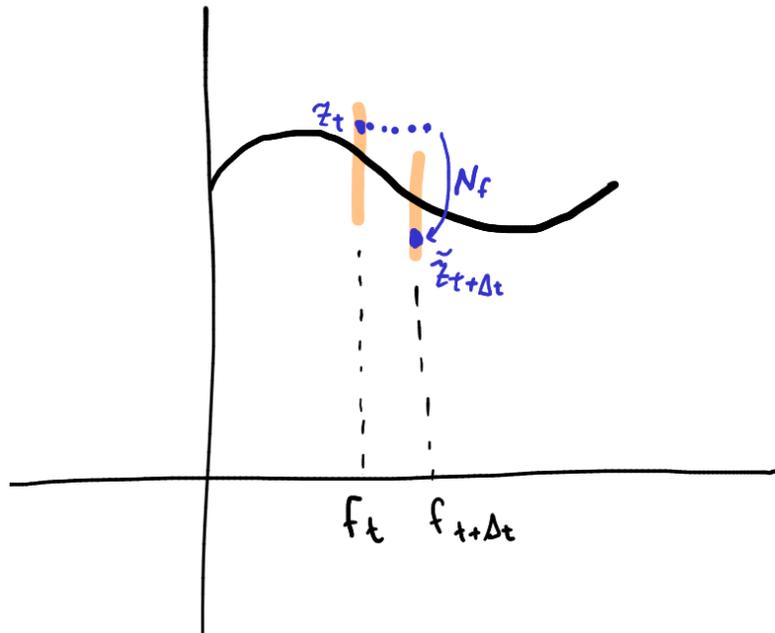


Figura 1.1: Esta figura pretende ilustrar el método de homotopia, en naranja supongamos que están los ceros aproximados de la raíz exacta z_t de f_t y en azul nuestra aproximación.

1.4. El problema de encontrar raíces de polinomios

En diversas áreas de la ciencia que modelan la realidad surgen naturalmente problemas cuya solución equivale a encontrar ceros de alguna función polinomial. Lamentablemente, desde comienzos del siglo XIX sabemos gracias al trabajo de Abel y Galois que no es posible dar formulas en general para encontrar ceros de polinomios de grado mayor a 4. Un primer encare de Smale era tratar de encontrar puntos donde los valores sean arbitrariamente pequeños. Desde la publicación del artículo de Smale “The fundamental theorem of algebra and complexity theory” [23] en 1981 se inicia la teoría de la complejidad de encontrar una solución a un sistema polinomial a través del método de Smale (una variante del método de Newton que utiliza un método de homotopias). El trabajo monográfico se centra en el estudio del artículo de Armentano y Shub “Smale’s fundamental theorem of algebra reconsidered” [4] donde el objetivo es estudiar la complejidad media de este método mejorado.

Adelantémonos un poco como para dar un panorama rápido.

Queremos a partir de:

$$f(z) = \sum_{i=0}^d a_i z^i,$$

y un punto x_0 : hallar una manera veloz de computar una sucesión x_k que tienda a una raíz desconocida ζ de f y dar como output el primer x_K que sea un *cerro aproximado* para el método de Newton, es decir:

$$|N_f^n(x_K) - \zeta| \leq \left(\frac{1}{2}\right)^{2^n - 1} \cdot |x_K - \zeta|, \quad \forall n \geq 1,$$

en este caso N_f esta definido asi:

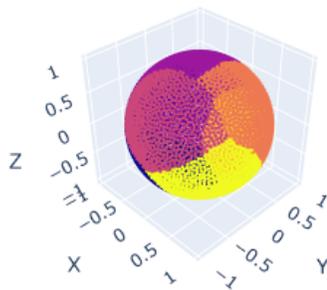
$$N_f(x) = x - \frac{f(x)}{f'(x)}.$$

Si conocemos un par (f_0, x_0) de manera que f_0 es un polinomio de grado d y además $f_0(x_0) = 0$, podemos considerar la homotopia entre f y f_0 mediante f_t con $0 \leq t \leq 1$ y buscar una forma de “continuar” x_0 sobre la curva $\{(f_t, x_t) : f_t(x_t) = 0\}$ y así obtener el x_K con las propiedades deseadas. A continuación veremos como podemos acotar $K(f)$, el numero de operaciones realizadas para encontrar dicho punto para un polinomio genérico. Para esto definiremos el numero de condición $\mu_{norm}(f, \zeta)$, que captura que tan sensible es la raíz a perturbaciones de los coeficientes del polinomio. Luego veremos que el integral de linea de μ_{norm} sobre la curva (f_t, z_t) nos proporciona una cota para la complejidad, lo que nos permite acotar de manera geométrica la complejidad, que es de naturaleza discreta. Este resultado permite estudiar la complejidad media: dada una medida de probabilidad ϕ sobre el espacio de los polinomios,

la complejidad media se define como el promedio de la complejidad

$$\mathbb{E}(K) = \int_{\mathcal{P}_d} K(f) \phi f$$

Veremos como podemos acotar la esperanza de K . Esta complejidad media en el artículo de Armentano y Shub [4] se describe en función de las cuencas de atracción de las raíces, es decir los conjuntos definidos por la raíz a la que convergen sus puntos por el método de homotopías. Finalizaremos haciendo aproximaciones mediante el método de quasi-Montecarlo en \mathbb{S}^2 sobre el área de las cuencas de atracción para el método de homotopías utilizado en el artículo.



Las cuencas de atracción para un polinomio aleatorio de grado 7.

Capítulo 2

Teoría de Kantorovich

2.1. Puntos Fijos

Definición. 2.1.1. Sea (X, d) un espacio métrico. Una función $f : X \rightarrow X$ es Lipschitz, con $\lambda \geq 0$ si existe $\lambda > 0$ que cumple:

$$d(f(x), f(y)) \leq \lambda \cdot d(x, y)$$

para todo $x, y \in X$. El mínimo valor λ para el cual se cumple esta desigualdad lo llamaremos constante de Lipschitz para f .

Necesariamente una función Lipschitz es uniformemente continua.

Definición. 2.1.2. Sea $f : X \rightarrow X$, y $x_0 \in X$, se definen las iteradas de x_0 por f a la sucesión $\{x_k\}_{k \in \mathbb{N}}$ definida inductivamente por:

$$x_{k+1} := f(x_k)$$

para todo natural k , es decir $x_k = f^k(x_0)$ donde $f^k = f \circ f \circ \dots \circ f$, k -veces.

Diremos que una función f Lipschitz, es contractiva, o que es una contracción si su constante cumple $\lambda < 1$.

Una de las mejores cosas que nos va a pasar en la vida para las aplicaciones contractivas es el siguiente teorema:

Teorema. 2.1.3. Sea (X, d) un espacio métrico completo, y sea $f : X \rightarrow X$ una contracción con constante de Lipschitz λ . Entonces valen las siguientes afirmaciones.

1. Para cualquier $x_0 \in X$, la sucesión de las iteradas $f^k(x_0)$ converge a un punto fijo x^* . Además el punto fijo es único.
2. Para todo $k \geq 0$, se tiene

$$d(x_k, x^*) \leq \frac{\lambda^k}{1 - \lambda} d(x_0, x_1).$$

3.

$$\frac{1}{1+\lambda}d(x_0, x_1) \leq d(x_0, x^*) \leq \frac{1}{1-\lambda}d(x_0, x_1).$$

Demostración. De la definición de contracción resulta que:

$$d(x_k, x_{k+1}) = d(f(x_{k-1}), f(x_k)) \leq \lambda d(x_{k-1}, x_k) \leq \dots \leq \lambda^k d(x_0, x_1)$$

Con esto y utilizando la desigualdad triangular:

$$\begin{aligned} d(x_k, x_{k+p}) &\leq d(x_k, x_{k+1}) + d(x_{k+1}, x_{k+2}) + \dots + d(x_{k+p-1}, x_{k+p}) \\ &\leq (\lambda^k + \lambda^{k+1} + \dots + \lambda^{k+p})d(x_0, x_1) = \lambda^k(1 + \lambda + \dots + \lambda^p)d(x_0, x_1) \\ &\leq \frac{\lambda^k}{1-\lambda}d(x_0, x_1). \end{aligned}$$

Esta ultima cota la obtenemos de la serie geométrica $\sum_{i=0}^p \lambda^{k+i} \leq \sum_{i=0}^{\infty} \lambda^{k+i}$

Esto ultimo en particular muestra que la sucesión es de Cauchy, como el espacio es completo y por ser una contracción x_k converge a un único x^* . \square

Observación. 2.1.4. Por el Teorema anterior se sabe que el punto fijo x^* se encuentra en la bola de centro x_0 y radio $d(x_0, x_1)/(1-\lambda)$. Con este dato se puede hacer un estudio de la complejidad del algoritmo, es decir fijado $\varepsilon > 0$, sabiendo $d(x_0, x_1)$ y λ se tiene que: $d(x_{k_\varepsilon}, x^*) < \varepsilon$ si

$$k_\varepsilon > \frac{1}{\log \lambda} \left(\log \left(\frac{1-\lambda}{d(x_0, x_1)} \right) + \log(\varepsilon) \right).$$

Teorema. 2.1.5. Sea (X, d) un espacio métrico completo y sea B la bola de centro x_0 y radio $r > 0$. Supongamos además que $f : B \rightarrow X$ es una contracción de constante λ , que además verifica:

$$d(x_0, f(x_0)) < (1-\lambda)r.$$

Entonces f tiene un único punto fijo en B .

Demostración. sea $0 < r_0 < r$ de manera que $d(x_0, f(x_0)) \leq (1-\lambda)r_0$, para $x \in \overline{B}$ tenemos que:

$$d(f(x), x_0) \leq d(f(x), f(x_0)) + d(f(x_0), x_0) \leq \lambda r_0 + (1-\lambda)r_0 = r_0.$$

Esto nos dice que $\overline{f(B)} \subset \overline{B}$. Por el teorema anterior ganamos. \square

2.2. Aproximaciones sucesivas y cálculos aproximados

Ahora nos vamos a centrar en cálculos contemplando errores, en vez de trabajar con orbitas por f de un punto x_0 vamos a considerar las pseudo-orbitas: sucesiones $\{\tilde{x}_k\}$ que satisfacen una condición un poco mas relajada:

$$d(\tilde{x}_{k+1}, f(\tilde{x}_k)) \leq \varepsilon$$

para $k = 1, 2, 3, \dots$. Una pregunta central ahora sera: ¿que podemos decir sobre $\{\tilde{x}_k\}$?

En los problemas que se estudian en la ciencia, es inevitable que hallan errores y no solo por errores de medición (al realizar un experimento las mediciones tienen sus márgenes) o errores en el modelo sino por errores en los cálculos matemáticos como los realizados usualmente por las computadoras que se someten a redondeos resultando en inexactitudes.

Definición. 2.2.1. Decimos que la sucesión $\{x_k\}$ es η -sombreada por $y \in X$ cuando $d(f^k(y), x_k) \leq \eta$.

El siguiente resultado expresa que toda pseudo-orbita de una contracción es “sombreada”.

Proposición. 2.2.2. Sea (X, d) un espacio métrico completo $f : X \rightarrow X$ una λ -contracción con punto fijo $x^* \in X$ y $\varepsilon > 0$. Entonces toda ε -pseudo-orbita \tilde{x}_k es $\frac{\varepsilon}{1-\lambda}$ -sombreada por $\tilde{x}_0 \in X$. En particular:

$$d(\tilde{x}_k, x^*) \leq \frac{\varepsilon}{1-\lambda} + \lambda^k d(\tilde{x}_0, x^*).$$

Demostración. Aplicaremos la desigualdad triangular a una pseudo-orbita arbitraria:

$$d(\tilde{x}_1, f(\tilde{x}_0)) \leq \varepsilon,$$

$$d(\tilde{x}_2, f^2(\tilde{x}_0)) \leq d(\tilde{x}_2, f(\tilde{x}_1)) + d(f(\tilde{x}_1), f^2(\tilde{x}_0)) \leq (1 + \lambda)\varepsilon,$$

...

$$d(\tilde{x}_k, f^k(\tilde{x}_0)) \leq (1 + \lambda + \dots + \lambda^{k-1})\varepsilon.$$

Acotando por la serie geométrica se ve que \tilde{x}_k es $\frac{\varepsilon}{1-\lambda}$ -sombreada.

Además:

$$\begin{aligned} d(\tilde{x}_k, x^*) &\leq d(\tilde{x}_k, f^k(\tilde{x}_0)) + d(f^k(\tilde{x}_0), x^*) \\ &\leq (1 + \lambda + \dots + \lambda^{k-1})\varepsilon + \lambda^k d(\tilde{x}_0, x^*). \end{aligned}$$

□

2.3. Convergencia Cuadrática

Ya probamos que la velocidad de convergencia de x_k para estas λ -contracciones es del orden $C \cdot \lambda^k$. En análisis Numérico a este tipo de convergencia se la clasifica como orden de convergencia lineal ya que la cantidad de cifras significativas de la aproximación x_k aumenta linealmente con k . Ahora vamos a asumir (como en Dedieu [12]) que f es de clase C^2 definida en un espacio de Banach \mathbb{E} (espacio vectorial métrico completo), con el fin de probar que en este contexto puede aumentarse la velocidad de convergencia.

Teorema. 2.3.1. (Convergencia Cuadrática). *Sea \mathbb{E} un espacio de Banach y sea $f : \mathbb{E} \rightarrow \mathbb{E}$ de clase C^2 . Supongamos que $x^* \in X$ es un punto fijo tal que $Df(x^*) = 0$. Sean $C > 0$ y $r > 0$ tales que: $\|D^2f(x)\| < 2C$ para todo $x \in B_r(x^*)$ y $2Cr \leq 1$.*

Entonces para todo $x_0 \in B_r(x^)$ la sucesión $x_{k+1} = f(x_k)$ satisface*

$$\|x_k - x^*\| \leq \left(\frac{1}{2}\right)^{2^k - 1} \cdot \|x_0 - x^*\|.$$

Formula de Taylor en espacio de Banach. Sean \mathbb{E} y \mathbb{F} espacios de Banach. Sea $f : U \rightarrow \mathbb{F}$ de clase C^k con $U \subset \mathbb{E}$ abierto. Sean $x, y \in U$ si el segmento $[x, y] \subset U$, entonces la formula de Taylor con resto de Lagrange en el punto x , esta dada por la siguiente expresión:

$$\begin{aligned} f(y) &= f(x) + Df(x)(y-x) + \frac{1}{2}D^2f(x)(y-x)^2 + \dots \\ &\dots + \frac{1}{(n-1)!}D^{n-1}f(x)(y-x)^{n-1} + R_n \end{aligned}$$

R_n es el resto de Lagrange:

$$R_n = \int_0^1 \frac{(1-t)^{n-1}}{(n-1)!} D^n f(x + t(y-x))(y-x)^n dt.$$

Recordamos que D^n denota la derivada n -ésima de f en x , la cual es una aplicación multilineal simétrica de \mathbb{E}^n en \mathbb{F}

$$D^p f(x)v = \sum_{i_1, \dots, i_p=1} \frac{\partial^p f}{\partial x_{i_1} \dots \partial x_{i_p}}(x) v_{i_1} \dots v_{i_p}.$$

Demostración. Sea $x \in B_r(x^*)$. Usando la formula de Taylor en el punto x , obtenemos:

$$\begin{aligned} \|f(x) - f(x^*)\| &\leq \left\| \int_0^1 (1-t) D^2 f(x^* + t(x-x^*))(x-x^*)^2 dt \right\| \\ &\leq \int_0^1 (1-t) \|D^2 f(x^* + t(x-x^*))(x-x^*)^2\| dt \end{aligned}$$

$$\leq 2C \|x - x^*\|^2 \int_0^1 (1-t) dt = C \|x - x^*\|^2.$$

Con esto podemos concluir:

$$\|f(x) - x^*\| \leq C \|x - x^*\|^2, \forall x \in B_r(x^*).$$

Como $Cr < 1$ sabemos que si $x_0 \in B_r(x^*)$ entonces $x_{k+1} = f(x_k)$ esta dentro de $B_r(x^*)$ para todo $k \in \mathbb{N}$.

Vamos a hacer inducción en k para concluir.

$$\begin{aligned} \|x_{k+1} - x^*\| &\leq C \|x_k - x^*\|^2 \\ &\leq C \left[\left(\frac{1}{2} \right)^{2^k - 1} \|x_0 - x^*\| \right]^2 \leq 2Cr \left(\frac{1}{2} \right)^{2^{k+1} - 1} \|x_0 - x^*\| \\ &\leq \left(\frac{1}{2} \right)^{2^{k+1} - 1} \|x_0 - x^*\|. \end{aligned}$$

□

Proposición. 2.3.2. Sea \mathbb{E} un espacio de Banach y sea $f : \mathbb{E} \rightarrow \mathbb{E}$ de clase C^2 con x^* punto fijo tal que $Df(x^*) = 0$. Sean $C > 0$, $r > 0$ y $\varepsilon > 0$ tales que:

1. $\|D^2f(x)\| < 2C, \forall x \in B_r(x^*)$.
2. $2Cr \leq 1$.
3. $4\varepsilon < r$.

