UNIVERSIDAD DE CANTABRIA

DEPTO. DE MATEMÁTICAS, ESTADÍSTICA Y COMPUTACIÓN

TESIS DOCTORAL

SOBRE EL PROBLEMA 17 DE SMALE: TEORÍA DE LA INTERSECCIÓN Y GEOMETRÍA INTEGRAL

CARLOS BELTRÁN ÁLVAREZ

SANTANDER, 2006

Universidad de Cantabria

Departamento de Matemáticas, Estadística y Computación

Programa de Doctorado Matemáticas y sus Aplicaciones

Sobre el Problema 17 de Smale: Teoría de la Intersección y Geometría Integral

Carlos Beltrán Álvarez

Dirigida por Luis Miguel Pardo Vasallo

Agradecimientos

Mi profesor, tutor y amigo Luis Miguel Pardo ha sido una fuente constante de inspiración, conocimiento y entusiasmo. Su comprensión profunda de las Matemáticas y su capacidad para dirigir una investigación son cualidades que despiertan mi admiración y que alimentan el agradecimiento profundo del que quiero dejar constancia en estas líneas. Gracias por su consejo y por su ejemplo, haber aprendido y trabajado con él será siempre el mejor de los recuerdos.

Gracias también a mi familia y amigos, y muy especialmente a Jara por su comprensión y por la confianza en mi depositada. Para ellos reservo, con cariño, la dedicatoria de estas páginas.

Al profesor Michael Shub, por darme la oportunidad de continuar trabajando y aprendiendo al término de esta etapa, y por haberme animado en varias ocasiones con su actitud y sus palabras, le debo también el más sincero agradecimiento. Su opinión y su presencia han inspirado muchos de mis esfuerzos durante los últimos años.

Los profesores Jean Pierre Dedieu, Jean Claude Yakoubsohn y Marc Giusti me han apoyado siempre y han hecho más sencilla y agradable la tarea de incorporarme al mundo de la investigación, abriéndome de par en par las puertas de sus laboratorios y despachos. Gracias por todo ello, serán siempre modelos a imitar por su calidad científica y su dedicación personal.

Son muchas otras las personas a las que quiero expresar mi gratitud por el apoyo y la ayuda recibidas durante el periodo que termina en la redacción de esta memoria: Los profesores que en múltiples ocasiones han respondido mis preguntas con paciencia (Juan Antonio Cuesta, Jesús Araujo), los que me han animado con su presencia en mis primeros logros (Jose Luis Montaña, Joos Heintz, Miguel Lobo, Tomás Recio), los que me han demostrado amistad y apoyo en toda circunstancia (Jorge San Martín, David Castro, Jose Enrique Morais), y otros muchos que han hecho de mis primeros años como investigador una época maravillosa. Mi más sincero agradecimiento.

The magician does not doubt that the same causes will always produce the same effects, that the performance of the proper ceremony, accompanied by the appropriate spell, will inevitably be attended by the desired result [...] Thus the analogy between the magical and the scientific conceptions of the world is close. In both of them the succession of events is assumed to be perfectly regular and certain, being determined by immutable laws, the operation of which can be foreseen and calculated precisely.

Sir James Frazer, The Golden Bough.

Índice general

Introducción I

Contexto histórico del problema I
Estado del Arte IV
Principales aportaciones de esta memoria XI
Problema 17 de Smale: El caso homogéneo XI
Problema 17 de Smale: El caso afín XVI
El condicionamiento lineal singular XVII
El condicionamiento no–lineal XX
Estimaciones discretas XXIV
Perspectivas, problemas a medio y largo plazo XXV

1. Preliminares 1

1.1. Elementos de Integración Geométrica 1
1.1.1. Volúmenes en espacios Euclídeos 1
1.1.2. Integración y Topología en variedades diferenciables . 4
1.1.3. La Fórmula de la Co–área 7
1.2. Volumen y distancia en el espacio proyectivo complejo 11
1.3. Notaciones de Geometría Algebraica. 15
1.4. El espacio de los sistemas de ecuaciones polinomiales 16
1.5. El condicionamiento en el Álgebra Lineal 24
1.5.1. El condicionamiento lineal generalizado 24
1.5.2. Estabilidad en el cálculo de núcleos e inversas 26
1.5.3. Teorema del número de condicionamiento 29
1.6. El condicionamiento de sistemas de ecuaciones polinomiales . 32
1.7. El método de Newton para sistemas de ecuaciones. El caso cero–dimensional 34
1.7.1. El operador de Newton proyectivo 34
1.7.2. El operador de Newton afín 36
1.8. El método de Newton para sistemas de ecuaciones. El caso de dimensión positiva. 37

2. El Condicionamiento de Matrices singulares: Un Análisis de Probabilidad 41
2.1. Introducción y resultados principales 41
2.2. Teoría de la Intersección y Geometría Integral en el espacio proyectivo complejo 43
2.3. Volumen de tubos en el espacio proyectivo complejo 50
2.3.1. Algunos resultados de Geometría Proyectiva 50
2.3.2. La cota para el volumen del tubo 58
2.4. Tubos extrínsecos 63
2.5. Distribución de probabilidad del condicionamiento en el Álgebra Lineal 65
2.5.1. Condicionamiento de las matrices de corango dado. . . 66
2.5.2. Condicionamiento de matrices simétricas y por bloques. 70
2.5.3. Un resultado general. 72

3. El Condicionamiento Generalizado de Sistemas de Ecuaciones Polinomiales: Un Análisis de Probabilidad 75
3.1. Introducción y conceptos básicos 75
3.1.1. Teorema del Número de Condicionamiento No–Lineal. 76
3.1.2. El control de la convergencia del operador de Newton en el caso de dimensión positiva 77
3.1.3. Separación y estabilidad del conjunto de soluciones . . 77
3.1.4. Algunas estimaciones de probabilidad 79
3.2. Demostración del Teorema del Número de Condicionamiento 82
3.3. El método de Newton en dimensión positiva 85
3.3.1. Demostración de la Proposición 3.1.2 88
3.4. La distancia a las variedades singulares de corango dado . . . 88
3.5. Integración en la variedad de incidencia generalizada 91
3.5.1. Demostración del Teorema 3.1.6 96
3.5.2. Algunas consecuencias 96
3.6. Distribuciones de probabilidad de los números de condicionamiento generalizados 102
3.7. Estabilidad en media del conjunto de soluciones 104
3.7.1. El valor esperable de $\mu_{\mathrm{av}}^{(m)}$ 105
3.7.2. El valor esperable de $\mu_{\mathrm{worst}}^{(m)}$ 107
3.8. El valor esperable del radio de convergencia del operador de Newton en dimensión positiva 113

4. Una Solución Probabilista al Problema 17 de Smale 115
4.1. Introducción. 115
4.1.1. Resultados principales. 116
4.1.2. Descripción explícita del conjunto questor 120
4.2. Notaciones y resultados previos 122
4.2.1. Algunas acciones unitarias 125

4.2.2. Homotopía y condicionamiento 126
4.2.3. La homotopía lineal. 127
4.2.4. Una estimación de volumen para círculos máximos . . 129
4.3. Una serie de reducciones geométricas 132
4.3.1. De círculos máximos a pares de sistemas 132
4.3.2. De pares de sistemas a la fibra en e_0 134
4.3.3. De la fibra en e_0 al espacio de matrices cuadradas . . 136
4.3.4. Del espacio de matrices cuadradas a los sistemas de ecuaciones lineales sub–determinados 140
4.4. El último paso argumental 143
4.4.1. Demostración del Teorema 4.4.1 147
4.5. Pares ε–eficientes para todo $\varepsilon > 0$ 148
4.5.1. Demostración del Teorema 4.1.8 150

5. El Problema 17 de Smale: El Caso Afín 153
5.1. Introducción 153
5.2. De ceros aproximados proyectivos a afines 156
5.3. El tamaño medio de las soluciones 160
5.3.1. Demostración del Teorema 5.1.1 163
5.4. Eficacia del algoritmo afín 164
5.4.1. Demostración del Teorema 5.1.2 164

6. Estimaciones Discretas 167
6.1. Introducción 167
6.2. Geometría de los Números 171
6.3. El caso de los conjuntos proyectivos 176
6.4. Estimaciones discretas para los números de condicionamiento 187
6.4.1. El condicionamiento lineal 187
6.4.2. El condicionamiento no–lineal 188
6.5. El problema 17 de Smale 191
6.5.1. Homotopía y conjuntos semi–algebraicos 191
6.5.2. Demostración del Teorema 6.1.5. 198

A. La Estructura Riemanniana del Espacio Proyectivo 201

Bibliografía 206

Introducción

Esta memoria nace del encuentro entre varios elementos de la Teoría de la Intersección Geométrica y de la Geometría Integral. La conjunción de ambas ha permitido abrir nuevos frentes en la comprensión de un problema central en varias ramas de las Matemáticas: El análisis de la complejidad de algoritmos de resolución de sistemas de ecuaciones polinomiales multivariadas.
La aportación más sugerente de esta memoria es ofrecer una respuesta positiva y probabilista al Problema 17 de los propuestos por Steve Smale en su lista de Problemas para el Siglo XXI, [119].
Este relevante resultado se obtiene como consecuencia de un análisis minucioso del comportamiento y propiedades de los números de condicionamiento de sistemas de ecuaciones polinomiales. Presentar el entramado que conduce a la resolución del Problema 17 de Smale y a los otros resultados descritos en la memoria exige también un intento de exponer el contexto científico en que se enmarca. Por eso hemos estructurado esta introducción en las siguientes secciones:

- Contexto histórico del problema.
- Estado del arte.
- Principales aportaciones de la memoria.
- Perspectivas, problemas a medio y largo plazo.

Contexto histórico del Problema

El problema de la resolución de sistemas de ecuaciones polinomiales no es un problema reciente en la Historia de las Matemáticas. Desde el Papyrus Rhine a los textos hindúes de la tradición china, el problema de resolver ecuaciones cuadráticas univariadas aparece en forma de ejercicios enunciados y resueltos por los autores (anónimos en la mayoría de los casos). La tradición griega (recogida en Euclides) también exhibe su conocimiento de la resolución de ecuaciones cuadráticas.
En el texto de Al-Khuarizmi encontramos ecuaciones univariadas de grado 2 resueltas a la manera de "Al-Uqlidisi".

No será hasta el Renacimiento italiano cuando la primera extensión del problema alcance rango de problema matemático: De las ecuaciones cuadráticas univariadas a las cúbicas, las cuárticas y las quínticas. El concepto de grado de un polinomio, la resolución mediante radicales de cúbicas y cuárticas y el establecimiento del problema de las quínticas puede seguirse en las obras o recuerdos históricos de autores como Scipio del Ferro, Annibal della Nave, Fiore, Tartaglia, Cardano, Ferrari, Bombelli, Viéte, Descartes...

En el Siglo XVIII, Issac Newton también intentó atacar el problema de la resolución por radicales de la ecuación de quinto grado. Fracasa en el análisis del problema y, en cambio, introduce una nueva forma de "resolver" ecuaciones polinomiales: El operador de Newton. De hecho, en [98], Newton enfrenta la "resolución" de una ecuación polinomial cúbica por un método que no sea por radicales. Por ejemplo, trata el problema de resolución de la ecuación $Y^3 - 2Y - 5$. Sobre la dificultad a la hora de aplicar su método, Newton escribe:

Quicquid laboris hic est, istud in Operatione sustituendi quantitates unas pro aliis reperietur.

Esto es, "Qué (escaso) trabajo hay aquí, en la operación de sustituir unas cantidades por otras". De este modo, Newton quería expresar de algún modo que, en su opinión, su método tiene "menos trabajo" que otros. Claramente Newton, interesado en los cálculos, está intuyendo el concepto de complejidad. En este punto debemos señalar que Newton parece reconocer que su "método" para coeficientes numéricos (Numeralis Aequationem) no siempre funciona; pero resta importancia al problema remitiéndolo a los "analistas": *Sicut Analystis notum est.* Aún así, lo recomienda:

Id quod varie perficias, at frequentem modum maxime expeditum puto, praesertim ubi Numeri Coefficientes constant ex pluribus Figuris.

Esto es, "Ello puede hacerse diversamente, yo juzgo el presente modo el más expedito, sobre todo, en el caso de que los mismos coeficientes consten de varias figuras (i.e. cifras)". Newton había desarrollado también su método para el caso que ahora llamaríamos no–arquimediano, esto es cuando los coeficientes no son números sino variables.

Otros dos grandes pilares en la historia de la resolución de ecuaciones algebraicas son, ya en el siglo XIX, N. Abel y E. Galois. Si bien Abel es el primero en descubrir la inexistencia de un método para resolver ecuaciones quínticas por radicales, Galois no solo caracteriza el fenómeno (mediante la introducción de la Teoría de Grupos), sino que además diseña un procedimiento algorítmico del que afirma:

Si maintenant vous me donez une équation que vous avez choisi à votre gué et que vous désirez connaitre si elle est ou non resoluble per radicaux,

je n'aurrai rien à y faire que de vous indiquer le moyen de répendre à votre question... sans vouloir charger ni moi ni personne de la faire. En un mot, les calculs sont impracticables.

Así, tanto Newton como Galois comprenden las dificultades de sus respectivas filosofías y conceptos de "resolución":

- Newton observa que su método parece requerir "menos cálculos", aunque elude el problema de saber si un punto es o no un buen punto inicial.
- Galois sabe que su algoritmo funciona siempre, pero la complejidad de su ejecución es tal que lo hace "impracticable".

Tanto Newton como Galois están preludiendo un problema relevante que se prolongará a lo largo del siglo XIX bajo el término genérico "Teoría de la Eliminación" y que encontrará su punto culminante (en el siglo XIX) en las obras de L. Kronecker ([84]) y D. Hilbert ([72]).
El problema ha dejado de ser un sencillo problema de tratamiento de ecuaciones polinomiales univariadas para convertirse en el diseño algorítmico y/o comprensión de la fenomenología de la resolución de sistemas de ecuaciones polinomiales multivariadas. El problema contempla las diversas vertientes que pueden surgir cuando se manipulan polinomios multivariados.
Así, sea $f = [f_1, \ldots, f_s]$ una lista de polinomios multivariados, de modo que $f_i \in \mathbb{C}[X_1, \ldots, X_n]$. Denotaremos por $V_{\mathbb{C}^n}(f)$ el conjunto algebraico afín de sus ceros comunes, es decir,

$$V_{\mathbb{C}^n}(f) := \{x \in \mathbb{C}^n : f_i(x) = 0,\ 1 \leq i \leq s\},$$

o también el conjunto de soluciones $(x_1, \ldots, x_n) \in \mathbb{C}^n$ del sistema de ecuaciones

$$\begin{cases} f_1(X_1, \ldots, X_n) = 0 \\ \vdots \\ f_s(X_1, \ldots, X_n) = 0 \end{cases}$$

El análisis de los conjuntos algebraicos $V_{\mathbb{C}^n}(f)$ se estructura en dos preguntas algorítmicas básicas:

Problema de Eliminación: HN (Hilbert Nullstellensatz): Diseñar un algoritmo eficiente tal que dado un sistema $f \in \mathbb{C}[X_1, \ldots, X_n]^s$, decida si $V_{\mathbb{C}^n}(f)$ es vacío o no.

Problema de Resolución: Diseñar un algoritmo eficiente tal que dado un sistema $f \in \mathbb{C}[X_1, \ldots, X_n]^s$, del cual conocemos que $V_{\mathbb{C}^n}(f)$ es no–vacío, aporte alguna información sobre $V_{\mathbb{C}^n}(f)$.

Ambos problemas tienen un carácter dual, como observaría el propio Kronecker, en el sentido de que los algoritmos de resolución pueden usarse (y, de hecho, se han usado eficazmente) para responder al Problema de Eliminación y viceversa. Aquí el término "eficiente" puede aún tomarse de manera difusa como el "trabajo" que requiere la ejecución de un algoritmo.

Ni los algoritmos imaginados por Hilbert, ni los diseñados por Kronecker serán algoritmos eficientes. Tampoco lo serán los que desarrollen sus continuadores entre finales del siglo XIX y el primer cuarto del siglo XX (Könz, Macaulay, Hermann), y todos estos problemas caen en el olvido histórico: Se consideran problemas imposibles e impracticables.

Estado del arte

A mediados del Siglo XX, la nueva potencia de cálculo proporcionada por los ordenadores y el desarrollo de nuevas herramientas algebraicas vuelve a convertir la resolución eficiente de sistemas de ecuaciones polinomiales en uno de los retos más importantes de las Matemáticas Computacionales. Son dos las ramas de la computación científica que se enfrentan a él: Una intenta aproximarse al problema desde la perspectiva simbólico–algebraica, y otra desde la perspectiva del análisis numérico. El objetivo es diseñar y analizar algoritmos óptimos, que resuelvan los problemas dados utilizando la menor cantidad de recursos posible. A continuación resumiremos brevemente los logros que estas ciencias ya han alcanzado.

Como hemos indicado ya, Galois había desarrollado un algoritmo pero la complejidad computacional (o sea, los recursos computacionales) requerida por su algoritmo era excesiva. Simplemente, aunque sabía cómo hacer algo, no tenía la capacidad de hacerlo en la práctica. Se encontraba así con una de las materias de estudio centrales en Complejidad Computacional: La *intratabilidad*. En la época en que vivió Galois, no estaban bien establecidas nociones como algoritmo o complejidad. Este paso, tan importante en la historia de las Matemáticas, vino de la mano de científicos como Gödel, Church y Turing. Sus trabajos culminaron en la creación de los primeros ordenadores, y la noción de Máquina de Turing se convirtió en la medida estándar de complejidad computacional.

La intratabilidad es uno de los aspectos más llamativos de los estudios de complejidad. Un problema matemático es intratable si los recursos computacionales que requiere su resolución son tan grandes que no hay posibilidad de resolverlo en la práctica. Además, la intratabilidad es independiente del algoritmo utilizado, pues es intrínseca al problema. Por ejemplo, los problemas matemáticos que requieren tiempo exponencial (en el tamaño del input) para ser resueltos son por naturaleza intratables: No hay esperanza de resolverlos eficientemente en ningún ordenador, presente o futuro, mediante ningún algoritmo, conocido o por descubrir.

Normalmente, nos referimos a los problemas cuyo tiempo de ejecución es polinomial en el tamaño del input como problemas tratables. Hay una gran cantidad de problemas para la cual carecemos de información suficiente para decidir si son o no tratables. Esa "tierra de nadie" suele llamarse la Frontera de la Intratabilidad (cf. [49]). El Nullstellensatz de Hilbert pertenece a esa frontera. Esto significa, simplemente, que nadie ha diseñado nunca un algoritmo que lo resuelva en tiempo polinomial en el número de variables, pero tampoco nadie ha demostrado que ese algoritmo no pueda existir.
Las estrategias para estudiar la complejidad computacional del Nullstellensatz de Hilbert se pueden dividir en dos grandes grupos: Sintácticas y semánticas. Resumimos brevemente a continuación ambas perspectivas.
Las estrategias sintácticas se caracterizan por el hecho de que los polinomios son considerados como listas de coeficientes (en codificación densa) de ciertos espacios vectoriales de alta dimensión. Entonces se trabaja con ellos como si fueran vectores, y se aplican métodos de Álgebra Lineal para obtener respuestas. Históricamente, el primer algoritmo sintáctico se remonta a Hilbert y su estudiante Hermann (cf. [70]). Estos autores redujeron el problema de decisión a la consistencia de un sistema de ecuaciones lineales, del modo que sigue. El Teorema de los Ceros de Hilbert (cf. [72]) establece que dada una lista de polinomios $f_1, \ldots, f_s \in \mathbb{C}[X_1, \ldots, X_n]$ de grado a lo más d, la variedad algebraica compleja que definen, $V_{\mathbb{C}^n}(f_1, \ldots, f_s) \subseteq \mathbb{C}^n$, es vacía si y solo si existen polinomios $g_1, \ldots, g_s \in \mathbb{C}[X_1, \ldots, X_n]$ tales que:

$$1 = g_1 f_1 + \cdots + g_s f_s. \tag{1}$$

Las igualdades como (1) se suelen llamar *identidades de Bézout*. Por el trabajo de Hermann, sabemos que existe una función $D(d, n)$ que depende solo del número de variables y del máximo de los grados, tal que, si existe alguna expresión de la forma (1), entonces existe otra expresión equivalente donde los grados de los polinomios g_i están acotados por $D(d, n)$. Así, la existencia de estos polinomios g_i (y, por tanto, la consistencia del sistema de ecuaciones) es equivalente a la consistencia de un sistema lineal con

$$s\binom{D(d,n)+n}{n}.$$

incógnitas y ecuaciones. El tiempo que tardamos en decidir este sistema es, por supuesto, polinomial en esta cantidad. Por lo tanto, cotas finas de la función $D(d, n)$ implican cotas finas superiores de complejidad para la resolución del Nullstellensatz mediante estas técnicas. Los estudios sobre cotas superiores de esa función suelen llamarse Nullstellensatz efectivo (véanse [16, 19, 81, 82, 62, 83] y las referencias en ellos incluidas). Las cotas conocidas sobre $D(d, n)$ se pueden resumir como sigue,

$$d^{n-1} \leq D(d, n) \leq d^n.$$

Concluimos que este intento de resolución no es eficiente ni aplicable pues la complejidad obtenida es del orden de

$$\binom{d^n + n}{n} \approx d^{n^2}.$$

Otra de las estrategias sintácticas para el problema de Hilbert consiste en el uso de técnicas de reescritura. De entre ellas, la más conocida y utilizada es la de las bases de Gröbner. Desde los estudios iniciales [73] y [17], una larga lista de trabajos han analizado este tipo de algoritmos (véanse por ejemplo [8, 25, 94, 128] y las referencias en ellos contenidas). Las bases de Gröbner están implementadas en la mayoría de los programas matemáticos en el mercado. La programación más eficiente hasta la fecha se debe a Faugère (las series *FGb*). En términos de complejidad computacional, se conocen cotas inferiores que son desalentadoras: Mayr y Meyer, en su trabajo [92], demostraron que el problema de calcular bases de Gröbner requiere unos recursos muy elevados, y necesariamente tiempo exponencial (en realidad, el resultado obtenido por Mayr y Meyer es mucho más fuerte, pues demuestra que la resolución por bases de Gröbner es EXPSPACE–completo). De nuevo la complejidad excesiva parece suponer una barrera en el problema de Hilbert.

Una tercera estrategia sintáctica utiliza los conceptos de Complejidad Estructural. Esto es, los problemas se clasifican en distintas clases de complejidad, y el estudio de la complejidad de un problema consiste en colocarlo en la clase apropiada donde este problema es completo. En [15], Blum, Shub & Smale demostraron que HN es completo en la clase $\mathrm{NP}_{\mathbb{C}}$, esto es, tiempo polinomial no–determinista en el modelo abstracto de Máquinas de Turing complejas (véase [14]). Otros autores han estudiado la complejidad de HN en el contexto más realista de Máquinas de Turing clásicas. En [80] (véanse también [106, 105]), el autor demuestra que HN está en la clase de complejidad PH (Polynomial Hierarchy).

A medio camino entre las estrategias sintácticas y semánticas están los intentos de atacar el caso sparse, esto es, cuando el número de monomios no–nulos es pequeño. En el libro de Sturmfels [125], se encuentran algunas técnicas específicamente dedicadas a tratar este caso (véase también [90, 91]). Ninguna de estas estrategias logra bajar la barrera de la exponencialidad en el número de variables.

Las estrategias semánticas tratan de utilizar las particulares propiedades de los objetos geométricos relacionados con la eliminación (esto es, las variedades algebraicas). Los polinomios son entonces vistos como aplicaciones, y la complejidad trata de relacionarse con parámetros intrínsecos de los problemas (invariantes geométricos, aritméticos y topológicos), estudiados durante años por los geómetras y algebristas. Estas ideas se han utilizado tanto para cotas superiores de complejidad (como veremos) como para cotas inferiores, véase por ejemplo [93] y las referencias en él incluidas. El objetivo

de los trabajos de cotas superiores de complejidad fue diseñar un algoritmo cuya complejidad fuese polinomial en alguno de los invariantes semánticos del problema. El grupo de trabajo TERA alcanzó este objetivo durante la década de los 90. Dos trabajos iniciaron esta corriente de pensamiento: En [100] se proponían las bases para un programa de investigación al respecto; en [52] se mostraba el primer algoritmo semántico para Teoría de la Eliminación. El programa propuesto en [100] fue completado en la serie de artículos [51, 50, 53]. Resumimos brevemente los logros de este proyecto. Primero, reformulamos el Nullstellensatz de Hilbert en los términos siguientes:

Problema 1 *Diseñar un algoritmo eficiente que realiza la siguiente función: Dada una lista de polinomios* $f_1, \ldots, f_s, g \in \mathbb{C}[X_1, \ldots, X_n]$ *de grados acotados por* d*, decidir si el polinomio* g *se anula en algún punto de la variedad* $V_{\mathbb{C}^n}(f_1, \ldots, f_s) \subseteq \mathbb{C}^n$*.*

Ésta es la forma usual en que se presentan los problemas de eliminación en Teoría de la Eliminación. También es la forma en que se suelen plantear los problemas NP–completos (véanse [69, 100] y sus referencias). El motivo por el que escribimos el problema de este modo es para resaltar el papel distinto que tiene el polinomio g. A partir de una lista de polinomios $f_1, \ldots, f_s$ queremos computar información sobre su variedad de ceros comunes $V_{\mathbb{C}^n}$. Esta información, que es lo que solemos llamar solución del sistema, ha de permitirnos responder futuras respuestas sobre la variedad. Como la pregunta de si un polinomio dado g se anula o no en algún punto de $V_{\mathbb{C}^n}$. En [100], se propone la creación de un algoritmo que satisfaga las dos siguientes condiciones siguientes:

- Su complejidad debe estar acotada por parámetros intrínsecos (semánticos) relacionados con la lista $f_1, \ldots, f_s$.
- Su output debe contener suficiente información para responder cualquier pregunta de eliminación como las planteadas en el Problema 1.

No es nuestro propósito describir con detalle las características de este algoritmo. Nos basta con señalar que el procedimiento obtenido tiene complejidad polinomial en una cantidad intrínseca al sistema: El máximo de los grados de las variedades intermedias (en el sentido de [66]). Ésta cantidad es, en casi todos los casos, igual al número de Bézout, esto es, el producto de los grados de los polinomios, que es típicamente una cantidad exponencial en el número de incógnitas. Pese a las mejoras conseguidas por [57, 54, 68], esta barrera no ha sido superada. Algunas de estas ideas fueron trasladadas al problema de conteo de soluciones reales de sistemas con coeficientes reales en [3, 4, 5, 6]. El algoritmo fue implementado por Lecerf y Salvy. Esta implementación, incluyendo algunas variaciones técnicas, fue presentada en [54]. El algoritmo semántico al que nos referimos supone un

avance categórico con respecto a las demás técnicas simbólicas que hemos descrito (muy brevemente) en párrafos anteriores: Su complejidad pasa a depender no de la forma concreta en que se escriban los polinomios del input, sino de parámetros intrínsecos a la variedad que queremos describir. La nueva cota de complejidad aún es, no obstante, exponencial en el número de incógnitas en la mayoría de los casos.
La experiencia TERA y el comportamiento del algoritmo Kronecker conducen a dos preguntas centrales:

- ¿Es el número de Bézout $\mathcal{D}$ una barrera para la complejidad de resolución de los sistemas de ecuaciones polinomiales?

- En caso afirmativo, ¿qué significado tiene la existencia de dicha barrera?

Un intento de responder a estas preguntas viene de la mano del concepto de resolución universal y de los resultados expuestos en [67, 101, 20]. En pocas palabras, un algoritmo universal es un procedimiento que permite responder a todas las preguntas de eliminación que conciernen a la variedad solución. Todos los algoritmos conocidos en Teoría de la Eliminación (tanto sintácticos como semánticos) son universales. En [20] se demuestra que *todo procedimiento universal requiere tiempo exponencial en el número de variables*. Esto es, si queremos obtener, a partir de un sistema, toda la información necesaria para responder a las preguntas de Teoría de la Eliminación, tenemos que estar dispuestos a pagar un alto coste computacional.
El concepto de resolución universal también aparece en algunos algoritmos numéricos. Por ejemplo, los algoritmos implementados por J. Verschelde y sus colaboradores (véase por ejemplo [129]), que pretenden aproximar *todas* las soluciones de un sistema de ecuaciones dado, son algoritmos universales y necesariamente funcionarán en tiempo exponencial. En efecto, el número de soluciones es genéricamente igual al número de Bézout, luego aproximarlas todas exigirá una complejidad, al menos, del orden de $\mathcal{D}$.
El hecho de que los procedimientos universales requieran una complejidad exponencial en el número de incógnitas no tiene por qué ser visto como un resultado negativo. Simplemente, nos hace darnos cuenta de la necesidad de buscar algoritmos *no–universales* de resolución de sistemas de ecuaciones polinomiales. Debemos rebajar el nivel de exigencia para obtener un tiempo de ejecución más razonable. Un algoritmo no–universal devuelve información parcial sobre el conjunto de soluciones de un sistema. Esta idea conlleva la aparición de una larga serie de nuevos problemas y preguntas. La primera de ellas es, sin embargo, muy difícil de responder: ¿Qué clase de información sobre la variedad solución tenemos a partir de la información parcial calculada por un algoritmo no–universal? Es necesaria mucha más experiencia con algoritmos no–universales antes de responder a esta pregunta.

Un ejemplo inicial de algoritmo semántico no–universal fue dado en [99]. Sin embargo, el tiempo de ejecución de este algoritmo vuelve a ser exponencial en el número de variables. La búsqueda de algoritmos de resolución no–universales conduce de modo natural al desarrollo de algoritmos numéricos.

Durante la primera mitad de los años 90, M. Shub y S. Smale sentaron las bases para una nueva concepción del análisis numérico. Se centraron en el problema de la resolución numérica de sistemas de ecuaciones polinomiales en la serie de artículos [112, 113, 114, 115, 116]. Otros muchos autores siguieron esa línea en trabajos como [14, 32, 33, 30, 34, 88, 89, 90, 91, 87, 134, 29, 138, 137, 136, 135, 2]. Algunos otros trabajos habían precedido la serie de artículos citada de Shub & Smale (véanse [117, 77, 102, 78, 118, 111])

La teoría de *ceros aproximados* de Shub & Smale busca desde sus orígenes la posibilidad de evitar el crecimiento exponencial de la complejidad de resolución. En pocas palabras, un cero aproximado de un sistema de ecuaciones es un punto que satisface la siguiente condición: Las sucesivas iteraciones del operador de Newton, iniciando en el cero aproximado, convergen a velocidad doblemente exponencial hacia una solución exacta del sistema. La información contenida en un cero aproximado no es únicamente una aproximación de la solución: Como se demuestra en [22], se puede obtener a partir de ella información simbólica de las soluciones del sistema, aunque en la presente memoria no nos ocuparemos de esta relación simbólico–numérica.

La información no–universal que buscamos ahora es la que está contenida en un cero aproximado del sistema objetivo. Esto es, el diseño de algoritmos que aproximen algún punto en la variedad solución consiste en la búsqueda de procedimientos de resolución no–universales que proporcionan la información contenida en un cero aproximado del sistema. Este tipo de algoritmos no busca por tanto responder a todas las preguntas de la Eliminación y por ello son susceptibles de poder evitar la cota exponencial de complejidad.

En todo caso, el concepto numérico de resolución puede enunciarse como sigue,

Problema de Resolución Numérica: Dado un sistema de ecuaciones polinomiales $f = [f_1, \ldots, f_n] \in \mathbb{C}[X_1, \ldots, X_n]^n$ definiendo una variedad $V_{\mathbb{C}^n}(f)$ de dimensión cero, hallar un cero aproximado para algún cero $\zeta \in V_{\mathbb{C}^n}(f)$.

Shub & Smale propusieron un método de homotopía para la búsqueda de ceros aproximados. Este método funciona, a grandes rasgos, como sigue. Sea f un sistema que queremos resolver, y sea g otro sistema que *sabemos* resolver. Sea z_0 la solución de g que ya conocemos. Consideremos una partición $\{h_i\}_{i=0\ldots k}$ del segmento entre g y f, de forma que $h_0 = g$ y $h_k = f$. Entonces, construimos una serie de puntos $\{z_i\}_{i=1\ldots k}$ como sigue: Para cada $i = 1 \ldots k$, z_i es el resultado de aplicar el operador de Newton correspondi-

ente a h_i con punto inicial z_{i-1}. El procedimiento tiene éxito si z_k es un cero aproximado de $h_k = f$. El número natural $k > 0$ suele llamarse el número de pasos de homotopía y su control es el elemento clave del algoritmo, pues de él dependen tanto la complejidad como la garantía (probabilista) de que z_k sea realmente un cero aproximado de f.
En su serie de profundos y novedosos artículos [112, 113, 115, 116], Shub & Smale demostraron que el procedimiento que acabamos de describir puede ser cuantitativamente estudiado y analizado. Es más, culminaron su estudio con un impresionante resultado que se puede escribir como sigue:

Teorema 0.0.1 (Shub & Smale, [115]) *Consideremos fijados una lista de grados $(d_1, \ldots, d_n)$ y un número positivo $\varepsilon > 0$. Entonces, existe un sistema polinomial g_ε y una solución ζ_ε de ese sistema tal que el método de homotopía encuentra un cero aproximado de casi todo sistema de ecuaciones $f = [f_1, \ldots, f_n]$, $\deg(f_i) = d_i$, (con probabilidad de éxito $1-\varepsilon$), en un número de pasos de homotopía del orden de $O(N^4\varepsilon^{-1})$.*

Aquí, como en el resto de esta Introducción, N es la talla del input (esto es, el número de coeficientes en codificación densa). El anterior resultado representa un avance sin precedentes en el estudio de los sistemas de ecuaciones: Por primera vez se insinúa que es posible encontrar algorítmicamente, en tiempo polinomial, información precisa sobre el conjunto solución de un sistema de ecuaciones. Nos dice que, si somos capaces de encontrar el par inicial $(g_\varepsilon, \zeta_\varepsilon)$, entonces podremos calcular ceros aproximados de sistemas en tiempo polinomial *con alta probabilidad*. Sin embargo, el resultado de Shub & Smale no puede transformarse en un algoritmo explícito. El motivo es que se trata puramente de un resultado existencial: No se da ninguna pista sobre cómo encontrar o calcular el par $(g_\varepsilon, \zeta_\varepsilon)$. Shub & Smale demostraron la *existencia* de un buen par inicial, pero no podemos saber cuál es. La falta de esta información explícita llevó a estos autores al enunciado de su Conjetura (véase [115]) en la cual proponen un sistema inicial determinado, y una solución de dicho sistema, como candidatos a satisfacer la tesis del teorema anterior. Sin embargo, nadie ha podido demostrar la verdad o la falsedad de esa conjetura a día de hoy.
Así pues, la pregunta sobre cómo resolver sistemas de ecuaciones (i.e. cómo calcular ceros aproximados de estos sistemas) se reduce a la búsqueda de un buen "par inicial" (g, z_0) que permita desarrollar la homotopía descrita por Shub & Smale con garantías de funcionamiento. Stephen Smale, ganador de la Medalla Fields en 1966, introdujo esta búsqueda en su lista de problemas matemáticos para el siglo XXI (cf. [119]).

Problema 17 de Smale:
Can a zero of n complex polynomial equations in n unknowns be found approximately, on the average, in polynomial time with a uniform algorithm?

En otras palabras,

Existe un algoritmo que calcule ceros aproximados de sistemas de ecuaciones polinomiales en tiempo polinomial en promedio?

En este enunciado, "uniform algorithm" quiere decir simplemente que no dependamos de información a priori, esto es, que no necesitemos un "oráculo" que nos proporcione información sobre el par (g, z_0) inicial. Smale pide además un resultado de naturaleza probabilista. Esto no es un hecho casual: Una filosofía común a muchos procedimientos del análisis numérico consiste en demostrar que, asumiendo una cierta probabilidad (pequeña) de fracaso, podemos encontrar soluciones de la mayor parte de los problemas en tiempo polinomial. Esto es, sostenemos la tesis siguiente:

Paradigma de algoritmo numérico: Dado un espacio de inputs $\mathcal{I}$ con una distribución de probabilidad fijada, un algoritmo numérico con función de recursos $\varphi : \mathbb{R}^2 \longrightarrow \mathbb{R}$ es un algoritmo $\mathcal{P}$ con espacio de inputs $\mathcal{I} \times (0, 1)$ tal que se tiene:

- Para cada $\varepsilon > 0$, la probabilidad de que un input $i \in \mathcal{I}$ sea resuelto por $\mathcal{P}\,|_{\mathcal{I}\times\{\varepsilon\}}$ es al menos $1 - \varepsilon$.
- Para cada $\varepsilon > 0$, y para cada input $i \in \mathcal{I}$ el algoritmo $\mathcal{P}\,|_{\mathcal{I}\times\{\varepsilon\}}$ se detiene después de $\varphi(|i|, \varepsilon^{-1})$ pasos, donde $|i|$ es el tamaño de i.

La función de recursos φ indica la clase de complejidad a la que pertenece el algoritmo en cuestión. Así, si φ está acotada por algún polinomio $f(X, Y)$, diremos que el algoritmo es polinomial. También podemos hablar, como en el caso de Teoría de Complejidad Clásica, de algoritmos no–deterministas, probabilistas con probabilidad de error acotada, etc.

Principales aportaciones de esta memoria

Estructuramos esta sección en varias subsecciones, cada una de las cuales contiene resultados de una naturaleza, aunque todos están profundamente relacionados entre sí.

Una solución probabilista al Problema 17 de Smale: El caso homogéneo

El resultado más sugerente de esta memoria consiste en una respuesta probabilista positiva al Problema 17 de Smale, tanto para el caso homogéneo como para el caso afín. Consideremos primero sistemas de ecuaciones $f = [f_1, \ldots, f_n]$ tales que para todo $i, 1 \leq i \leq n$, el polinomio f_i es homogéneo de grado d_i en las variables $X_0, \ldots, X_n$. Primero observamos que, en estas circunstancias, un punto $x \in \mathbb{C}^{n+1}$ es solución de f si y solo si lo es λx, para

todo $\lambda \in \mathbb{C},\ \lambda \neq 0$. Por ello, buscamos soluciones proyectivas del sistema f (esto es, puntos proyectivos $x \in \mathbb{P}_n(\mathbb{C})$ en los que se anulan todos los f_i, $1 \leq i \leq n$). El conjunto de soluciones proyectivas de f, denotado $V(f)$, es genéricamente un conjunto discreto. Hay una noción estable de cero aproximado proyectivo (descrita por primera vez en [110]), que podemos enunciar brevemente como sigue: Un punto proyectivo $z \in \mathbb{P}_n(\mathbb{C})$ es un cero aproximado de f con cero (exacto) asociado $\zeta \in V(f)$ si la secuencia de iteraciones del operador de Newton proyectivo está bien definida y converge cuadráticamente (esto es, a velocidad doblemente exponencial) al cero exacto ζ, donde la convergencia se mide en términos de la distancia tangente d_T, definida como la tangente de la distancia Riemanniana en $\mathbb{P}_n(\mathbb{C})$. La distancia tangente viene a medir la tangente del ángulo entre dos puntos proyectivos. Si la distancia d_T entre dos puntos proyectivos es pequeña, el ángulo que los separa es también pequeño. Además, si tenemos un cero aproximado proyectivo, obtener una ε–aproximación del cero exacto (en términos de d_T) requiere tan sólo

$$O(\log|\log\varepsilon|)$$

iteraciones de Newton. El objetivo que tenemos es la búsqueda de esos ceros aproximados. Entonces, demostraremos el siguiente resultado (cf. [11, 12]).

Teorema 0.0.2 (Solución Probabilista al Problema 17 de Smale) *Sea $\varepsilon > 0$ un número real positivo, representando una probabilidad de fracaso. Entonces, existe un algoritmo que encuentra un cero aproximado proyectivo de todo sistema de ecuaciones homogéneo (salvo por esa probabilidad de fracaso ε), realizando un número de operaciones aritméticas $O(N^5\varepsilon^{-2})$, donde N es la talla del input.*

En este resultado, como en los que se expondrán a continuación, hablaremos de "probabilidad" en los mismos términos en que Smale establece el enunciado de su problema. Expliquemos este término con más detalle. El conjunto de los sistemas de ecuaciones de grado fijado tiene una estructura de espacio vectorial. Se suele dotar a dicho espacio con un producto Hermitiano $\langle\cdot,\cdot\rangle_\Delta$ asociado a la lista de grados, que tiene varias características deseables (véase la Sección 1.4 para una descripción detallada). La más importante, a la que nos referiremos como "invariancia unitaria", es la siguiente: Si consideramos una transformación unitaria en el espacio de soluciones, el cambio de variables correspondiente define una transformación isométrica en el espacio de sistemas. La estructura de espacio de Hilbert así definida en el espacio de sistemas fue llamada por Shub & Smale estructura de Kostlan (y nosotros seguiremos su notación en este punto), aunque es conocida y utilizada por los geómetras desde principios del siglo XX.
Una segunda observación es que el conjunto de soluciones de un sistema no varía si multiplicamos dicho sistema por un escalar no–nulo. Por este motivo, las probabilidades y esperanzas relacionadas con el conjunto de soluciones de

un sistema se suelen medir en la esfera (para el producto Hermitiano $\langle \cdot, \cdot \rangle_\Delta$) con la estructura de variedad Riemanniana heredada de la del espacio de sistemas. Otra forma de medir probabilidades, totalmente equivalente a ésta, es medirlas en el espacio proyectivo asociado al espacio vectorial de sistemas, con la estructura de variedad Riemanniana heredada de la de éste. Así pues, cuando hablemos de probabilidades y esperanzas en el espacio de sistemas nos referiremos siempre a probabilidades y esperanzas en la esfera de los sistemas (o, equivalentemente, en el espacio proyectivo complejo asociado), con la estructura correspondiente heredada del producto Hermitiano $\langle \cdot, \cdot \rangle_\Delta$.
El algoritmo del teorema anterior se obtiene mediante la búsqueda de pares iniciales buenos para la homotopía de Newton, como ha sido descrita en la sección anterior. Podemos dar una noción general como sigue:
Un esquema de deformación homotópica de Newton (que abreviaremos como NHD) con par inicial $(g, z_0) \in \mathcal{H}_{(d)} \times \mathbb{P}_n(\mathbb{C})$ y función de recursos $\varphi : \mathcal{H}_{(d)} \times \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ es un esquema algorítmico basado en la siguiente estrategia:

Input: $f \in \mathcal{H}_{(d)}$, $\varepsilon \in \mathbb{R}^+$.

- *Realizar* $\varphi(f, \varepsilon)$ *"pasos de homotopía" siguiendo el segmento* $(1-t)g + tf$, $t \in [0, 1]$, *comenzando en* (g, z_0), *donde* z_0 *es un cero aproximado proyectivo de* g *asociado con algún cero* $\zeta_0 \in V(g)$.

Output:

O bien *failure*, o bien
un cero aproximado proyectivo $z_1 \in \mathbb{P}_n(\mathbb{C})$ de f.

Un algoritmo basado en NHD es un algoritmo que construye una poligonal con $\varphi(f, \varepsilon)$ vértices, de modo que el vértice inicial es (g, z_0) y el vértice final es el punto (f, z_1) para algún $z_1 \in \mathbb{P}_n(\mathbb{C})$, siendo z_1 el output del algoritmo. La poligonal es construida por "pasos de homotopía" que van de un vértice al siguiente. Por lo tanto, $\varphi(f, \varepsilon)$ es el número de pasos de homotopía realizados por el algoritmo. Hay distintos modos de realizar esos pasos de homotopía, entre los cuales se encuentra el operador proyectivo de Newton, como ha sido descrito en [110, 112, 88].
El número real ε se utiliza normalmente para controlar el número de pasos (a través de la función $\varphi(f, \varepsilon)$) y la probabilidad de fracaso del algoritmo (esto es, la probabilidad de que un input $f \in \mathcal{H}_{(d)}$ no sea resuelto en $\varphi(f, \varepsilon)$ pasos con par inicial (g, z_0)).
Una buena elección del par inicial (g, z_0) debe garantizar que el número de pasos $\varphi(f, \varepsilon)$ será razonablemente pequeño con alta probabilidad. Estableceremos este concepto en la siguiente definición, inspirada en el paradigma de algoritmo numérico antes mencionado.

Definición 0.0.3 *Sea* $\varepsilon > 0$ *un número real positivo. Decimos que un par inicial* (g, ζ), *donde* g *es un sistema homogéneo y* ζ *una solución de* g, *es*

ε–eficiente para NHD *si el esquema* NHD *con par inicial* (g, ζ) *y función de recursos*

$$\varphi(f, \varepsilon) := 10^8 n^5 N^3 d^4 \varepsilon^{-2}, \ \ \forall\, f \in \mathbb{S}_\Delta,\ \varepsilon > 0$$

satisface la siguiente propiedad:

$$\text{Prob}[f \textit{ sistema}: \text{ NHD } \textit{encuentra un cero}$$

$$\textit{aproximado proyectivo de } f\,] \geq 1 - \varepsilon,$$

donde Prob *significa probabilidad. Aquí,* N *es la talla del input, y* d *es el máximo de los grados de los polinomios de* f*.*

En esta definición hemos elegido una función de recursos concreta, para facilitar la lectura y comprensión de las ideas, pero dicha función podría ser sustituida por cualquier otro polinomio. Obsérvese que, una vez conocido un par ε–eficiente, para algún valor pequeño de ε, podemos resolver la mayoría de sistemas (probabilidad de éxito $1 - \varepsilon$) en un tiempo polinomial en el tamaño del input.
En estos términos, el resultado Teorema 0.0.1 de Shub & Smale se lee como sigue: Para cada $\varepsilon > 0$, existe un par $(g_\varepsilon, \zeta_\varepsilon)$ que es ε–eficiente. La debilidad principal de este enunciado es, como hemos indicado, la falta de pistas sobre cómo se construye el par $(g_\varepsilon, \zeta_\varepsilon)$. Evidentemente, sin un modo de calcular este par inicial, el esquema que acabamos de escribir no puede ser llevado a la práctica. De hecho, no podemos garantizar que sea un algoritmo, pues ni siquiera sabemos si g_ε es calculable o si ζ_ε es computable. Otro aspecto mejorable es la dependencia del par $(g_\varepsilon, \zeta_\varepsilon)$ del valor ε.
A lo largo de esta memoria (véase el Capítulo 4) mostraremos una solución a estos problemas. El método que proponemos es probabilístico y no asume ningún conocimiento a priori, y es capaz de resolver la mayor parte de los sistemas de ecuaciones polinomiales. Comenzamos con la siguiente noción.

Definición 0.0.4 *Una clase* $\mathcal{G} \subseteq \mathcal{H}_{(d)} \times \mathbb{P}_n(\mathbb{C})$ *se llama conjunto questor para pares eficientes si para todo número real* $\varepsilon > 0$*, la probabilidad de que un par elegido al azar* $(g, \zeta) \in \mathcal{G}$ *sea* ε*–eficiente es mayor o igual que*

$$1 - \varepsilon.$$

Entonces, demostraremos el siguiente resultado.

Teorema 0.0.5 *Para toda lista de grados* $(d) = (d_1, \ldots, d_n)$*, existe un conjunto questor para pares eficientes,* $\mathcal{G}_{(d)}$*, que resuelve la mayor parte de los sistemas en* $\mathcal{H}_{(d)}$ *en tiempo polinomial en el tamaño del input* N*.*

La existencia de un conjunto questor $\mathcal{G}_{(d)} \subseteq \mathcal{H}_{(d)} \times \mathbb{P}_n(\mathbb{C})$ como el que acabamos de describir proporciona otra variación, de naturaleza probabilista, de los algoritmos basados en NHD. En efecto, fijemos un conjunto

questor $\mathcal{G}_{(d)}$ (que no depende del $\varepsilon > 0$ que elijamos). Entonces, tenemos el siguiente esquema basado en NHD.

Input: $f \in \mathcal{H}_{(d)}$, $\varepsilon \in \mathbb{R}^+$.

- *Elegir al azar* $(g, \zeta) \in \mathcal{G}_{(d)}$.
- *Realizar un número polinomial (en* $\varepsilon^{-1}, n, N, d$*) de pasos de homotopía siguiendo el segmento* $(1-t)g + tf$, $t \in [0,1]$, *empezando en* (g, ζ).

Output:

o bien failure, *o bien*
un cero aproximado proyectivo $z \in \mathbb{P}_n(\mathbb{C})$ *de* f.

De todas maneras, un resultado existencial como el Teorema 4.1.5 no proporciona una solución al problema principal de [115], pues seguiríamos siendo incapaces de diseñar un algoritmo concreto. Por ello, mostraremos una clase tratable algorítmicamente $\mathcal{G}_{(d)}$, y demostraremos que es un conjunto questor para pares eficientes. La exposición concreta de la clase que proponemos requiere algunas notaciones previas, por lo que emplazamos al lector a leerla en la introducción del Capítulo 4.
La existencia de este conjunto questor para pares eficientes $\mathcal{G}_{(d)}$, garantiza que, dada una lista de grados (d) y un real positivo ε, es posible encontrar, mediante un sencillo método probabilístico, un par ε–eficiente. A continuación, exponemos un resultado que garantiza que podemos encontrar pares que sean ε–eficientes *para todo* $\varepsilon > 0$ (véase el Teorema 4.1.8).

Teorema 0.0.6 *Para un par elegido al azar* $(g, \zeta) \in \mathcal{G}_{(d)}$, *la siguiente propiedad se satisface con probabilidad al menos* $3/4$*:* (g, ζ) *es* ε*–eficiente para todo* ε, $0 < \varepsilon < 1/2$.

Como consecuencia, queda totalmente eliminada la dependencia del par inicial del valor de ε. Podemos resumir el resultado obtenido del modo que sigue,

Teorema 0.0.7 *Fijados una lista de grados y un número de incógnitas, existe un sistema polinomial* g *y una solución* ζ *de ese sistema tal que para todo* $\varepsilon > 0$, *se tiene:*

$$\text{Prob}[f \text{ sistema homogéneo}: \text{ La homotopía comenzando en } (g, \zeta)$$

$$\text{con } cN^5\varepsilon^{-2} \text{ pasos encuentra un cero aprox. proyectivo de } f] \geq 1 - \varepsilon,$$

donde N *es la talla del input y* $c > 0$ *es una constante universal. Además, el par* (g, ζ) *puede ser encontrado mediante un sencillo método probabilístico.*

En otras palabras, podemos encontrar mediante un método probabilístico un "buen par inicial". Y una vez que tenemos un ese par (g, ζ), *casi todo* sistema f (probabilidad de fracaso ε) es resuelto por el método de homotopía con el par inicial (g, ζ) en un número de pasos polinomial en la talla del input y en ε^{-1}. Por tanto, nos ajustamos al paradigma de algoritmo numérico que sugeríamos antes. En los capítulos 4 y 5 escribiremos los detalles sobre cómo se elige el par (g, ζ) y cuál es la potencialidad del resultado.
Este enunciado se entiende fácilmente si elegimos un $\varepsilon > 0$ concreto. Por ejemplo, si elegimos $\varepsilon = N^{-1}$, tenemos que con probabilidad de éxito al menos $1 - N^{-1}$, encontramos ceros aproximados proyectivos de sistemas en tiempo $O(N^7)$, donde N es el tamaño del input. Como ejemplo de aplicación, si tratamos con sistemas cúbicos en n incógnitas, el algoritmo de homotopía calcula un cero aproximado de casi todos los sistemas cúbicos (probabilidad de éxito $1 - \frac{1}{n^2}$) realizando un número de operaciones aritméticas del orden de $O(n^{26})$. Ésta es la primera vez que se describe un algoritmo que calcule ceros aproximados de sistemas cúbicos en tiempo polinomial en el número de incógnitas.