Entonces $\forall x_0 \in B_r(x^*)$, la pseudo-orbita \tilde{x}_k comenzando en x_0 es 2ε -sombreada por $\{f^k(x_0)\}$, y se tiene:

$$\|\tilde{x}_k - x^*\| \leq 2\varepsilon + \left(\frac{1}{2} \right)^{2^k - 1} \|x_0 - x^*\|.$$

Demostración. Ejercicio. □

Gracias a este ultimo resultado sabemos que el orden de convergencia cuadrática se mantiene si consideramos nuestros cálculos sujetos a errores de cálculo.

Capítulo 3

Teoría α de Smale

3.1. Método de Newton

En general encontrar ceros de una función que no es lineal puede ser engorroso o incluso imposible de realizar con total exactitud, una estrategia para tratar de aproximarnos a un cero es suponer que localmente la función es lineal, esto es lo que hace el método de Newton.

Sea $f : \mathbb{E} \rightarrow \mathbb{F}$; sistema de ecuaciones, $x \in \mathbb{E}$ donde \mathbb{E} y \mathbb{F} son espacios de Banach.

$$f(x) + Df(x)(y - x) = 0$$

es la ecuación de la aproximación lineal a f en el punto $(x, f(x)) \in \mathbb{E} \times \mathbb{F}$ y la idea es hallar la raíz y de esta ecuación lineal, es decir:

$$y = x - Df(x)^{-1}f(x)$$

y así definimos el operador de Newton como

$$N_f(x) := x - Df(x)^{-1}f(x) \tag{3.1}$$

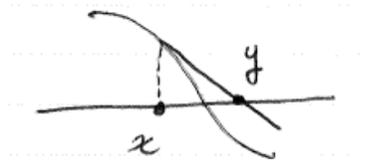


figura ilustrativa

$N_f(x)$ está definido siempre que $Df(x)$ sea invertible, es decir siempre que f no esté dentro de la variedad discriminante: Σ_f .

Definición. 3.1.1. Definimos el método de Newton como sigue:

1. $x_0 \in \mathbb{E}$.
2. $x_{k+1} := N_f(x_k)$.

Una primera observación es que $N_f(x) = x$ si y solo si $f(x) = 0$. Además:

$$DN_f(x) : \mathbb{E} \rightarrow \mathbb{F}$$

queda definido por:

$$DN_f(x)\dot{x} = \dot{x} + Df(x)^{-1}D^2f(x)\dot{x} \cdot Df(x)^{-1}f(x) - Df(x)^{-1}Df(x)\dot{x}$$

$$DN_f(x) = Df(x)^{-1}D^2f(x) \cdot Df(x)^{-1}f(x).$$

Si $f(\zeta) = 0$ tenemos $N_f(\zeta) = \zeta$ y $DN_f(\zeta) = 0$, esto quiere decir que los ceros regulares de f son puntos fijos súper-atradores del método de Newton, mas precisamente el polinomio de Newton de grado 2 de N_f desarrollado en un entorno del punto ζ se ve así:

$$N_f(\zeta + z) = \zeta + D^2N_f(\zeta)(\zeta - z)^2 + r(z).$$

Observación. 3.1.2. El método de Newton para resolver $f(z) = 0$, cuando f es un polinomio complejo en una variable, es una aproximación de Euler a la ecuación diferencial ordinaria:

$$dz/dt = -\frac{1}{2}\text{grad}|f|^2$$

Para el caso de polinomios complejos: tenemos que

Lema 3.1.3. $\frac{1}{2}\text{grad}|f|^2 = f\bar{f}'$.

Demostración. Por un lado tenemos: $|f|^2 : \mathbb{R}^2 \rightarrow \mathbb{R}$, $|f|^2 = \langle f, f \rangle$, con esto:

$$\text{grad}|f|^2 = 2(\langle f, f_x \rangle, \langle f, f_y \rangle)$$

Ahora si escribimos $f(x, y) = u(x, y) + iv(x, y)$ entonces: $f' = u_x + iv_x$ entonces $\bar{f}' = u_x - iv_x$. Calculamos $f\bar{f}' = uu_x + v_xv + i(-uv_x + vu_x)$, por las ecuaciones de Cauchy Riemann nos queda:

$$\begin{aligned} f\bar{f}' &= (uu_x + v_xv) + i(uu_y + vv_y) \\ &= \langle f, f_x \rangle + i\langle f, f_y \rangle. \end{aligned}$$

□

Y con esto:

$$-\frac{1}{2}grad|f|^2 = p(z) \frac{-f(z)}{f'(z)}$$

si definimos $p(z) = f'(z)\overline{f'(z)}$. Como consecuencia las campos $-1/2grad|f|^2$ y $\frac{-f(z)}{f'(z)}$ difieren solo por re-escalar por p . Si hacemos una aproximación de Euler de uno podemos encontrar una aproximación del otro cambiando el tamaño del paso por p . Una aproximación de Euler de una ecuación ordinaria $y'(t) = \phi(t, y(t))$ fijando $y(t_0) = y_0$ consiste en fijar un tamaño de paso h y definir inductivamente $t_{n+1} = t_n + h$, $y_{n+1} = y_n + h\phi(t_n, y_n)$. Decimos que y_n es una aproximación de $y(t_n)$.

Considerando todo esto, si fijamos z_0 entonces:

$$z_{n+1} = z_n - h_n \frac{f(z_n)}{f'(z_n)} \quad n = 1, 2, 3, \dots$$

con $h_n > 0$ es una aproximación de Euler de $\frac{-f(z)}{f'(z)}$, y si $h_n = 1$ recuperamos el método de Newton. en [23] se estudia el caso en que $h_n = h$ es constante para n y además se asume $0 < h < 1$.

Proposición. 3.1.4. *Sea $\zeta \in U$ abierto de \mathbb{E} y $f : U \rightarrow \mathbb{F}$ de clase C^2 tal que $f(\zeta) = 0$ con $Df(\zeta)$ invertible. Sea $r > 0$ tal que $\overline{B_r(\zeta)} \subset U$ y sea:*

$$K(f, \zeta, r) := \sup_{x \in \overline{B_r(\zeta)}} \|Df(\zeta)^{-1}D^2f(x)\|$$

Si se cumple que

$$2rK(f, \zeta, r) \leq 1$$

entonces para todo $x_0 \in \overline{B_r(\zeta)}$ esta definido la sucesión $x_{k+1} = N_f(x_k)$, converge y además:

$$\|x_k - \zeta\| \leq \left(\frac{1}{2}\right)^{2^k - 1} \|x_0 - \zeta\|$$

Demostración. Veamos que N_f esta definido en $\overline{B_r(\zeta)}$, probaremos que $Df(x)$ es invertible para todo $x \in \overline{B_r(\zeta)}$. Para esto basta probar que $Df(\zeta)^{-1}Df(x)$ es invertible.

Aplicando Taylor de primer orden en ζ obtenemos:

$$Df(\zeta)^{-1}Df(x) = Id_{\mathbb{E}} + \int_0^1 Df(\zeta)^{-1}Df(\zeta - t(x - \zeta))(x - \zeta)dt,$$

entonces:

$$\|Df(\zeta)^{-1}Df(x) - Id_{\mathbb{E}}\| \leq K(f, \zeta, r) \leq \frac{1}{2},$$

de donde resulta que $Df(\zeta)^{-1}Df(x)$ es invertible y $\|Df(\zeta)^{-1}Df(x)\| \leq 2$.

Del desarrollo de Taylor de orden 2 de f en x obtenemos:

$$Df(\zeta)^{-1}Df(\zeta) = Df(\zeta)^{-1}Df(x) + Df(\zeta)^{-1}Df(x)(\zeta - x) \\ + \int_0^1 (1-t)D^2f(x+t(\zeta-x))(\zeta-x)^2 dt.$$

Entonces multiplicando a izquierda por $Df(x)^{-1}Df(\zeta)$ obtenemos:

$$Df(x)^{-1}f(x) = x - \zeta - Df(x)^{-1}Df(\zeta) \int_0^1 (1-t)D^2f(x+t(\zeta-x))(x-\zeta)^2 dt.$$

De donde resulta que

$$\|N_f(x) - \zeta\| \leq \|Df(x)^{-1}Df(\zeta)\| \cdot K(f, \zeta, r) \cdot \|x - \zeta\|^2 \int_0^1 (1-t) dt \\ \leq K(f, \zeta, r) \|x - \zeta\|^2.$$

Es decir: $\|N_f(x) - \zeta\| \leq K(f, \zeta, r) \|x - \zeta\|^2$. La prueba sigue por inducción de manera similar al resultado de convergencia cuadrática. \square

La proposición anterior (autoría de Kantoróvich), nos da un criterio para decir si en un entorno de ζ el método de Newton converge a velocidad cuadrática, a partir de $K(f, \zeta, r)$ que es un supremo en $\overline{B_r(\zeta)}$, mas adelante Smale mejora este resultado utilizando valores que solo dependen del punto x_0 con el que trabajamos.

La idea ahora es dar un criterio sencillo para decir que la sucesión de Newton converge a un cero que a priori no conocemos.

Definición. 3.1.5.

$$\beta(f, x_0) := \|N_f(x_0) - x_0\| = \|Df(x_0)^{-1}f(x_0)\|.$$

(Si $Df(x_0)$ no es invertible entonces $\beta(f, x_0) = +\infty$). Este valor en esencia es: cuán grande es el primer paso del método de Newton a partir de x_0 .

Teorema. 3.1.6. (Kantoróvich) Sea $x_0 \in U$, $\overline{B_r(x_0)} \subset U$. Supongamos que:

1. $Df(x_0)$ es isomorfismo,
2. $2\beta(f, x_0) \leq r$,
3. $2\beta(f, x_0) \cdot K(f, x_0, r) \leq 1$.

Entonces existe $\zeta \in \overline{B_r(x_0)}$ tal que $f(\zeta) = 0$, $Df(\zeta)$ es un isomorfismo y $\|x_0 - \zeta\| \leq C_K = 1,63\dots$ es una constante universal. Además la sucesión $x_{k+1} = N_f(x_k)$ esta bien definida y satisface:

$$\|x_k - \zeta\| \leq C_K \left(\frac{1}{2}\right)^{2^k - 1} \beta(f, x_0)$$

y por lo tanto:

$$2 \cdot \beta(f, x_1) \cdot K(f, x_1, r/2) \leq \beta(f, x_0) \cdot 2 \cdot K(f, x_0, r) \leq 1$$

Continuaremos argumentando por inducción para $\overline{B_{r/2^k}(x_k)}$:

$$\beta(f, x_k) = \|x_{k+1} - x_k\| \leq \frac{1}{2^{2^k-1}} \cdot \beta(f, x_0).$$

Esta última cota nos dice que $\{x_k\}$ es de Cauchy y por lo tanto converge a ζ utilizando la desigualdad triangular, tenemos que:

$$\|\zeta - x_0\| \leq \sum_{k=0}^{+\infty} \frac{1}{2^{2^k-1}} \beta(f, x_0) = C_K \cdot \beta(f, x_0) \leq r.$$

Veamos ahora que $Df(\zeta)$ no es cero:

$$\begin{aligned} \|Df(x_0)^{-1}Df(\zeta) - Id\| &\leq K(f, x_0, r) \cdot \|\zeta - x_0\| \\ &\leq C_K \cdot \beta(f, x_0) \cdot K(f, x_0) < 1 \end{aligned}$$

y así probamos que $Df(\zeta)$ debe ser invertible. Luego como N_f es continua y esta definida en $\overline{B_r(x_0)}$ se tiene que

$$N_f(\zeta) = \lim_k N_f(x_k) = \lim_k x_{k+1} = \zeta.$$

□

3.2. Teoría γ de Smale

Vamos a considerar: \mathbb{E}, \mathbb{F} espacios de Banach, $U \subset \mathbb{E}$ abierto y $f : U \rightarrow \mathbb{F}$ analítica:

$$f(y) = f(x) + \sum_{k=1}^{+\infty} \frac{D^k f(x)}{k!} (y-x)^k$$

el radio de convergencia de la serie esta dado por:

$$R(f, x)^{-1} = \limsup_{k \rightarrow \infty} \left\| \frac{D^k(x)}{k!} \right\|^{1/k}$$

Asumiremos para desarrollar la teoría que $B(x, (1 - \sqrt{2}/2)R(x, f)) \subset U$

Definición. 3.2.1. Sea $x \in U$ tal que $x \notin \Sigma_f$. Definimos la función γ :

$$\gamma(f, x) := \sup_{k > 1} \left\| Df(x)^{-1} \cdot \frac{D^k f(x)}{k!} \right\|^{1/(k-1)}$$

(declarando el valor $+\infty$ puede extenderse esta definición a Σ_f).

Observación. 3.2.2. Si f es una función cuadrática, entonces $\gamma(f, x) = \frac{1}{2}K(f, x, r)$.

Lema 3.2.3.

$$\frac{1}{\gamma(f, x)} \leq R(f, x)$$

Demostración. Comenzamos observando que

$$\begin{aligned} R(f, x)^{-1} &\leq \limsup \|Df(x)\|^{1/k} \left\| Df(x)^{-1} \frac{D^k f(x)}{k!} \right\|^{1/k} \\ &\leq \limsup_k \left\| Df(x)^{-1} \cdot \frac{D^k f(x)}{k!} \right\|^{1/(k-1)} \leq \sup_{k>1} \left\| Df(x)^{-1} \frac{D^k f(x)}{k!} \right\| = \gamma(f, x). \end{aligned}$$

□

El siguiente teorema nos da una condición suficiente para saber si dada una raíz conocida ζ , tenemos que el algoritmo de Newton comenzando en x_0 converge a ella.

Teorema. 3.2.4. (Teorema γ) Sea $\zeta \in U$ tal que $f(\zeta) = 0$ y $Df(\zeta)$ es isomorfismo y sea $x_0 \in U$ tal que:

$$\|x_0 - \zeta\| \leq C_\gamma \cdot \frac{1}{\gamma}$$

donde C_γ es una constante definida por $C_\gamma = \frac{3-\sqrt{7}}{2} = 0,17712434\dots$ Bajo estas condiciones x_0 es un cero aproximado, es decir:

$$\|x_k - \zeta\| \leq \left(\frac{1}{2}\right)^{2^k - 1} \cdot \|x_0 - \zeta\|.$$

La demostración de este resultado requiere algunos lemas:

Lema 3.2.5. el polinomio $\psi(u) = 2u^2 - 4u + 1$ decrece de 1 a 0 en el intervalo $[0, 1 - \frac{\sqrt{2}}{2}]$.

Lema 3.2.6. Dados $x, x_1 \in U$ tales que $u := \|x_1 - x\| \gamma(f, x) \leq 1 - \frac{1}{\sqrt{2}}$, entonces $Df(x)^{-1}D(x_1)$ es invertible y además:

$$\|Df(x_1)^{-1}Df(x)\| \leq \frac{(1-u)^2}{\psi(u)}.$$

Demostración. Primero tenemos que $\|x_1 - x\| \leq \frac{1}{R(f, x)} \left(1 - \frac{1}{\sqrt{2}}\right)$ Entonces:

$$Df(x_1) = Df(x) + \sum_{j=1}^{+\infty} \frac{D^{j+1}f(x)}{j!} (x_1 - x)^j$$

multiplicando por $Df(x)^{-1}$ se tiene:

$$\begin{aligned} \|Df(x)^{-1}Df(x_1) - Id\| &\leq \sum_{j=1}^{+\infty} \left\| Df(x)^{-1} \frac{D^{j+1}f(x)}{j!} (x_1 - x)^j \right\| \\ &\leq \sum_{j=1}^{+\infty} (j+1) \left\| Df(x)^{-1} \frac{D^{j+1}f(x)}{(j+1)!} \right\| \cdot \|x_1 - x\|^j \\ &= \sum_{j=1}^{+\infty} (j+1) \left(\left\| Df(x)^{-1} \frac{D^{j+1}f(x)}{(j+1)!} \right\|^{1/j} \cdot \|x_1 - x\| \right)^j. \end{aligned}$$

Esto implica que podemos escribir:

$$\begin{aligned} \|Df(x)^{-1} \cdot Df(x_1) - Id\| &\leq \sum_{j=1}^{+\infty} (j+1) (\gamma(f, x) \|x - x_1\|)^j = \sum_{j=1}^{+\infty} (j+1) u^j \\ &= \left(\frac{1}{1-u} \right)^2 - 1 = \frac{1}{(1-u)^2} - 1. \end{aligned}$$

La condición $u < 1 - \frac{1}{\sqrt{2}}$ implica que $\frac{1}{(1-u)^2} - 1 < 1$. Así probamos que $Df(x_1)$ es invertible, y además:

$$\begin{aligned} \|Df(x_1)^{-1} \cdot Df(x)\| &\leq \\ \frac{1}{1 - \left(\frac{1}{(1-u)^2} - 1 \right)} &= \frac{(1-u)^2}{2(1-u)^2 - 1} = \frac{(1-u)^2}{2u^2 - 4u + 1} = \frac{(1-u)^2}{\psi(u)}. \end{aligned}$$

□

Lema 3.2.7. Sea $\zeta \in U$, $f(\zeta) = 0$, $Df(\zeta)$ invertible y sea $x \in U$ tal que $u = \|x - \zeta\| \gamma(f, \zeta) < \frac{5-\sqrt{17}}{4}$.