Una solución probabilista al Problema 17 de Smale: El caso afín

El caso homogéneo es el más ampliamente estudiado por Shub & Smale y es también en el que desarrollamos las herramientas técnicas más importantes de resolución numérica. Sin embargo, obviamente el caso no–homogéneo tiene también un gran interés, y se enuncia de hecho en términos más clásicos: Dado un sistema de ecuaciones $f = [f_1, \ldots, f_n]$ en las variables $X_1, \ldots, X_n$, encontrar un cero aproximado (afín) de f. Esto es, se trata del enunciado denominado *Problema de Resolución Numérica* hace algunos párrafos. Denotemos por $V_{\mathbb{C}^n}(f)$ el conjunto de soluciones afines de un sistema f. Un cero aproximado afín de un sistema $f \in \mathbb{C}[X_1, \ldots, X_n]^n$ con cero (exacto) asociado $\zeta \in \mathbb{C}^n$ es un punto $z \in \mathbb{C}^n$ tal que las sucesivas iteraciones del operador de Newton afín están bien definidas y convergen (en términos de la norma usual $\|\cdot\|_2$) a velocidad doblemente exponencial hacia el cero ζ. Esto es, a partir de un cero aproximado podemos obtener aproximaciones (en términos de la distancia euclídea) del cero exacto muy rápidamente.
El salto de la resolución de sistemas homogéneos a la resolución de sistemas afines no es elemental: Este proceso requiere controlar la distribución de probabilidad de la norma de las soluciones afines de los sistemas de ecuaciones. En efecto, podemos enunciar en términos sencillos el nuevo problema como sigue: Sea f el sistema que queremos resolver. Consideremos un cero aproximado proyectivo z' del sistema homogéneo asociado a f (obtenido igualando el grado de todos los monomios en cada polinomio, al multiplicarlo por una incógnita extra X_0 con el grado apropiado). Si $\zeta := (\zeta_0 : \cdots : \zeta_n)$ es una solución (proyectiva) del sistema homogéneo asociado a f, con $\zeta_0 \neq 0$,

entonces

$$\varphi_0^{-1}(\zeta) := \left(\frac{\zeta_1}{\zeta_0}, \ldots, \frac{\zeta_n}{\zeta_0}\right) \in V_{\mathbb{C}^n}(f)$$

es una solución afín de f. Una estrategia inicial para encontrar un cero aproximado afín de un sistema f puede ser la siguiente: Primero buscamos un cero aproximado proyectivo del sistema homogéneo asociado, $z' \in \mathbb{P}_n(\mathbb{C})$, y luego consideramos, si existe, el punto afín $z = \varphi_0^{-1}(z') \in \mathbb{C}^n$. Por supuesto, $\varphi_0^{-1}(z')$ no tiene por qué ser un cero aproximado afín de f. ¿Cómo podemos, a partir del conocimiento de z', obtener un cero aproximado afín z? Intuitivamente, dado que la distancia proyectiva es (esencialmente) el ángulo entre dos puntos, podemos esperar que la distancia euclídea $\|z - \zeta\|_2$ esté controlada por $d_T(z', \varphi_0^{-1}(\zeta))$ y por la norma de z, ζ. En el Capítulo 5, demostraremos que es así.
Con ello, el problema de obtener un cero aproximado afín a partir de un cero aproximado proyectivo se reduce, módulo algunos cálculos técnicos, al problema de controlar la norma (euclídea) de las soluciones de un sistema de ecuaciones. El resultado técnico principal del Capítulo 5 es una cota para la distribución de probabilidad de dicha norma. Éste es un resultado con interés geométrico y aritmético propio, y puede escribirse como lo enunciamos a continuación.

Teorema 0.0.8 *Sea $\delta > 0$ un número real, y consideremos fijada una lista de grados $(d_1, \ldots, d_n)$. La probabilidad de que un sistema $f \in \mathbb{C}[X_1, \ldots, X_n]^n$ elegido al azar tenga alguna solución afín $\zeta \in V_{\mathbb{C}^n}(f)$ con $\|\zeta\|_2 \geq \delta$ es a lo sumo*

$$\frac{\mathcal{D}\sqrt{\pi n}}{\delta},$$

donde $\mathcal{D} := \prod_{i=1}^n d_i$ es el número de Bézout asociado a la lista de grados.

Utilizando este resultado, en el Capítulo 5 extenderemos los resultados de la sección anterior al caso afín. Esto es, demostraremos el siguiente resultado:

Teorema 0.0.9 (Respuesta al Problema 17, caso afín) *Sea $\varepsilon > 0$ un número real positivo, representando una probabilidad de fracaso. Entonces, existe un algoritmo que encuentra un cero aproximado afín de todo sistema de ecuaciones polinomiales (salvo por esa probabilidad de fracaso ε), realizando un número de operaciones aritméticas $O(N^5\varepsilon^{-2})$, donde N es la talla del input.*

Sobre la distribución de probabilidad del condicionamiento lineal singular

En el proceso que conduce a la solución propuesta del Problema 17 de Smale, hay una pieza que juega un papel esencial: El número de condicionamiento para sistemas de ecuaciones, y su distribución de probabilidad. Para llegar

a comprender en su totalidad este concepto, conviene primero hablar del caso lineal. El número de condicionamiento juega un papel central en el análisis del error de los procesos numéricos que involucran cálculo matricial. El problema relativo a su distribución de probabilidad se puede expresar como sigue. Sea $\mathcal{P}$ un procedimiento de Análisis Numérico cuyo espacio de posibles inputs es un subconjunto $\mathcal{C} \subseteq \mathcal{M}_n(\mathbb{C})$ del espacio de matrices complejas. Esta situación llevará asociado un número de condicionamiento $\kappa_{\mathcal{C}}$ adaptado al procedimiento $\mathcal{P}$ y al espacio de inputs $\mathcal{C}$. La distribución de probabilidad de $\kappa_{\mathcal{C}}$ vendrá dada por el comportamiento de la cantidad

$$\frac{vol[\{A \in \mathcal{C} : \kappa_{\mathcal{C}}(A) > \varepsilon^{-1}\}]}{vol[\mathcal{C}]}, \tag{2}$$

donde $\varepsilon > 0$ recorre los números reales positivos, y *vol* representa alguna forma apropiada de medir volúmenes en $\mathcal{C}$. Hay varios trabajos dedicados a estudiar el caso de que $\mathcal{C}$ sea el espacio total de matrices, véanse por ejemplo [117, 103, 40, 42]. En el caso de matrices con coeficientes reales (que no es el que nos ocupa en esta memoria), destacamos los trabajos [26, 27] (compárese con [42, 117]).

En todos los ejemplos que conocemos, el número de condicionamiento $\kappa_{\mathcal{C}}$ es una cantidad que sólo depende de la clase proyectiva de la matriz. Por ejemplo, el número de condicionamiento clásico del Álgebra Lineal, κ, puede verse como una función

$$\kappa : \quad \mathbb{P}(\mathcal{M}_n(\mathbb{C})) \quad \longrightarrow [0, \infty].$$

Por ello, estudiamos su distribución de probabilidad con la métrica Gaussiana o, equivalentemente, en el espacio proyectivo complejo. Una observación crucial viene de la mano del Teorema de Eckart–Young (atribuido también en esta memoria a Schmidt y Mirsky, siguiendo el criterio de [122]): El número de condicionamiento de una matriz es exactamente el inverso de la distancia de dicha matriz al conjunto de matrices singulares. Por lo tanto, en el caso de que $\mathcal{C}$ sea el espacio total de matrices proyectivas $\mathbb{P}(\mathcal{M}_n(\mathbb{C}))$ y $\kappa_{\mathcal{C}}$ sea el condicionamiento usual κ_D, la tarea de estimar la cantidad de la ecuación (2) consiste en estudiar el volumen de un *tubo* alrededor de una variedad proyectiva particular: El conjunto de las matrices singulares. Este problema concreto fue resuelto a principios de la década de los 90 por Smale, Renegar, Demmel y Edelman en [117, 103, 40, 41, 42] (hay muchos estudios anteriores sobre tubos, una recopilación bibliográfica puede encontrarse en la introducción del Capítulo 2). En [103, 40] se propone una cota general para el volumen de un tubo en torno a cualquier variedad proyectiva, con lo que se pueden atacar diversos casos de la situación descrita en la expresión (2). En el Capítulo 2 mejoraremos ligeramente las cotas obtenidas por Renegar y Demmel. Un tubo en el espacio proyectivo complejo se define como sigue: Para un conjunto cualquiera $T \subseteq \mathbb{P}_n(\mathbb{C})$ y un número positivo $\varepsilon > 0$,

el tubo de radio ε en torno a T es el conjunto de puntos T_ε cuya distancia proyectiva a algún punto de T es menor que ε. Entonces, demostraremos el siguiente resultado (que está incluido, como el resto de resultados de esta sección, en los artículos [9, 10]),

Teorema 0.0.10 *Sea $V \subseteq \mathbb{P}_n(\mathbb{C})$ una variedad proyectiva equidimensional de dimensión compleja m. Sea $\varepsilon > 0$ un número real positivo. Entonces,*

$$\frac{vol[V_\varepsilon]}{vol[\mathbb{P}_n(\mathbb{C})]} \leq 2\deg(V)\left(\frac{e\ n\ \varepsilon}{n-m}\right)^{2(n-m)}.$$

donde $\deg(V)$ *es el grado de V, vol es la medida natural en el espacio proyectivo complejo y e es la base del logaritmo neperiano.*

Sin embargo, muchas veces el conjunto $\mathcal{C}$ en el que nos interesa estudiar la distribución de probabilidad del condicionamiento no es el total, sino una variedad algebraica proyectiva del espacio $\mathbb{P}(\mathcal{M}_n(\mathbb{C}))$. En este caso, el Teorema de Schidt-Mirsky-Eckart-Young vuelve a reducir el problema al cálculo del volumen de un tubo, pero ahora no nos interesa el volumen de ese tubo en el espacio total, sino el volumen (para la dimensión apropiada) de la *intersección de ese tubo con la variedad algebraica* $\mathcal{C}$. Se trata por tanto de un problema geométrico totalmente nuevo, para el que no conocemos ningún estudio anterior. En el Capítulo 2 proponemos una solución a este problema. Más concretamente, demostraremos el siguiente resultado,

Teorema 0.0.11 *Sean $V, V' \subseteq \mathbb{P}_n(\mathbb{C})$ dos variedades proyectivas equidimensionales de dimensiones respectivas $m > m' \geq 1$. Sea $\varepsilon > 0$ un número positivo. Sea V'_ε el tubo de radio ε en torno a V'. Entonces, se tiene:*

$$\frac{vol[V'_\varepsilon \cap V]}{vol[V]} \leq 2\deg(V')\left(\frac{en}{n-m'}\right)^{2(n-m')}\left[e\frac{n-m'}{m-m'}\ \varepsilon\right]^{2(m-m')},$$

donde $\deg(V')$ *es el grado de V' y vol es la medida natural de volumen en la variedad V.*

La técnica de estudio de volúmenes de tubos que hemos desarrollado trata de dar una respuesta a situaciones como la planteada en (2). Como ejemplo de aplicación, sea $\mathcal{C} := \Sigma^{n-1}_{\mathcal{M}} \subseteq \mathbb{P}(\mathcal{M}_n(\mathbb{C}))$ la clase formada por todas las matrices proyectivas singulares complejas.
Para una matriz singular, se puede definir un número de condicionamiento de un modo muy natural: Dada una matriz $A \in \mathbb{P}(\mathcal{M}_n(\mathbb{C}))$, de rango $n-1$, se define $\kappa_D^{(n-1)}(A) := \|A\|_F\|A^\dagger\|_2$, donde $\|\cdot\|_F$ es la norma de Frobenius y † denota inversa de Moore–Penrose. En la Sección 1.5 veremos que el número de condicionamiento singular $\kappa_D^{(n-1)}(A)$ controla la estabilidad en el cálculo de núcleos e inversas generalizadas.

Nótese que $\Sigma_{\mathcal{M}}^{n-1}$ es una variedad proyectiva de codimensión compleja 1 en $\mathbb{P}(\mathcal{M}_n(\mathbb{C}))$. Tiene por tanto asociada de manera natural una forma de volumen *vol* dada por la estructura Riemanniana de su parte lisa. El resultado que proporciona nuestra técnica se puede resumir como sigue.

Teorema 0.0.12 *Con las notaciones que acabamos de introducir,*

$$\frac{vol[\{A \in \Sigma_{\mathcal{M}}^{n-1} : \kappa_D^{(n-1)}(A) > \varepsilon^{-1}\}]}{vol[\Sigma_{\mathcal{M}}^{n-1}]} \leq (n^{10/3}\varepsilon)^6.$$

Además, la esperanza de $\kappa_D^{(n-1)}$ *en el espacio de matrices singulares satisface la siguiente desigualdad:*

$$E_{\Sigma_{\mathcal{M}}^{n-1}}[\kappa_D^{(n-1)}] \leq 2n^{10/3}.$$

En el Capítulo 2 extenderemos ampliamente este tipo de resultados a matrices rectangulares de rango dado y matrices con ciertas estructuras particulares (simétricas, por bloques...).

Sobre la distribución de probabilidad del condicionamiento no–lineal

En el caso de sistemas de ecuaciones polinomiales hay también una noción natural de número de condicionamiento μ_{norm}, inicialmente definida por Shub & Smale en [112]. Esencialmente, el condicionamiento $\mu_{\text{norm}}(f, \zeta)$ de un sistema f en un punto ζ controla el condicionamiento como aplicación lineal de la matriz diferencial $d_\zeta f$ (véase la Sección 1.6 para una definición completa). El número μ_{norm} tiene numerosas propiedades, algunas de las cuales serán discutidas a lo largo de esta memoria. En [113], Shub & Smale obtenían cotas superiores para la distribución de probabilidad del número de condicionamiento μ_{norm} (véase [91] para algunos resultados en el caso sparse). Esos resultados son esenciales para el diseño de algoritmos numéricos eficientes para la resolución de problemas en geometría algebraica, como se verá en los capítulos 4 y 5 de esta memoria; de hecho, conocer la distribución de los problemas mal condicionados es un buen punto de partida para programar algoritmos que los resuelvan, y proporciona información útil sobre la precisión necesaria de las operaciones que realicemos, el error del resultado que podamos obtener,la probabilidad de éxito y la complejidad total de nuestro algoritmo. Otros muchos autores se han interesado por el condicionamiento de los sistemas de ecuaciones; entre ellos citamos [30, 34, 37, 91, 55].
Consideremos fijada una lista de grados (d). El condicionamiento μ_{norm}, descrito por Shub & Smale para el caso cero–dimensional, está fuertemente ligado a la distancia "en la fibra" de un sistema a la variedad $\Sigma_{(d)}^{n-1}$ (esto es, el conjunto de los sistemas con alguna solución singular).

Para extender esta noción al caso de dimensión positiva, esto es, el caso de $m < n$ ecuaciones homogéneas en las incógnitas $X_0, \ldots, X_n$, debemos considerar la variedad $\Sigma_{(d)}^{m-1}$ de los sistemas que tienen alguna solución de rango no maximal. Los números de condicionamiento generalizados que proponemos están ligados a un refinamiento natural de $\Sigma_{(d)}^{m-1}$ descrito como sigue: Sea $\Sigma_{(d)}^{r} \subseteq \Sigma_{(d)}^{m-1}$ la variedad proyectiva formada por todos los sistemas polinomiales que poseen alguna singularidad de rango menor o igual que r (véase [1], por ejemplo). Entonces, tenemos la cadena

$$\Sigma_{(d)}^{m-1} \supseteq \cdots \supseteq \Sigma_{(d)}^{1}.$$

Como demostraremos en el Teorema 3.1.1 (Teorema del Número de Condicionamiento), los números de condicionamiento $\mu_{\mathrm{norm}}^{(r)}$ que analizaremos están relacionados con la distancia "en la fibra" de un sistema dado a cada una de las variedades $\Sigma_{(d)}^{r}$. Dicho de otra manera, para un sistema de m ecuaciones polinomiales homogéneas $f \in \mathbb{C}[X_0, \ldots, X_n]^m$ y una solución $\zeta \in \mathbb{P}_n(\mathbb{C})$ de f, podríamos definir el número de condicionamiento $\mu_{\mathrm{norm}}^{(r)}(f, \zeta)$ como el inverso de la distancia de f al conjunto de sistemas g tales que:

$$g(\zeta) = 0, \quad rank(d_\zeta g) \leq r,$$

donde $d_\zeta g$ es la aplicación diferencial de g en ζ.
El número de condicionamiento $\mu_{\mathrm{norm}}^{(r)}$ tiene distintas interpretaciones en función del número de ecuaciones m y del valor de r:

1. Si elegimos $r = m = n$ (caso cero–dimensional), entonces el condicionamiento $\mu_{\mathrm{norm}}^{(n)}$ controla la complejidad de los métodos de homotopía para la resolución de sistemas de ecuaciones polinomiales (véase [112]). También controla la separación entre dos soluciones distintas (resultado debido a J.P. Dedieu, [31])

2. Si elegimos $r = m < n$ (caso de dimensión positiva), entonces el condicionamiento $\mu_{\mathrm{norm}}^{(m)}$ se relaciona con el radio de convergencia del operador de Newton en dimensión positiva (véase [34, 116] y la Sección 3.1.2). También controla la estabilidad del conjunto solución, en el sentido de [30, 37] (véase por ejemplo el Teorema 3.1.4).

3. Si elegimos $r < m = n$, el condicionamiento $\mu_{\mathrm{norm}}^{(r)}$ describe la distribución de los distintos tipos de singularidades de los sistemas de ecuaciones polinomiales. Más concretamente, describe la probabilidad de que un sistema esté cerca (distancia "en la fibra") de otro sistema con una solución singular de corango dado. Esto es consecuencia del Teorema del Número de Condicionamiento al que nos hemos referido arriba.

Por estos motivos, tiene un fuerte interés el estudio de la distribución de probabilidad de los distintos números de condicionamiento, para los distintos valores posibles de r, m, n.
En el Capítulo 3 demostraremos cotas superiores para la distribución de probabilidad de estos números en la situación más general posible, abarcando los condicionamientos de sistemas de ecuaciones para cualesquiera valores de r, m, n. Comenzamos con el siguiente teorema (véase el Teorema 3.6.1 para una versión más precisa), que estima la distribución de probabilidad del condicionamiento generalizado $\mu^{(r)}_{\text{norm}}$ en su forma más general, abarcando a la vez el caso singular (esto es, caso $r < m$) y el caso sub–determinado (esto es, $m < n$). Se trata de un resultado que generaliza ampliamente la cota obtenida por Shub & Smale en [113].

Teorema 0.0.13 *Sea $(d) = (d_1, \ldots, d_m)$ tal que $d_i > 1$ para algún i, $1 \leq i \leq m$. Sea $\varepsilon > 0$ un número real, y sea ϕ_ε la función que lleva a cada sistema f sobre el volumen del conjunto de soluciones ζ de f tales que $\mu^{(r)}_{\text{norm}}(f, \zeta) > \varepsilon^{-1}$. Entonces, tenemos:*

$$\mathrm{E}[\phi_\varepsilon] \leq 2\pi e^{1/3} \mathrm{vol}[\mathbb{P}_{n-m}(\mathbb{C})] \mathcal{D}\left(\sqrt{Nmr(n+1)}\ \varepsilon\right)^{2(m-r+1)(n-r+2)}.$$

También mostraremos cotas para el caso de la distancia absoluta a las variedades $\Sigma^r_{(d)}$. En efecto, demostraremos el siguiente resultado (véase el Corolario 3.1.5, y también [13]).

Teorema 0.0.14 *Sea $1 \leq r \leq n-1$ un número natural y sea $\mathrm{dist_r}$ la función que lleva cada sistema de ecuaciones cero–dimensional a la distancia que le separa de la variedad $\Sigma^r_{(d)}$. Entonces, la esperanza de la función $\frac{1}{\mathrm{dist_r}}$ satisface la siguiente desigualdad:*

$$E\left[\frac{1}{\mathrm{dist_r}}\right] \leq 8N(r+1)^3\ d\left[\prod_{i=1}^{n}(d_i+1)\right]^{\frac{1}{2(n-r)^2}},$$

donde $d := \max\{d_i : 1 \leq i \leq m\}$ es el máximo de los grados que aparecen en la lista (d).

El condicionamiento generalizado $\mu^{(r)}_{\text{norm}}$ se define para un sistema f y una solución del mismo ζ. Nos centramos a continuación en el caso $r = m$. Para todo sistema f, podemos considerar los dos números definidos como sigue,

$$\mu^{(m)}_{\text{worst}}(f) := \sup_{\zeta \in V(f)} \mu^{(m)}_{\text{norm}}(f, \zeta),$$

$$\mu^{(m)}_{\text{av}}(f) := \mathrm{E}_{V(f)}[\mu^{(m)}_{\text{norm}}].$$

Esto es, $\mu^{(m)}_{\text{worst}}(f)$ es el caso peor de los condicionamientos en la variedad solución de f, y $\mu^{(m)}_{\text{av}}(f)$ es el valor esperable de dicho condicionamiento en

la variedad solución de f. Como hemos indicado en el segundo de los items escritos arriba (véase también la introducción del Capítulo 3), el condicionamiento $\mu^{(m)}_{\text{norm}}$ controla la estabilidad del conjunto solución ante perturbaciones del input. En palabras sencillas, podemos decir que $\mu^{(m)}_{\text{av}}$ controla la estabilidad *media* del conjunto de soluciones de f, y $\mu^{(m)}_{\text{worst}}$ controla la estabilidad del caso peor (o estabilidad *total*) del conjunto de soluciones de f. A continuación, un resultado que acota la esperanza del condicionamiento $\mu^{(m)}_{\text{av}}$.

Teorema 0.0.15 *Sea $(d) = (d_1, \ldots, d_m)$ tal que $d_i > 1$ para algún i, $1 \leq i \leq m$. Entonces, se satisface la siguiente desigualdad:*

$$\mathrm{E}[\mu^{(m)}_{\text{av}}] \leq 3m\sqrt{nN}.$$

En el caso $m = 1$, podemos incluso obtener una igualdad (véase el Teorema 3.7.1).
En cuanto al condicionamiento del caso peor $\mu^{(m)}_{\text{worst}}$, también obtenemos una cota, aunque mucho peor que la que acabamos de exponer:

Teorema 0.0.16 *Sea $(d) = (d_1, \ldots, d_m)$ tal que $d_i > 1$ para algún i, $1 \leq i \leq m$. Entonces, se satisface la siguiente desigualdad:*

$$\mathrm{E}[\mu^{(m)}_{\text{worst}}] \leq \frac{\mathcal{D}^{1/4}}{d^{3/2}}[10N^{1/2}mn^{1/2}d^{3/2}]^{\frac{n-m+2}{2}}.$$

Podemos escribir la conclusión de estos dos resultados como sigue: El valor esperado del condicionamiento $\mu^{(m)}_{\text{av}}(f)$ es muy pequeño (parecido a la raíz cuadrada del tamaño del input $N + 1$). Por lo tanto, para un sistema de ecuaciones elegido al azar, *podemos esperar que la mayor parte del conjunto de soluciones sea extremadamente estable.* Sin embargo, si pedimos que la totalidad del conjunto de soluciones sea estable, nos encontramos con que la situación puede ser totalmente distinta: Solamente podemos obtener una cota exponencial para el valor esperable de $\mu^{(m)}_{\text{worst}}$. Puede suceder por lo tanto que la mayoría de los sistemas de ecuaciones tengan algunas soluciones muy inestables (sobre todo si el valor de $n - m$ es grande), aunque sepamos que la mayoría de esas soluciones son muy estables.
Resaltamos a continuación otro de los resultados que obtendremos, que viene a combinar el resultado del Teorema 0.0.16 con el item 2 de los expuestos arriba.

Teorema 0.0.17 *Para todo sistema de ecuaciones $f \in \mathbb{C}[X_1, \ldots, X_n]^m$, $m < n$, se tiene que existe un radio $R(f) \geq 0$ tal que para todo punto en un entorno de radio $R(f)$ del conjunto solución $V_{C^n}(f)$, el operador de Newton afín de dimensión positiva converge a velocidad doblemente exponencial hacia un punto de la variedad solución. Además, se tiene:*

1. *Para casi todo sistema f, el radio $R(f)$ es estrictamente mayor que* 0. *Esto es,*
$$Prob[f \text{ sistema} : R(f) > 0] = 1.$$

2. *Para una lista de grados fijada $(d) = (d_1, \ldots, d_m)$, la esperanza del radio de convergencia $R(f)$, cuando recorremos el espacio de sistemas, satisface la siguiente desigualdad:*
$$\mathrm{E}\,[R(f)] \geq \frac{c}{\mathcal{D}^{1/4}[10m\sqrt{nN}d^{3/2}]^{\frac{n-m+2}{2}}},$$
donde $\mathcal{D} := \prod_{i=1}^{m} d_i$ es el número de Bézout asociado a la lista de grados (d), $c > 1/10$ es una constante universal, $d := \max_{1\leq i\leq m}\{d_i\}$ es el máximo de los grados y N es el tamaño del input (en codificación densa).

Este resultado significa lo siguiente: Para casi todas las variedades intersección completa $V \subseteq \mathbb{C}^n$, existe un tubo V_R de radio $R > 0$ tal que todos los puntos de V_R son ceros aproximados de V. Además, podemos proporcionar una cota inferior para el valor esperable del radio R.

Estimaciones discretas

El Capítulo 6 de esta memoria es de una naturaleza totalmente distinta a los capítulos anteriores. Hasta ahora, hemos hablado de probabilidades y esperanzas considerando el espacio de sistemas como un conjunto continuo, y en las demostraciones de los resultados que hemos descrito se utilizan técnicas de integración en variedades Riemannianas, siempre con el cómodo soporte que nos proporcionan el cálculo y la topología cuando tratamos problemas continuos. Sin embargo, la realidad computacional nos obliga a trabajar en toda circunstancia con números que se pueden escribir con un número finito de dígitos, y el comportamiento en el caso discreto de las distintas propiedades analizadas en el continuo podría ser mucho peor, a priori. Sin un resultado que lo garantice, podría suceder que la distribución de probabilidad del condicionamiento de las matrices con coeficientes racionales fuera mucho peor que en el caso del total (dado que el conjunto de las matrices racionales tiene medida nula dentro del espacio de todas las matrices complejas). El mismo problema podría aparecer en el caso no–lineal.
Por ello incluimos el Capítulo 6, que contiene las versiones discretas de los principales resultados expuestos en los capítulos 2, 3 y 4. La herramienta fundamental es el conteo de puntos enteros en conjuntos semi–algebraicos, siguiendo la línea de [28, 21, 23]. Incluimos algunas mejoras técnicas a los resultados de conteo ya existentes, que nos permiten obtener cotas aún más precisas. La correcta exposición de estos resultados exige una amplia recopilación de varias notaciones y conceptos, por lo que nos parece excesivo incluirlos en esta introducción. Remitimos al lector a la introducción del Capítulo 6.