Entonces tenemos:

$$\|N_f^k(x) - \zeta\| \leq \left(\frac{u}{\psi(u)} \right)^{2^k - 1} \cdot \|x - \zeta\|.$$

Hay que tener presente que $\frac{u}{\psi(u)} \in (0, 1)$ si $u < \frac{5-\sqrt{17}}{4}$.

Demostración.

$$\begin{aligned} Df(x) &= Df(\zeta) + \sum_{j=1}^{+\infty} \frac{D^{j+1}f(\zeta)}{(j+1)!} (x - \zeta)^j \\ f(x) &= f(\zeta) + Df(\zeta)(x - \zeta) + \sum_{j=1}^{+\infty} \frac{D^{j+1}f(\zeta)}{(j+1)!} (x - \zeta)^{j+1} \end{aligned}$$

Las ecuaciones anteriores se obtienen mediante el desarrollo de Taylor. Ahora por hipótesis $Df(x)$ es invertible, tenemos:

$$N_f(x) - \zeta = (x - \zeta) - Df(x)^{-1}f(x) = Df(x)^{-1} \cdot (Df(x)(x - \zeta) - f(x))$$

utilizando (5) y (6):

$$\begin{aligned} N_f(x) - \zeta &= Df(x)^{-1} \left[\sum_{j=1}^{\infty} \left(\frac{1}{j!} - \frac{1}{(j+1)!} \right) D^{j+1}f(\zeta)(x - \zeta)^{j+1} \right] \\ &= Df(x)^{-1} Df(\zeta) \cdot \left[\sum_{j=1}^{+\infty} j \cdot Df(\zeta)^{-1} \frac{D^{j+1}f(\zeta)}{(j+1)!} (x - \zeta)^{j+1} \right]. \end{aligned}$$

Tomamos norma y acotamos:

$$\|N_f(x) - \zeta\| \leq \|Df(x)^{-1} \cdot Df(\zeta)\| \cdot \sum_j^{+\infty} j \cdot \left\| Df(\zeta)^{-1} \frac{D^{j+1}f(\zeta)}{(j+1)!} \right\| \cdot \|x - \zeta\|^j \cdot \|x - \zeta\|.$$

Ahora utilizamos el lema anterior ($u < \frac{5-\sqrt{17}}{4} < 1 - \frac{1}{\sqrt{2}}$).

$$\|N_f(x) - \zeta\| \leq \frac{(1-u)^2}{\psi(u)} \left(\sum_{j=1}^{+\infty} j \cdot u^j \right) \cdot \|x - \zeta\| = \frac{(1-u)^2}{\psi(u)} u \left(1 - \frac{1}{1-u} \right)' \|x - \zeta\|$$

(La sumatoria sale de acotar termino a termino por $\gamma(f, \zeta)$) Entonces:

$$\|N_f(x) - \zeta\| \leq \frac{u}{\psi(u)} \cdot \|x - \zeta\|,$$

y la idea sería seguir por recurrencia.

Primero $x_1 = N_f(x)$ satisface las hipótesis: $u \leq \frac{5-\sqrt{17}}{4}$ implica $\frac{u}{\psi(u)} \leq 1$:

$$\|x_1 - \zeta\| \cdot \gamma(f, \zeta) \leq \left(\frac{u}{\psi(u)} \right) \cdot \|x - \zeta\| \gamma(f, \zeta) < \|x - \zeta\| \gamma(f, \zeta) < \frac{5 - \sqrt{17}}{4}.$$

En particular, observar que $\|N_f(x)\zeta\| \leq \|x - \zeta\|$.

Ahora el paso inductivo:

$$\left\| N_f^{k+1}(x) - \zeta \right\| \leq \left(\frac{\tilde{u}}{\psi(\tilde{u})} \right) \left\| N_f^k(x) - \zeta \right\|$$

con $\tilde{u} = \left\| N_f^k(x) - \zeta \right\| \gamma(f, \zeta)$. Por hipótesis inductiva, tenemos:

$$\left\| N_f^{k+1}(x) - \zeta \right\| \leq \left(\frac{\gamma(f, \zeta)}{\psi(\tilde{u})} \right) \left(\frac{u}{\psi(u)} \right)^{2^{k+1}-2} \|x - \zeta\|^2,$$

seguimos:

$$\begin{aligned} \left\| N_f^{k+1}(x) - \zeta \right\| &\leq \left(\frac{\gamma(f, \zeta)}{\psi(\tilde{u})} \right) \left(\frac{\psi(u)}{u} \right) \left(\frac{u}{\psi(u)} \right)^{2^{k+1}-1} \|x - \zeta\|^2 \\ &= \frac{\psi(u)}{\psi(\tilde{u})} \left(\frac{u}{\psi(u)} \right)^{2^{k+1}-1} \|x - \zeta\| \leq \left(\frac{u}{\psi(u)} \right)^{2^{k+1}-1} \|x - \zeta\|. \end{aligned}$$

La última desigualdad se debe a que $\psi(u) \leq \psi(\tilde{u})$ ya que ψ es decreciente y $u = \|x - \zeta\| \geq \|N_f^k(x) - \zeta\| \geq \gamma(f, \zeta) = \tilde{u}$. \square

Prueba del Teorema γ de Smale 3.2.4. Teniendo que $\frac{u}{\psi(u)} \leq \frac{1}{2}$ si $u \leq \frac{3-\sqrt{7}}{2}$ podemos aplicar el lema anterior y concluir con la demostración del teorema. \square

3.3. Ceros aproximados

Definición. 3.3.1. Diremos que x_0 es un cero aproximado de ζ si la sucesión $x_{k+1} = N_f(x_k)$ esta definida para todo $k \in \mathbb{N}$, ζ es raíz de f y además:

$$\|x_k - \zeta\| \leq \left(\frac{1}{2} \right)^{2^k-1} \cdot \|x_0 - \zeta\|$$

Observación. 3.3.2. El teorema γ implica que $B(\zeta, \frac{5-\sqrt{17}}{2\gamma})$ esta en la cuenca de atracción de ζ para N_f . Para garantizar que es un cero aproximado el teorema requiere que el punto este a menos de $\frac{3-\sqrt{7}}{2\gamma}$.

Corolario. 3.3.3. Dos raíces no degeneradas ζ y ζ' de f están a distancia:

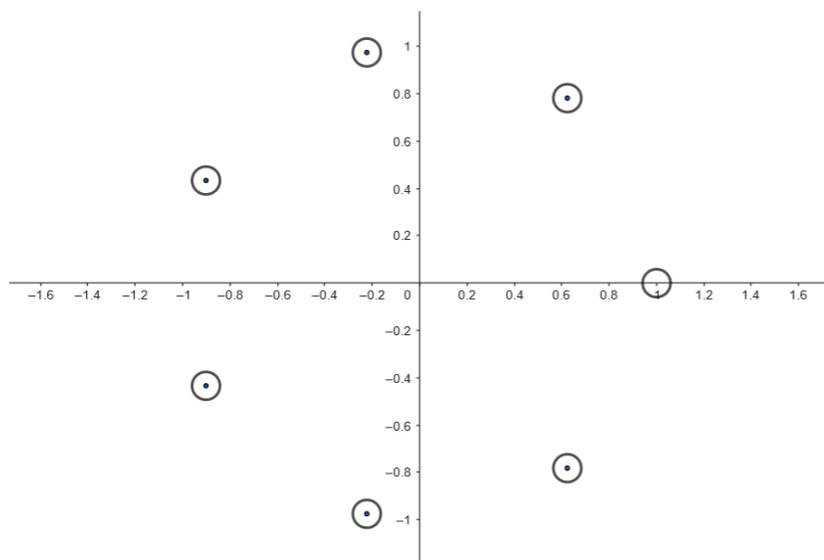
$$\|\zeta - \zeta'\| \geq \frac{5 - \sqrt{7}}{4} \cdot \max \left\{ \frac{1}{\gamma(f, \zeta)}, \frac{1}{\gamma(f, \zeta')} \right\}.$$

Ejemplo 3.3.4. Consideraremos $f : \mathbb{C} \rightarrow \mathbb{C}$ $f(z) = z^d - 1$. Las raíces de f son los complejos ζ^i con $\zeta = e^{\frac{2\pi}{d}}$.

Veamos que $f^{(k)}(\zeta) = d \dots (d-k) \zeta^{d-k}$ y con esto podemos decir que:

$$\begin{aligned} \gamma(f, \zeta) &= \sup_{k=2, \dots, d} \left[\frac{d! \zeta^{1-k}}{d \cdot (d-k)! \cdot k} \right]^{1/(k-1)} \leq \sup_{k=2, \dots, d} \left[\frac{(d-1) \dots (d-k+1)}{k!} \right]^{1/(k-1)} \\ &\leq (d-1) \sup_{k=2, \dots, d-1} \left(\frac{1}{k!} \right)^{1/(k-1)} \leq \frac{d-1}{2}. \end{aligned}$$

Entonces sabemos por el teorema γ que $\|z - \zeta\| < \frac{3-\sqrt{7}}{d-1}$ implica que z es un cero aproximado de ζ .



Acá se deja apreciar para el caso de $z^7 - 1$.

3.4. Teoría α de Smale

La idea ahora es saber si x_0 es un cero aproximada de alguna raíz ζ a priori desconocida. Para esto trabajaremos solo con información que depende de x_0 y f .

Definición. 3.4.1. Recordar que: $\beta(f, x) = \|N_f(x) - x\| = \|Df(x)^{-1}f(x)\|$. Definiremos la función α de la siguiente manera:

$$\alpha(f, x) := \beta(f, x) \cdot \gamma(f, x)$$

Teorema. 3.4.2 (α Smale). Si

$$\alpha(f, x) < \alpha_0$$

entonces x es un cero aproximado de f , donde α_0 es una constante universal. Además, x se encuentra a una distancia menor a $2\beta(f, x)$ de la raíz de f a la que converge.

Demostremos este teorema en tres etapas:

1. Acotar $\|DN_f(y)\| < 2\alpha(f, y)$.
2. Estimar $\alpha(f, y)$ en términos de $\alpha(f, x)$ donde $y \in \overline{B_r(x)}$ que permita concluir que f es una contracción en esta bola.
3. Aplicar el teorema de aproximaciones sucesivas.

Lema 3.4.3. $\|DN_f(x)\| \leq 2\alpha(f, x)$ para todo $x \in U$.

Demostración. Recordar que $DN_f(x) = Df(x)^{-1}D^2f(x)Df(x)^{-1}f(x)$, y acá podemos acotar por γ y β de esta manera:

$$\|DN_f(x)\| \leq 2\gamma(f, x) \cdot \beta(f, x) = 2\alpha(f, x).$$

□

Lema 3.4.4. Sean $x, x_1 \in U$, $u = \|x - x_1\| \gamma(f, x) < 1 - \frac{1}{\sqrt{2}}$. Entonces para $k \geq 2$:

1. $\left\| Df(x_1)^{-1} \frac{D^k f(x)}{k!} \right\| \leq \frac{1}{\psi(u)} \left(\frac{\gamma(f, x)}{1-u} \right)^{k-1},$
2. $\|Df(x)^{-1}f(x)\| \leq \beta(f, x) + \frac{\|x_1 - x\|}{1-u},$
3. $\beta(f, x_1) \leq \frac{1-u}{\psi(u)} ((1-u)\beta(f, x) + \|x_1 - x\|),$
4. $\gamma(f, x_1) \leq \frac{\gamma(f, x)}{(1-u)\psi(u)},$
5. $\alpha(f, x_1) \leq \frac{(1-u)\alpha(f, x) + u}{\psi(u)^2}.$

Demostración. Para probar (1) usamos el desarrollo de Taylor $D^k f(x)$ y multiplicamos a izquierda por $Df(x_1)^{-1}$

$$Df(x_1)^{-1} \frac{D^k f(x)}{k!} = Df(x_1)^{-1} Df(x) \sum_{j=0}^{+\infty} Df(x)^{-1} \frac{D^{k+1} f(x)}{j!k!} (x_1 - x)^j$$

y de esto resulta:

$$\begin{aligned} \left\| Df(x_1)^{-1} \frac{D^k f(x_1)}{k!} \right\| &\leq \frac{(1-u)^2}{\psi(u)} \sum_{j=0}^{+\infty} \frac{(k+k)!}{k!j!} \gamma(f, x)^{k+j-1} \|x_1 - x\|^j \\ &= \frac{1}{\psi(u)} \left(\frac{\gamma(f, x)}{1-u} \right)^{k-1} \end{aligned}$$

y probamos (1).

Utilizando: $\sum_{j=0}^{+\infty} \frac{(k+j)!}{k!j!} u^j = \frac{1}{(1-u)^{k-1}}$.

Ahora desarrollando $f(x)$ se obtiene

$$Df(x)^{-1} f(x_1) = Df(x)^{-1} f(x) + (x_1 - x) + \sum_{k=2}^{+\infty} Df(x)^{-1} \frac{D^k f(x)}{k!} (x_1 - x)^k$$

y con esto tenemos:

$$\begin{aligned} \|Df(x)^{-1} f(x_1)\| &\leq \beta(f, x) + \|x_1 - x\| + \sum_{k=2}^{+\infty} \left\| Df(x)^{-1} \frac{D^k f(x)}{k!} \right\| \|x_1 - x\|^k \\ &\leq \beta(f, x) + \|x_1 - x\| + \sum_{k=2}^{+\infty} \gamma(f, x)^{k-1} \|x_1 - x\|^k \\ &\leq \beta + \|x_1 - x\| \left(1 + \sum_{k=2}^{+\infty} u^{k-1} \right) \\ &= \beta(f, x) + \frac{\|x_1 - x\|}{1-u} \end{aligned}$$

y probamos (2).

De la definición de β , resulta:

$$\begin{aligned} \beta(f, x_1) &= \|Df(x_1)^{-1} f(x_1)\| \leq \|Df(x_1)^{-1} Df(x)\| \|Df(x)^{-1} f(x_1)\| \\ &\leq \frac{(1-u)^{-1}}{\psi(u)} \left(\beta(f, x) + \frac{\|x_1 - x\|}{1-u} \right) = \frac{1-u}{\psi(u)} (\beta(f, x)(1-u) + \|x_1 - x\|) \end{aligned}$$

y probamos (3).

Vamos a por (4):

$$\gamma(f, x_1) = \sup_{k \geq 2} \|D^k f(x_1)\|^{1/(k-1)} \leq \sup_{k \geq 2} \left(\frac{1}{\psi(u)} \right)^{1/(k-1)} \frac{\gamma(f, x)}{1-u} = \frac{\gamma(f, x)}{\psi(u)(1-u)}.$$

Y por ultimo, poniendo en uso (3) y (4):

$$\alpha(f, x_1) = \gamma(f, x_1)\beta(f, x_1) \leq \frac{\gamma(f, x)}{\psi(u)^2} ((1-u)\beta(f, x) + \|x_1 - x\|) = \frac{(1-u)\alpha(f, x_1) + u}{\psi(u)^2}.$$

□

Proposición. 3.4.5. *Sea $r > 0$, $\alpha_0 \leq 1$, $x \in U$ tales que:*

1. $u_0 = r\gamma(f, x) < 1 - \frac{1}{\sqrt{2}}$,
2. $\alpha(f, x) \leq \alpha_0$,
3. $\lambda := 2 \left(\frac{(1-u_0)\alpha_0 + u_0}{\psi(u_0)^2} \right) < 1$,
4. $\alpha_0 + \lambda u_0 \leq u_0$.

Entonces $N_f(x)|_{\overline{B_r(x)}} : \overline{B_r(x)} \rightarrow \overline{B_r(x)}$ es una contracción y por lo tanto existe único $\zeta \in \overline{B_r(x)}$, tal que $f(\zeta) = 0$ y además para todo $x_0 \in \overline{B_r(x)}$ se tiene que $N_f^k(x_0) \rightarrow \zeta$.

Demostración. Veamos primero $\|DN_f(y)\| < \lambda$ para $y \in \overline{B_r(x)}$: Observamos que si $x_1 \in \overline{B_r(x)}$ entonces se cumple que $\|x - x_1\| \gamma(f, x) \leq r\gamma(f, x) < 1 - \frac{1}{\sqrt{2}}$, esto hace que estemos en las hipótesis del teorema anterior:

$$\|DN_f(x_1)\| \leq 2\alpha(f, x) \leq 2 \frac{(1-u)\alpha(f, x) + u}{\psi(u)^2} \leq 2 \frac{(1-u)\alpha_0 + u_0}{\psi(u_0)^2} = \lambda < 1.$$

Esto implica que N_f es una λ -contracción en la bola cerrada.

Para ver que $N_f(\overline{B_r(x)}) \subset \overline{B_r(x)}$ observar que $\lambda r + \|x - N_f(x)\| \leq r$ si y solo si:

$$\lambda r \gamma(f, x) + \|N_f(x) - x\| \leq r \gamma(f, x)$$

si y solo si:

$$\lambda u_0 + \alpha(f, x) \leq u_0$$

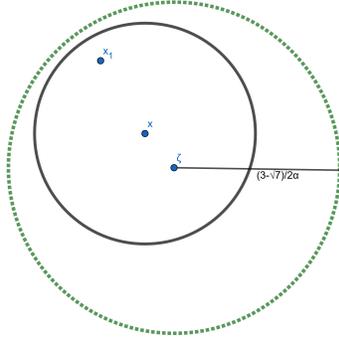
y por hipótesis esto es verdad ya que $\alpha(f, x) \leq \alpha_0$ y esto ultimo es consecuencia de (3) y (4). □

Prueba del Teorema α 3.4.2. $u_0 = 0,06, \alpha_0 = 0,04$ hacen que $\lambda = 0,3316367... < 1/2$. Además para todo $x_0 \in \overline{B(x, \frac{u_0}{\gamma(f, x)})}$ y para $\zeta \in \overline{B_r(x)}$ tal que $f(\zeta) = 0$ tenemos que:

$$\|x_0 - \zeta\| \leq \|x_0 - x\| + \|x - \zeta\| \leq \frac{2u_0}{\gamma(f, x)}$$

Entonces:

$$\|x_0 - \zeta\| \cdot \gamma(f, \zeta) \leq \frac{2u_0}{\gamma(f, x)} \gamma(f, \zeta) \leq \frac{2u_0}{(1-x_0)\psi(u_0)} = 0,1663967... < C_r < \frac{3 - \sqrt{7}}{2}$$



es decir que

$$\overline{B\left(x, \frac{u_0}{\gamma(f, x)}\right)} \subset \overline{B\left(\zeta, \frac{3 - \sqrt{7}}{2}\right)}$$

lo que nos dice que x es un cero aproximado. \square

Teorema. 3.4.6 (α Robusto). *existen u_0 y α_0 constantes universales tales que: Si $x \in U$ satisface $\alpha(f, x) \leq \alpha_0$ entonces existe un único ζ con $f(\zeta) = 0$ tal que $\|\zeta - x\| \leq \frac{u_0}{\gamma(f, x)}$. Además*

$$\overline{B\left(x, \frac{u_0}{\gamma(f, x)}\right)} \subset \overline{B\left(\zeta, \frac{3 - \sqrt{7}}{2} \frac{1}{\gamma(f, \zeta)}\right)}$$

y N_f es una contracción de $\overline{B\left(x, \frac{u_0}{\gamma(f, x)}\right)}$ de constante menor a $1/2$.

Capítulo 4

Complejidad del Teorema Fundamental del Álgebra

En [23] Smale empieza mencionando la discusión si en verdad el teorema fundamental del álgebra había sido probado satisfactoriamente en sus primeras pruebas (resulta que recién Gauss, dio una prueba que nos deja contentos). A continuación veremos una prueba que inspira un algoritmo para hallar raíces: el método de homotopia: también conocido como Newton global, consiste en sucesivas aplicaciones del método de Newton. Este es el prototipo del principal método para hallar soluciones a sistemas de ecuaciones en análisis numérico. Acá vamos a tratar de abordar el problema de la complejidad del algoritmo, hasta ahora hemos trabajado con resultados acerca de la convergencia y precisión.

Teorema. 4.0.1 (Teorema Fundamental del Álgebra). *Cualquier polinomio complejo f de grado d no constante de coeficientes complejos tiene una raíz $\zeta \in \mathbb{C}$, : $f(\zeta) = 0$.*

La herramienta principal que utilizaremos es:

Teorema. 4.0.2 (Función inversa). *Sea f un polinomio de coeficientes complejos y sea $z \in \mathbb{C}$ con $f'(z) \neq 0$. Entonces existe $\delta > 0$ y un mapa complejo diferenciable:*

$$f_z^{-1} : D_\delta(f(z)) \rightarrow \mathbb{C}$$

con $f_z^{-1}(f(z)) = z$ y $f(f_z^{-1}(w)) = w$ para todo $w \in D_\delta(f(z))$.

Para nuestra demostración vamos a usar 2 lemas:

Lema 4.0.3. *Un polinomio f de grado d tiene a lo sumo d raíces.*

Demostración. Por inducción en el grado, veremos solo el paso inductivo: Si tenemos $f(\zeta) = 0$. Entonces haciendo la división de polinomios obtenemos: $f(z) = g(z)(z - \zeta)$ donde g es un polinomio de grado $d - 1$, por hipótesis inductiva g tiene a lo sumo $d - 1$ raíces, concluimos que f tiene a lo sumo $(d - 1) + 1$ raíces. \square

Lema 4.0.4. sea $f(z) = a_0 + a_1z + \dots + a_dz^d$ con $a_d \neq 0$, si ζ es raíz de f , entonces:

$$|\zeta| < 2 \max_{1 \leq k \leq d} \left(\left| \frac{a_{d-k}}{a_d} \right|^{1/k}, 1 \right).$$

Demostración. sean: $b = \max_{1 \leq k \leq d} \left(\left| \frac{a_{d-k}}{a_d} \right|^{1/k}, 1 \right)$ y $g(z) = \frac{1}{a_d b^d} f(bz)$ tenemos que:

$$g(z) = \sum_{i=0}^d c_i z^i$$

con $c_d = 1$, y $\|c_i\| \leq 1$ para todo $i = 0, \dots, d$. Entonces:

$$\left| \sum_{i=0}^{d-1} c_i z^i \right| < \sum_{k=0}^{d-1} |z^k| = \frac{|z|^d - 1}{|z| - 1} \leq |z|^d$$

La última desigualdad se cumple para $\|z\| \geq 2$, o sea que $g(z) \neq 0$ en este caso implicando $f(z) \neq 0$ para $\|z\|/b \geq 2$. \square

Definición. 4.0.5. Decimos que un mapa $f : \mathbb{C} \rightarrow \mathbb{C}$ es propio cuando la preimagen $f^{-1}(S)$ de un compacto $S \subset \mathbb{C}$ es también un conjunto compacto.

Una consecuencia directa del lema anterior es:

Corolario. 4.0.6. Un polinomio no constante es propio.

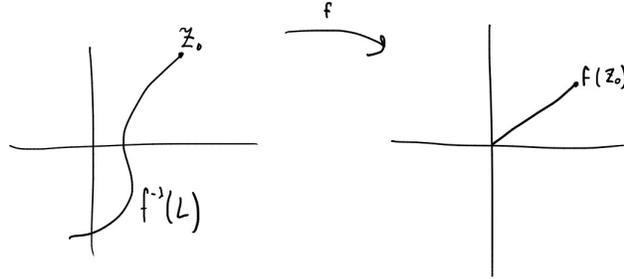
Definición. 4.0.7. Diremos que θ es un punto crítico de un polinomio f si $f'(\theta) = 0$. Diremos que $f(\theta)$ es un valor crítico si θ es un punto crítico. Como f' es de grado $d - 1$ existen a lo sumo $d - 1$ valores críticos para f .

Demostración del Teorema Fundamental del Álgebra. En lo que sigue consideramos el caso en que 0 no es valor crítico, de ser así existen raíces. Diremos que un punto z_0 es un buen punto si el segmento:

$$L_{z_0} = \{tf(z_0) : 0 \leq t \leq 1\}$$

no contiene valores críticos.

Para ver que efectivamente existen buenos puntos utilizamos el siguiente argumento: existe un punto z_1 que no es crítico de f , por el teorema de la función inversa f manda un entorno de z_0 en un entorno de $f(z_0)$, esta imagen infinitos puntos que determinan infinitos segmentos que no se intersecan (sin ser en el 0), como tenemos finitos valores críticos alguno de estos segmentos no contiene un valor crítico.



Consideramos ahora L_{z_0} siendo z_0 un buen punto. Sabemos que $f^{-1}(L_{z_0})$ es cerrado por continuidad y acotado por ser f propio es decir es compacto. Por función inversa, para cada $z \in L_{z_0}$ tenemos un $\delta_z > 0$ de manera que esta definida la inversa en el disco $D_{\delta_z}(z)$. Utilizando la compacidad podemos obtener un $\delta > 0$ que nos sirva para todo $z \in L_{z_0}$. Sea n el primer entero tal que $n > |f(z_0)|/\delta$.

Definimos ahora w_i para $i = 0, \dots, n$ por:

$$w_i = \frac{(n-i)f(z_0)}{n}$$

notar que $|w_i - w_{i-1}| < \delta$ para $i = 1, \dots, n$ así que:

$$w_i \in D_\delta(w_{i-1})$$

Ahora vamos a definir inductivamente:

$$z_i = f_{z_{i-1}}^{-1}(w_i)$$

para $i = 1, \dots, n$. Por construcción se tiene que $f(z_i) = w_i$ y con esto obtenemos que $f(z_n) = w_n = 0$, es decir z_n es raíz del f . \square

Observación. 4.0.8. $f_{z_0}^{-1} : D_\delta(z_0)$ puede extenderse a la unión:

$$D_\delta(z_0) \cup D_\delta(z_1) \cup \dots \cup D_\delta(z_n)$$

Mas aun, puede definirse en un δ - entorno de L .

4.1. Método de Homotopía

Ahora encaramos el siguiente problema, a partir del *input* :

$$(\varepsilon, a_0, \dots, a_d), \quad \varepsilon > 0, \quad f(z) = \sum_{i=0}^d a_i z^i$$

dar el *output* : x con la propiedad:

$$|f(x)| < \varepsilon$$

el siguiente algoritmo esta fuertemente inspirado en la prueba que acabamos de ver y aporta un entendimiento de algunos de los métodos e ideas muy recurrentes en el análisis numérico. Conocido como “método de homotopía”, “Newton global” o “de continuación”.

Comenzamos: supondremos que se cumple $0 < \varepsilon < 1$, $z_0 \in \mathbb{C}$ y f un polinomio no constante, consideraremos:

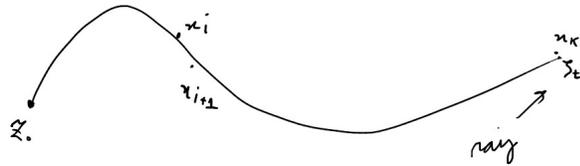
$$f_t(z) = f(z) - tf(z_0)$$

para $0 \leq t \leq 1$. Ahora, Observar que f_t es un polinomio del mismo grado que f . Sea ζ_t la curva compleja con $f_t(\zeta_t) = 0$, $\zeta_1 = z_0$ y $f'(\zeta_t) \neq 0$ para todo $t \in (0, 1]$

Vamos a describir una secuencia finita de puntos t_i con $t_0 = 1$ y $t_0 > t_1 > \dots > t_k = 0$ para “seguir de cerca” o “aproximar” a la curva ζ_t . Definiremos $x_0 = z_0$ y continuamos:

$$x_i = N_{f_{t_i}}(x_{i-1})$$

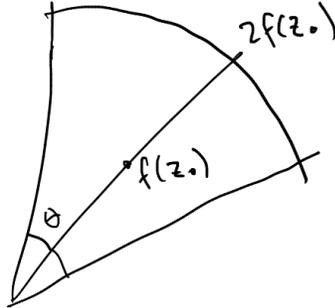
para $i \leq k$



En el libro “Complexity and real computation: A manifesto” [8] se describe una función θ_{f_0, z_0} que mide que tan cerca esta nuestra curva ζ_t de un punto crítico de f mediante argumentos geométricos y análisis complejo rudimentario. Luego se trata de acotar la complejidad utilizando la teoría γ .

Definición. 4.1.1. Dado $z_0 \in \mathbb{C}$ se define el conjunto cuña $W_{f, z_0, \theta}$ de ángulo θ alrededor de $f(z_0)$ como:

$$W_{f, z_0, \theta} = \left\{ w \in \mathbb{C} : 0 < |w| < 2|f(z_0)|, \arg \frac{w}{f(z_0)} < \theta \right\}.$$



A partir de este conjunto vamos a definir $\theta = \theta(f, z_0)$ como el máximo ángulo que hace que $W_{f, z_0, \theta}$ no contenga valores críticos. Es evidente que en el caso que $f'(z) \neq 0$ se puede “extender” W siempre que no halla críticos en el borde. Luego definimos:

$$M = \frac{25}{25 + \sin \theta_{f, z_0}}.$$

Con esto tenemos todos los ingredientes para enunciar el siguiente teorema:

Teorema. 4.1.2. *Si definimos $t_i = M^i$, $x_i = N_{f_{t_i}}(x_{i-1})$ con $x_0 = z_0$, Entonces se verifica que*

$$|f(x_i)| \leq 2M^i |f(z_0)|.$$

Además, si

$$k > \left(1 + \frac{25}{\sin \theta_{f, z_0}}\right) (\ln |f(z_0)| + \ln 1/\varepsilon + f)$$

Entonces: $|f(x_i)| < \varepsilon$

Demostración. Ver:BCSS[8] □

Observación. 4.1.3. Este teorema nos dice que el numero de pasos necesarios para alcanzar una precisión en el Output de menos de ε es mayor cuando $f(z_0)$ esta cerca de un valor crítico. Esto ultimo, la cercanía con los valores críticos es lo que captura el numero de condición.

Capítulo 5

Homogeneización y Numero de Condición

Acá vamos a hacer un estudio más contemporáneo y en vez de θ_{f_0, z_0} vamos a trabajar directamente con el numero de condición como en Shub [16] donde se acota la complejidad para el método aplicado a sistemas polinomiales. Complexity of Bézout's Theorem es una serie de artículos iniciados por Shub y Smale en los 80 y continuado por varios colaboradores que profundizan en resultados afines a la resolución de sistemas polinomiales mediante el método de homotopia.

Para este enfoque trabajamos con polinomios homogéneos, que definiremos a continuación, sin dejar de mencionar cómo los relacionamos con nuestros queridos polinomios univariados:

Hasta ahora veníamos trabajando con polinomios de la forma

$$p(z) = \sum_{k=0}^d a_k z^k.$$

Definimos el polinomio homogeneizado de p como $h_p : \mathbb{C}^2 \rightarrow \mathbb{C}$:

$$h_p(z, w) = \sum_{k=0}^d a_k z^k w^{d-k}, \quad (z, w) \in \mathbb{C}^2.$$

Observación. 5.0.1. Si ξ es raíz de p entonces cualquier múltiplo de $(\xi, 1)$ es raíz de h_p pues:

$$h_p(\lambda \cdot (\xi, 1)) = \lambda^d \cdot p(\xi) = 0.$$

Por eso las raíces de h_p las podemos considerar en el *espacio proyectivo*: $\mathbb{P}(\mathbb{C}^2)$

El espacio $\mathbb{P}(\mathbb{C}^2)$ es el cociente de $\mathbb{C}^2 - \{(0, 0)\}$ por la relación de equivalencia $(x, y) \sim (z, w)$ si existe $\lambda \in \mathbb{C}^*$ de manera que $(x, y) = (\lambda z, \lambda w)$.

Observación. 5.0.2. La geometría proyectiva surge de las finas artes, se caracteriza por el siguiente axioma: dos rectas cualesquiera se intersecan en al menos un punto. A diferencia de la geometría euclídea que puede formularse como la geometría de regla y compás, la proyectiva se queda solo con la regla. Además de ser un objeto de estudio interesante en si mismo, la geometría proyectiva tiene aplicaciones a otras geometrías tanto euclídeas como no-euclídeas (Ver Coxeter [11]).

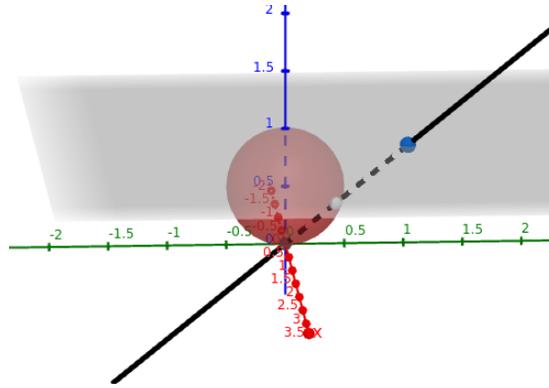
A nuestros efectos prácticos vamos a considerar a la esfera \mathbb{S}^2 como modelo de $\mathbb{P}(\mathbb{C}^2)$.

Como en Bezout III (Shub y Smale[22]), consideramos \mathbb{S}^2 la esfera en $\mathbb{C}^2 \times \mathbb{R} = \mathbb{R}^3$ de radio $1/2$ centrada en $(0, 0, 1/2)$. Para $z \in \mathbb{C}$ denotamos por \hat{z} el punto en \mathbb{S}^2 definido por la proyección estereográfica en $(0, 0, 0)$ sobre el plano $\{(x, y, z) : z = 1\}$:

$$\hat{z} = \frac{(z, 1)}{1 + |z|^2}$$

La norma en \mathbb{R}^3 restringida a la esfera induce una distancia en $\mathbb{P}(\mathbb{C})$: Sean $z_1, z_2 \in \mathbb{C}$,

$$\|\hat{z}_1 - \hat{z}_2\| = \frac{|z_1 - z_2|}{(1 + |z_1|)^{1/2}(1 + |z_2|)^{1/2}}$$



En rojo \mathbb{S}^2 , el punto azul $(z, 1)$ y el blanco \hat{z} .

Tenemos el mapa i_d entre los polinomios complejos de grado d y los homogéneos de grado d :

$$i_d : \mathcal{P}_d \rightarrow \mathcal{H}_d,$$

$$i_d(p)(z, w) = z_0^d p\left(\frac{z}{w}\right)$$

es un isomorfismo entre los polinomios mónicos de grado d y los homogéneos del mismo grado.

También tenemos: $j : \mathbb{C} \rightarrow \mathbb{P}(\mathbb{C}^2)$, que manda ceros de un polinomio en su homogeneizado. En el otro sentido podemos tomar: $h(\cdot, \cdot) \rightarrow h(\cdot, 1)$.