Perspectivas, problemas a medio y largo plazo

Los resultados expuestos a lo largo de esta memoria suponen algunos pasos en el camino hacia la comprensión de los problemas de resolución y complejidad de sistemas de ecuaciones polinomiales. La existencia de un algoritmo probabilístico que resuelve la gran mayoría de sistemas de ecuaciones (con coeficientes complejos) en tiempo polinomial en el tamaño del input sugiere muchas preguntas, algunas de ellas "viejas", que pueden ser vistas bajo el filtro de los nuevos resultados. Tratar de comprender la naturaleza de estas cuestiones es un objetivo muy ambicioso, que consideramos puede marcar la línea de investigación que continuará de modo natural el trabajo ya realizado. A continuación explicamos con cierto detalle algunas de los problemas abiertos que cobran nuevo sentido a la luz de los avances de esta memoria.

- **¿Cuál es el conjunto de los sistemas que pueden resolverse mediante el método de homotopía propuesto?** La respuesta intuitiva, que supondría una respuesta casi total al problema de aproximación numérica de soluciones de sistemas de ecuaciones, es: Los problemas que pueden ser resueltos son aquellos que están bien condicionados, en términos del número de condicionamiento μ_{norm}. Pero esta afirmación intuitiva parece muy difícil de demostrar. Tal vez un camino para empezar a buscar es analizar con detalle la uniformidad de la condición de ser buen par inicial para la homotopía, y tratar comprender más a fondo la estructura geométrica de la variedad de sistemas mal condicionados.

- **¿Cuántos ceros aproximados podemos encontrar?** De nuevo es una pregunta muy natural, que permanece sin respuesta. Normalmente, como ya hemos dicho, los sistemas de ecuaciones tienen un número enorme de soluciones (exponencial en el número de incógnitas). Por tanto, hemos argumentado en esta introducción que es un objetivo (muy probablemente) imposible describirlas todas en tiempo polinomial. Sin embargo, aún es deseable aproximar tantas soluciones como queramos, en función de nuestras necesidades y nuestra capacidad de cálculo. Los teoremas principales de esta introducción garantizan que, probabilísticamente, podemos encontrar *algún* cero de la mayoría de sistemas de ecuaciones polinomiales. Pero no tenemos resultados concretos, por el momento, sobre cómo se podrían encontrar distintas soluciones (más allá de probar distintos pares iniciales, y comprobar si las soluciones obtenidas por ellos son diferentes entre sí). Tenemos algunos elementos que permiten intuir el modo de resolver este problema, mediante el uso de simetrías en el espacio de soluciones. Sin embargo, aún tenemos que comprender más a fondo este problema para dar una respuesta definitiva.

- **¿Qué sucede con los sistemas de ecuaciones reales y soluciones reales?** Todas las reflexiones hechas en esta memoria, y también los resultados encontrados por Shub & Smale en su serie de artículos se desarrollan en el caso de coeficientes complejos, y búsqueda también de soluciones complejas. Sin embargo, muchas veces es nuestra voluntad conocer alguna solución real de los sistemas que encontremos. Por lo tanto, un objetivo muy deseable es demostrar que la probabilidad de resolver sistemas con coeficientes reales es también muy grande, y que lo mismo sucede si buscamos únicamente soluciones reales. De nuevo, esta cuestión es, hasta donde sabemos, una pregunta muy difícil: Los argumentos utilizados para el caso complejo no pueden ser utilizados directamente para el caso real. Esto debe ser hecho, probablemente, con técnicas y argumentos nuevos. Una respuesta a las preguntas anteriormente planteadas sería, naturalmente, de gran ayuda para resolver este problema.

- **¿Cuál es la naturaleza de las soluciones singulares?** Esto es, ¿podemos utilizar métodos de homotopía generalizados para encontrar soluciones singulares de sistemas? La geometría de las singularidades es, en general, mucho más complicada que la geometría de las soluciones regulares. Sin embargo, algunos avances en la dirección del estudio de las estructuras geométricas correspondientes ya están incluidos en esta memoria, como el estudio de la distribución de probabilidad de las singularidades de corango dado en los casos lineal y no–lineal. También hay algunos trabajos sobre la generalización del Método de Newton para el caso singular, como [56, 55] que podrían indicar el camino a seguir para resolver este problema.

- **¿Qué se puede decir del problema sobre–determinado?** El enunciado estándar de un problema NP completo es en forma de sistema de ecuaciones con $n + 1$ ecuaciones y n incógnitas. Este caso parece ser totalmente diferente del estudiado en estas páginas, dado que la mayoría de los sistemas con esta estructura no tienen solución alguna. Sin embargo, no dejamos de preguntarnos si se pueden modificar los argumentos que hemos utilizado para tratar el caso sobre –determinado de manera efectiva. Una respuesta a esta pregunta tendría, seguramente, la estructura de un algoritmo numérico. Por tanto, tendríamos que aceptar una cierta probabilidad de fracaso: Difícilmente podríamos obtener una respuesta global. Debido a ello, se trata aparentemente de un problema menos duro que la Conjetura de Cook, aunque una respuesta positiva al mismo conllevaría un avance considerable en la comprensión del problema de P versus NP.

Capítulo 1

Preliminares

Los resultados, definiciones y conceptos que aparecen en este capítulo no son completamente originales. Hemos intentado que esta memoria sea un escrito auto–contenido, lo que exige la recopilación de ciertos conocimientos previos, que serán englobados en este capítulo. Algunos de ellos son bien conocidos, y se pueden encontrar en libros de texto de la licenciatura. Es el caso de los conceptos sobre integración en variedades, la estructura Riemanniana del espacio proyectivo complejo o algunas propiedades del condicionamiento lineal. También introduciremos algunos otros conceptos más específicos que serán utilizados en los siguientes capítulos. La mayor parte de las notaciones utilizadas en la memoria están aquí recopiladas, aunque hemos procurado recordar las más específicas en el momento de su uso, para facilitar la lectura.

1.1. Elementos de Integración Geométrica

El cálculo de volúmenes y de integrales en variedades Riemannianas es una herramienta que aparece de forma natural en el estudio de los problemas de nuestro entorno. No es nuestra intención abarcar en unas pocas líneas un tema tan amplio como éste, pero sí que introduciremos los conceptos básicos y algunas herramientas que usaremos a lo largo de esta memoria.

1.1.1. Volúmenes en espacios Euclídeos

Sea W un espacio vectorial real de dimensión n. Siguiendo el criterio de [14], diremos que una forma de volumen ω en W es el valor absoluto de una forma n–multilineal alternada $\omega' : W^n \longrightarrow \mathbb{R}$ (esta definición difiere de la usual, en la que no se utiliza el valor absoluto). Esto es, $\omega(v_1, \ldots, v_n) := |\omega'(v_1, \ldots, v_n)|$, donde ω' es n–multilineal y

$$\omega'(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = -\omega'(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n), \forall i \neq j.$$

Es bien sabido que una forma de volumen ω en W queda determinada a partir del valor $\omega(v_1, \ldots, v_n)$ de ω en una base cualquiera $[v_1, \ldots, v_n]$ de W.

En efecto, para cualquier n–tupla $[w_1, \ldots, w_n]$ de vectores de W, tal que $w_j = \sum_{i=1}^{n} a_{ij} v_i$ donde $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{R})$ es una matriz real, se tiene que

$$\omega(w_1, \ldots, w_n) = |det(A)| \omega(v_1, \ldots, v_n).$$

Así, podemos considerar la forma de volumen natural $dx_1 \cdots dx_n$ en $\mathbb{R}^n$ definida como $dx_1 \cdots dx_n(e_1, \ldots, e_n) = 1$, para la base usual $e_1, \ldots, e_n$ de $\mathbb{R}^n$.
Fijada una forma de volumen ω en un espacio vectorial W de dimensión n, y dados un conjunto de vectores $v_1, \ldots, v_n$ de V, llamamos volumen del paralelepípedo formado por $v_1, \ldots, v_n$ al número real $\omega(v_1, \ldots, v_n)$.

Definición 1.1.1 *Sea W un espacio vectorial real de dimensión n con producto interior $\langle \cdot, \cdot \rangle$. Entonces, W tiene una forma de volumen natural ω dada por*

$$\omega(b_1, \ldots, b_n) = 1,$$

para toda base ortonormal $\{b_1, \ldots, b_n\}$ de W. Sea $\{v_1, \ldots, v_n\} \subseteq W$ una colección de vectores de W. Entonces, el volumen (en W) del paralelepípedo formado por $v_1, \ldots, v_n$ con respecto a ω es

$$\omega(v_1, \ldots, v_n) = |\det(v_{ij})|,$$

donde (v_{ij}) es la matriz formada por las coordenadas de los vectores v_i en cualquier base ortonormal de W.
Además, si $L : W_1 \longrightarrow W_2$ es una aplicación lineal entre espacios vectoriales de igual dimensión con productos interiores $\langle \cdot, \cdot \rangle_1$, $\langle \cdot, \cdot \rangle_2$, definimos el determinante de L, como

$$\det(L) := |\det(A)|,$$

donde A es una matriz coordenada de L en bases ortonormales de W_1 y W_2.

Se comprueba de forma inmediata que la definición anterior no depende de las bases ortonormales de W, W_1 y W_2 escogidas. Además, se comprueba que si $v_1, \ldots v_n \in W_1$ es un conjunto de vectores, $L : W_1 \longrightarrow W_2$ y si $\det(L) \neq 0$, entonces se tiene que

$$\omega_{W_2}(L(v_1), \ldots, L(v_n)) = \det(L) \omega_{W_1}(v_1, \ldots, v_n),$$

donde ω_{W_i}, $i = 1, 2$ es la forma de volumen asociada al producto interior de W_i.

Lema 1.1.2 *Sea W un espacio vectorial de dimensión n con producto interior $\langle \cdot, \cdot \rangle$. Sea $\{v_1, \ldots, v_s\}$ un conjunto de vectores en W, con $s \leq n$, y sea $W' := \langle v_1, \ldots, v_s \rangle$ el espacio vectorial generado por esa familia de vectores. Finalmente, sea $M := (v_{ij}) \in \mathcal{M}_{n \times s}(\mathbb{R})$ la matriz cuyas columnas son las*

coordenadas de los vectores v_i en cualquier base ortonormal de W. Entonces, el volumen en W' del paralelepípedo formado por $v_1, \ldots, v_s$ es igual a:

$$\det(M^t M)^{1/2},$$

donde M^t es la matriz traspuesta de M.

Demostración.– Sea $b_1, \ldots b_n$ la base ortonormal de W en que está definida la matriz M. Sea $b'_1, \ldots, b'_n$ otra base ortonormal de W, tal que los vectores $b'_1, \ldots, b'_s$ generan el subespacio W'. Sea

$$(b'_1 \ \cdots \ b'_n)O = (b_1 \ \cdots \ b_n)$$

el cambio de base, donde $O \in \mathcal{M}_n(\mathbb{R})$ es una matriz ortonormal. Entonces, en la base $b'_1, \ldots, b'_n$, las coordenadas de los vectores $v_1, \ldots, v_s$ son las columnas de la matriz

$$OM \in \mathcal{M}_{n\times s}(\mathbb{R}).$$

Por la forma en que hemos definido la nueva base $b'_1, \ldots, b'_n$, tenemos que

$$OM = \begin{pmatrix} P \\ 0 \end{pmatrix},$$

donde $P \in \mathcal{M}_s(\mathbb{R})$. Por lo tanto, en la base ortonormal $b'_1, \ldots, b'_s$ de W', las coordenadas de los vectores $v_1, \ldots, v_s$ son las columnas de la matriz P. Por definición, por tanto,

$$\omega_{W'}(v_1, \ldots, v_s)^2 = \det(P)^2 = \det(P^t P) = \det\left((P^t \ 0)\begin{pmatrix} P \\ 0 \end{pmatrix}\right) =$$

$$\det((OM)^t(OM)) = \det(M^t O^t OM) = \det(M^t M),$$

y el lema queda demostrado. ∎

En el caso de un espacio vectorial complejo, la estrategia general para medir volúmenes pasa por considerar el espacio vectorial real asociado. Sea W un espacio vectorial complejo de dimensión (compleja) n con producto hermitiano $\langle \cdot, \cdot \rangle$. Denotemos por $W_{\mathbb{R}}$ el espacio vectorial asociado a W, que es un espacio vectorial real de dimensión $2n$. Tenemos además la aplicación

$$\begin{array}{rccc} \varphi_{\mathbb{R}} : & W_{\mathbb{R}} & \longrightarrow & W \\ & (\alpha_1, \beta_1, \ldots \alpha_n, \beta_n) & \mapsto & (\alpha_1 + \sqrt{-1}\beta_1, \ldots, \alpha_n + \sqrt{-1}\beta_n). \end{array}$$

Además, el espacio $W_{\mathbb{R}}$ lleva asociado un producto interior $\langle \cdot, \cdot \rangle_{\mathbb{R}}$, dado como sigue:

$$\langle x, y \rangle_{\mathbb{R}} := \mathbb{R}e(\langle \varphi_{\mathbb{R}}(x), \varphi_{\mathbb{R}}(y) \rangle_{\mathbb{R}},$$

con su correspondiente forma de volumen asociada, $\omega_{W_{\mathbb{R}}}$. Dados n vectores (complejos) $v_1, \ldots, v_n \in W$, llamaremos volumen del paralelepípedo definido por $v_1, \ldots, v_n$ a la cantidad

$$\omega_{W_{\mathbb{R}}}(\varphi_{\mathbb{R}}^{-1}(v_1), \varphi_{\mathbb{R}}^{-1}(\sqrt{-1}v_1), \ldots, \varphi_{\mathbb{R}}^{-1}(v_n), \varphi_{\mathbb{R}}^{-1}(\sqrt{-1}v_n))$$

Tenemos el siguiente resultado, que es la versión compleja del Lema 1.1.2.

Lema 1.1.3 *Sea W un espacio vectorial complejo de dimensión n con producto hermitiano $\langle \cdot, \cdot \rangle$, y sean $v_1, \ldots, v_s$ vectores en W, con $s \leq n$. Sea $W' := \langle v_1, \ldots, v_s \rangle$ el espacio vectorial generado por esa familia de vectores. Finalmente, sea $M := (v_{ij}) \in \mathcal{M}_{n \times s}(\mathbb{R})$ la matriz cuyas columnas son las coordenadas de los vectores v_i en cualquier base unitaria de W. Entonces, el volumen en W' del paralelepípedo formado por $v_1, \ldots, v_s$ es igual a:*

$$\det(M^* M),$$

donde M^* es la matriz conjugada traspuesta de M.

Demostración.– Sea $M = M_1 + \sqrt{-1} M_2$ la descomposición de M en partes real e imaginaria. Se comprueba que $\det(M^* M) \in \mathbb{R}$ satisface:

$$\det(M^* M) = det(A + \sqrt{-1} B),$$

donde $A := M_1^t M_1 + M_2^t M_2,\ B = M_1^t M_2 - M_2^t M_1$. Por otro lado, por la definición y el Lema 1.1.2, el volumen en W' del paralelepípedo formado por $v_1, \ldots, v_s$ vale

$$\det \begin{pmatrix} A & B \\ -B & A \end{pmatrix}^{1/2}.$$

Finalmente, el lema se deduce de la conocida igualdad

$$|det(A + \sqrt{-1} B)|^2 = \det \begin{pmatrix} A & B \\ -B & A \end{pmatrix} \in \mathbb{R},$$

para todo par de matrices cuadradas reales A, B. ∎

1.1.2. Integración y Topología en variedades diferenciables

Incluimos aquí un pequeño resumen de las propiedades básicas de la integración en variedades. Para nosotros, "diferenciable" significará siempre "infinitamente diferenciable", tanto al referirnos a variedades como a aplicaciones entre variedades. Nos remitimos a [48] para la terminología básica sobre variedades diferenciables. Como habitualmente, para cada variedad diferenciable M supondremos que el espacio topológico subyacente satisface las siguientes hipótesis:

- El conjunto M es un espacio topológico Hausdorff.
- M es localmente compacto (esto es, para todo elemento $p \in M$, existe un abierto $U \subseteq M$ que contiene a p y un compacto $K \subseteq M$ tal que $U \subseteq K$).
- Se puede escribir M como unión contable de compactos.

- M es paracompacto (esto es, además de Hausdorff, todo recubrimiento abierto $\mathcal{C}$ de M tiene un refinamiento que es localmente finito, en el sentido de que todo punto $p \in M$ tiene un entorno U tal que el número de elementos de $\mathcal{C}$ que intersecan a U es finito).

- La topología de M tiene una base contable.

Para cada $p \in M$, denotaremos por T_pM el espacio tangente a M en p. Para cada aplicación diferenciable entre variedades $F : M \longrightarrow M'$ y para cada $p \in M$, denotaremos por $d_pF : T_pM \longrightarrow T_{F(p)}M'$ la aplicación diferencial entre los respectivos espacios tangentes. También denotaremos por $\dim(M)$ la dimensión (real o compleja, según el contexto) de M.

Recordemos que una variedad Riemanniana real (compleja) es una variedad diferenciable real (compleja) en la que además hay una forma bilinear simétrica y definida positiva (Hermitiana) $\langle \cdot, \cdot \rangle_p$ en cada espacio tangente T_pM, que varía de forma diferenciable. Esto permite definir de modo natural los conceptos de ortogonalidad y ortonormalidad de vectores tangentes en un punto p. También queda asignada así una norma $\| \cdot \|_p$ en cada espacio tangente T_pM, del modo natural: $\|v\|_p := \langle v, v \rangle_p^{1/2}$. Como es habitual, diremos que una aplicación diferenciable entre variedades $F : M \longrightarrow M'$ es isometría en un punto $p \in M$ si la diferencial d_pF es una isometría de espacios vectoriales (esto es, identifica los productos internos de T_pM y $T_{F(p)}N$). También diremos que F es isometría si lo es en todo punto $p \in M$.

Una *función de densidad* en una variedad diferenciable M de dimensión m viene dada por una medida positiva (absolutamente continua con respecto a la medida de Lebesgue) en cada abierto de $\mathbb{R}^m$ asociado a un atlas de M, de forma que se verifique la ley de composición de cartas (véase [48, págs. 135,136] para más detalles). Así, se definen una σ–Álgebra y una medida positiva en la variedad M a través de las cartas, lo que permite definir volúmenes e integrales en M siguiendo el esquema clásico de Teoría de la Medida (véase por ejemplo [107]). Una variedad Riemanniana tiene asociada de modo natural una función de densidad, y por tanto definiciones de volumen e integral. También quedan así establecidos los conceptos de conjuntos medibles, funciones medibles e integrables, conjuntos de medida cero y demás nociones relacionadas. Con los axiomas topológicos que hemos asumido, se tiene además que la variedad Riemanniana M es σ–finita con respecto a la medida asociada, esto es, M puede escribirse como unión numerable de conjuntos con medida finita. Por ello, podemos aplicar cuando lo deseemos el Teorema de Fubini para integración en producto de variedades (véase por ejemplo [107]). Siguiendo la práctica usual, diremos que una propiedad se satisface en casi todo punto si se satisface en el complementario de un conjunto de medida cero. En general, para una función integrable $f : M \longrightarrow [-\infty, +\infty]$ definida en una variedad Riemanniana M,

denotaremos su integral en M mediante

$$\int f\ dM, \quad \text{o bien} \quad \int_{x\in M} f(x)\ dM.$$

Para un subconjunto medible $A \subseteq M$, denotaremos por $\nu_M[A]$ el volumen de A. En el caso de que M sea $\mathbb{R}^k$ con la estructura usual, $\nu_{\mathbb{R}^k}$ es la medida de Lebesgue y será denotada como es habitual por $\mathscr{L}^m$.
Si $N \subseteq M$ es una subvariedad diferenciable de M, entonces N hereda de M una estructura de variedad Riemanniana, y tiene por tanto una medida positiva asociada a ella de modo natural.
Este modo de definir volúmenes e integrales, que sigue el criterio de [48], permite utilizar los teoremas clásicos de integración de Lebesgue, como el Teorema de Convergencia Monótona. Además, no necesitamos imponer a la variedad ninguna condición de orientabilidad.
Algunos resultados de Topología Diferencial serán utilizados con cierta frecuencia en esta memoria. Entre ellos se encuentra el Teorema de Morse–Sard, un elegante y potente resultado que puede encontrarse, por ejemplo, en [74, pg. 69]. Recordemos que, dada una aplicación diferenciable entre variedades diferenciables $F : M \longrightarrow N$, se dice que $x \in M$ es un punto singular si la aplicación diferencial d_xF no es sobreyectiva. Un punto regular es un elemento de M que no es un punto singular. Además, un elemento $y \in N$ se dice que es un valor singular si $F^{-1}(y)$ contiene algún punto singular. Un valor regular es un elemento de N que no es un valor singular.

Teorema 1.1.4 (Morse–Sard) *Sea $F : M \longrightarrow N$ una aplicación diferenciable. Entonces, el conjunto de valores singulares de F tiene medida cero en N.*

Es un hecho bien conocido, consecuencia del Teorema de la Función Implícita, que la anti–imagen de un valor regular mediante una aplicación $F : M \longrightarrow N$ es una variedad diferenciable. Sin embargo, un resultado más fuerte y útil puede demostrarse con el mismo esfuerzo, como se hace notar en [74, pg. 22], y escribiremos a continuación. Para ello, recordemos la noción de transversalidad.

Definición 1.1.5 (Transversalidad) *Sea $F : M \longrightarrow N$ una aplicación diferenciable entre variedades diferenciables. Sea $W \subseteq N$ una subvariedad diferenciable de N. Decimos que F es transversal a W si para todo par de puntos $x \in M$, $y \in W$ tales que $F(x) = y$ se tiene:*

$$T_yW + d_xF(T_xM) = T_yN.$$

En otras palabras, si el espacio tangente en y a W y la imagen de la diferencial d_xF generan todo el espacio tangente en y a N.

Entonces, se tiene el siguiente resultado (véase [74] o también [59]).

Corolario 1.1.6 *Sea $F : M \longrightarrow N$ una aplicación diferenciable entre variedades diferenciables. Sea $W \subseteq N$ una subvariedad diferenciable de N, tal que F es transversal a W. Entonces, el conjunto $F^{-1}(W)$ es una subvariedad diferenciable de M, y la codimensión de $F^{-1}(W)$ en M es igual a la codimensión de W en N. Esto es,*

$$\dim(M) - \dim(F^{-1}(W)) = \dim(N) - \dim(W).$$

A continuación escribimos el célebre Teorema de Transversalidad de Thom, que también utilizaremos, en su versión débil. Puede encontrarse por ejemplo en [38] o [59]. Este resultado viene a decir que la condición de transversalidad es casi siempre cierta, en cierto sentido.

Teorema 1.1.7 (Thom) *Sean M, N, T variedades diferenciables. Sea también $W \subseteq N$ una subvariedad diferenciable de N. Sea $j : T \longrightarrow \mathcal{C}^\infty(M, N)$ una aplicación cualquiera, tal que la aplicación*

$$\begin{array}{rccc} \Theta: & T \times M & \longrightarrow & N \\ & (t, x) & \mapsto & j(t)(x) \end{array}$$

es diferenciable. Supongamos además que Θ es transversal a W. Entonces, para casi todo $t \in T$ se tiene que $j(t)$ es transversal a W.

1.1.3. La Fórmula de la Co–área

A lo largo de esta memoria utilizaremos con frecuencia la Fórmula de la Co–área, una poderosa herramienta debida a H. Federer (véase [45, pág. 258], [95, pág. 31]) que generaliza el Teorema de Fubini y el Teorema del Cambio de Variable simultáneamente. Se trata de un resultado técnico de la Teoría de Integración Geométrica con numerosas aplicaciones en distintos campos de las matemáticas. El enunciado de Federer es muy general, nosotros utilizaremos una versión simplificada como la que se describe en [14, pág. 241] (véase también [45, págs. 249, 280–282]). La correcta exposición de este resultado requiere algunas definiciones y resultados previos.

Definición 1.1.8 *Sean M y N dos variedades Riemannianas, de dimensiones respectivas $m \geq n$. Sea $x \in M$ un punto cualquiera y sea $F : M \longrightarrow N$ una aplicación diferenciable tal que la diferencial $d_x F$ es sobreyectiva.*
Definimos el espacio horizontal en x, $H_x \subseteq T_x M$, como el complemento ortogonal de $Ker(d_x F)$ en el espacio tangente $T_x M$. Esto es,

$$H_x := (Ker(d_x F))^\perp \subseteq T_x M.$$

Obsérvese que, dado que estamos suponiendo que $d_x F$ es sobreyectiva, la dimensión de H_x es igual a $n = \dim(N)$. Esto permite definir el concepto de jacobiano normal, como sigue.

Definición 1.1.9 *Sean M y N dos variedades Riemannianas, de dimensiones respectivas $m \geq n$. Sea $x \in M$ un punto cualquiera y sea $F : M \longrightarrow N$ una aplicación diferenciable tal que la diferencial d_xF es sobreyectiva. Sea $[v_1, \ldots, v_n]$ una base ortonormal del espacio horizontal en x, H_x. Definimos el jacobiano normal de F en x, denotado NJ_xF, como el volumen en $T_{F(x)}N$ del paralelepípedo definido por el conjunto de vectores*

$$d_xF(v_1), \ldots, d_xF(v_n).$$

En otras palabras, tenemos:

$$NJ_xF := \left|\det\left(d_xF\mid_{H_x}\right)\right|,$$

donde $d_xF\mid_{H_x}$ se identifica con cualquiera de sus matrices coordenadas, en bases ortonormales.
Si d_xF no es sobreyectiva, o si $m < n$, se define $NJ_xF := 0$.

El siguiente resultado permite calcular explícitamente el jacobiano normal en multitud de ocasiones.