Ahora ganamos que el espacio de nuestro problema es un compacto y tiene medida finita $\frac{\pi^{d-1}}{\Gamma(d)}$

Es de interés ahora estudiar un poco mas en detalle el siguiente conjunto:

$$\mathcal{V} = \{(f, x) \in \mathbb{P}(\mathcal{H}_d) \times \mathbb{P}(\mathbb{C}^2) : f(x) = 0\}$$

Proposición. 5.0.3. \mathcal{V} es una variedad conexa de $\mathbb{P}(\mathcal{H}_d) \times \mathbb{P}(\mathbb{C}^2)$ de codimensión 1 y el espacio tangente a \mathcal{V} en un punto (f, x) es:

$$T_{(f,x)} = \{(h, w) \in T_f \times T_x : h(x) + Df(x)(w) = 0\}$$

Demostración. Vamos a utilizar el mapa evaluación: $ev : \mathcal{H}_d \times \mathbb{C}^2 \rightarrow \mathbb{C}$ definido por

$$ev(f, x) = f(x).$$

Observamos que el diferencial evaluado en (h, w) nos queda:

$$Dev_{(f,x)}(h, w) = h(x) + Df(x)w.$$

Tomando $w = 0$ y variando solamente h podemos ver que 0 es un valor regular de ev . Sea ahora $\mathcal{V}' = \{(f, x) \in \mathcal{H}_d \times (\mathbb{C}^2 - \{0\}) : f(x) = 0\}$. Tenemos que por el teorema de función inversa \mathcal{V}' es una variedad suave y su codimensión es 1 y su espacio tangente es el espacio vectorial formado por los (h, w) tales que $h(x) + Df(x)w = 0$.

Vamos a ver ahora que \mathcal{V}' es conexa. Para $x \in \mathbb{C}^2 - \{0\}$ definimos $\mathcal{V}'_x = \{(f, x) \in \mathcal{V}'\}$. Los $f \in \mathcal{H}_d$ que aparecen en la primera coordenada de \mathcal{V}'_x forman un espacio vectorial de \mathcal{H}_d de codimensión 1. Esto prueba que \mathcal{V}'_x es conexa. Ahora veremos que dados dos puntos $x, x' \neq 0$ hay un camino que conecta \mathcal{V}'_x y $\mathcal{V}'_{x'}$. Sea $f \in \mathcal{V}'_x$ y nos tomamos L_t una familia de mapas lineales invertibles $L_t : \mathbb{C}^2 \rightarrow \mathbb{C}^2$ para $0 \leq t \leq 1$ de manera que $t \rightarrow L_t$ es una función continua, $L_0(x) = x$, y $L_1(x) = x'$. Con esto tenemos que

$$(f_t, x_t) := (f \circ L_t^{-1}, L_t x)$$

es un camino que conecta \mathcal{V}'_x con $\mathcal{V}'_{x'}$.

Consideremos ahora la acción del grupo $\mathbb{C}^* \times \mathbb{C}^*$ sobre $\mathcal{H}_d \times \mathbb{C}^2 - \{0\}$ dada por $(\lambda, \mu) \cdot (f, x) = (\lambda f, \mu x)$. Esta acción es libre (\cdot es inyectiva) y deja invariante a \mathcal{V}' . Cocientar por esta acción $\mathcal{H}_d \times \mathbb{C}^2 - \{0\}$ da como resultado $\mathbb{P}(\mathcal{H}_d) \times \mathbb{P}(\mathbb{C}^2)$ y el cociente de \mathcal{V}' es \mathcal{V} , esto prueba que \mathcal{V} es conexa. Para ver que \mathcal{V} es una variedad suave de $\mathbb{P}(\mathcal{H}_d) \times \mathbb{P}(\mathbb{C}^2)$ consideramos:

$$\mathcal{V}'' = \mathcal{V}' \cap S^1(\mathcal{H}_d) \times S^1(\mathbb{C}^2)$$

Donde $S^1(\mathcal{H}_d)$ es la esfera unidad de \mathcal{H}_d con la norma definida a continuación:

Definición. 5.0.4. Para $g \in \mathcal{H}_d$

$$\|g\| = \left(\sum_{k=0}^d |a_k|^2 \binom{d}{k}^{-1} \right)^{1/2}$$

es la norma de Bombieri-Weyl, que sale del producto:

$$\langle h, g \rangle = \sum_{k=0}^d a_k \bar{b}_k \binom{d}{k}^{-1}$$

esta definida por la estructura unitaria de Bombieri-Weyl.

Sabemos que \mathcal{V}' es transversal a $S^1(\mathcal{H}_d) \times S^1(\mathbb{C}^2)$ (la intersección es trasversal) y por lo tanto \mathcal{V}'' es una variedad diferenciable. La variedad \mathcal{V} puede verse ahora como el cociente de \mathcal{V}'' por la acción libre de $S^1 = \{\lambda \in \mathbb{C} : |\lambda| = 1\}$. Como S^1 es compacto entonces \mathcal{V} es una variedad diferenciable. Ahora, se tiene que el cociente de una variedad diferenciable por la acción libre de S^1 (grupo compacto) es una variedad diferenciable y por lo tanto podemos concluir que \mathcal{V} es una variedad diferenciable.

La codimensión de \mathcal{V} se mantiene al cocientar. □

La idea en adelante es dar resultados para el método global de Newton para polinomios homogéneos.

Definición. 5.0.5. Para polinomios homogéneos definimos el método de Newton de la siguiente manera:

$$N_f(x) = x - (Df(x)|_{x^\perp})^{-1} f(x).$$

donde $(Df(x)|_{x^\perp})$ es la restricción de $Df(x)$ al complemento ortogonal de x .

La homotopia que consideraremos a partir de un $\zeta \in \mathbb{C}^2$ para un polinomio homogéneo f de grado d es f_t :

$$f_t(z) = f(z) - (1-t) \frac{\langle z, \zeta \rangle^d}{\langle \zeta, \zeta \rangle^d} f(\zeta). \quad (5.1)$$

La aproximación numérica de la curva (ζ_t, f_t) que cumple $f_t(\zeta_t) = 0$, con $\zeta_0 = \zeta$ dada una partición del intervalo unidad $0 = t_0 < t_1 < \dots < t_k = 1$:

$$z_{i+1} = N_{f_{i+1}}(z_i).$$

A continuación definimos el número de condición para polinomios homogéneos.

Definición. 5.0.6. Para $(g, x) \in \mathbb{P}(\mathcal{H}_d) \times \mathbb{P}(\mathbb{C}^2)$, $g(z, w) = \sum_{i=0}^d a_k z^k w^{d-k}$ definimos:

$$\mu_{norm}(g, x) := \frac{d^{1/2} \|g\| |x|^{d-1}}{|(Dg(x)|_{x^\perp})|}$$

En el caso en que $(Dg(x)|_{x^\perp})$ no sea invertible, declararemos $\mu_{norm}(g, x) = \infty$ ($(Dg(x)|_{x^\perp})$ es la restricción de $Dg(x)$ al complemento ortogonal de x).

$$\mu(g) := \max_{\zeta: g(\zeta)=0} \mu(g, \zeta)$$

Observación. 5.0.7. El numero de condición tiene el cometido de codificar que tan sensible es la raíz a perturbaciones de los coeficientes del polinomio. Mide la sensibilidad del error en el output relativo al error de input.

Observación. 5.0.8. En la elección de la homotopia f_t que definimos en (5.1) lo que hacemos es deformar el polinomio para forzar que al comienzo ζ sea una raíz. En estudios anteriores como, en particular en los pioneros trabajos de Shub y Smale (la serie de papers sobre la complejidad del Teorema de Bézout) [17, 18, 19, 21, 20] lo que se estilaba era partir de un polinomio fijo, con sus raíces conocidas y bien condicionadas para luego continuar alguna de ellas a cualquier polinomio. La cuestión de encontrar el mejor polinomio en términos de condicionamiento ha sido estudiada en profundidad. En Bezout III [17] se prueba que tener un polinomio bien codiciado implica que sus raíces $\{\zeta_1, \dots, \zeta_d\}$, se distribuyen en la esfera de manera esparza en el sentido que *potencial logarítmico* es grande:

$$\sum_{0 \leq i < j \leq d} \log(\|\zeta_i - \zeta_j\|).$$

se llama polinomios de *Fekete* a los que cumplen que sus raíces alcanzan el máximo de la función que acabamos de definir.

Vale la pena mencionar que Beltran y Pardo en [6] prueban el problema 17 1 de Smale de manera probabilística partiendo de un polinomio aleatorio con una raíz conocida.

En el estudio de la complejidad y en algunos problemas afines como el de Fekete (Ver Beltran [7]), nos interesa estudiar las raíces de polinomios a menos de transformaciones unitarias. Para eso introducimos el producto interno de Bombieri-Weyl en los polinomios homogéneos que nos dice un poco mas de como se comportan estas funciones que el producto usual: esta estructura hace que $\langle z^\alpha, z^\beta \rangle = 0$ para $\alpha, \beta \in \mathbb{N}^2$ con tal que la suma de sus coordenadas es igual a d y $\alpha \neq \beta$.

Teorema. 5.0.9. El producto Bombieri-Weil es invariante por transformaciones unitarias, es decir: sea U transformación unitaria $U : \mathbb{C}^2 \rightarrow \mathbb{C}^2$ entonces:

$$\langle h \circ U^{-1}, g \circ U^{-1} \rangle_W = \langle h, g \rangle_W$$

para $f, g \in \mathcal{H}_d$

$$\langle U\zeta, U\zeta' \rangle_{\mathbb{C}^2} = \langle \zeta, \zeta' \rangle_{\mathbb{C}^2}$$

Demostración. Primero definimos el núcleo de $\xi = (z, w) \in \mathbb{C}^2$ como:

$$K_\xi(\cdot) = \langle \cdot, \xi \rangle_{\mathbb{C}^2}^d$$

Observamos que K_ξ es un polinomio homogéneo de grado d :

$$K_\xi(x, y) = \langle (x, y), (z, w) \rangle_{\mathbb{C}^2}^d = (x\bar{z} + y\bar{w})^d = \sum_{i=0}^d \bar{z}^i \bar{w}^{d-i} x^i y^{d-i} \binom{d}{i}$$

que cumple

$$\langle p, K_\xi \rangle_W = \sum_{i=0}^d \frac{a_i z^i w^{d-i} \binom{d}{i}}{\binom{d}{i}} = p(\xi)$$

(por esto es que los llamamos núcleos reproductores).

En particular, $\langle K_\xi, K_\zeta \rangle_W = K_\xi(\zeta)$. Observar también que:

$$K_{U\xi}(\cdot) = \langle \cdot, U\xi \rangle_{\mathbb{C}^2}^d = \langle U^{-1}\cdot, \xi \rangle_{\mathbb{C}^2} = K_\xi \circ U^{-1}(\cdot) = UK_\xi(\cdot)$$

Ahora podemos probar lo siguiente:

$$\langle UK_\xi, UK_\zeta \rangle_W = \langle K_{U\xi}, K_{U\zeta} \rangle_W = \langle U\xi, U\zeta \rangle_{\mathbb{C}^2}^d = \langle \xi, \zeta \rangle_{\mathbb{C}^2}^d = \langle K_\xi, K_\zeta \rangle_W.$$

Hasta este punto probamos que para los núcleos se cumple la propiedad deseada, concluiremos la demostración probando que el conjunto de núcleos es un generador de los polinomios.

Sea $f \in \{K_\xi, \xi \in \mathbb{C}^2\}^\perp$ es decir $\langle f, K_\xi \rangle = 0$ para cualquier ξ . Esto implica que f es el polinomio nulo, ya que $\langle f, K_\xi \rangle = f(\xi) = 0$. Por lo tanto $\{K_\xi, \xi \in \mathbb{C}^2\}$ genera \mathcal{H}_d : un polinomio cualquiera puede escribirse como suma de núcleos:

$$f = \sum_{i=1}^k \alpha_i K_{\xi_i}, g = \sum_{i=1}^l \beta_i K_{\zeta_i}$$

así que se verifica por linealidad que $\langle Uf, Ug \rangle = \langle f, g \rangle$. □

Observación. 5.0.10. μ_{norm} cumple con las siguientes propiedades:

1. $\mu_{norm}(\lambda g, x) = \mu_{norm}(g, x)$ para todo $\lambda \in \mathbb{C} - 0$;
2. $\mu_{norm}(g, \lambda x) = \mu_{norm}(g, x)$ para todo $\lambda \in \mathbb{C} - 0$;
3. $\mu_{norm}(Ug, Ux) = \mu_{norm}(g, x)$ para todo U transformación unitaria de \mathbb{C}^2 ;
4. $\mu_{norm}(g, x) \geq 1$ ($= \infty$ para raíces dobles).

5.1. Cota Geodésica para la complejidad del método de Homotopia

En esta sección vamos a enunciar el Teorema de cota Geodésica de la complejidad. Por un lado tenemos un problema de naturaleza discreta que es dar el número de operaciones requeridas por el método de homotopia para dar un cero aproximado: quien halla encarado un problema de este estilo sabe que en general los problemas discretos son más engorrosos de resolver con exactitud que digamos, calcular una integral. En esta sección veremos una manera de construir inductivamente una partición $\{t_0, \dots, t_K\}$ para nuestra homotopia f_t (que a grandes rasgos dependerá en cada paso t_i de un número de condición $\mu_{norm}(f_{t_i})$) y así acotar K mediante el integral del número de condición sobre la curva (f_t, ζ_t) . Con este objetivo en mente, quedaran algunos resultados intermedios de naturaleza más bien técnica sin demostración pero tratando de dar cierta intuición de por qué deberían funcionar.

Comenzamos con algunos resultados previos:

Teorema. 5.1.1. *Sean $g \in \mathbb{P}(\mathcal{H}_d)$, $\zeta, \eta \in \mathbb{P}(\mathbb{C}^2)$. Fijado $\varepsilon > 0$ existe una constante $C > 0$ tal que, de cumplirse:*

$$d(\zeta, \eta) < \frac{C}{d^{3/2} \mu_{norm}(g, \zeta)}$$

Entonces:

$$\mu_{norm}(g, \eta) \leq (1 + \varepsilon) \mu_{norm}(g, \zeta)$$

Demostración.

Notar que $u \leq d^{3/2} \mu_{norm}(g, \eta) \cdot d(\eta, \zeta) \leq C$. Luego $r_0 \sim d(\eta, \zeta)$, en la definiciones de K es ≤ 1 y $\left(\frac{\|\eta\|}{\|\zeta\|}\right)^{d-1} \sim (1 + d(\eta, \zeta))^{d-1} < e^c$. □

Observación. 5.1.2. Este teorema nos permite cuantificar las variaciones locales en el número de condición, es decir, variaciones pequeñas en el dominio de f generan variaciones pequeñas en el número de condición.

Teorema. 5.1.3. *Dado $\varepsilon > 0$, hay una constante $C > 0$ tal que: para $f, g \in \mathbb{P}(\mathcal{H}_d)$ y $\zeta \in \mathbb{P}(\mathbb{C}^2)$ con*

$$d(f, g) < \frac{C}{d^{1/2} \mu_{norm}(f, \zeta)},$$

entonces:

$$\mu_{norm}(g, \zeta) \leq (1 + \varepsilon) \mu_{norm}(f, \zeta).$$

Demostración. Por la Proposición 5b) de [17] sección I-3

$$\mu_{norm}(f, \zeta) \leq \frac{\mu_{norm}(f, \zeta)(1 + d(f, g))}{1 - D^{1/2} d(f, g) \mu_{norm}(f, \zeta)}$$

$$\leq \frac{\mu_{norm}(f, \zeta) \left(1 + \frac{C}{D^{1/2} \mu_{norm}(f, \zeta)}\right)}{1 - C}$$

Recordando que $\mu_{norm}(f, \zeta) \geq 1$.

□

Teorema. 5.1.4. *Dado ε hay una constante $C > 0$ tal que: Si $f, g \in \mathbb{P}(\mathcal{H}_d)$ y $\zeta, \eta \in \mathbb{P}(\mathbb{C}^2)$ y*

$$1. \ d(f, g) < \frac{C}{d^{1/2} \mu_{norm}(f, \zeta)}$$

$$2. \ d(\zeta, \eta) < \frac{C}{d^{3/2} \mu_{norm}(f, \zeta)}$$

Entonces:

$$\frac{1}{1 + \varepsilon} \mu_{norm}(g, \eta) \leq \mu_{norm}(f, \zeta) \leq (1 + \varepsilon) \mu_{norm}(g, \eta)$$

Demostración. aplicar los 2 teoremas anteriores para la igualdad de la izquierda. Por otro lado: dada la igualdad de la izquierda aplicar de nuevo los teoremas para probar la de la derecha, aplicando el teorema para g, η en lugar de f, ν . □

Proposición. 5.1.5. *Dado $\varepsilon > 0$ existe una constante $C > 0$ con la siguiente propiedad: Sea (f_t, ζ_t) una curva C^1 en la variedad V , para $t_0 \leq t \leq t_1$ vamos a definir una partición $\{s_i\}$ de $[t_0, t_1]$:*

- $s_0 = t_0$,
- s_i es el primer valor $t \in (s_{i-1}, t_1)$ que cumple:

$$\int_{s_{i-1}}^{s_i} \left(\|\dot{f}\| + \|\dot{\zeta}\| \right) dt = \frac{C}{d^{3/2} \mu_{norm}(f_{s_{i-1}}, \zeta_{s_{i-1}})}$$

o t_1 en caso contrario (el ultimo intervalo pude integrar menos).