Lema 1.1.10 *Sea $F : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ una aplicación diferenciable, y supongamos que $m \geq n$. Sea $x \in \mathbb{R}^m$ un punto cualquiera y sea $D \in \mathcal{M}_{n\times m}(\mathbb{R})$ la matriz de la diferencial de F en x en cualesquiera bases ortonormales de $\mathbb{R}^m$ y $\mathbb{R}^n$. Entonces, se tiene la siguiente igualdad:*

$$NJ_xF = \det(DD^t)^{1/2}.$$

Sea ahora $F : \mathbb{C}^m \longrightarrow \mathbb{C}^n$ una aplicación diferenciable, y supongamos que $m \geq n$. Sea $x \in \mathbb{C}^m$ un punto cualquiera y sea $D \in \mathcal{M}_{n\times m}(\mathbb{C})$ la matriz de la diferencial de F en x en cualesquiera bases unitarias de $\mathbb{C}^m$ y $\mathbb{C}^n$. Entonces, se tiene la siguiente igualdad:

$$NJ_xF = \det(DD^*).$$

Demostración.– Se trata de un resultado de Álgebra Lineal elemental. Haremos el caso real. Así, sea $A \in \mathcal{M}_{n\times m}$ una matriz con coordenadas reales. Sea $Ker(A) \subseteq \mathbb{R}^m$ su núcleo, y sea $H = Ker(A)^\perp \subseteq \mathbb{R}^m$ el complemento ortogonal de $Ker(A)$ en $\mathbb{R}^m$. Basta comprobar que, si A es sobreyectiva, se tiene:

$$\left|\det\left(A\mid_H\right)\right| = \det(AA^t)^{1/2},$$

donde $A\mid_H$ se identifica con su matriz coordenada en bases ortonormales cualesquiera. Esta última igualdad es inmediata a partir de la Descomposición en Valores Singulares de A. El caso complejo es muy parecido, teniendo en cuenta que cada vector de la base cuenta doble, pues él y su conjugado generan un espacio real de dimensión 2. ∎

Una de las propiedades más importantes del jacobiano normal es la siguiente "regla de la cadena".

Proposición 1.1.11 *Sean M_1, M_2, M_3 variedades Riemannianas, y sean $F : M_1 \longrightarrow M_2$ y $G : M_2 \longrightarrow M_3$ dos aplicaciones diferenciables. Sea $x \in M_1$ un punto y supongamos además que se satisface al menos una de las dos condiciones siguientes:*

- *La aplicación F es una isometría en x.*
- $\dim(M_2) = \dim(M_3)$.

Entonces, también se satisface la siguiente igualdad:

$$NJ_x(G \circ F) = NJ_{F(x)}G \; NJ_xF.$$

Demostración.– Sean m_1, m_2, m_3 las dimensiones respectivas de M_1, M_2 y M_3. Trivialmente, basta demostrar el enunciado en el caso de que $m_1 \geq m_2 \geq m_3$. Denotemos $y := F(x)$, $z := G(y) = G \circ F(x)$. Sean $\varphi_i : \mathbb{R}^{m_i} \longrightarrow U_i \subseteq M_i$ cartas locales de M_i, $i = 1, 2, 3$, en x, y, z respectivamente. Es decir, tenemos la siguiente situación.

$$\begin{array}{ccccc} U_1 & \xrightarrow{F} & U_2 & \xrightarrow{G} & U_3 \\ {\scriptstyle\varphi_1}\uparrow & & {\scriptstyle\varphi_2}\uparrow & & {\scriptstyle\varphi_3}\uparrow \\ \mathbb{R}^{m_1} & & \mathbb{R}^{m_2} & & \mathbb{R}^{m_3} \end{array}$$

Exigimos además que $\varphi_1(0) = x, \varphi_2(0) = y, \varphi_3(0) = z$, y que $d_0\varphi_i$, $i = 1, 2, 3$ sea una isometría lineal. Dado que $d_0\varphi_1$ y $d_0\varphi_2$ son isometrías lineales, se deduce fácilmente de la definición que

$$NJ_xF = NJ_0(\varphi_2^{-1} \circ F \circ \varphi_1),$$

y de igual modo

$$NJ_yG = NJ_0(\varphi_3^{-1} \circ G \circ \varphi_2), \qquad NJ_x(G \circ F) = NJ_0(\varphi_3^{-1} \circ G \circ F \circ \varphi_1),$$

Sean D_1, D_2, D_3 las matrices respectivas de las aplicaciones

$$d_0(\varphi_2^{-1} \circ F \circ \varphi_1), d_0(\varphi_3^{-1} \circ G \circ \varphi_2), d_0(\varphi_3^{-1} \circ G \circ F \circ \varphi_1)$$

en bases naturales. Por el Lema 1.1.10, deducimos que

$$NJ_xF = \det(D_1D_1^t)^{1/2}, \;\; NJ_yG = (D_2D_2^t)^{1/2}, \;\; NJ_x(G \circ F) = (D_3D_3^t)^{1/2}.$$

Además, la regla de la cadena para la diferencial nos dice que

$$D_3 = D_2D_1.$$

Por lo tanto,

$$NJ_x(G \circ F) = (D_3D_3^t)^{1/2} = \det(D_2D_1D_1^tD_2^t)^{1/2},$$

de donde se deduce de inmediato la proposición, pues si d_xF es isometría entonces $D_1D_1^t = Id_{m_1}$, y si $m_2 = m_3$ entonces D_2 es cuadrada, y se tiene que

$$\det(D_2D_1D_1^tD_2^t) = \det(D_2)\det(D_1D_1^t)\det(D_2^t) = \det(D_1D_1^t)\det(D_2D_2^t).$$

■

Una consecuencia inmediata de la Proposición 1.1.11 es el siguiente resultado, que será utilizado numerosas veces a lo largo de esta memoria.

Corolario 1.1.12 *Sean M, N dos variedades Riemannianas cualesquiera, y sea $F : M \longrightarrow N$ una aplicación diferenciable. Sean $x_1, x_2 \in M$ dos puntos regulares de F. Supongamos que existen isometrías $\varphi_M : M \longrightarrow M$ y $\varphi_N : N \longrightarrow N$ tales que $\varphi_M(x_1) = x_2$ y*

$$F \circ \varphi_M = \varphi_N \circ F.$$

Entonces, tenemos:

$$NJ_{x_1}F = NJ_{x_2}F.$$

Además, si existe una aplicación inversa $G : N \longrightarrow M$ localmente definida alrededor de $F(x) \in N$, entonces

$$NJ_xF = \frac{1}{NJ_{F(x)}G}.$$

Con las notaciones de la definición de jacobiano normal, sea $y \in N$ un valor regular de la aplicación F (esto es, F es transversal a $\{y\}$). Entonces, el Corolario 1.1.6 nos garantiza que $F^{-1}(y)$ es una variedad Riemanniana de dimensión $m - n$.
Más aún, el Teorema 1.1.4 nos garantiza que casi todo punto y de N es un punto regular de F. Por tanto, para casi todo punto y de N, podemos considerar integrales de funciones en la fibra $F^{-1}(y)$. Esta observación dota de sentido al siguiente resultado, debido a Federer. Véanse [45, 14].

Teorema 1.1.13 (Fórmula de la Co–área) *Sean M y N dos variedades Riemannianas. Sean m y n sus dimensiones respectivas, con $m \geq n$. Sea $F : M \longrightarrow N$ una aplicación diferenciable sobreyectiva tal que la diferencial d_xF es sobreyectiva para casi todo punto $x \in M$. Sea $\phi : M \longrightarrow [0, \infty]$ una función integrable. Entonces, se tienen las dos siguientes igualdades:*

$$\int_{x\in M} \phi(x)\ dM = \int_{y\in N}\int_{x\in F^{-1}(y)} \phi(x)\frac{1}{NJ_xF}\ dF^{-1}(y)\ dN,$$

$$\int_{x\in M} \phi(x)NJ_xF\ dM = \int_{y\in N}\int_{x\in F^{-1}(y)} \phi(x)\ dF^{-1}(y)\ dN.$$

Sea $\mathcal{L}$ el espacio de los círculos máximos en la esfera unidad $S^1(\mathbb{R}^k)$. Es bien sabido que $\mathcal{L}$ tiene una estructura de variedad Riemanniana, con una estructura Riemanniana natural invariante por transformaciones ortogonales (véase por ejemplo [108]). Consideremos esta estructura normalizada de modo que el volumen del total sea igual a 1. Entonces, se tiene el siguiente resultado, que es una famosa igualdad de Geometría Integral debida a Santaló, y puede deducirse directamente del Teorema 1.1.13.

Teorema 1.1.14 (Santaló, [108]) *Sea $\phi : S^1(\mathbb{R}^k) \longrightarrow [-\infty, +\infty]$ una función integrable. Entonces, se tiene la siguiente igualdad:*

$$\int_{L\in\mathcal{L}} \int_{f\in L} \phi(x)\ dL\ d\mathcal{L} = 2\pi \frac{\int_{x\in S^1(\mathbb{R}^k)} \phi(x)\ dS^1(\mathbb{R}^k)}{vol[S^1(\mathbb{R}^k)]},$$

donde $vol[S^1(\mathbb{R}^k)]$ es el volumen de la esfera $S^1(\mathbb{R}^k)$.

1.2. Volumen y distancia en el espacio proyectivo complejo

En general, dado un espacio vectorial complejo W con producto hermitiano $\langle\cdot,\cdot\rangle_W : W \times W \longrightarrow \mathbb{C}$, denotaremos por $(W, \langle\cdot,\cdot\rangle_W)$ el espacio hermitiano correspondiente. La norma en $(W, \langle\cdot,\cdot\rangle_W)$ se denotará normalmente $\|\cdot\|_W$. En el caso de que $W = \mathbb{C}^{n+1}$, denotamos simplemente por $\langle\cdot,\cdot\rangle_2$ el producto hermitiano usual, y por $\|\cdot\|_2$ la norma usual asociada. Diremos que un conjunto de vectores $S = \{v_1, \ldots, v_s\} \in W$ son ortogonales si $\langle v_i, v_j\rangle_W = 0,\ i \neq j$. Diremos que S es una familia unitaria si además de ser ortogonales, sus vectores satisfacen $\|v_i\|_W = 1$ para $1 \leq i \leq s$. También utilizaremos los términos ortogonal y ortonormal para el caso de espacios vectoriales reales. Una de las propiedades más importantes del producto hermitiano usual $\langle,\rangle_2$ en $\mathbb{C}^{n+1}$ es la invariancia unitaria. Esto es, para todo par de puntos $\underline{x}, \underline{y} \in \mathbb{C}^{n+1}$ y para toda matriz unitaria U de tamaño $n+1$, se tiene:

$$\langle \underline{x}, \underline{y}\rangle_2 = \langle U\underline{x}, U\underline{y}\rangle_2.$$

El grupo de matrices unitarias y en general el grupo de matrices jugarán un papel clave en varios de los resultados que obtendremos; por ello conviene fijar algunas notaciones. Dados dos naturales positivos $n_1, n_2 \in \mathbb{N}$, denotaremos por $\mathcal{M}_{n_1\times n_2}(\mathbb{C})$ el espacio vectorial complejo formado por las matrices de talla $n_1 \times n_2$ con coeficientes complejos. En el caso de que $n_1 = n_2$ (matrices cuadradas), escribiremos simplemente $\mathcal{M}_{n_1}(\mathbb{C})$. Consideraremos en $\mathcal{M}_{n_1\times n_2}(\mathbb{C})$ el producto interno y la norma de Frobenius, definidos como sigue. Para todo par de matrices $M_1, M_2 \in \mathcal{M}_{n_1\times n_2}(\mathbb{C})$, definimos

$$\langle M_1, M_2\rangle_F := tr(M_1 M_2^*) = tr(M_2 M_1^*), \quad \|M_1\|_F := \langle M_1, M_1\rangle_F^{1/2},$$

donde $tr(\cdot)$ es la traza y M_i^* es la matriz conjugada traspuesta de M_i, $i = 1, 2$. Este producto interior y esta norma son los que resultan de considerar las matrices en $\mathcal{M}_{n_1 \times n_2}(\mathbb{C})$ como vectores en $\mathbb{C}^{n_1 n_2}$. Como es habitual, la norma como operador lineal de $M \in \mathcal{M}_{n_1 \times n_2}(\mathbb{C})$ se denotará $\|M\|_2$. Esto es,

$$\|M\|_2 := \sup_{v \in \mathbb{C}^{n_2}} \|Mv\|_2.$$

Dado un natural $n \geq 0$, denotaremos por $\mathcal{U}_{n+1}$ el grupo de las matrices unitarias de talla $n+1$. Esto es,

$$\mathcal{U}_{n+1} := \{U \in \mathcal{M}_{n+1} : UU^* = Id_{n+1}\},$$

donde U^* es la matriz conjugada traspuesta de U y Id_{n+1} la matriz unidad. Es bien sabido que $\mathcal{U}_{n+1}$ es una subvariedad diferenciable real de $\mathcal{M}_{n+1}(\mathbb{C})$ de dimensión (real) $(n+1)^2$. La estructura Riemanniana que consideraremos en $\mathcal{U}_{n+1}$ es la heredada de $\mathcal{M}_{n+1}(\mathbb{C})$, pero normalizada de modo que el volumen de $\mathcal{U}_{n+1}$ es igual a 1. Dado un conjunto medible $T \subseteq \mathcal{U}_{n+1}$, denotaremos por $\nu_{\mathcal{U}_{n+1}}[T]$ su volumen. Para todo elemento $U \in \mathcal{U}_{n+1}$, las dos aplicaciones siguientes son isometrías:

$$\begin{array}{rccc} U_L : & \mathcal{U}_{n+1} & \longrightarrow & \mathcal{U}_{n+1}, \\ & U' & \mapsto & UU' \end{array} \qquad \begin{array}{rccc} U_R : & \mathcal{U}_{n+1} & \longrightarrow & \mathcal{U}_{n+1} \\ & U' & \mapsto & U'U. \end{array}$$

En esta memoria nos referiremos normalmente a la acción por la izquierda U_L y denotaremos simplemente por $U = U_L : \mathcal{U}_{n+1} \longrightarrow \mathcal{U}_{n+1}$ esa aplicación. Denotaremos mediante $B_{\mathbb{C}^{n+1}}(\underline{x}, \varepsilon)$ la bola abierta de radio ε centrada en $\underline{x}$. Esto es,

$$B_{\mathbb{C}^{n+1}}(\underline{x}, \varepsilon) := \{\underline{y} \in \mathbb{C}^{n+1} : \|\underline{x} - \underline{y}\|_2 < \varepsilon\}.$$

También escribiremos $S^{\varepsilon}(\mathbb{C}^{n+1}) = \partial B_{\mathbb{C}^{n+1}}(0, \varepsilon)$ para denotar la esfera de radio ε en $\mathbb{C}^{n+1}$. En otras palabras,

$$S^{\varepsilon}(\mathbb{C}^{n+1}) := \{\underline{x} \in \mathbb{C}^{n+1} : \|\underline{x}\|_2 = \varepsilon\}.$$

En el caso de la esfera centrada en 0 de radio 1, escribiremos simplemente $S(\mathbb{C}^{n+1}) := S^1(\mathbb{C}^{n+1})$. Consideraremos la variedad diferenciable $S(\mathbb{C}^{n+1})$ equipada con la estructura Riemanniana heredada de la de $\mathbb{C}^{n+1}$. Sea $\mathbb{P}_n(\mathbb{C})$ el espacio proyectivo complejo de dimensión (compleja) n. Esto es, $\mathbb{P}_n(\mathbb{C}) := \mathbb{P}(\mathbb{C}^{n+1})$ es el conjunto de las clases de $\mathbb{C}^{n+1} \setminus \{0\}$ mediante la relación de equivalencia

$$\underline{x} \sim \underline{y} \Leftrightarrow \exists \lambda \in \mathbb{C} : \underline{x} = \lambda \underline{y}, \quad \underline{x}, \underline{y} \in \mathbb{C}^{n+1} \setminus \{0\}.$$

(véase por ejemplo [48] para definiciones y conceptos básicos del espacio proyectivo complejo). Para todo conjunto $A \subseteq \mathbb{P}_n(\mathbb{C})$ y para toda matriz

unitaria $U \in \mathcal{U}_{n+1}$ consideraremos el conjunto UA desplazado de A por U. En otras palabras,

$$UA := \{y \in \mathbb{P}_n(\mathbb{C}) : \exists x \in A : Ux = y\}.$$

También consideraremos la proyección canónica

$$\begin{array}{rccc} \pi: & \mathbb{C}^{n+1} \setminus \{0\} & \longrightarrow & \mathbb{P}_n(\mathbb{C}) \\ & \underline{x} & \mapsto & \{\underline{y} : \underline{y} = \lambda \underline{x},\ \lambda \in \mathbb{C}\}. \end{array}$$

A partir de un producto hermitiano cualquiera en $\mathbb{C}^{n+1}$ se define de modo natural una estructura de variedad Riemanniana en $\mathbb{P}_n(\mathbb{C})$. Este paso está descrito con detalle en el Apéndice A, aunque con el objetivo de facilitar la lectura recordamos aquí los hechos más importantes, para el caso del producto hermitiano usual.
Consideramos la restricción $p := \pi\,|_{S(\mathbb{C}^{n+1})}: S(\mathbb{C}^{n+1}) \longrightarrow \mathbb{P}_n(\mathbb{C})$, denominada normalmente la fibración de Hopf. Entonces, la estructura Riemanniana usual de $\mathbb{P}_n(\mathbb{C})$ es la única estructura que existe tal que p es una submersion Riemanniana, esto es, p es una submersión y para todo $\underline{x} \in S(\mathbb{C}^{n+1})$, $d_{\underline{x}}p$ define una isometría entre el espacio horizontal H_x de p en x, $H_x = (d_{\underline{x}}p)^{-1}(0)$, y $T_x\mathbb{P}_n(\mathbb{C})$ (véase por ejemplo [48, Prop. 2.28, ex. 2.29]).
En estas circunstancias, el espacio tangente en cada punto $x \in \mathbb{P}_n(\mathbb{C})$ se identifica con $x^\perp := \{y \in \mathbb{C}^{n+1} : \langle x, y\rangle_2 = 0\}$, equipado con el producto hermitiano heredado de $\mathbb{C}^{n+1}$. Esta simple observación facilita el cálculo de diferenciales y jacobianos. Una forma más concreta de expresar esta idea es la siguiente: Elijamos un representante concreto $\underline{x}$ de x, tal que $\|\underline{x}\|_2 = 1$. Consideremos el difeomorfismo

$$\begin{array}{rccc} \varphi_{\underline{x}}: & x^\perp & \longrightarrow & \mathbb{P}_n(\mathbb{C}) \setminus x^\perp. \\ & y & \mapsto & \underline{x} + y \end{array} \tag{1.1}$$

Entonces, se tiene que $\varphi_{\underline{x}}$ es una isometría en 0. De este modo, podemos identificar totalmente $T_x\mathbb{P}_n(\mathbb{C})$ con $x^\perp$.
El espacio $\mathbb{P}_n(\mathbb{C})$ hereda la invariancia unitaria de su correspondiente afín. En otras palabras, para toda matriz $U \in \mathcal{U}_{n+1}$, la siguiente aplicación es una isometría.

$$\begin{array}{ccc} \mathbb{P}_n(\mathbb{C}) & \longrightarrow & \mathbb{P}_n(\mathbb{C}) \\ x & \mapsto & Ux. \end{array}$$

Los elementos de $\mathbb{P}_n(\mathbb{C})$ son representados normalmente por sus *coordenadas homogéneas*, que se definen como sigue: Si $x \in \mathbb{P}_n(\mathbb{C})$ es la clase proyectiva del punto $\underline{x} = (x_0, \ldots, x_n)$, las coordenadas homogéneas de x son $(x_0 : \cdots : x_n)$. Así, se tiene la siguiente igualdad

$$(x_0 : \cdots : x_n) = (\lambda x_0 : \cdots : \lambda x_n), \quad \forall\, 0 \neq \lambda \in \mathbb{C}$$

Frecuentemente consideraremos el punto proyectivo complejo e_0 definido como sigue,

$$e_0 := (1 : 0 : \cdots : 0) \in \mathbb{P}_n(\mathbb{C}).$$

Es bien sabido que la distancia Riemanniana, también conocida como distancia de Fubini–Study, entre dos puntos cualesquiera en el espacio proyectivo complejo viene dada por la fórmula (cf. por ejemplo [14]):

$$d_R(x, y) := \arccos \frac{|\langle \underline{x}, \underline{y} \rangle_2|}{\|\underline{x}\|_2 \|\underline{y}\|_2},$$

donde $\underline{x}, \underline{y}$ son respectivamente representantes afines cualesquiera de x e y. Denotaremos por $d_{\mathbb{P}}$ la distancia proyectiva, que se define como el seno de la distancia Riemanniana. Es decir,

$$d_{\mathbb{P}}(x, y) = \sin d_R(x, y) = \sqrt{1 - \frac{|\langle \underline{x}, \underline{y} \rangle_2|^2}{\|\underline{x}\|_2^2 \|\underline{y}\|_2^2}}.$$

A lo largo de esta memoria, utilizaremos con mucha más frecuencia la distancia proyectiva $d_{\mathbb{P}}$ que la distancia Riemanniana, por motivos estéticos. Denotaremos por $B_{\mathbb{P}}(x, \varepsilon) \subseteq \mathbb{P}_n(\mathbb{C})$ la bola abierta de radio ε centrada en x con respecto a $d_{\mathbb{P}}$. En otras palabras,

$$B_{\mathbb{P}}(x, \varepsilon) := \{y \in \mathbb{P}_n(\mathbb{C}) : d_{\mathbb{P}}(x, y) < \varepsilon\}.$$

Dada una subvariedad diferenciable compleja $M \subset \mathbb{P}_n(\mathbb{C})$ de dimensión compleja m, la consideraremos naturalmente equipada con la estructura riemanniana heredada de $\mathbb{P}_n(\mathbb{C})$.
Para un subconjunto medible $A \subseteq \mathbb{P}_n(\mathbb{C})$, denotaremos simplemente por $\nu_n[A]$ el volumen de A. Esto es,

$$\nu_n[A] := \nu_{\mathbb{P}_n(\mathbb{C})}[A].$$

El espacio proyectivo complejo tiene un volumen conocido, que será denotado por ϑ_n y viene dado por la fórmula

$$\vartheta_n := \nu_n[\mathbb{P}_n(\mathbb{C})] = \frac{\pi^n}{n!} = \frac{\pi^n}{\Gamma(n+1)}, \quad n \geq 0 \tag{1.2}$$

donde Γ es la función Gamma de Euler. Si consideramos $\mathbb{P}_m(\mathbb{C})$, $m < n$, como una subvariedad de $\mathbb{P}_n(\mathbb{C})$ (esto es, un subespacio proyectivo lineal de dimensión compleja m de $\mathbb{P}_n(\mathbb{C})$), entonces su volumen como subvariedad coincide con su volumen como espacio proyectivo complejo.
Desde [124], tenemos una fórmula explícita para el volumen de $B_{\mathbb{P}}(x, \varepsilon)$ (véase [24] para una referencia más moderna). En efecto, se tiene:

$$\nu_n[B_{\mathbb{P}}(x, \varepsilon)] = \vartheta_n \varepsilon^{2n}.$$

1.3. Notaciones de Geometría Algebraica.

Una variedad algebraica proyectiva, o simplemente una variedad proyectiva, es un subconjunto del espacio proyectivo complejo $\mathbb{P}_n(\mathbb{C})$ que se puede expresar como el conjunto de soluciones proyectivas de un conjunto finito de polinomios homogéneos. Recomendamos al lector las referencias [109, 96] para generalidades sobre variedades proyectivas.
Se puede definir una topología en $\mathbb{P}_n(\mathbb{C})$ cuyos conjuntos cerrados son exactamente las variedades proyectivas. Esta topología se conoce como la topología de Zariski. Con estas definiciones, decimos que un subconjunto de $\mathbb{P}_n(\mathbb{C})$ es una variedad casi–proyectiva si es un subconjunto abierto Zariski de una variedad proyectiva (véase [109] para más detalles al respecto). Dado un subconjunto $A \subseteq \mathbb{P}_n(\mathbb{C})$, la clausura Zariski de A es la variedad proyectiva más pequeña que contiene a A.
Sea $V \subseteq \mathbb{P}_n(\mathbb{C})$ una variedad casi–proyectiva. Denotaremos por $\dim(V)$ su dimensión de Krull, y diremos que $n - \dim(V)$ es la codimensión de V. Un punto liso $a \in V$ es un punto tal que el germen V_a de V en a es una subvariedad diferenciable compleja de $\mathbb{P}_n(\mathbb{C})$ de dimensión (compleja) igual a $\dim(V)$. De otro modo, un punto $a \in V$ es liso si existe un entorno U_a (para la topología usual) de A en V tal que U_a es una variedad diferenciable de dimensión $\dim(V)$. Denotamos por $Reg(V)$ el conjunto de todos los puntos simples de V. La clausura Zariski de $Reg(V)$ es igual a la unión de todas las componentes irreducibles de la clausura Zariski de V de dimensión igual a $\dim(V)$. Con estas notaciones, existe una variedad proyectiva $V_1 \subseteq \mathbb{P}_n(\mathbb{C})$ tal que $\dim(V_1) < \dim(V)$ y tal que se satisface la siguiente igualdad:

$$Reg(V) \setminus V_1 = V \setminus V_1.$$

Se dice que V y $Reg(V)$ son genéricamente iguales. Sea $V \subseteq \mathbb{P}_n(\mathbb{C})$ una variedad casi–proyectiva de dimensión m. Entonces, $Reg(V) \subseteq \mathbb{P}_n(\mathbb{C})$ es una variedad diferenciable compleja de dimensión m, equipada con la estructura Riemanniana heredada de la de $\mathbb{P}_n(\mathbb{C})$. Para todo subconjunto $A \subseteq V$ tal que $A \cap Reg(V)$ es medible en $Reg(V)$, definimos el volumen m–dimensional complejo de A mediante la siguiente identidad,

$$\nu_m[A] := \nu_{Reg(V)}[A \cap Reg(V)].$$

De modo similar, podemos considerar la integral $\int_A f \; dV$ de una función integrable $f : A \longrightarrow \mathbb{R}$.
La noción de grado geométrico (o simplemente grado) de una variedad proyectiva $V \subseteq \mathbb{P}_n(\mathbb{C})$ es una noción clásica que viene desde los orígenes, en el siglo XIX, de la Teoría de Eliminación. La principal propiedad que satisface cualquier noción de grado es la desigualdad de Bézout (véase [66]). Distintas pruebas de esta desigualdad se pueden encontrar también en [130, 47].
Sea $W \subseteq \mathbb{P}_n(\mathbb{C})$ un subconjunto abierto Zariski de una variedad proyectiva

irreducible $V \subseteq \mathbb{P}_n(\mathbb{C})$ de dimensión m. El grado de W, $\deg(W)$ se define como la cantidad

$$\text{máx}\{\sharp(L \cap W) : L \subseteq \mathbb{P}_n(\mathbb{C}) \text{ lineal}, \ \dim(L) = n - m, \sharp(L \cap W) < +\infty\}.$$

Se tiene que $\deg(W) = \deg(V)$ para cualquier abierto Zariski W de la variedad irreducible V. Si $V \subseteq \mathbb{P}_n(\mathbb{C})$ es una variedad cualquiera, $\deg(V)$ se define como la suma de los grados de sus componentes irreducibles. De modo similar, para todo subconjunto constructible $C \subset \mathbb{P}_n(\mathbb{C})$ podemos definir $\deg(C)$ como la suma de los grados de sus componentes irreducibles localmente cerradas (véase [66] para algunas ideas al respecto). Esta noción de grado satisface la desigualdad de Bézout para subconjuntos localmente cerrados (con respecto a la topología de Zariski) de $\mathbb{P}_n(\mathbb{C})$ (véase [66]), es decir:

$$\deg(W_1 \cap W_2) \leq \deg(W_1) \deg(W_2),$$

para W_1 y W_2 subconjuntos localmente cerrados. El siguiente resultado, que se sigue inmediatamente de la definición de grado que acabamos de introducir, muestra la relación entre el grado y el grupo de matrices unitarias.