Entonces

$$k \leq \max \left\{ 1, \frac{(1 + \varepsilon)}{C} d^{3/2} \int_{t_0}^{t_1} \mu_{norm}(f_t, \zeta_t) \left(\|\dot{f}\| + \|\dot{\zeta}\| \right) dt \right\},$$

y además

$$\mu_{norm}(f_t, \zeta_t) \leq (1 + \varepsilon)^k \mu_{norm}(f_0, \zeta_0).$$

Demostración.

$$\begin{aligned} & \int_{s_{i-1}}^{s_i} \mu_{norm}(f_t, \zeta_t) \left(\|\dot{f}\| + \|\dot{\zeta}\| \right) dt \geq \\ & \geq \frac{1}{(1 + \varepsilon)} \int_{s_{i-1}}^{s_i} \mu_{norm}(f_{s_{i-1}}, \zeta_{s_{i-1}}) \left(\|\dot{f}\| + \|\dot{\zeta}\| \right) dt \geq \frac{1}{1 + \varepsilon} \frac{C}{d^{3/2}} \end{aligned}$$

5.1. COTA GEODÉSICA PARA LA COMPLEJIDAD DEL MÉTODO DE HOMOTOPIA45

por lo tanto:

$$\int_{s_0}^{s_k} \mu_{norm}(f_t, \zeta_t) \left(\|\dot{f}_t\| + \|\dot{\zeta}_t\| \right) dt \geq \frac{(k-1)C}{(1+\varepsilon)d^{3/2}}$$

y además:

$$(k-1) \frac{C}{(1+\varepsilon)d^{3/2}} \leq \int_{t_0}^{t_1} \mu_{norm}(f_t, \zeta_t) \left(\|\dot{f}_t\| + \|\dot{\zeta}_t\| \right) dt.$$

Esto nos dice que:

$$k-1 \leq \frac{(1+\varepsilon)d^{3/2}}{C} \int_{t_0}^{t_1} \mu_{norm}(f_t, \zeta_t) \left(\|\dot{f}_t\| + \|\dot{\zeta}_t\| \right) dt.$$

□

Otro teorema de esta seguidilla:

Lema 5.1.6. Sean:

$u < \frac{3-\sqrt{7}}{4}$, $f \in \mathcal{H}_d$ y $\zeta \in \mathbb{P}(\mathbb{C}^2)$ raíz de f . Si se cumple la desigualdad:

$$d(x, \zeta) \leq \frac{u}{d^{3/2} \mu_{norm}(f, \zeta)}.$$

Entonces:

$$d(N_f(x), \zeta) \leq \frac{4u}{\psi(2u)} d(x, \zeta).$$

Demostración. $\tan d(x, \zeta) \leq 2d(x, \zeta)$ usamos Lema 1 P263 de BCSS [8]

$$\begin{aligned} d(N_f(x), \zeta) &\leq \tan d(N_f(x), \zeta) \leq \frac{2u}{\psi(2u)} \tan d(x, \zeta) \\ &\leq \frac{4u}{\psi(2u)} d(x, \zeta). \end{aligned}$$

□

Para el siguiente teorema utilizaremos u_0 la solución de $\frac{4u}{\psi(2u)} = 1/2$, $u_0 = \frac{16-\sqrt{232}}{16}$.

Teorema. 5.1.7 (Cero aproximado). Sea $f \in \mathbb{P}(\mathcal{H}_d)$, $\zeta \in \mathbb{P}(\mathbb{C}^2)$ con $f(\zeta) = 0$. De cumplirse $d(x, \zeta) < \frac{u_0}{d^{3/2} \mu_{norm}(f, \zeta)}$ entonces:

$$d(N_f^k(x), \zeta) \leq \left(\frac{1}{2} \right)^{2^k - 1} d(x, \zeta),$$

es decir, x es un cero aproximado de ζ .

Demostración. Por inducción: supongamos que:

$$\begin{aligned} d(N_f^k(x), \zeta) &\leq \left(\frac{4u_0}{\psi(2u_0)} \right)^{2^k-1} d(x, \zeta) \\ &\leq \frac{\left(\frac{4u_0}{\psi(2u_0)} \right)^{2^k-1} u_0}{d^{3/2} \mu_{norm}(f, \zeta)} \end{aligned}$$

Sea $u_1 = \left(\frac{4u_0}{\psi(2u_0)} \right)^{2^k-1} u_0$. Y ahora notar que $u_1 \leq u_0$ y por lo tanto $\psi(2u_1) \geq \psi(2u_0)$.

$$\begin{aligned} d(N_f^{k+1}(x), \zeta) &\leq \frac{4u_1}{\psi(2u_1)} d(N_f^k(x), \zeta) \leq \frac{4u_1}{\psi(2u_0)} d(N_f^k(x), \zeta) \\ &\leq 4 \frac{\left(\frac{4u_0}{\psi(2u_0)} \right)^{2^k-1} u_0}{\psi(2u_0)} \left(\frac{4u_0}{\psi(2u_0)} \right)^{2^k-1} d(x, \zeta) \leq \left(\frac{4u_0}{\psi(2u_0)} \right)^{2^{k+1}-1} d(x, \zeta). \end{aligned}$$

□

Recordemos que estamos trabajando con: (f_t, ζ_t) de clase C^1 , $\frac{d}{dt}(f_t, \zeta_t)$ su vector tangente y $\left\| \frac{d}{dt}(f_t, \zeta_t) \right\|$ el largo de su vector tangente.

Teorema. 5.1.8 (Cota Geodésica de la complejidad). *Existe una constante $C_1 > 0$ tal que si (f_t, ζ_t) con $t_0 \leq t \leq t_1$ es una curva C^1 en \mathcal{V} , entonces:*

$$C_1 d^{3/2} \int_{t_0}^{t_1} \mu_{norm}(f_t, \zeta_t) \left\| \frac{d}{dt}(f_t, \zeta_t) \right\| dt$$

pasos del método global de Newton son suficientes para “continuar” un cero aproximado x_0 de f_0 a x_1 un cero aproximado de f_1 .

Demostración. Elegimos $C < \mu_0$ y ε suficientemente pequeño para satisfacer hipótesis de 5.1.4 y 5.1.7, $u = 2C(1 + \varepsilon) < \frac{3-\sqrt{7}}{4}$ y $\frac{4y}{\psi(2u)} < \frac{1}{2(1+\varepsilon)}$. Por lo tanto, si:

- $d(f, g) < \frac{C}{d^{1/2} \mu_{norm}(f, \zeta)}$,
- $d(\zeta, \eta) \leq \frac{C}{d^{3/2} \mu_{norm}(f, \zeta)}$,
- $f(\zeta) = 0, g(\zeta) = 0$,
- $d(x, \zeta) \leq \frac{C}{d^{3/2} \mu_{norm}(f, \zeta)}$.

Entonces:

$$d(N_g(x), \eta) \leq \frac{C}{d^{3/2} \mu_{norm}(g, \eta)}.$$

Por el teorema anterior $N_g(x)$ es un cero aproximado de g con cero asociado η .

Ahora aplicamos la proposición 5.1.5 para producir s_0, \dots, s_k de manera que:

$$d(x_0, \zeta_{t_0}) < \frac{C}{d^{3/2} \mu_{norm}(f_{t_0}, \zeta_{t_0})}.$$

Entonces: $x_i = N_{f_{s_i}}(x_{i-1})$ es un cero aproximado de f_{s_i} con cero aproximado ζ_{s_i} :

$$d(x_i, \zeta_{s_i}) < \frac{C}{d^{3/2} \mu_{norm}(f_{s_i}, \zeta_{s_i})}.$$

□

Corolario. 5.1.9. *Existe una constante $C_2 > 0$ de manera que si (f_t, ζ_t) es un camino C^1 en \mathcal{V} entonces*

$$C_2 d^{3/2} \int_{t_0}^{t_1} \mu_{norm}^2(f_t, \zeta_t) \|\dot{f}_t\| dt$$

pasos del método global de Newton son suficientes para continuar un cero aproximado x_0 de f_{t_0} con cero asociado ζ_0 a un cero aproximado x_1 de f_{t_1} con cero asociado ζ_1

Demostración. $\|\dot{\zeta}_t\| \leq \mu_{norm} \|\dot{f}_t\|.$

□

$$\|(\dot{f}_t, \dot{\zeta}_t)\|_k^2 = \mu_{norm}(f, \zeta)^2 \left(\|\dot{f}\|^2 + \|\dot{\zeta}\|^2 \right).$$

Es la métrica del numero de condición.

Observación. 5.1.10. No hay unicidad de geodésica pero este $\mathcal{W} = \mathcal{V} - \Sigma'$ es completo con la métrica $\|\cdot\|_k$

5.2. Complejidad promedio $\mathbb{E}(K)$

Ahora a lo que nos compete: tenemos estudiado cuanto tarda a partir de un punto z_0 y un polinomio f dar un cero aproximado el algoritmo de Smale: $K(f, z_0)$ la cota de la complejidad geodésica. Ahora nos preguntamos:

¿Qué podemos decir de la complejidad promedio?

¿Tomado un polinomio “al azar” es esperable dar una cota de la complejidad?

Estas preguntas van en la linea del siguiente problema:

Problema 17 de Smale 1. *¿Es posible encontrar un cero aproximado de un sistema polinomial, en promedio, en tiempo polinomial con un algoritmo uniforme?*

No esta demás mencionar que Lairez [14] responde afirmativamente al *Problema 17*. De todas maneras, en este trabajo monográfico retomamos ideas de [1] que plantean problemas sobre las cuencas de atracción del método de homotopias para polinomios homogéneos.

Definición. 5.2.1. *Definimos el promedio, o esperanza de una función $\phi : \mathbb{P}(\mathcal{H}_d) \rightarrow \mathbb{R}$, como:*

$$\mathbb{E}_{\mathbb{P}(\mathcal{H}_d)}(\phi) = \frac{1}{\text{Vol}(\mathbb{P}(\mathcal{H}_d))} \int_{f \in \mathbb{P}(\mathcal{H}_d)} \phi(f) df = \frac{\Gamma(N)}{\pi^{N-1}} \int_{f \in \mathbb{P}(\mathcal{H})} \phi(f) df.$$

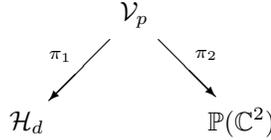
Como $\phi(\lambda(f)) = \phi(f)$ con $\lambda \in \mathbb{C} - \{0\}$ también computamos esta medida en \mathcal{H}_d :

$$\mathbb{E}_{\mathcal{H}_d} = \frac{1}{\pi} \int_{\mathcal{H}_d} \phi(f) e^{-\|f\|^2} df$$

Recordemos que estamos trabajando en: la variedad solución!

$$\mathcal{V}_p = \{(f, x) \in \mathcal{H}_d \times \mathbb{P}(\mathbb{C}^2) : f(x) = 0\}$$

este objeto es medio campista (central) de nuestro estudio. Esta equipado con dos proyecciones:



Luego de eso, vamos a los bifos a calcular la complejidad promedio. Ya vimos que con el método de Newton global puede definirse una partición de $[0, 1]$ para aproximar ζ_t . Definiendo

$$x_{t+\Delta t} = N_{f_{t+\Delta t}}(x_t)$$

Donde

$$N_f(x) = x - (Df(x)|_{x^\perp})^{-1} f(x)$$

El resultado principal de Shub en [16] (lo que probamos en la sección anterior) es que Δ_t puede ser elegido de manera que $t_0 = 0$, $t_k = t_{k-1} + \Delta_{t_k}$, x_{t_k} es un cero aproximado de f_{t_k} con cero asociado ζ_{t_k} , y $t_K = 1$ para:

$$K(f, g, \zeta) \leq C d^{3/2} \int_0^1 \mu(f_t, \zeta) \|(\dot{f}_t, \dot{\zeta}_t)\|_{(f_t, \zeta_t)} dt$$

Aquí C es una constante universal.

$$\|(\dot{f}_t, \dot{\zeta}_t)\| = (\|\dot{f}_t\|_{f_t}^2 + \|\dot{\zeta}_t\|_{\zeta_t}^2)^{1/2}$$

es la norma del vector tangente a la curva proyectada en (f_t, ζ_t) en $\mathcal{V}_{\mathbb{P}}$.

En $\mathcal{V}_{\mathbb{P}}$ tenemos que $\left\| \dot{\zeta}_t \right\|_{\zeta_t} \leq \mu(f_t, \zeta_t) \left\| \dot{f}_t \right\|_{f_t}$ y con esto podemos escribir ahora la siguiente cota:

$$K(f, g, \zeta) \leq CD^{3/2} \int_0^1 \mu(f_t, \zeta_t) \left\| \dot{f}_t \right\|_{f_t} dt.$$

Donde C es quizá una constante universal diferente. Finalmente en el caso de la homotopía lineal tenemos $\left\| \dot{f}_t \right\|_{f_t} = \frac{\sin(\theta) \|f_0\| \|f_1\|}{\|f_t\|^2}$ con θ el ángulo entre los extremos de homotopía f_0 y f_1 . Con esto tenemos:

$$K(f, g, \zeta) \leq CD^{3/2} \sin(\theta) \|f_0\| \|f_1\| \int_0^1 \frac{\mu(f_t, \zeta_t)^2}{\|f_t\|^2} dt.$$

Vamos a considerar el mapa:

$$\Phi : \mathcal{H}_d \times \mathbb{P}(\mathbb{C}^2) \times [0, 1] \rightarrow \mathcal{V}$$

Definido por:

$$\Phi(f, \zeta, t) = (f_t, \zeta_t)$$

donde

$$f_t(z) = f(z) - (1-t) \frac{K_{\zeta}(z)}{K_{\zeta}(\zeta)} f(\zeta).$$

Proposition 1. Para casi todo $f \in \mathcal{H}_d$ el conjunto de $\zeta \in \mathbb{P}(\mathbb{C}^2)$ tal que Φ esta bien definida para todo $t \in [0, 1]$ tiene medida total. Mas aun, para todo $\zeta \in \mathbb{P}(\mathbb{C}^2)$, el conjunto de los $f \in \mathcal{H}_d$ tal que Φ esta bien definido $\forall t \in [0, 1]$ tiene medida total.

Observación. 5.2.2. En la prueba de este resultado (ver Armentano, Shub [4]) se verifica que el complemento de estos conjuntos para los cuales esta bien definida la homotopía es una variedad algebraica real. Con esto estamos preparando el terreno para un posible estudio de las cuencas de atracción del método.

Definimos un poco mas de notación: $\mathcal{T}(f, \zeta) = K(f, f_0, \zeta)$ y $\mathcal{T}_{\zeta}(f) = \mathcal{T}(f, \zeta)$ entonces el costo promedio de este algoritmo satisface:

Proposición. 5.2.3.

$$\mathbb{E}_{\mathcal{H}_d}(\mathcal{T}_{\zeta}) = \mathbb{E}_{\mathcal{H}_d \times \mathbb{P}(\mathbb{C}^2)}(\mathcal{T}) \leq (I)$$

Con

$$(I) = \frac{Cd^{3/2}}{(2\pi)^{d+1} \text{Vol}(\mathbb{P}(\mathbb{C}^2))} \int_{f \in \mathcal{H}_d} \int_{\zeta \in \mathbb{P}(\mathbb{C}^2)} \int_{t \in [0, 1]} A(f, \zeta, t) df d\zeta dt$$

Donde:

$$A(f, \zeta, t) = \frac{\mu(f_t, \zeta_t)^2}{\|f_t\|^2} \times (|f|^2 - |\zeta|^{-d} |f(\zeta)|^2)^{1/2} |f(\zeta)|^2 e^{|f|^2/2}$$

Teorema. 5.2.4. [Teorema 1 de [4]]

$$(I) = \frac{Cd^{3/2}}{(2\pi)^{d+1}} \int_{h \in \mathcal{H}_d} \left[\sum_{\eta: h(\eta)=0} \frac{\mu^2(h, \eta)}{|h|^2} \Theta(h, \eta) \right] e^{-|h|^2/2} dh$$

donde:

$$\Theta(h, \eta) = \frac{1}{\text{vol}(\mathbb{P}(\mathbb{C}^2))} \int_{\zeta \in B(h, \eta)} \theta_h(\zeta) d\zeta$$

con

$$\begin{aligned} \theta_h(\zeta) &= (|h|^2 - |\zeta|^{-2d} |h(\zeta)|^2)^{1/2} |\zeta|^{-d} |h(\zeta)| \mathcal{I}(|\zeta|^{-2d} |h(\zeta)|^2)/2 \\ \mathcal{I}(\alpha) &= \int_0^1 e^{(1-t^{-2})\alpha t^{-3}} dt \end{aligned}$$

Observación. 5.2.5. La prueba del teorema 5.2.4 involucra complicadas formulas con enormes cancelaciones.

Observación. 5.2.6. Un polinomio homogeneo $h : \mathbb{C}^2 \rightarrow \mathbb{C}$ no necesariamente define una función de $\mathbb{P}(\mathbb{C}^2)$ a \mathbb{C} pero la función \hat{h} que manda $\zeta \in \mathbb{P}(\mathbb{C}^2)$:

$$\zeta \rightarrow \left| \frac{h(\zeta)}{|\zeta|^d} \right|$$

si esta bien definida.

5.3. Cuencas de atracción

Simulaciones realizadas dan pie a la conjetura de que las d cuencas de atracción del método global de Newton para un polinomio de grado d con d raíces diferentes tienen la misma medida. De probar que las cuencas tienen una medida igual o similar, todo parece apuntar gracias a las cotas integrales del Teorema 1 que involucran las cuencas, que podría darse una nueva prueba del problema 17 de Smale para este tipo de homotopias.

Observación. 5.3.1. En el caso mas general, para sistemas polinomiales entender mejor el tamaño de las cuencas de atracción puede llevar a una prueba del problema 17. El promedio del número de condición cumple:

$$\frac{1}{(2\pi)^N} \int_{h \in \mathcal{H}_{(d)}} \sum_{\eta/h(\eta)=0} \frac{\mu^2(h, \eta)}{\|h\|^2} e^{-\|h\|^2/2} dh \leq \frac{e(n+1)}{2} \mathcal{D}$$

donde $\mathcal{D} = d_1 \cdot d_2 \cdot \dots \cdot d_n$ es el número de Bezout. La cuestión sería ver como el factor $\Theta(h, \eta)$ afecta el integral de 5.2.3. Si el integral en las \mathcal{D} cuencas esta razonablemente balanceado el factor \mathcal{D} en la integral podrían cancelarse.

Ahora veremos un resultado sobre la geometría de las cuencas de atracción para polinomios complejos: Vamos a utilizar el siguiente teorema del análisis complejo (ver “Complex analysis” [15] para la demostración):

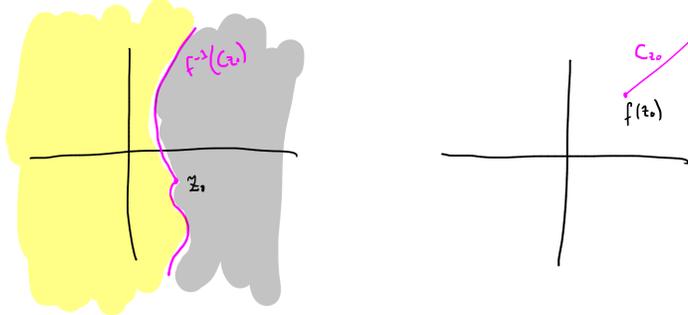
Teorema 5.3.2. *Sea $f(z)$ una función analítica en el punto z_0 , $f(z_0) = w_0$ de manera que $f(z) - w_0$ tiene un cero de orden n en z_0 . Si $\varepsilon > 0$ es suficientemente pequeño, existe un δ tal que para todo a con $|a - w_0| < \delta$, la función $f(z) - a$ tiene exactamente n raíces en el disco $|z - z_0| < \varepsilon$.*

Sea $f \in \mathcal{P}_d$ y $z_0 \in \mathbb{C}$ un punto crítico de f , es decir $f'(z_0) = 0$. Consideramos el siguiente conjunto:

$$C_{z_0} = \{t \cdot f(z_0), t > 1\}.$$

Sabemos que el método global de Newton no converge a una raíz partiendo de un $\zeta_0 \in f^{-1}(C_{z_0})$ y por el teorema que acabamos de enunciar sabemos que localmente $f^{-1}(C_{z_0})$ divide a \mathbb{C} en 2 regiones suponiendo que $f(z) - f(z_0)$ tiene un cero de orden 2.

Figura 5.1:



Podemos suponer sin pérdida de generalidad (casos de medida 0 en \mathcal{P}_d) que no hay más críticos en este segmento C_{z_0} y al tener necesariamente 2 preimágenes cada punto, $f^{-1}(C_{z_0})$ divide a \mathbb{C} . Si repetimos esta división para los demás puntos críticos de f obtenemos d regiones abiertas de \mathbb{C} que tienen que ser las d cuencas de atracción de las d raíces que tiene f .

Ejemplo 5.3.3. *Veamos las cuencas de $z^d - 1$. Se puede probar que partiendo de $\zeta \in \mathbb{C}$ tal que $\arg(\zeta) = \frac{i\pi}{d}$ el método de Newton no converge: Consideremos el segmento C_{-1} correspondiente al valor crítico -1 . Su preimagen por f son los d rayos que salen del origen, las bisectrices entre las raíces de la unidad.*

Esto que acabamos de mencionar es para el caso de la homotopia entre polinomios univariados. Las simulaciones las realizamos para el caso homogéneo:

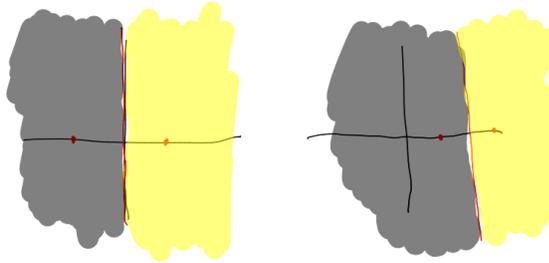
$$h_t(z) = f(z) - (1-t) \frac{\langle z, \zeta \rangle^d}{\langle \zeta, \zeta \rangle^d} f(\zeta).$$

Observamos que si nos restringimos a $z \in \zeta + \zeta^\perp$ recuperamos el método de Smale para el polinomio univariado:

$$h_t(z) = f(z) - (1-t)f(\zeta).$$

En el caso de $p \in \mathcal{P}_d$ las áreas de las cuencas de atracción (con la medida inducida por la proyección estereográfica de la uniforme en la esfera) no se preservan por traslaciones. (Basta mirar el caso de un polinomio con dos raíces, donde las cuencas están separadas por la mediatriz.) Sin embargo, si miramos el polinomio homogeneizado, y mirarnos la esfera de Riemann, en ese caso, sí las áreas coinciden.

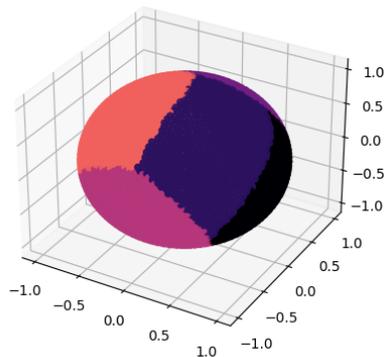
Figura 5.2: En este ejemplo al considerar las cuencas de un polinomio de grado 2 con raíces reales simétricas (izquierda) y otro con raíces reales asimétricas (derecha), al pasar a la esfera claramente las áreas no son iguales.



Esto es una ventaja de usar el $\mathbb{P}(\mathbb{C}^2)$ o la esfera de Riemann como dominio, donde las transformaciones unitarias se identifican con las isomerías de la esfera de Riemann.

Una posible ventaja de trabajar con la homotopia para los homogeneizados es que al considerar las áreas en la esfera de Riemann, parecen ser muy similares.

Figura 5.3:



5.4. Simulaciones

Concluimos este trabajo con simulaciones en *Phyton* con el objetivo de observar las áreas de las cuencas de atracción para un polinomio complejo en una variable, mediante el método de continuación de su polinomio homogeneizado.

Sorteamos un polinomio de grado $d = 7$, los coeficientes que obtuvimos: $f(z) = \sum_{i=0}^7 a_i z^i$ con:

$$\begin{aligned} a_0 &= 1,8586983152757441 - i1,3557586148003529 \\ a_1 &= 3,188447869061666 + i1,2294120018138717 \\ a_2 &= -2,4443752786266253 - i4,309971940571411 \\ a_3 &= 3,5774177358770007 + i9,034179896544229 \\ a_4 &= 10,827012601519256 + i7,087767595519535 \\ a_5 &= -1,5556902716958274 - i2,085327025630292 \\ a_6 &= 2,7145398197260455 + i1,5119078821808167 \end{aligned}$$

Raíces y cuencas: Utilizamos el metodo Quasi-Montecarlo. Primero obtenemos una aproximación de los puntos de Fekete para $N = 5000$, es decir $X = \{x_1, \dots, x_N\}$, puntos en \mathbb{S}^2 y mediante la estereografica, hacemos la homotopia para el homogeneizado de f a partir de cada punto y lo asignamos a la cuenca correspondiente.

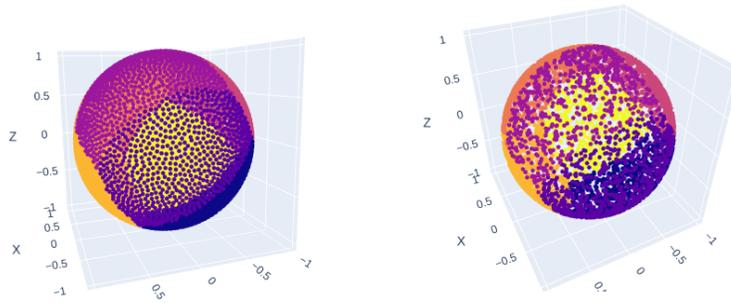
Observamos que los tamaños de las cuencas siempre están muy cerca de

$$1/d = 0,1428.$$

Raiz ζ_i	$B(f, \zeta_i)$ Q-Montecarlo	$B(f, \zeta_i)$ Montecarlo
$-1,8258408916002673 - i2,1314297736542613$	0,1444	0,1488
$-0,9785573456036832 + i0,2131261424641405$	0,144	0,146
$-0,15950130589302938 + i0,28648550391315275$	0,141	0,1382
$0,07242902038788969 - i0,5200502493060052$	0,1428	0,1402
$0,34449487515904714 + i0,329613837578822$	0,1408	0,1414
$0,40131197094334803 + i1,4243551929335154$	0,1454	0,1456
$1,3408828619850752 - i0,8505552875792256$	0,1416	0,1398

Cuadro 5.1: Estimados del área de $B(f, \zeta_i)$ mediante el computo de: $\#(B(f, \zeta_i) \cap X)/N$, con X equiviariado en el caso de Quasi-Montecarlo y con distribución gaussiana en Montecarlo.

Figura 5.4:



Observación. 5.4.1. Cabe mencionar que en el artículo de Armenano y Beltran [3] se prueba que los polinomios aleatorios con la distribución utilizada en las simulaciones están *bien condicionados* 5.0.8, o sea que en promedio, al sortear un polinomio sus raíces “se repelen”: están a buena distancia entre si en la esfera, esta propiedad se ve reflejada en el *potencial Logaritmico*.

5.5. Codigo

Código de las simulaciones en Phyton

April 4, 2024

```
[1]: import numpy as np
from numpy.polynomial import polynomial
import math

import numpy as np
from numpy.polynomial.polynomial import polyval2d

import numpy as np
import numpy.linalg as la
import time as t
import matplotlib.pyplot as plt
from statistics import median
from mpl_toolkits.mplot3d import Axes3D
from mpl_toolkits.mplot3d.art3d import Poly3DCollection
from scipy.spatial import ConvexHull

import plotly.graph_objects as go
```

$$\gamma(f, x) := \sup_{k>1} Df(x)^{-1} \cdot \frac{D^k f(x)^{1/(k-1)}}{k!}$$

```
[2]: def gammasmale(p,z):
    q=polynomial.polyder(p)

    poly = np.polynomial.Polynomial(p)
    d=poly.degree()
    a=[]
    for i in range(d):
        if i>1:
            p=polynomial.polyder(p)
            r=abs(np.polyval(p,z)/(math.factorial(i)*np.polyval(q,z)))*(1/(i-1))
            a.append(r)
    return max(a)
```

$$\beta(f, x_0) := N_f(x_0) - x_0 = Df(x_0)^{-1}f(x_0).$$

Esta funcion captura que tan pronunciado es la primera iteracion del metodo de Newton

```
[3]: def betasmale(p,z):
      q=polynomial.polyder(p)
      b=abs(np.polyval(p,z)/np.polyval(q,z))
      return b
```

$$\alpha(f, x) := \beta(f, x) \cdot \gamma(f, x)$$

```
[4]: def alphasmale(p,z):
      return gammasmale(p,z)*betasmale(p,z)
```

$$N_f(x) := x - Df(x)^{-1}f(x)$$

```
[5]: def newton(f, z):
      x = z-(np.polyval(f,z))/(np.polyval(np.polyder(f),z)) #x=z-f(z)/f'(z)
      return x
```

$$x_i = N_{f_i}(x_{i-1})$$

```
[6]: def smale(f,xi, pasos): #Newton global para polinomios complejos de una variable
      a=[]
      f_t=np.polyadd(f,-np.polyval(f,xi))
      x_t=xi
      a.append(xi)
      i=0
      for i in range(pasos):
          f_t=np.polyadd(f,-(1-(1+i)/pasos)*np.polyval(f,xi)) #f_t=f-(1-t)f(x_i)
          x_t=newton(f_t,x_t)
          a.append(x_t)
      return a
```

```
[7]: #el producto interno que vamos a usar en C^2
def pint(x,y):
    return x[0]*y[0].conjugate()+x[1]*y[1].conjugate()
```

La primera de estas funciones sorteas una polinomio $f \in \mathcal{P}_d(\mathbb{C}^2)$ osea $f(z) = \sum_{i=0}^d a_i z^i$, la distribucion utilizada para $Re(a_i)$ y $Im(a_i)$ es una Normal Gaussiana $(0, \binom{d}{i})$

La segunda funcion manda $f \rightarrow h_f \in \mathbb{C}^2$ $h_z(z, w) = \sum_{i=0}^d a_i z^i w^{d-1}$

y que nunca falte el nucleo reproductor:

$$K_\zeta(z) = \langle z, \zeta \rangle^d,$$

con $z, \zeta \in \mathbb{C}^2$

```
[8]: #Esta funcion sortea un polinomio de grado N

def random_poly(d):
    c = np.zeros(d+1)
    for i in range(d+1):
        c[i] = np.sqrt(math.comb(d,i))
    a = np.random.randn(d+1)*c + 1j*np.random.randn(d+1)*c
    return a

#esta da un homogeneizado a partir de un polinomio univariado
def homogeneizado(f):
    d=f.shape[0]
    h_p=np.zeros((d,d), complex)
    for i in range(d):
        h_p[i,d-1-i]=f[i]
    return h_p

#nos da el nucleo reproductor como polinomio homogeneo
def nucleoreproductor(z,d):
    k_z=h_p=np.zeros((d,d), complex)
    for i in range(d):
        k_z[i,d-1-i]=((z[0].conjugate())**i)*((z[1].conjugate())**(d-1-i))*math.
        ↪comb(d-1,i)
    return k_z
```

```
[9]: def homoderz(h,d):

    h_z=np.zeros((d,d), complex)
    for i in range(1,d):
        h_z[i-1,d-1-i]=i*h[i,d-1-i]
    return h_z

def homoderw(h,d):

    h_w=np.zeros((d,d), complex)
    for i in range(0,d-1):
        h_w[i,d-2-i]=(d-1-i)*h[i,d-1-i]
    return h_w

def Diff(z,w,h,d):
    c = np.zeros(2)
    return polyval2d(z,w,homoderz(h,d)),polyval2d(z,w,homoderw(h,d))
```

```
[10]: def Diffinv(z,w,h,d):
    Dif=Diff(z,w,h,d)
    return (1/(Dif[0]-Dif[1]*((z/w).conjugate()))).conjugate()
```

```

def NewtonHI(z,w,h,d):
    c=polyval2d(z, w, h)#.conjugate()
    z_1=Diffinv(z,w,h,d)*c
    Y=np.zeros(2,complex)
    Y[0]=z_1
    Y[1]=-z_1*((z/w).conjugate())
    return Y

def NewtonH(z,w,h,d):
    Y=NewtonHI(z,w,h,d)
    Z=np.zeros(2,complex)
    Z[0]=z-Y[0]
    Z[1]=w-Y[1]
    #trate de normalizar para ver si funcionaba asi pero no se dio

    #norma=Z[0]*Z[0].conjugate()+Z[1]*Z[1].conjugate()
    #norma=norma**(1/2)
    #W=np.zeros(2,complex)
    #W[0]=Z[0]/norma
    #W[1]=Z[1]/norma
    return Z

def continuarH(z,w,h,d,pasos):
    orb=[]
    Z_t=[z,w]
    K_zeta=nucleoreproductor([z,w],d)
    K=polyval2d(z, w, K_zeta)
    f_zeta=polyval2d(z, w, h)

    for i in range(pasos):

        h_t=h-K_zeta*(1/K)*(f_zeta)*(1-i/pasos)
        #print(i,polyval2d(Z_t[0],Z_t[1],h_t),1-(i+1)/pasos)
        Z_t=NewtonH(Z_t[0],Z_t[1],h_t,d)

    return Z_t

```

```

[11]: # Funciones auxiliares

# Función log-energy. La entrada es una configuración de puntos en la esfera w
def log_energy(w):
    w=np.array(w)
    n=len(w)
    f = 0
    for i in range(n):
        Di = w - w[i,:]

```

```

    di = np.sqrt(np.sum(Di**2,axis=1))
    di[i] = 1
    f += np.sum(-np.log(di))
return f

# Plotea la configuración de puntos z en la esfera unidad
def plot_points(z):
    pi = np.pi
    cos = np.cos
    sin = np.sin
    phi, theta = np.mgrid[0.0:pi:100j, 0.0:2.0*pi:100j]
    xx = sin(phi)*cos(theta)
    yy = sin(phi)*sin(theta)
    zz = cos(phi)
    fig = plt.figure()
    ax = fig.add_subplot(111, projection='3d')
    ax.plot_surface(xx, yy, zz, rstride=1, cstride=1, color='c', alpha=0.5,
→linewidth=0.5)
    ax.scatter(z[:,0],z[:,1],z[:,2], color='red')