Proposición 1.3.1 *Sea $V \subseteq \mathbb{P}_n(\mathbb{C})$ una variedad proyectiva equi–dimensional de dimensión m. Sea $L \subseteq \mathbb{P}_n(\mathbb{C})$ un subespacio proyectivo lineal cualquiera de dimensión $n - m$. Entonces, tenemos:*

$$\deg(V) = \max_{U \in \mathcal{U}_{n+1}} \{\sharp(UL \cap V) : \sharp(UL \cap V) < +\infty\}.$$

1.4. El espacio de los sistemas de ecuaciones polinomiales

A lo largo de esta memoria, trataremos con frecuencia el espacio de sistemas de ecuaciones. En esta ocasión, utilizaremos dos estructuras Riemannianas distintas, y trabajaremos en el espacio vectorial de sistemas, en el espacio proyectivo asociado y en la esfera correspondiente a cada una de las estructuras que vamos a utilizar.
Comenzamos introduciendo algunas notaciones para fijar estos conceptos. Sean $n, l \in \mathbb{N}$ dos números naturales positivos, y consideremos elegido un orden cualquiera en el conjunto

$$\mathcal{E}_l := \{\alpha := (\alpha_0, \ldots, \alpha_n) \in \mathbb{N}^{n+1} : \alpha_0 + \cdots + \alpha_n = l\}.$$

Entonces, denotamos por H_l el siguiente conjunto:

$$H_l := \prod_{\alpha \in \mathcal{E}_l} \mathbb{C}.$$

Obsérvese que H_l puede identificarse con el conjunto de polinomios homogéneos de grado l en $\mathbb{C}[X_0, \ldots, X_n]$. En efecto, para un elemento $j = (a_\alpha)_{\alpha \in \mathcal{E}_l} \in H_l$, podemos considerar el polinomio homogéneo

$$\sum_{\alpha=(\alpha_0,\ldots,\alpha_n)\in\mathcal{E}_l} a_\alpha X_0^{\alpha_0} \cdots X_n^{\alpha_n}.$$

Otra forma de interpretar H_l es como el conjunto de polinomios (no necesariamente homogéneos) de grado l en $\mathbb{C}[X_1, \ldots, X_n]$. En efecto, para un elemento $j = (a_\alpha)_{\alpha \in \mathcal{E}_l} \in H_l$, podemos considerar el polinomio homogéneo

$$\sum_{\alpha=(\alpha_0,\ldots,\alpha_n)\in\mathcal{E}_l} a_\alpha X_1^{\alpha_1} \cdots X_n^{\alpha_n}.$$

Esto es, podemos considerar los elementos en H_l como polinomios homogéneos en $n+1$ variables o como polinomios cualesquiera en n variables, en ambos casos, dados por su lista de coeficientes en codificación densa y para el orden (monomial) asociado al elegido en el conjunto $\mathcal{E}_l$. Ambas interpretaciones de este espacio serán importantes en esta memoria. Obsérvese que H_l es un espacio vectorial complejo de dimensión

$$\dim(H_l) = \binom{n+l}{l}.$$

Por lo tanto, tenemos definido el producto hermitiano usual en H_l, que denotaremos como de costumbre por $\langle \cdot, \cdot \rangle_2$.
Sean ahora $n \geq m \in \mathbb{N}$ dos números naturales positivos, y sea $(d) := (d_1, \ldots, d_m) \in \mathbb{N}^m$ una m–tupla de números naturales positivos. Definimos el espacio $\mathcal{H}_{(d)}^m$ como sigue:

$$\mathcal{H}_{(d)}^m := \prod_{i=1}^m H_{d_i}.$$

Muy habitualmente nos referiremos al caso cero–dimensional, esto es, cuando $m = n$. En ese caso, escribiremos simplemente $\mathcal{H}_{(d)} = \mathcal{H}_{(d)}^n$, para simplificar la notación.
Podemos interpretar $\mathcal{H}_{(d)}^m$ de dos maneras distintas:

- Como el espacio de sistemas de m ecuaciones polinomiales homogéneas de grados respectivos $d_1, \ldots, d_m$, con coeficientes complejos e incógnitas $X_0, \ldots, X_n$. Equivalentemente, podemos ver $\mathcal{H}_{(d)}^m$ como el espacio de aplicaciones polinomiales $f : \mathbb{C}^{n+1} \longrightarrow \mathbb{C}^m$, de forma que $f = [f_1, \ldots, f_m]$ y f_i es una aplicación polinomial homogénea de grado d_i. De este modo, tenemos:

$$\begin{array}{rccc} f: & \mathbb{C}^{n+1} & \longrightarrow & \mathbb{C}^m, \\ & x & \mapsto & (f_1(x), \ldots, f_m(x)) \end{array}$$

  manda cada $x \in \mathbb{C}^{n+1}$ al vector de los polinomios f_i evaluados en x.

- Como el espacio de sistemas de m ecuaciones polinomiales (no necesariamente homogéneas) de grados respectivos $d_1, \ldots, d_m$, con coeficientes complejos e incógnitas $X_1, \ldots, X_n$. Equivalentemente, podemos ver $\mathcal{H}^m_{(d)}$ como el espacio de aplicaciones polinomiales $f : \mathbb{C}^n \longrightarrow \mathbb{C}^m$, de forma que $f = [f_1, \ldots, f_m]$ y f_i es una aplicación polinomial (no necesariamente homogénea) de grado d_i. De este modo, tenemos:

$$\begin{array}{rccc} f: & \mathbb{C}^n & \longrightarrow & \mathbb{C}^m, \\ & x & \mapsto & (f_1(x), \ldots, f_m(x)) \end{array}$$

A partir de ahora, denotaremos los elementos de $\mathcal{H}^m_{(d)}$ por las letras $f, g, h \ldots$, y diremos que f, g, h son sistemas de ecuaciones, o simplemente sistemas. Escribiremos $f = [f_1, \ldots, f_m] \in \mathcal{H}^m_{(d)}$ cuando queramos señalar que $f_i \in H_{d_i}, 1 \leq i \leq m$ son los polinomios que contiene el sistema f. La primera de las interpretaciones mencionadas permite la siguiente construcción: Sea $A \in \mathcal{M}_{n+1}(\mathbb{C})$ una matriz cualquiera, y sea $f \in \mathcal{H}^m_{(d)}$ un sistema. Entonces, podemos considerar el sistema

$$\begin{array}{rccc} f \circ A: & \mathbb{C}^{n+1} & \longrightarrow & \mathbb{C}^m, \\ & x & \mapsto & (f_1(Ax), \ldots, f_m(Ax)) \end{array}$$

esto es la aplicación polinomial obtenida al componer f con A. Entonces, $f \circ A$ es un elemento de $\mathcal{H}^m_{(d)}$.
Denotamos por $N+1$ la dimensión compleja de $\mathcal{H}^m_{(d)}$. Es decir

$$N + 1 := \sum_{i=1}^{m} \binom{n + d_i}{d_i}.$$

Así, podemos identificar $\mathcal{H}^m_{(d)} \equiv \mathbb{C}^{N+1}$. Una vez fijada una lista de grados $(d) = (d_1, \ldots, d_m) \in \mathbb{N}^m$, denotaremos por d y $\mathcal{D}$ el máximo de los grados y el número de Bézout asociado a (d), respectivamente. Esto es,

$$d := \max_{1 \leq i \leq m} \{d_i\}, \quad \mathcal{D} := \prod_{i=1}^{m} d_i.$$

Podemos considerar $\mathcal{H}^m_{(d)}$ equipado con el producto hermitiano usual de $\mathbb{C}^{N+1}$, que puede también ser definido como sigue: Sean $f := [f_1, \ldots, f_m]$ y $g := [g_1, \ldots, g_m]$ dos sistemas de ecuaciones en $\mathcal{H}^m_{(d)}$. Entonces,

$$\langle f, g \rangle_2 := \sum_{i=1}^{m} \langle f_i, g_i \rangle_2.$$

Denotaremos mediante $(\mathcal{H}^m_{(d)}, can)$ el espacio hermitiano formado por $\mathcal{H}^m_{(d)}$ con el producto usual. También consideraremos el espacio proyectivo asociado $(\mathbb{P}(\mathcal{H}^m_{(d)}), can)$, equipado con la estructura Riemanniana usual, heredada

del producto usual de la manera descrita en la Sección 1.2 (o con más detalle, en el Apéndice A). Finalmente, consideraremos también la esfera de centro 0 y radio 1 en $\mathcal{H}^m_{(d)}$ para la norma usual, que será denotada simplemente por $\mathbb{S}$.

Sin embargo, la estructura Riemanniana usual, que acabamos de describir, será utilizada tan sólo en resultados técnicos intermedios, y no es la estructura principal que utilizaremos para nuestros resultados. Este lugar corresponde a la estructura comúnmente llamada de Kostlan. Se trata de una estructura utilizada en la literatura clásica que tiene numerosas propiedades, re–descubierta en varias ocasiones por Bombieri, Kostlan y Shub & Smale (véase por ejemplo [113, 112, 14, 30, 88, 37]). Describimos a continuación con detalle esta estructura.

Para cada i, $1 \leq i \leq m$, consideramos el producto hermitiano en H_{d_i} definido como sigue: Para cada $j_1, j_2 \in H_{d_i}$,

$$j_1 = (a_\alpha)_{\alpha \in \mathcal{E}_{d_i}}, \;\; j_2 = (b_\alpha)_{\alpha \in \mathcal{E}_{d_i}}, \quad a_\alpha, b_\alpha \in \mathbb{C}, \; \forall \alpha \in \mathcal{E}_{d_i}$$

definimos:

$$\langle j_1, j_2 \rangle_{\Delta_i} := \sum_{\alpha \in \mathcal{E}_{d_i}} \binom{d_i}{\alpha}^{-1} a_\alpha \overline{b_\alpha},$$

donde $\overline{b_\alpha}$ es el conjugado complejo b_α y $\binom{d_i}{\alpha}$ es el conocido como coeficiente multinomial. Esto es,

$$\binom{d_i}{\alpha} = \frac{d_i!}{\alpha_0! \cdots \alpha_n!} \in \mathbb{N}.$$

El producto hermitiano que acabamos de definir induce un producto hermitiano en $\mathcal{H}^m_{(d)}$ definido del modo descrito a continuación. Para dos elementos cualesquiera $f := [f_1, \ldots, f_m], g := [g_1, \ldots, g_m]$ de $\mathcal{H}^m_{(d)}$, definimos

$$\langle f, g \rangle_\Delta := \sum_{i=1}^m \langle f_i, g_i \rangle_{\Delta_i}.$$

La propiedad principal del producto hermitiano $\langle \cdot, \cdot \rangle_\Delta$ es la invariancia unitaria, que podemos expresar como sigue. Sean $f, g \in \mathcal{H}^m_{(d)}$ dos sistemas. Sea $U \in \mathcal{U}_{n+1}$ una matriz unitaria cualquiera. Consideremos los elementos $f \circ U, g \circ U \in \mathcal{H}^m_{(d)}$. Entonces, se tiene la siguiente igualdad (véase por ejemplo [14, Teor. 1, pág 218]):

$$\langle f \circ U, g \circ U \rangle_\Delta = \langle f, g \rangle_\Delta.$$

Este producto hermitiano induce una norma $\| \cdot \|_\Delta$ y una estructura Riemanniana en el espacio $\mathcal{H}^m_{(d)}$, llamada por Shub & Smale *estructura de Kostlan*. Dado que es la estructura que utilizaremos más frecuentemente, denotaremos simplemente por $\mathcal{H}^m_{(d)}$ el espacio con la estructura de Kostlan. También

denotaremos por $\mathbb{S}_\Delta$ la esfera para la norma $\|\cdot\|_\Delta$ en $\mathcal{H}^m_{(d)}$, y consideraremos $\mathbb{S}_\Delta$ equipada con la estructura Riemanniana heredada de $\mathcal{H}^m_{(d)}$. Finalmente, el espacio proyectivo complejo $\mathbb{P}(\mathcal{H}^m_{(d)})$ hereda también una estructura Riemanniana de $\mathcal{H}^m_{(d)}$, mediante el proceso que hemos descrito en el Apéndice A, que nos permite identificar, para cada sistema proyectivo $f \in \mathbb{P}(\mathcal{H}^m_{(d)})$, el espacio tangente $T_f\mathbb{P}(\mathcal{H}^m_{(d)})$ con el ortogonal

$$f^\perp := \{g \in \mathcal{H}^m_{(d)} : \langle f, g\rangle_\Delta = 0\},$$

equipado con el producto hermitiano heredado de $\mathcal{H}^m_{(d)}$. Sea $\underline{f}$ un representante afín cualquiera de f, tal que $\|\underline{f}\|_\Delta$=1, y consideremos la carta afín

$$\begin{array}{rccc} \varphi_{\underline{f}}: & f^\perp & \longrightarrow & \mathbb{P}(\mathcal{H}^m_{(d)}) \setminus f^\perp. \\ & g & \mapsto & \underline{f} + g \end{array} \tag{1.3}$$

Entonces, $\varphi_{\underline{f}}$ es un difeomorfismo. Además, $\varphi_{\underline{f}}$ es una isometría en 0. Denotaremos simplemente por $\mathbb{P}(\mathcal{H}^m_{(d)})$ el espacio proyectivo con la estructura Riemanniana que acabamos de describir, que será llamada también estructura de Kostlan.

Gran parte de los resultados de esta memoria consisten en el análisis de ciertas propiedades relacionadas con los sistemas $f \in \mathcal{H}^m_{(d)}$. Además, estas propiedades son siempre invariantes bajo multiplicación por un escalar no nulo. Por ejemplo, el conjunto de soluciones de un sistema f y el conjunto de soluciones del sistema asociado λf, para cualquier $\lambda \in \mathbb{C}\setminus\{0\}$, coinciden. Por este motivo, es natural el uso en $\mathcal{H}^m_{(d)}$ de cualquier medida de probabilidad que sea simétrica respecto al origen (es decir, invariante bajo multiplicación por escalar no nulo). En particular, por ejemplo, la medida Gaussiana. Otra forma de expresar esto es calculando las probabilidades y esperanzas en el espacio proyectivo $\mathbb{P}(\mathcal{H}^m_{(d)})$ o la esfera asociada $\mathbb{S}_\Delta$. De ahora en adelante, tendremos esto en cuenta a la hora de calcular probabilidades y esperanzas en el espacio de sistemas. Por ello, las siguientes expresiones son equivalentes para toda función integrable $\phi : \mathbb{P}(\mathcal{H}^m_{(d)}) \longrightarrow [-\infty, +\infty]$:

$$\mathrm{E}_{f\in\mathcal{H}^m_{(d)}}[\phi(f)] \equiv \mathrm{E}_{f\in\mathbb{S}_\Delta}[\phi(f)] \equiv \mathrm{E}_{f\in\mathbb{P}(\mathcal{H}^m_{(d)})}[\phi(f)], \tag{1.4}$$

donde E significa esperanza.

También denotaremos por Δ la matriz (que se suele llamar matriz de Kostlan) asociada al producto hermitiano $\langle\cdot,\cdot\rangle_\Delta$. Esto es, Δ es la matriz diagonal cuyas entradas vienen dadas por la fórmula

$$\Delta := Diagonal\left(\binom{d_i}{\alpha}^{-1/2}\right)_{\substack{1\le i\le m \\ \alpha\in\mathcal{E}_{d_i}}},$$

respetando el orden elegido para cada $\mathcal{E}_{d_i}$. De este modo, se tiene la siguiente igualdad:

$$\langle f, g\rangle_\Delta = \langle \Delta f, \Delta g\rangle_2, \ \ \forall f, g \in \mathcal{H}^m_{(d)}.$$

El siguiente resultado es una consecuencia trivial del Lema A.0.7 del Apéndice A, y se obtiene directamente a partir de la definición de la estructura Riemanniana del proyectivo (véase por ejemplo [23]). Nos permite relacionar las estructuras de Kostlan y canónica.

Lema 1.4.1 *Con las notaciones anteriores, las dos aplicaciones siguientes son isometrías.*

$$\begin{array}{ccccccc} (\mathcal{H}^m_{(d)}, can) & \longrightarrow & \mathcal{H}^m_{(d)}, & \mathbb{S} & \longrightarrow & \mathbb{S}_\Delta. \\ f & \mapsto & \Delta^{-1} f & f & \mapsto & \Delta^{-1} f \end{array}$$

También es isometría la siguiente aplicación:

$$\begin{array}{ccc} (\mathbb{P}(\mathcal{H}^m_{(d)}), can) & \longrightarrow & \mathbb{P}(\mathcal{H}^m_{(d)}), \\ f & \mapsto & \Delta^{-1} f \end{array}$$

donde denotamos por $\Delta^{-1}f$ la clase proyectiva de cualquier punto $\Delta^{-1}\underline{f}$, con $\underline{f}$ un representante afín cualquiera de f.

Para un sistema dado $f = [f_1, \ldots, f_m] \in \mathcal{H}^m_{(d)}$, denotaremos por $V(f)$ el conjunto de soluciones proyectivas de f.

$$V(f) := \{x \in \mathbb{P}_n(\mathbb{C}) : f_i(x) = 0,\ 1 \leq i \leq m\} \subseteq \mathbb{P}_n(\mathbb{C}).$$

Por tanto, el conjunto de ceros de un sistema $f \in \mathcal{H}^m_{(d)}$ (ó $f \in \mathbb{P}(\mathcal{H}^m_{(d)})$) es una variedad proyectiva $V(f)$.
A veces estamos más interesados en el estudio de las variedades afines. Dado un sistema $f \in \mathcal{H}^m_{(d)}$, la variedad afín asociada a f es un conjunto $V_{\mathbb{C}^n}(f) \subseteq \mathbb{C}^n$ definido como sigue:

$$V_{\mathbb{C}^n}(f) := \{(x_1, \ldots, x_n) \in \mathbb{C}^n : (1 : x_1 \cdots : x_n) \in V(f)\} \subseteq \mathbb{C}^n.$$

Dicho de otra manera, $V_{\mathbb{C}^n}(f)$ es la pre–imagen de $V(f)$ mediante la aplicación

$$\varphi_0 := \varphi_{(1,0,\ldots,0)} : \begin{array}{ccc} \mathbb{C}^n & \longrightarrow & \mathbb{P}_n(\mathbb{C}). \\ (x_1, \ldots, x_n) & \mapsto & (1 : x_1 : \ldots : x_n) \end{array} \tag{1.5}$$

El conjunto de soluciones afines $V_{\mathbb{C}^n}(f)$ es genéricamente una variedad de dimensión $n - m$, al igual que $V(f)$. Los conjuntos de soluciones $V_{\mathbb{C}^n}(f)$ y $V(f)$ se relacionan del modo que sigue:

$$x \in V_{\mathbb{C}^n}(f) \Longrightarrow \varphi_0(x) \in V(f),$$

$$z = (z_0 : \cdots : z_n) \in V(f), z_0 \neq 0 \Longrightarrow \varphi_0^{-1}(z) \in V_{\mathbb{C}^n}(f).$$

Dado un elemento $f \in \mathcal{H}_{(d)}^m$, podemos considerar por tanto su conjunto de soluciones proyectivas o afines. Esta elección depende de cómo interpretemos f: Si lo interpretamos como sistema de ecuaciones homogéneas en $X_0, \ldots, X_n$, entonces normalmente buscaremos soluciones proyectivas, esto es, soluciones en $V(f)$. Si por el contrario queremos interpretar f como un sistema de ecuaciones (no necesariamente homogéneas) en $X_1, \ldots, X_n$, entonces las soluciones que buscamos son los puntos de $V_{\mathbb{C}^n}(f)$.
La *variedad de incidencia* W que definimos a continuación tiene un papel esencial en diversos sitios de esta memoria,

$$W := \{(f, \zeta) \in \mathbb{P}(\mathcal{H}_{(d)}^m) \times \mathbb{P}_n(\mathbb{C}) : \zeta \in V(f)\} \subseteq \mathbb{P}(\mathcal{H}_{(d)}^m) \times \mathbb{P}_n(\mathbb{C}). \quad (1.6)$$

En el Capítulo 4, consideraremos W como un subconjunto de $\mathbb{S}_\Delta \times \mathbb{P}_n(\mathbb{C})$, sin embargo no cambiamos la notación por evitar un recargo de la misma. Con frecuencia utilizaremos las proyecciones canónicas sobre la primera y la segunda componentes:

$$p_1 : W \longrightarrow \mathbb{P}(\mathcal{H}_{(d)}^m), \quad p_2 : W \longrightarrow \mathbb{P}_n(\mathbb{C}).$$

Para todo punto $\zeta \in \mathbb{P}_n(\mathbb{C})$, se denotará por V_ζ el conjunto de sistemas que se anulan en el punto ζ. Esto es,

$$V_\zeta := \{f \in \mathbb{P}(\mathcal{H}_{(d)}^m) : \zeta \in V(f)\} \subseteq \mathbb{P}(\mathcal{H}_{(d)}^m).$$

Además, consideraremos los siguientes dos subespacios de $\mathcal{H}_{(d)}^m$:

$$L_{e_0} := \{g = [g_1, \ldots, g_m] \in \mathcal{H}_{(d)}^m : g_i = X_0^{d_i - 1} \sum_{j=1}^{n} a_{ij} X_j, 1 \leq i \leq m\} \subseteq \mathcal{H}_{(d)}^m,$$

$$L_{e_0}^\perp := \{g \in \mathcal{H}_{(d)}^m : \langle g, f \rangle_\Delta = 0, \ \forall f \in L_{e_0}\} \subseteq \mathcal{H}_{(d)}^m.$$

En otras palabras, L_{e_0} es el conjunto de sistemas que se anulan en e_0 y que son lineales en las variables $X_1, \ldots, X_n$. El conjunto $L_{e_0}^\perp$ puede verse como el conjunto de sistemas de orden 2 en e_0, esto es, los sistemas g tales que tanto g como la diferencial de g se anulan en e_0.
Hay dos maneras diferentes de interpretar la diferencial de un sistema $f \in \mathcal{H}_{(d)}^m$ en un punto, dependiendo de que interpretemos f como aplicación homogénea con dominio $\mathbb{C}^{n+1}$ o como aplicación polinomial (no nec. homogénea) con dominio en $\mathbb{C}^n$. Distinguimos estos conceptos del modo siguiente: Si estamos refiriéndonos a f como aplicación homogénea, escribiremos $d_x f$ para la diferencial de f en x, donde $x \in \mathbb{C}^{n+1}$ es un punto cualquiera. Si, por el contrario, estamos considerando f como aplicación con dominio en $\mathbb{C}^n$, entonces denotaremos por $\partial_x f$ la diferencial de f en x, donde $x \in \mathbb{C}^n$ es un punto cualquiera. Generalizando esta idea, para todo natural $k \geq 0$,

denotaremos por $d_x^{(k)} f, \partial_x^{(k)} f$ la diferencial k–ésima de f vista como aplicación k–lineal, según consideremos a f como aplicación homogénea o no, respectivamente. Podemos escribir:

$$\partial_x^{(k)} f = d_{(1,x)}^{(k)}(f\mid_{\{1\}\times\mathbb{C}^n}),\ \ \forall\, x \in \mathbb{C}^n,\ k \geq 0. \tag{1.7}$$

Dado un sistema $f \in \mathbb{P}(\mathcal{H}_{(d)}^m)$ y un punto cualquiera $x \in \mathbb{P}_n(\mathbb{C})$, denotamos por $T_x f$ la restricción de la diferencial $d_x f$ al ortogonal de x. En otras palabras,

$$T_x f := (d_x f)\mid_{x^\perp},$$

donde consideramos elegidos representantes tales que $\|f\|_\Delta = \|x\|_2 = 1$. La expresión $T_x f$ será considerada indistintamente como aplicación o como matriz en el espacio $\mathcal{M}_{m\times n}(\mathbb{C})$, suponiendo para esto último que elegimos una base ortonormal cualquiera de $\zeta^\perp$. En particular, en el caso $\zeta = e_0 := (1:0:\cdots:0)$, podemos identificar

$$T_{e_0} f \equiv \partial_{e_0} f \equiv \begin{pmatrix} \frac{\partial f_1}{\partial X_1}(e_0) & \cdots & \frac{\partial f_1}{\partial X_n}(e_0) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial X_1}(e_0) & \cdots & \frac{\partial f_m}{\partial X_n}(e_0) \end{pmatrix} \in \mathcal{M}_{m\times n}(\mathbb{C}), \tag{1.8}$$

para algún representante de norma 1 de f.
También utilizaremos con cierta frecuencia la siguiente aplicación,

$$\begin{array}{rccc} \psi_{e_0}: & L_{e_0} & \longrightarrow & \mathcal{M}_{m\times n}(\mathbb{C}), \\ & g & \mapsto & \Delta(d)^{-1/2} T_{e_0} g \end{array}$$

donde estamos utilizando la notación

$$\Delta(d)^{-1/2} := Diag(d_1^{-1/2}, \ldots, d_m^{-1/2}) \in \mathcal{M}_m(\mathbb{R}).$$

Se comprueba fácilmente que ψ_{e_0} es una isometría de espacios de Hilbert (véase por ejemplo [14, Lemma 17, page 235]).
Para cada valor positivo de $r \in \mathbb{N}$, $1 \leq r \leq m$, consideraremos el conjunto $\Sigma_{(d)}^r \subseteq \mathbb{P}(\mathcal{H}_{(d)}^m)$ definido como sigue:

$$\Sigma_{(d)}^r := \{f \in \mathbb{P}(\mathcal{H}_{(d)}^m) : \exists \zeta \in \mathbb{P}_n(\mathbb{C}), f(\zeta) = 0, rank(d_\zeta f) \leq r\}.$$

En la Sección 3.4 demostraremos que $\Sigma_{(d)}^r$ es una variedad proyectiva, y estimaremos su dimensión y su grado. En el caso cero–dimensional, esto es, el caso de que m y n sean iguales, denotaremos simplemente mediante $\Sigma_{(d)} := \Sigma_{(d)}^{n-1}$ el conjunto de sistemas que tienen alguna solución singular. Las variedades

$$\mathbb{P}(\mathcal{H}_{(d)}^m) = \Sigma_{(d)}^m \supseteq \Sigma_{(d)}^{m-1} \supseteq \cdots \supseteq \Sigma_{(d)}^1$$

forman una estratificación del espacio de sistemas $\mathbb{P}(\mathcal{H}_{(d)}^m)$, estableciendo una clasificación inicial de los tipos de singularidades que aparecen en los sistemas de ecuaciones.

1.5. El condicionamiento en el Álgebra Lineal

El número de condicionamiento del Álgebra Lineal fue definido por primera vez por A. Turing en [127], y sus interesantes propiedades fueron estudiadas, en un primer momento, por von Neumann y Goldstine en [97], y, más adelante, por Wilkinson en [132]. Se han realizado numerosos estudios sobre el número de condicionamiento, entre los que podemos resaltar [40, 58, 71, 126, 132, 104]. La importancia del condicionamiento matricial radica en que puede utilizarse para controlar el error relativo de las soluciones de un problema en presencia de pequeños errores en el enunciado del mismo. De un modo más técnico, sea A una matriz regular de tamaño n. Hay dos definiciones posibles de número de condicionamiento, que difieren ligeramente:

$$\kappa(A) := \|A\|_2\|A^{-1}\|_2, \quad \text{o bien} \quad \kappa_D(A) := \|A\|_F\|A^{-1}\|_2,$$

donde $\|\cdot\|_2$ representa la norma como operador lineal y $\|\cdot\|_F$ la norma de Frobenius. Ambos números κ y κ_D son esencialmente equivalentes, en el sentido siguiente:

$$\kappa(A) \leq \kappa_D(A) \leq n\kappa(A), \quad \forall A \in \mathcal{M}_n(\mathbb{C}).$$

Durante esta memoria utilizaremos la segunda de las dos definiciones, por adaptarse mejor a los cálculos que realizamos. Escribamos brevemente la propiedad principal de κ_D como cantidad que controla el error relativo en las solución de sistemas de ecuaciones lineales. Esto es, dados dos sistemas

$$Ax = b, \;\; A'x' = b,$$

se satisface la siguiente desigualdad:

$$\frac{\|x - x'\|_2}{\|x'\|_2} \leq \kappa_D(A)\frac{\|A - A'\|_2}{\|A\|_F}.$$

En otras palabras, el error relativo en las soluciones queda acotado por el error relativo en la matriz inicial y un factor escalar: El número de condicionamiento. Un resultado similar se obtiene si consideramos que el vector independiente b puede también ser inexacto.

1.5.1. El condicionamiento lineal generalizado

William Kahan, G.W. Stewart y J. Sun han estudiado el condicionamiento de las matrices singulares. Como referencias para el lector, señalamos [76] y [122] para propiedades generales de este condicionamiento. Para definir correctamente los conceptos de esta sección, es necesario recordar la Descomposición en Valores Singulares de una matriz. Se trata ésta de una

descomposición matricial clásica que tiene sus orígenes en el siglo XIX, de la manos de científicos como Eugenio Beltrami, Camille Jordan, James Joseph Sylvester, Erhard Schmidt y Hermann Weyl. Su historia puede leerse con detalle en el survey [121]. Sean $1 \leq n_1 \leq n_2$ dos números naturales, y sea $A \in \mathcal{M}_{n_1 \times n_2}(\mathbb{C})$ el espacio de las matrices complejas de tamaño $n_1 \times n_2$. Sea A una matriz, $A \in \mathcal{M}_{n_1 \times n_2}(\mathbb{C})$. Entonces, existe una descomposición en valores singulares de A (véase por ejemplo [122]), esto es

$$A = U \ (D \ 0) \ V^*. \tag{1.9}$$

donde:

- Las matrices $U \in \mathcal{U}_{n_1}$ y $V \in \mathcal{U}_{n_2}$ son unitarias de tamaños respectivos n_1 y n_2, y V^* es la conjugada traspuesta de V.
- La matriz $D := Diag(\sigma_1, \cdots, \sigma_{n_1}) \in \mathcal{M}_{n_1}(\mathbb{R})$ es la matriz de los valores singulares de A, $\sigma_1 \geq \cdots \geq \sigma_{n_1} \geq 0$. Éstos dependen únicamente de A, y no de la descomposición (1.9) particular elegida.
- La expresión $(D \ 0) \in \mathcal{M}_{n_1 \times n_2}(\mathbb{R})$ denota la matriz de n_1 filas y n_2 columnas obtenida añadiendo a D una matriz idénticamente nula de talla $n_1 \times (n_2 - n_1)$.

Definición 1.5.1 (Condicionamiento Lineal Generalizado) *Sea $A \in \mathcal{M}_{n_1 \times n_2}(\mathbb{C}) \setminus \{0\}$ una matriz cualquiera. Consideramos una Descomposición en Valores Singulares de A,*

$$A = U \ (D \ 0) \ V^*, \quad D := Diag(\sigma_1, \cdots, \sigma_{n_2}).$$

Para todo número natural r, $2 \leq r \leq n_1$, definimos el condicionamiento generalizado de A:

$$\kappa_D^{(r)}(A) := \frac{\|A\|_F}{\sqrt{\sigma_r^2 + \cdots + \sigma_{n_2}^2}},$$

donde $\|A\|_F := \sqrt{\sigma_1^2 + \cdots + \sigma_{n_2}^2}$ es la norma de Frobenius de A. Esta definición también es válida para matrices proyectivas $A \in \mathbb{P}(\mathcal{M}_{n_1 \times n_2}(\mathbb{C}))$, al ser invariante si se multiplica por un escalar no nulo. Para el caso $r = 1$, se tiene:

$$\kappa_D^{(1)}(A) = 1,$$

para toda matriz afín o proyectiva $A \neq 0$.

En el caso de que $n_1 = n_2 = n$, el condicionamiento generalizado $\kappa_D^{(n)}$ tal y como lo acabamos de definir coincide con el condicionamiento usual para matrices cuadradas, $\kappa_D(A) := \|A\|_F \|A^{-1}\|_2$, $A \in \mathcal{M}_n(\mathbb{C})$. De hecho, se deduce trivialmente a partir de la definición el siguiente resultado

Lema 1.5.2 *El condicionamiento generalizado* $\kappa_D^{(r)}(A)$ *de una matriz* $A \in \mathcal{M}_{n_1\times n_2}(\mathbb{C})$ *tal que* $rank(A) = r$ *satisface:*

$$\kappa_D^{(r)}(A) = \|A\|_F \|A^\dagger\|_2,$$

donde $A^\dagger$ *es la inversa generalizada de Moore–Penrose de* A*.*

Demostración.– En efecto, sea $A = U\ (D\ 0)\ V^*$ la descomposición en valores singulares de A, donde $D = diag(\sigma_1, \cdots, \sigma_r, 0, \ldots, 0) \in \mathcal{M}_{n_1\times n_2}(\mathbb{C})$, con $\sigma_1 \geq \ldots \geq \sigma_r > 0$. Sea $D^\dagger$ la matriz definida como sigue:

$$D^\dagger := diag(\sigma_1^{-1}, \cdots, \sigma_r^{-1}, 0, \ldots, 0).$$

Entonces, $A^\dagger$ viene dada por la fórmula siguiente (véase por ejemplo [122, pp. 102-104]).

$$A^\dagger = V \begin{pmatrix} D^\dagger \\ 0 \end{pmatrix} U^*,$$

donde la expresión $\binom{D^\dagger}{0} \in \mathcal{M}_{n_2\times n_1}(\mathbb{C})$ denota la matriz de n_2 filas y n_1 columnas obtenida añadiendo a $D^\dagger$ una matriz idénticamente nula de talla $(n_2 - n_1) \times n_1$. Por lo tanto, tenemos la igualdad $\|A^\dagger\|_2 = \sigma_r^{-1}$ y el lema queda demostrado. ∎

1.5.2. Estabilidad en el cálculo de núcleos e inversas

El siguiente resultado muestra la importancia que tiene el número de condicionamiento generalizado como medida de la estabilidad del cálculo de la inversa de Moore–Penrose (véase [122, Corol. 3.10, p. 145]).

Proposición 1.5.3 *Sean* $A, A' \in \mathcal{M}_{n_1\times n_2}(\mathbb{C})$ *dos matrices de igual rango* r*. Entonces, tenemos:*

$$\frac{\|A^\dagger - (A')^\dagger\|_F}{\|(A')^\dagger\|_2} \leq \sqrt{2}\ \kappa_D^{(r)}(A) \frac{\|A - A'\|_F}{\|A\|_F}.$$

Además, el número $\kappa_D^{(r)}(A)$ controla también la estabilidad del cálculo de núcleos de matrices singulares. Para demostrar esto, introducimos algunos conceptos y resultados. El Teorema de Existencia de la Descomposición en Valores Singulares tiene como caso particular el siguiente resultado.

Lema 1.5.4 *Sean* L *y* L' *dos subespacios lineales de* $\mathbb{C}^n$ *de dimensión* m*. Entonces, existen bases ortonormales* $\{v_1, \ldots, v_m\}$ *de* L *y* $\{w_1, \ldots, w_m\}$ *de* L'*, y números reales* $1 \geq \lambda_1 \geq \cdots \geq \lambda_m \geq 0$ *tales que:*

$$\langle v_i, w_j\rangle_2 = \lambda_i \delta_{ij}.$$

Existen diferentes definiciones (equivalentes en esencia) para la distancia entre dos subespacios de igual dimensión. Utilizaremos la siguiente, que es aceptada de forma general (véase por ejemplo [58, p. 76]).

Definición 1.5.5 *Sean $L_{\mathbb{R}}, L'_{\mathbb{R}} \subseteq \mathbb{R}^n$ dos subespacios lineales reales de dimensión m. Entonces, definimos la distancia proyectiva entre $L_{\mathbb{R}}$ y $L'_{\mathbb{R}}$ como sigue:*

$$dist(L_{\mathbb{R}}, L'_{\mathbb{R}}) = \|\pi_{L_{\mathbb{R}}} - \pi_{L'_{\mathbb{R}}}\|_2,$$

donde $\pi_{L_{\mathbb{R}}}$ (resp. $\pi_{L'_{\mathbb{R}}}$) es la proyección ortogonal sobre $L_{\mathbb{R}}$ (resp. $L'_{\mathbb{R}}$), y $\|\pi_{L_{\mathbb{R}}} - \pi_{L'_{\mathbb{R}}}\|_2$ es la norma de esa aplicación como operador lineal.
La distancia entre dos subespacios complejos $L, L' \subseteq \mathbb{C}^n$ de dimensión m se define de igual manera:

$$dist(L, L') = \|\pi_L - \pi_{L'}\|_2.$$

Nota 1.5.6 *En [58] y [131] pueden encontrarse algunas propiedades de la distancia que acabamos de definir. Nos interesan en particular las siguientes:*

- *Sea θ el mayor ángulo principal (en el sentido de [58, p. 603]) entre L y L'. Entonces,:*

  $$dist(L, L') = \sin\theta.$$

  En particular, en las notaciones del Lema 1.5.4, podemos escribir:

  $$dist(L, L') = \sqrt{1 - \lambda_m^2}.$$

- *Si $\dim(L) = \dim(L') = 1$, entonces $dist(L, L') = d_{\mathbf{P}}(L, L')$ donde $d_{\mathbf{P}}(L, L')$ es la distancia proyectiva entre los puntos proyectivos que definen L y L'.*

El siguiente resultado relaciona el condicionamiento generalizado $\kappa_D^{(r)}$ con la estabilidad de las soluciones de sistemas de ecuaciones lineales cuadrados singulares.

Proposición 1.5.7 *Sean $A, A' \in \mathcal{M}_n(\mathbb{C})$ dos matrices cuadradas, de forma que $rank(A) = rank(A') = r$. Sean $L := Ker(A)$ y $L' := Ker(A')$ los núcleos respectivos de A y A', que son subespacios complejos de dimensión $m = n - r$. Entonces, tenemos:*

$$dist(L, L') \leq \kappa_D^{(r)}(A) \frac{\|A' - A\|_2}{\|A\|_F}.$$

Demostración.– Sean $\{v_1, \ldots, v_m\}$, $\{w_1, \ldots, w_m\}$, y $1 \geq \lambda_1 \geq \cdots \geq \lambda_m \geq 0$ los vectores y escalares definidos en el Lema 1.5.4, generando respectivamente L y L'. Hemos visto que

$$dist(L, L') = \sqrt{1 - \lambda_m^2}.$$

Por otro lado, tenemos la siguiente igualdad:

$$\kappa_D^{(r)}(A)\frac{\|A'-A\|_2}{\|A\|_F} = \frac{\|A'-A\|_2}{\sigma_r},$$

donde σ_r es el valor singular más pequeño no nulo de A. Por lo tanto, basta con demostrar que

$$\sqrt{1-\lambda_m^2} \leq \frac{\|A'-A\|_2}{\sigma_r}.$$

Observamos que se satisface la siguiente igualdad:

$$L^\perp = \{w \in \mathbb{C}^n : V^*w \in \langle e_1, \ldots, e_r\rangle\},$$

donde $A = UDV^*$ es la Descomposición en Valores Singulares de A, y $\langle e_1, \ldots, e_r\rangle$ es el subespacio de $\mathbb{C}^n$ generado por los primeros r vectores de la base canónica. Además, $A^\dagger = VD^\dagger U^*$. Por lo tanto, tenemos que para todo vector $w \in L^\perp$, se satisface:

$$A^\dagger Aw = V\begin{pmatrix} Id_r & 0 \\ 0 & 0 \end{pmatrix}V^*w = w.$$

Deducimos las siguientes desigualdades para todo vector $w \in L^\perp$:

$$\|w\|_2 = \|A^\dagger Aw\|_2 \leq \|A^\dagger\|_2\|Aw\|_2, \Longrightarrow \|Aw\|_2 \geq \frac{\|w\|_2}{\|A^\dagger\|_2}.$$

Primero, supongamos que $\lambda_m = 0$. Entonces, $w_m \in L^\perp$. Por lo tanto:

$$\|(A'-A)w_m\|_2 = \|Aw_m\|_2 \geq \frac{1}{\|A^\dagger\|_2}, \Longrightarrow \|(A'-A)\|_2 \geq \frac{1}{\|A^\dagger\|_2}.$$

Por el Lema 1.5.2, concluimos que $\frac{1}{\|A^\dagger\|_2} = \sigma_r$. Luego en ese caso tenemos que:

$$\sqrt{1-\lambda_m^2} = 1 \leq \|(A'-A)\|_2\|A^\dagger\|_2 = \frac{\|A'-A\|_2}{\sigma_r},$$

de donde se sigue el resultado.
Supongamos ahora que $\lambda_m \neq 0$. Sea $w'_m = \frac{w_m}{\lambda_m}$. Entonces, tenemos:

$$\langle v_m, w'_m - v_m\rangle_2 = \frac{1}{\lambda_m}\langle v_m, w_m\rangle_2 - \|v_m\|_2^2 = 1 - 1 = 0.$$

Definimos $\delta w = w'_m - v_m$, y $\delta A = A' - A$. Se tiene que:

$$\frac{\|\delta w\|_2^2}{\|w'_m\|_2^2} = \lambda_m^2\langle w'_m - v_m, w'_m - v_m\rangle_2 =$$

$$\lambda_m^2(\|w'_m\|_2^2 - \langle w'_m, v_m\rangle_2) - \lambda_m^2\langle v_m, w'_m - v_m\rangle_2 = 1 - \lambda_m^2,$$

y por lo tanto:

$$\frac{\|\delta w\|_2}{\|w'_m\|_2} = \sqrt{1-\lambda_m^2} = dist(L, L').$$

Concluimos que basta con demostrar que

$$\|\delta A\|_2 \geq \frac{\sigma_r \|\delta w\|_2}{\|w'_m\|_2}.$$

Ahora, sabemos que $\delta A w'_m + A\delta w = (A+\delta A)(v_m + \delta w) = A' w'_m = 0$, de donde se tiene:

$$\delta A \frac{w'_m}{\|w'_m\|_2} = \frac{-A\delta w}{\|w'_m\|_2},$$

Por lo tanto,

$$\|\delta A\|_2 \geq \frac{\|A\delta w\|_2}{\|w'_m\|_2}.$$

La demostración quedará completa si verificamos que

$$\|A\delta w\|_2 \geq \sigma_r \|\delta w\|_2.$$

Ahora, observamos que $\delta w \in L^{\perp}$, luego se tiene la siguiente desigualdad:

$$\|A\delta w\|_2 \geq \frac{1}{\|A^{\dagger}\|_2}\|\delta w\|_2.$$

Por el Lema 1.5.2, se tiene que $\frac{1}{\|A^{\dagger}\|_2} = \sigma_r$ y hemos terminado la demostración. ■

1.5.3. Teorema del número de condicionamiento

El condicionamiento de una matriz tiene una propiedad esencial, que permite analizarlo en términos geométricos: Coincide con la inversa de la distancia al conjunto de problemas "mal condicionados". En esta sección fijaremos esta idea tan natural, en que jugarán un papel esencial las distintas variedades de la estratificación por rango del espacio de matrices: Sea r un número natural, $1 \leq r \leq n_1$. Denotemos por $\Sigma^r_{\mathcal{M}}$ la variedad proyectiva de todas las matices complejas de rango a lo más r. Esto es,

$$\Sigma^r_{\mathcal{M}} := \{A \in \mathbb{P}(\mathcal{M}_{n_1 \times n_2}(\mathbb{C})) : rank(A) \leq r\}.$$

El siguiente resultado proporciona la dimensión y el grado de las variedades $\Sigma^r_{\mathcal{M}}$. La primera parte es [18, Prop. 1.1]. La igualdad en el grado puede leerse en [64, pp. 243-244] o en [47, p. 261].

Proposición 1.5.8 *Para todo natural $1 \leq r \leq n_1$, el conjunto $\Sigma^r_{\mathcal{M}}$ es una variedad proyectiva irreducible de $\mathbb{P}(\mathcal{M}_{n_1\times n_2}(\mathbb{C}))$ de codimensión $(n_2 - r)(n_1 - r)$. Además,*

$$\deg(\Sigma^r_{\mathcal{M}}) = \prod_{i=0}^{n_2-r-1} \frac{(n_1+i)!\ i!}{(r+i)!\ (n_1-r+i)!}.$$

Como consecuencia inmediata tenemos el siguiente resultado.

Corolario 1.5.9 *Con las notaciones de la Proposición 1.5.8, se tiene:*

$$\deg(\Sigma^r_{\mathcal{M}}) = \prod_{i=1}^{n_1-r}\prod_{j=1}^{n_2-r} \frac{r+i+j-1}{i+j-1}.$$

En particular,

$$\deg(\Sigma^r_{\mathcal{M}}) \leq \binom{n_1}{r}^{n_2-r}, \quad \deg(\Sigma^r_{\mathcal{M}}) \leq (r+1)^{(n_2-r)(n_1-r)}.$$

Los siguientes dos resultados, conocidos como teoremas del número de condicionamiento, se atribuyen normalmente a Eckart y Young. Sin embargo, preferimos atribuirlos a los cuatro autores que se nombran siguiendo el criterio de Stewart y Sun en [122]. Ambos resultados relacionan de un modo muy natural el condicionamiento de una matriz con el inverso de la distancia al conjunto de matrices de rango dado. Para el segundo de ellos hemos decidido incluir una demostración porque no es fácil encontrarla en la literatura, y porque será utilizada más adelante en esta memoria. Una demostración para el primero se puede encontrar por ejemplo en [122].

Teorema 1.5.10 (Schmidt–Mirsky–Eckart–Young) *Para toda matriz $A \in \mathcal{M}_{n_1\times n_2}(\mathbb{C})$, y para todo número natural $2 \leq r \leq n_1$, se tiene:*

$$\begin{array}{c}\min \\ rank(A') \leq r-1 \\ A' \in \mathcal{M}_{n_1\times n_2}(\mathbb{C})\end{array} \|A' - A\|_F = \sqrt{\sigma_r^2 + \cdots + \sigma_{n_2}^2},$$

donde $\sigma_r, \ldots, \sigma_{n_2}$ son los últimos valores singulares de A.

Teorema 1.5.11 (Schmidt–Mirsky–Eckart–Young) *Sea A una matriz proyectiva, $A \in \mathbb{P}(\mathcal{M}_{n_1\times n_2}(\mathbb{C}))$, y sea $2 \leq r \leq n_1$ un número natural. Entonces, se tiene:*

$$d_{\mathbf{P}}(A, \Sigma^{r-1}_{\mathcal{M}}) = \frac{1}{\kappa^{(r)}_D(A)}.$$

Demostración.– Sea $A \in \mathbb{P}(\mathcal{M}_{n_1 \times n_2}(\mathbb{C}))$. Denotamos por el mismo símbolo A un representante tal que $\|A\|_F = \sigma_1^2 + \cdots + \sigma_{n_2}^2 = 1$. Consideramos la Descomposición en Valores Singulares de A, $A = U(D\ 0)V^*$. Sea la matriz D' definida como sigue.

$$D' = Diag(\sigma_1, \ldots, \sigma_{r-1}, 0, \ldots, 0).$$

Entonces, la matriz $A' = U(D'\ 0)V^*$ tiene la siguiente propiedad:

$$|\langle A', A\rangle_F| = \sigma_1^2 + \ldots + \sigma_{r-1}^2 \in \mathbb{R}, \quad rank(A') = r-1,$$

$$\|A'\|_F^2 = \sigma_1^2 + \ldots + \sigma_{r-1}^2.$$

Por lo tanto, tenemos:

$$d_{\mathbb{P}}(A, \Sigma_{\mathcal{M}}^{r-1}) \leq d_{\mathbb{P}}(A, A') = \sqrt{1 - \frac{(\sigma_1^2 + \ldots + \sigma_{r-1}^2)^2}{\sigma_1^2 + \ldots + \sigma_{r-1}^2}} =$$

$$\sqrt{\sigma_r^2 + \cdots + \sigma_{n_2}^2} = \frac{1}{\kappa_D^{(r)}(A)}.$$

Sea ahora $A' \in \mathbb{P}(\mathcal{M}_{n_1 \times n_2}(\mathbb{C}))$ una matriz proyectiva cualquiera tal que $rank(A') \leq r-1$. Podemos elegir un representante de A', que denotamos por la misma letra, tal que:

$$\|A'\|_F^2 = 1 - (\sigma_r^2 + \cdots + \sigma_{n_2}^2), \quad \langle A', A\rangle_2 \in \mathbb{R}^{0,+}.$$

Entonces, por el Teorema 1.5.10 sabemos que

$$\sigma_r^2 + \cdots + \sigma_{n_2}^2 \leq \|A' - A\|_F^2 =$$

$$\langle A' - A, A' - A\rangle_F = 2 - (\sigma_r^2 + \cdots + \sigma_{n_2}^2) - 2\langle A', A\rangle_F,$$

y se tiene que:

$$|\langle A, A'\rangle_F| \leq \frac{2 - 2(\sigma_r^2 + \cdots + \sigma_{n_2}^2)}{2} = 1 - (\sigma_r^2 + \cdots + \sigma_{n_2}^2).$$

Concluimos que

$$d_{\mathbb{P}}(A, A') = \sqrt{1 - \frac{|\langle A', A\rangle_F|^2}{\|A\|_F^2 \|A'\|_F^2}} \geq \sqrt{1 - \frac{(1 - (\sigma_r^2 + \cdots + \sigma_{n_2}^2))^2}{1 - (\sigma_r^2 + \cdots + \sigma_{n_2}^2)}} = \frac{1}{\kappa_D^{(r)}(A)},$$

lo que termina la demostración del lema. ∎

En estas últimas secciones hemos resumido las propiedades principales del número de condicionamiento matricial: Su capacidad para controlar los errores cometidos ante perturbaciones del input y su relación con la distancia a las variedades $\Sigma_{\mathcal{M}}^r$. Esta última propiedad nos permitirá estudiar la distribución de probabilidad del condicionamiento en el Capítulo 2.

1.6. El condicionamiento de sistemas de ecuaciones polinomiales

El siguiente paso natural es preguntarse si hay una situación parecida que nos permita estudiar el caso de sistemas no–lineales. Naturalmente, en esta situación la dificultad es mayor y el número de estudios al respecto, considerablemente menor. No obstante, en los últimos años el interés de la comunidad científica en este campo ha aumentado notoriamente, a raíz de la serie de artículos [112, 113, 114, 115, 116], donde Shub & Smale propusieron y analizaron un número de condicionamiento para los sistemas de ecuaciones polinomiales cero–dimensionales. De entre los diversos estudios que han generalizado y analizado las propiedades del condicionamiento, podemos resaltar los trabajos [31, 30, 37, 88, 33, 34, 55]. La definición de Shub & Smale puede escribirse como sigue.