```

[12]: *#Gradiente de log_energy con respecto a xi i entre 0 y N-1.*

```

def part_log_grad(x,i) :
    x = np.array(x)
    n = len(x)
    grad = np.zeros(3)

    for j in range(n) :
        if j != i :
            grad += -(x[i]-x[j])/la.norm(x[j]-x[i])**2

    return grad

#Gradiente de log_energy
def log_grad(x) :
    x = np.array(x)
    n = len(x)

    grad = [part_log_grad(x,i) for i in range(n)]

    return np.array(grad)

#Proyeccion en la esfera:
def proy(w) :
    return w/la.norm(w)

```

[13]: *#Descenso por gradiente proyectado por coordenadas.*

```

def gd_proy_coord(X, alpha, s, tol = 10e-5) :

```

```

i = 0
X0 = np.array(X)
n = len(X)
it = 0
val = [log_energy(X)]
d = []
while True :

    aux1 = X0
    x = X0[i]

    while True :

        aux2 = x
        x_proy = proy(x-s*part_log_grad(X0,i))
        x = x + alpha*(x_proy-x)

        it+=1

        if la.norm(aux2 - x) < tol :
            d.append(la.norm(aux2 - x))
            break

    X0[i] = x

    val.append(log_energy(X0))
    i +=1

    if i==n :

        if la.norm(X0 - aux1) < tol :
            break
        else :
            i = 0

return X0, it, val, d

```

```

[14]: #Descenso por gradiente proyectado.
def gd_proy(X0, alpha, s, tol = 10e-5) :
    X = np.array(X0)
    it = 0
    val = []
    while True :

        val.append(log_energy(X))
        aux = X
        X_proy = X - s*log_grad(X)

```

```

X_proy = [proy(v) for v in X_proy]

X = X + alpha*(X_proy-X)

it+=1

if la.norm(X - aux) < tol or it > 100 :
    break

return X, it, val

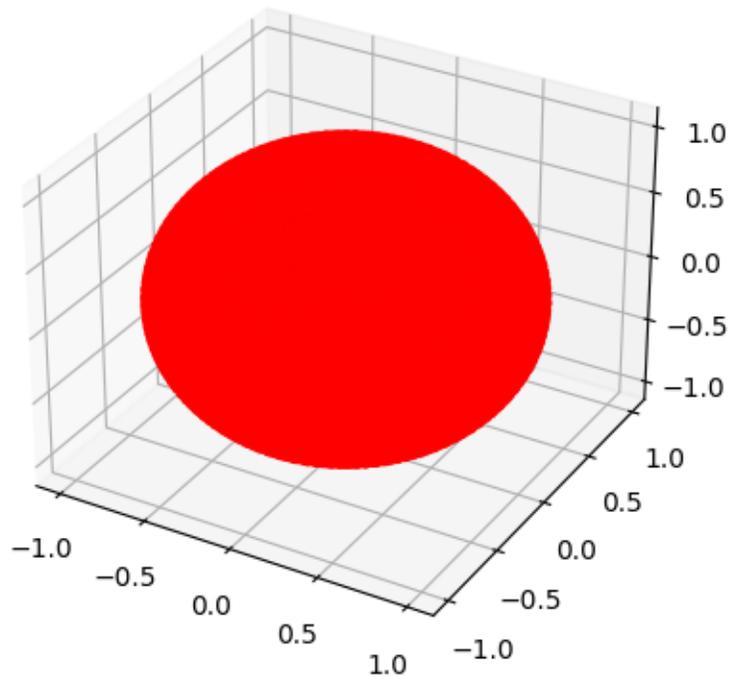
```

```

[18]: #De aca salen los puntos que vamos a usar para la simulacion, tarda un rato en
      →compilar, para 5000 pts con gd_proy(X0, 1, 1,10e-5) tarda 24 min
N=5000 # cantidad de puntos
#configuracion de puntos. Cada fila es un punto
X0 = np.random.randn(N,3)
X0 = X0 / np.linalg.norm(X0, axis=-1)[:, np.newaxis]
X, it, val = gd_proy(X0, 1, 1,10e-2) #10e-5
plot_points(X)
print('Energía logarítmica: ', log_energy(X))

```

Energía logarítmica: -4850062.985188703



Los puntos en $X = \{x_1, x_2, \dots, x_N\}$ estan en $S^2 = \{(x, y, z) : x^2 + y^2 + z^2 = 1\}$, pero los queremos

en $S = \{(x, y, z) : x^2 + y^2 + (z - 1/2)^2 = 1/4\}$.

Luego a cada punto le hacemos la continuacion por la homotopia y vamos a decir a cual raiz $\{\zeta_1, \zeta_2, \dots, \zeta_d\}$ tiende.

```
[19]: d=7 #fijamos el grado para la simulacion
f=random_poly(d)
poli=homogeneizado(f)

R=polynomial.polyroots(f) #estas son las raices del polinomio antes de
→homogeneizar
color_N=np.zeros(N)
N=X.shape[0]
for i in range(N):
    S=X[i][0]/2, X[i][1]/2, X[i][2]/2+1/2 #primero mandamos puntos de S^2 a S
    P_i=(S[0]+S[1]*1j)/S[2] #tengo puntos que sortie en la esfera S y los mando al
→plano afin (Z,1)
    Q_i=continuarH(P_i,1,poli,d+1,100)

    R_i=(R-Q_i[0]/Q_i[1]) #aca me fijo a que cuenca corresponde el punto
    color_N[i]=np.abs(R_i).argmin()
```

```
[20]: print(f,R)
```

```
[ 0.69149215-1.03805443j  2.71453982+1.51190788j -1.55569027-2.08532703j
 10.8270126 +7.0877676j   3.57741774+9.0341799j  -2.44437528-4.30997194j
  3.18844787+1.229412j    1.85869832-1.35575861j] [-1.82584089-2.13142977j
-0.97855735+0.21312614j -0.15950131+0.2864855j
 0.07242902-0.52005025j  0.34449488+0.32961384j  0.40131197+1.42435519j
 1.34088286-0.85055529j]
```

```
[30]: for i in range(d):
        print('a_',i, '=',f[d-i])
```

```
a_ 0 = (1.8586983152757441-1.3557586148003529j)
a_ 1 = (3.188447869061666+1.2294120018138717j)
a_ 2 = (-2.4443752786266253-4.309971940571411j)
a_ 3 = (3.5774177358770007+9.034179896544229j)
a_ 4 = (10.827012601519256+7.087767595519535j)
a_ 5 = (-1.5556902716958274-2.085327025630292j)
a_ 6 = (2.7145398197260455+1.5119078821808167j)
```

```
[35]: print('Abajo va el tamaño de cada cuenca (#B_i/N), observamos que esta siempre
→cerca de 1/d=',1/d)
for i in range(d):
    print('raiz',R[i], 'area estimada', (color_N==i).sum()/N)
```

```
Abajo va el tamaño de cada cuenca (#B_i/N), observamos que esta siempre cerca de
1/d= 0.14285714285714285
raiz (-1.8258408916002673-2.1314297736542613j) area estimada 0.1444
```

```

raiz (-0.9785573456036832+0.2131261424641405j) area estimada 0.144
raiz (-0.15950130589302938+0.28648550391315275j) area estimada 0.141
raiz (0.07242902038788969-0.5200502493060052j) area estimada 0.1428
raiz (0.34449487515904714+0.329613837578822j) area estimada 0.1408
raiz (0.40131197094334803+1.4243551929335154j) area estimada 0.1454
raiz (1.3408828619850752-0.8505552875792256j) area estimada 0.1416

```

```

[32]: x, y, z = X[:,0],X[:,1],X[:,2]

fig = go.Figure(data=[go.Scatter3d(x=x, y=y, z=z, mode='markers',
→marker=dict(size = 1.5 , color = color_N))])
fig.update_layout(scene=dict(
    xaxis_title='X',
    yaxis_title='Y',
    zaxis_title='Z',
    aspectmode='data'
))
fig.show()

```

```

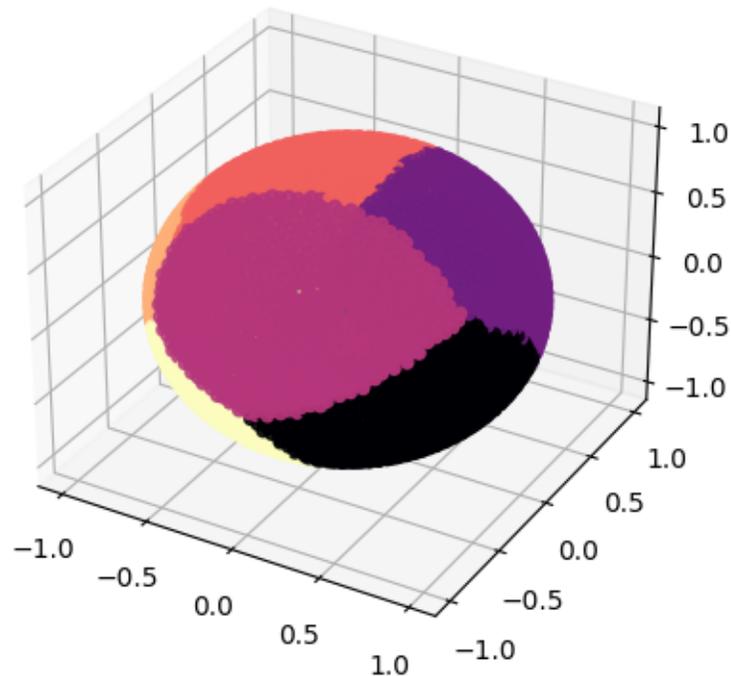
[36]: def plot_pointsC(z,colors):
    pi = np.pi
    cos = np.cos
    sin = np.sin
    phi, theta = np.mgrid[0.0:pi:100j, 0.0:2.0*pi:100j]
    xx = sin(phi)*cos(theta)
    yy = sin(phi)*sin(theta)
    zz = cos(phi)
    fig = plt.figure()
    ax = fig.add_subplot(111, projection='3d')
    ax.plot_surface(xx, yy, zz, rstride=1, cstride=1, color='c', alpha=0.5,
→linewidth=0.5)
    c=ax.scatter(z[:,0],z[:,1],z[:,2], cmap='magma_r',c=colors)

```

```

[37]: plot_pointsC(X,color_N)

```



```
[38]: !pip install pymanopt
```

```
Requirement already satisfied: pymanopt in /usr/local/lib/python3.10/dist-
packages (2.2.0)
Requirement already satisfied: numpy>=1.16.0 in /usr/local/lib/python3.10/dist-
packages (from pymanopt) (1.25.2)
Requirement already satisfied: scipy<1.10,>=1.0 in
/usr/local/lib/python3.10/dist-packages (from pymanopt) (1.9.3)
```

```
[39]: import pymanopt as pym
from pymanopt.manifolds import Product, Sphere

manilist = []
for _ in range(2):
    manilist.append(Sphere(3))
manifold = Product(manilist)
print(manifold.random_point())
```

```
[array([-0.5423877 ,  0.70619431, -0.45508811]), array([ 0.42106537,
 0.89202019, -0.16432873])]
```

Aca vamos a sortear con distribucion uniforme 5000 puntos en la esfera S^2

```
[40]: manilist = []
a = X
for i in range(5000):
    manilist.append(Sphere(3))
    manifold = Product(manilist)
    a[i]=manifold.random_point()[0]
    #x,y,z+=a[0],a[1],a[2]
```

```
[42]: R=polynomial.polyroots(f) #estas son las raices del polinomio antes de
    →homogeneizar
color_N=np.zeros(N)
N=X.shape[0]
for i in range(N):
    S=a[i,0]/2, a[i,1]/2, a[i,2]/2+1/2 #primero mandamos puntos de S^2 a S
    P_i=(S[0]+S[1]*1j)/S[2] #tengo puntos que sortie en la esfera S y los mando al
    →plano afin (Z,1)
    Q_i=continuarH(P_i,1,poli,d+1,100)

    R_i=(R-Q_i[0]/Q_i[1]) #aca me fijo a que cuenca corresponde el punto
    color_N[i]=np.abs(R_i).argmin()
```

```
[43]: x, y, z = a[:,0],a[:,1],a[:,2]

fig = go.Figure(data=[go.Scatter3d(x=x, y=y, z=z, mode='markers',
    →marker=dict(size = 1.5 , color = color_N))])
fig.update_layout(scene=dict(
    xaxis_title='X',
    yaxis_title='Y',
    zaxis_title='Z',
    aspectmode='data'
))
fig.show()
```

```
[45]: print('Abajo va el tamaño de cada cuenca (#B_i/N), observamos que esta siempre
    →cerca de 1/d=',1/d)
for i in range(d):
    print('cuenca Nro',i,':',(color_N==i).sum()/N)
```

Abajo va el tamaño de cada cuenca (#B_i/N), observamos que esta siempre cerca de 1/d= 0.14285714285714285

```
cuenca Nro 0 : 0.1488
cuenca Nro 1 : 0.146
cuenca Nro 2 : 0.1382
cuenca Nro 3 : 0.1402
cuenca Nro 4 : 0.1414
cuenca Nro 5 : 0.1456
cuenca Nro 6 : 0.1398
```

Bibliografía

- [1] D. Armentano. Complexity of path-following methods for the eigenvalue problem. *Foundations of Computational Mathematics*, 14(2):185–236, 2014.
- [2] D. Armentano, C. Beltrán, P. Bürgisser, F. Cucker, and M. Shub. Condition length and complexity for the solution of polynomial systems. *Foundations of Computational Mathematics*, 16(6):1401–1422, 2016.
- [3] D. Armentano, C. Beltrán, and M. Shub. Minimizing the discrete logarithmic energy on the sphere: the role of random polynomials. *Transactions of the American Mathematical Society*, 363(6):2955–2965, 2011.
- [4] D. Armentano and M. Shub. Smale’s fundamental theorem of algebra reconsidered. *Foundations of Computational Mathematics*, 14:85–114, 2014.
- [5] C. Beltrán. A continuation method to solve polynomial systems and its complexity. *Numerische Mathematik*, 117:89–113, 2011.
- [6] C. Beltrán and L. Pardo. Smale’s 17th problem: average polynomial time to compute affine and projective solutions. *Journal of the American Mathematical Society*, 22(2):363–385, 2009.
- [7] C. Beltrán Álvarez et al. Sobre el problema número 7 de smale. 2020.
- [8] L. Blum, F. Cucker, M. Shub, and S. Smale. Complexity and real computation: A manifesto. *International Journal of Bifurcation and Chaos*, 6(01):3–26, 1996.
- [9] P. Bürgisser and F. Cucker. On a problem posed by steve smale. *Annals of Mathematics*, pages 1785–1836, 2011.
- [10] P. Bürgisser and F. Cucker. Condition, volume 349 of *grundlehren der mathematischen wissenschaften*. Springer-Verlag, Berlin. doi, 10:978–3, 2013.
- [11] H. S. M. Coxeter. *Projective geometry*. Springer Science & Business Media, 2003.
- [12] J.-P. Dedieu. Points fixes. *Points fixes, zéros et la méthode de Newton*, pages 5–74, 2006.

- [13] J.-P. Dedieu, G. Malajovich, and M. Shub. Adaptive step-size selection for homotopy methods to solve polynomial equations. *IMA Journal of Numerical Analysis*, 33(1):1–29, 2013.
- [14] P. Lairez. A deterministic algorithm to compute approximate roots of polynomial systems in polynomial average time. *Foundations of computational mathematics*, 17(5):1265–1292, 2017.
- [15] D. Martin and L. Ahlfors. *Complex analysis*. New York: McGraw-Hill, 1966.
- [16] M. Shub. Complexity of bézout’s theorem vi: Geodesics in the condition (number) metric. *Foundations of Computational Mathematics*, 9(2):171–178, 2009.
- [17] M. Shub and S. Smale. Complexity of bézout’s theorem. i. geometric aspects. *Journal of the American Mathematical Society*, 6(2):459–501, 1993.
- [18] M. Shub and S. Smale. Complexity of bezout’s theorem ii volumes and probabilities. In *Computational algebraic geometry*, pages 267–285. Springer, 1993.
- [19] M. Shub and S. Smale. Complexity of bezout s theorem: Iii. condition number and packing. *Journal of complexity*, 9(1):4–14, 1993.
- [20] M. Shub and S. Smale. Complexity of bezout’s theorem v: polynomial time. *Theoretical Computer Science*, 133(1):141–164, 1994.
- [21] M. Shub and S. Smale. Complexity of bézout’s theorem iv: probability of success; extensions. *SIAM Journal on Numerical Analysis*, 33(1):128–148, 1996.
- [22] M. Shub and S. Smale. Complexity of bezout’s theorem. *Collected Papers Of Stephen Smale, The (In 3 Volumes)-Volume 3*, 3:1421, 2000.
- [23] S. Smale. The fundamental theorem of algebra and complexity theory. 1981.
- [24] S. Smale. Mathematical problems for the next century, mathematics: frontiers and perspectives. *Amer. Math. Soc., Providence, RI*, 271, 2000.