Definición 1.6.1 *Sea $f \in \mathbb{P}(\mathcal{H}^n_{(d)})$ un sistema de n ecuaciones homogéneas en incógnitas $X_0, \ldots, X_n$, y sea $\zeta \in \mathbb{P}_n(\mathbb{C})$ una solución de dicho sistema. Entonces, definimos*

$$\mu_{\text{norm}}(f, \zeta) := \|f\|_\Delta \|(T_\zeta f)^{-1} diag(\|\zeta\|_2^{d_i - 1} d_i^{1/2})\|_2.$$

En particular, si elegimos representantes tales que $\|f\|_\Delta = \|\zeta\|_2 = 1$, tenemos:

$$\mu_{\text{norm}}(f, \zeta) = \|(T_\zeta f)^{-1} \Delta(d)^{1/2}\|_2 = \frac{\kappa_D^{(n)}(\Delta(d)^{-1/2} T_\zeta f)}{\|\Delta(d)^{-1/2} T_\zeta f\|_F},$$

donde $\kappa_D^{(n)} \equiv \kappa_D$ es el condicionamiento matricial usual, definido en la Sección 1.5.

Los estudios más importantes relacionados con el condicionamiento no–lineal que acabamos de definir, debidos a Shub & Smale, proporcionan principalmente tres resultados:

- Un Teorema del Número de Condicionamiento, relacionándolo con el inverso de la distancia a una variedad proyectiva particular.
- Una cota para la distribución de probabilidad del citado condicionamiento.
- Un fortísimo resultado que relaciona el condicionamiento μ_{norm} con la complejidad de los métodos de homotopía para el cálculo de soluciones aproximadas. Éste es el punto de partida para la resolución del Problema 17 de Smale propuesta en el Capítulo 4 de esta memoria.

En [37] se propone una generalización del número de condicionamiento μ_{norm} a los sistemas de ecuaciones homogéneos con m ecuaciones e incógnitas $X_0, \ldots, X_n$, para el caso afín. Degót también demuestra un Teorema del Número de Condicionamiento afín para esos casos. Siguiendo la línea de [112], [30] y [37], y utilizando los conceptos de condicionamiento lineal generalizado de la Sección 1.5, podemos definir el condicionamiento generalizado de los sistemas de ecuaciones de un modo muy natural.

Definición 1.6.2 *Sean $n \geq m \geq r \geq 1$ tres números naturales positivos. Sea $f \in \mathbb{P}(\mathcal{H}_{(d)}^m)$ un sistema y $\zeta \in \mathbb{P}_n(\mathbb{C})$ una solución de dicho sistema. Elegimos representantes f y ζ tales que $\|f\|_\Delta = \|\zeta\|_2 = 1$. Definimos el condicionamiento generalizado como*

$$\mu_{\text{norm}}^{(r)}(f, \zeta) := \frac{\kappa_D^{(r)}(\Delta(d)^{-1/2} T_\zeta f)}{\|\Delta(d)^{-1/2} T_\zeta f\|_F},$$

donde $\kappa_D^{(r)}$ es el condicionamiento lineal generalizado definido en la Sección 1.5.

Nota 1.6.3 *Obsérvese que en el caso cero–dimensional (esto es, $m = n$), si tomamos también $r = n$ entonces el condicionamiento $\mu_{\text{norm}}^{(n)}$ de la Definición 1.6.2 se convierte en el condicionamiento usual de Shub & Smale, que hemos visto en la Definición 1.6.1. A lo largo de esta memoria, siempre que estemos en esa situación (esto es, siempre que $r = m = n$), denotaremos*

$$\mu_{\text{norm}} \equiv \mu_{\text{norm}}^{(n)}.$$

Otra aclaración importante es la siguiente: En el caso $r = m$ y si $d_\zeta f$ tiene rango máximo (equivalentemente, $T_\zeta f$ tiene rango máximo), el condicionamiento $\mu_{\text{norm}}^{(m)}(f, \zeta)$ toma la forma

$$\mu_{\text{norm}}^{(m)}(f, \zeta) = \|(\Delta(d)^{-1/2} T_\zeta f)^\dagger\|_2 =$$

$$\|(\Delta(d)^{-1/2} T_\zeta f)^* [(\Delta(d)^{-1/2} T_\zeta f)(\Delta(d)^{-1/2} T_\zeta f)^*]^{-1}\|_2 =$$

$$\|(T_\zeta f)^* \Delta(d)^{-1/2} [\Delta(d)^{-1/2} T_\zeta f (T_\zeta f)^* \Delta(d)^{-1/2}]^{-1}\|_2 =$$

$$= \|(T_\zeta f)^* [T_\zeta f (T_\zeta f)^*]^{-1} \Delta(d)^{1/2}\|_2 = \|(T_\zeta f)^\dagger \Delta(d)^{1/2}\|_2,$$

donde estamos suponiendo que $\|f\|_\Delta = \|\zeta\|_2 = 1$. Si dejamos f y ζ ser representantes cualesquiera, podemos escribir

$$\mu_{\text{norm}}^{(m)}(f, \zeta) = \|f\|_\Delta \|(T_\zeta f)^\dagger diag(\|\zeta\|_2^{d_i - 1} d_i^{1/2})\|_2, \tag{1.10}$$

que es una expresión muy similar a la de la Definición 1.6.1.

El condicionamiento generalizado $\mu_{\text{norm}}^{(r)}$ es invariante unitario, para todo $r \geq 1$. Esto es, para cualquier matriz unitaria $U \in \mathcal{U}_{n+1}$ y cualquier par $(f, \zeta) \in \mathcal{H}_{(d)}^m \times \mathbb{P}_n(\mathbb{C})$ se tiene la siguiente igualdad:

$$\mu_{\text{norm}}^{(r)}(f, \zeta) = \mu_{\text{norm}}^{(r)}(f \circ U, U^{-1}\zeta). \tag{1.11}$$

En el Capítulo 3, demostraremos un Teorema del Número de Condicionamiento para esta situación general que acabamos de describir. Además, obtendremos cotas para la distribución de probabilidad de ese condicionamiento que generalizan las obtenidas por Shub & Smale en [113] (véase también [14]) y también calcularemos otras distribuciones de probabilidad relacionadas.

1.7. El método de Newton para sistemas de ecuaciones. El caso cero–dimensional

Existe una larguísima relación de libros y estudios realizados sobre métodos de Newton para la resolución de ecuaciones polinomiales. Algunas referencias que recomendamos son [14, 34, 33, 35, 36, 78, 55, 56, 88, 134]. En [110], se describe por primera vez el método proyectivo de Newton, y la serie de artículos [112, 113, 114, 115, 116] proponen un método de homotopía lineal. Sería pretencioso tratar de explicar, en breves líneas, la importancia y el alcance que tiene el operador de Newton. Tampoco es nuestro objetivo incluir en estas páginas una detallada explicación de la naturaleza del mismo ni resumir todas las propiedades del operador. Nos limitaremos a exponer el operador de Newton proyectivo y afín, y a recordar los resultados más relevantes relacionados con él, obtenidos por Shub & Smale en [112, 113, 115], que serán necesarios para los resultados principales de esta memoria. Casi todos los resultados aquí presentes se encuentran resumidos en [14]. Recomendamos también al lector el libro de Dedieu [34].

1.7.1. El operador de Newton proyectivo

Comenzamos con el operador en el caso proyectivo, que trata de encontrar soluciones proyectivas de sistemas de ecuaciones homogéneos. Sea pues $(f, z) \in \mathcal{H}_{(d)}^n \times \mathbb{P}_n(\mathbb{C})$ un par. Sea $z^\perp := \{w \in \mathbb{C}^{n+1} : \langle w, z\rangle_2 = 0\}$ el espacio tangente de $\mathbb{P}_n(\mathbb{C})$ en z. Si la restricción de la diferencial de f en z, $T_z f := d_z f \mid_{z^\perp}$ es sobreyectiva, definimos la iteración de Newton de f en z como:

$$N_f(z) := z - (T_z f)^{-1} f(z) \in \mathbb{P}_n(\mathbb{C}).$$

Consideremos la aplicación $d_T : \mathbb{P}_n(\mathbb{C}) \times \mathbb{P}_n(\mathbb{C}) \longrightarrow \mathbb{R}$ dada como

$$d_T(z_1, z_2) := \tan(d_R(z_1, z_2)).$$

Es decir, d_T es la tangente de la distancia Riemanniana en $\mathbb{P}_n(\mathbb{C})$. Obsérvese que d_T no es exactamente una función distancia (falla la desigualdad triangular). Siguiendo a Shub & Smale, dejamos esta formulación por razones estéticas.

Definición 1.7.1 *Sea $\zeta \in V(f)$ una solución proyectiva de $f \in \mathcal{H}^n_{(d)}$. Decimos que $z \in \mathbb{P}_n(\mathbb{C})$ es un cero aproximado proyectivo de f con cero asociado ζ si la secuencia*

$$z_0 := z, \quad z_{i+1} := N_f(z_i)\ \forall i \geq 0$$

está definida, y

$$d_T(\zeta, z_i) \leq \frac{1}{2^{2^i-1}} d_T(\zeta, z), \forall i \geq 0.$$

El siguiente resultado, que puede encontrarse en [14], garantiza la convergencia de la secuencia de Newton bajo ciertas circunstancias:

Teorema 1.7.2 (Shub & Smale) *Sea $f \in \mathbb{P}(\mathcal{H}_{(d)})$ un sistema, $\zeta \in V(f)$. Sea $\gamma_0(f, \zeta)$ el número definido como sigue*

$$\gamma_0(f,\zeta) := \|\zeta\|_2 \max_{k\geq 1} \left\| (T_\zeta f)^{-1} \frac{d^{(k)}_\zeta f}{k!} \right\|_2^{\frac{1}{k-1}}. \tag{1.12}$$

Sea $z \in \mathbb{P}_n(\mathbb{C})$ un punto, tal que

$$d_T(z,\zeta)\gamma_0(f,\zeta) \leq \frac{3-\sqrt{7}}{2}.$$

Entonces, z es un cero aproximado proyectivo de f con cero asociado ζ.

Nota 1.7.3 *Como demuestran Shub & Smale en [112], se tiene la siguiente desigualdad,*

$$\gamma_0(f,\zeta) \leq \frac{d^{3/2}}{2}\mu_{\text{norm}}(f,\zeta).$$

Por tanto, otra condición suficiente para que z sea cero aproximado proyectivo de f con cero asociado ζ es:

$$d_T(z,\zeta)\mu_{\text{norm}}(f,\zeta) \leq \frac{3-\sqrt{7}}{d^{3/2}}.$$

Hemos establecido, por tanto, una noción de "cero aproximado proyectivo" y hemos dado condiciones suficientes para que un punto proyectivo sea cero aproximado proyectivo de un sistema dado. En el Capítulo 4 daremos un método de homotopía para encontrar ceros aproximados proyectivos de sistemas de ecuaciones.

1.7.2. El operador de Newton afín

El caso afín tiene una estructura muy similar al caso proyectivo, sin embargo hay una diferencia fundamental: En el estudio del método de Newton proyectivo, se trata de encontrar puntos cercanos (en términos de la distancia tangente d_T) a ceros proyectivos de los sistemas. La distancia tangente puede verse como una medida del ángulo que separa dos puntos proyectivos. Pero si queremos aproximar soluciones afines (soluciones en $V_{\mathbb{C}^n}(f)$), no basta con controlar el ángulo, debemos exigir que la distancia (en norma usual) que separa los puntos sea muy pequeña.
Con este objetivo nace el método de Newton para encontrar soluciones afines de sistemas de ecuaciones cero–dimensionales.
Sea pues $f \in \mathcal{H}_{(d)}$ un sistema y sea $z \in \mathbb{C}^n$ un punto afín. Si la diferencial de f en z, $\partial_z f$, es sobreyectiva, definimos la iteración de Newton de f en z como:

$$N_f(z) := z - (\partial_z f)^{-1} f(z) \in \mathbb{C}^n.$$

Definición 1.7.4 *Sea $\zeta \in V_{\mathbb{C}^n}(f)$ una solución de $f \in \mathcal{H}_{(d)}$. Decimos que $z \in \mathbb{C}^n$ es un cero aproximado afín de f con cero asociado ζ si la secuencia*

$$z_0 := z, \quad z_{i+1} := N_f(z_i)\ \forall i \geq 0$$

está definida, y

$$\|z_i - \zeta\|_2 \leq \frac{1}{2^{2^i-1}} \|z - \zeta\|_2, \forall i \geq 0.$$

El siguiente resultado, que puede encontrarse en [112] o en [34], garantiza la convergencia de la secuencia de Newton bajo ciertas circunstancias:

Teorema 1.7.5 (Shub & Smale) *Sea $f \in \mathcal{H}_{(d)}$ un sistema, $\zeta \in V_{\mathbb{C}^n}(f)$. Sea $\gamma(f, \zeta)$ el número definido como sigue*

$$\gamma(f, \zeta) := \max_{k \geq 1} \left\| (\partial_\zeta f)^{-1} \frac{\partial_\zeta^{(k)} f}{k!} \right\|_2^{\frac{1}{k-1}}.$$

Sea $z \in \mathbb{C}^n$ un punto, tal que

$$\|z - \zeta\|_2 \gamma(f, \zeta) \leq \frac{3 - \sqrt{7}}{2}.$$

Entonces, z es un cero aproximado afín de f con cero asociado ζ.

Hemos establecido por tanto una noción de "cero aproximado afín", y hemos dado condiciones suficientes para que un punto afín sea cero aproximado afín de un sistema dado. En el Capítulo 5 utilizaremos el método de homotopía del Capítulo 4 para encontrar ceros aproximados afines de sistemas en el caso afín.

1.8. El método de Newton para sistemas de ecuaciones. El caso de dimensión positiva.

Generalizamos a continuación las ideas de la subsección 1.7.2, al caso de ceros afines en dimensión positiva, esto es cuando el número de ecuaciones es menor que el número de incógnitas. Una variedad algebraica afín $V_{\mathbb{C}^n} \subseteq \mathbb{C}^n$ se llama intersección completa si tiene codimensión m y existe alguna lista de grados (d) y algún sistema $f \in \mathcal{H}^m_{(d)}$ tal que $V_{\mathbb{C}^n}(f) = V_{\mathbb{C}^n}$. Nos ocupamos por tanto de la búsqueda de soluciones afines. El caso $m = n$ es simplemente el caso de variedades algebraicas cero–dimensionales.
Uno de los problemas centrales de la matemática computacional es la resolución del siguiente problema

---

Aproximación de variedades intersección completa

Input:

- Una lista de polinomios $f = [f_1, \ldots, f_m] \ \in \mathcal{H}^m_{(d)}$ tales que $V_{\mathbb{C}^n}(f)$ es intersección completa de codimensión m.
- Un número real positivo $\varepsilon > 0$.

Output: Un punto $z \in \mathbb{C}^n$ en el tubo de radio ε alrededor de $V_{\mathbb{C}^n}(f)$. Esto es, un punto $z \in \mathbb{C}^n$ tal que

$$dist(z, V_{\mathbb{C}^n}(f)) := \text{ínf}\{\|z - \zeta\|_2 : \zeta \in V_{\mathbb{C}^n}(f)\} < \varepsilon.$$

Como indicaron Shub & Smale (véase por ejemplo el Teorema 1.7.5), el diseño de algoritmos eficientes para el caso cero–dimensional es consecuencia del comportamiento en media de cierta cantidad γ, asociada al sistema input $f \in \mathcal{H}^m_{(d)}$ y a la solución que queremos encontrar. Esta cantidad controla la convergencia del operador de Newton. ¿Qué situación nos encontramos en el caso de dimensión positiva? Veremos que los resultados son casi equivalentes, pero cambiando la inversión de matrices (que aparece en el caso cero–dimensional) por la inversión generalizada.
Sea $f \in \mathcal{H}^m_{(d)}$ un sistema de ecuaciones y sea $z \in \mathbb{C}^n$ un punto afín. El operador de Newton de f en z se define mediante la siguiente igualdad:

$$N_f(z) := z - (\partial_z f)^\dagger f(z).$$

donde $(\partial_z f)^\dagger$ es la pseudo–inversa de Moore–Penrose de $\partial_z f$. Con estas notaciones, decimos que un punto $z \in \mathbb{C}^n$ es un *cero aproximado afín* de f si la secuencia de iteraciones del operador de Newton aplicado a z existe, es convergente y además se satisface la siguiente desigualdad para todo número natural $0 \leq k \in \mathbb{N}$:

$$dist(N_f^{(k)}(z), V_{\mathbb{C}^n}(f)) \leq \frac{2}{2^{2^k-1}} dist(z, V_{\mathbb{C}^n}(f)). \tag{1.13}$$

Obsérvese que exigimos a $N_f^{(k)}(z)$ ser convergente, y de hecho converge a un cero $\zeta \in V_{\mathbb{C}^n}(f)$ de f. Más aún, en todos los usos que haremos de esta propiedad, la velocidad de convergencia a dicha solución es cuadrática. Esto es, se tendrá que

$$\|N_f^{(k)}(z) - \zeta\|_2 \leq \frac{2}{2^{2^k-1}}\|z - \zeta\|_2.$$

Obsérvese que, en caso de conocer un cero aproximado afín z de f, el problema de obtener una aproximación a distancia ε de la variedad $V(f)$ requiere tan sólo

$$O(\log|\log\varepsilon|)$$

iteraciones de Newton. En el caso de dimensión positiva, la convergencia del operador de Newton en un punto está garantizada por el γ–teorema que escribimos a continuación. Sea $f \in \mathcal{H}_{(d)}^m$ un sistema. Si $\zeta \in V_{\mathbb{C}^n}(f)$ es una solución regular de f, se define:

$$\gamma(f,\zeta) := \max_{k\geq 2}\left\|(\partial_\zeta f)^\dagger \frac{\partial_\zeta^{(k)} f}{k!}\right\|_2^{\frac{1}{k-1}}.$$

Si ζ es una solución singular de f se define $\gamma(f,\zeta) := +\infty$. Obsérvese que esta noción generaliza la dada en el Teorema 1.7.5. El γ–teorema en dimensión positiva es como sigue (véase [116, th. C1], [34, th. 134]).

Teorema 1.8.1 *Sea $f \in \mathcal{H}_{(d)}^m$ un sistema de ecuaciones polinomiales. Consideramos el conjunto*

$$\mathcal{V}_f := \{x \in \mathbb{C}^n : \exists\zeta \in V_{\mathbb{C}^n}(f),\ \|x-\zeta\|_2\gamma(f,\zeta) \leq u_0\},$$

donde u_0 es una constante universal (aproximadamente igual a 0,05992). Sea $x \in \mathcal{V}_f$ un punto afín, y sea $\zeta \in V_{\mathbb{C}^n}(f)$ una solución de f tal que

$$\|x-\zeta\|_2\gamma(f,\zeta) \leq u_0.$$

Entonces, la serie de Newton $x_k := N_f^k(x)$ converge a un cero $\zeta' \in V_{\mathbb{C}^n}(f)$, y para todo $k \geq 0$ se tiene la siguiente desigualdad:

$$\|x_k - \zeta'\|_2 \leq \frac{2}{2^{2^k-1}}\|x-\zeta\|_2.$$

Obsérvese que para un sistema $f \in \mathcal{H}_{(d)}^m$, el conjunto $\mathcal{V}_f$ tiene una expresión parecida a la de un entorno tubular del conjunto de soluciones de f, y el "radio" de este entorno en cada punto solución es exactamente

$$\frac{u_0}{\gamma(f,\zeta)}.$$

Resultará muy conveniente considerar el caso peor del número γ definido más arriba, cuando recorremos el conjunto de soluciones de f. Esto es, para $f \in \mathcal{H}^m_{(d)}$ definimos:

$$\gamma_{\mathrm{worst}}(f) := \sup_{\zeta \in V_{\mathbb{C}^n}(f)} \gamma(f,\zeta),$$

y demostramos el siguiente resultado, que es un corolario casi inmediato del Teorema 1.8.1.

Corolario 1.8.2 *Existe una constante universal* $0 < u_0 \sim 0{,}05992$ *con la siguiente propiedad: Para todo* $z \in \mathbb{C}^n$ *tal que*

$$dist(z, V_{\mathbb{C}^n}(f))\gamma_{\mathrm{worst}}(f) \leq u_0,$$

se tiene que z *es un cero aproximado afín de* $f \in \mathcal{H}^m_{(d)}$. *Esto es, para todo* $k \geq 0$,

$$dist(N_f^{(k)}(z), V_{\mathbb{C}^n}(f)) \leq \frac{2}{2^{2^k-1}} dist(z, V_{\mathbb{C}^n}(f)).$$

Demostración.– Sea $\zeta \in V_{\mathbb{C}^n}(f)$ tal que $dist(x, V_{\mathbb{C}^n}(f)) = \|x-\zeta\|_2$. Entonces, tenemos la siguiente cadena de desigualdades:

$$\|x-\zeta\|_2 \gamma(f,\zeta) \leq dist(x, V_{\mathbb{C}^n}(f))\gamma_{\mathrm{worst}}(f) \leq u_0.$$

Por el Teorema 1.8.1, existe una solución ζ' de f tal que la serie de Newton $x_k := N_f^{(k)}(x)$ satisface:

$$dist(x_k, V_{\mathbb{C}^n}(f)) \leq \|x_k - \zeta'\|_2 \leq \frac{2}{2^{2^k-1}}\|x-\zeta\|_2 = \frac{2}{2^{2^k-1}} dist(x, V_{\mathbb{C}^n}(f)),$$

como queríamos. ∎

Así pues, acabamos de ver que todo punto en el entorno tubular de radio $u_0\gamma_{\mathrm{worst}}(f)^{-1}$ alrededor del conjunto de soluciones de f es un cero aproximado afín de f. El punto crítico de este resultado es, por supuesto, que $\gamma_{\mathrm{worst}}(f)$ puede tener un valor infinito. Por ejemplo, si $V_{\mathbb{C}^n}(f)$ contiene algún punto singular ζ, entonces tenemos que $\gamma_{\mathrm{worst}}(f) \geq \gamma(f,\zeta) = +\infty$. En el Capítulo 3 veremos que, salvo para un conjunto de medida nula en el espacio de sistemas $\mathcal{H}^m_{(d)}$, se tiene que γ_{worst} es siempre finito. Además, relacionaremos γ con el número de condicionamiento $\mu^{(m)}_{\mathrm{norm}}$ de la Sección 1.6 y obtendremos cotas para el valor esperable de γ_{worst} y del radio de convergencia $u_0\gamma^{-1}_{\mathrm{worst}}$.

Capítulo 2

El Condicionamiento de Matrices singulares: Un Análisis de Probabilidad

2.1. Introducción y resultados principales

Como ya hemos discutido en la Sección 1.5 (véase el Teorema 1.5.11) y en la Introducción de esta Memoria, el estudio de la distribución y probabilidad del condicionamiento lineal no es otra cosa que el análisis del volumen de un tubo alrededor de una variedad algebraica proyectiva compleja particular: La que forman las matrices singulares.
En este capítulo establecemos cotas superiores e inferiores para el volumen de un tubo alrededor de una variedad proyectiva cualquiera. Dado un subconjunto cualquiera $T \subseteq \mathbb{P}_n(\mathbb{C})$, y para todo número positivo $\varepsilon > 0$, definimos el *tubo de radio ε alrededor de T* como el subconjunto $T_\varepsilon \subseteq \mathbb{P}_n(\mathbb{C})$ siguiente:

$$T_\varepsilon := \{z \in \mathbb{P}_n(\mathbb{C}) : d_{\mathbb{P}}(z, T) < \varepsilon\},$$

donde $d_{\mathbb{P}}$ es la distancia proyectiva definida en la Sección 1.2. Dicho de otra manera, T_ε es el conjunto (abierto) de puntos del espacio proyectivo complejo cuya distancia a algún punto de T es menor que ε.
Muchos autores han realizado estudios sobre tubos. Entre los que han trabajado en el espacio proyectivo complejo podemos citar a F. J. Flaherty, a R. A. Wolf o a A. Gray (véase respectivamente [46, 133, 60]). En la monografía [61] se pueden encontrar gran cantidad de referencias y resultados al respecto. El estudio más relevante de los citados es el debido a Alfred Gray. Gray estudia valores exactos para el volumen $\nu_n[V_\varepsilon]$ para el caso de que $V \subseteq \mathbb{P}_n(\mathbb{C})$ es una variedad proyectiva lisa y ε un número positivo menor que la distancia de T a su punto focal más próximo (véase [61] para una definición detallada). Lamentablemente, la fórmula no es válida para variedades singulares. Además, puede ser una tarea extremadamente difícil

el cálculo de la distancia focal a la que nos hemos referido. Por estas dos razones, los resultados de Gray no son fáciles de utilizar en la práctica.
Estas consideraciones llevaron a J. Renegar, y posteriormente a J. Demmel, a la búsqueda de otros resultados que, aunque menos precisos que los de Gray, tuviesen una mayor aplicabilidad y un campo de acción menos restrictivo. Así, primero Renegar en [103], y luego Demmel en [39, 40], acotaron superior e inferiormente el volumen de tubos alrededor de subvariedades algebraicas de $\mathbb{P}_n(\mathbb{C})$. En este capítulo obtendremos cotas más precisas para este problema, y también obtendremos algunos otros resultados con interés geométrico propio. Los resultados aquí incluidos pueden encontrarse en [9, 10].
Como en el Capítulo 1, para cada conjunto medible $A \subseteq \mathbb{P}_n(\mathbb{C})$ denotamos por $\nu_n[A]$ su volumen para la estructura Riemanniana usual de $\mathbb{P}_n(\mathbb{C})$. Además, recordemos que para cada natural $k \geq 0$ hemos denotado por ϑ_k el volumen del espacio $\mathbb{P}_k(\mathbb{C})$ (véase la identidad (1.2)). El siguiente resultado es una versión del resultado técnico 2.3.6 que demostraremos más adelante.

Teorema 2.1.1 *Sea $V \subseteq \mathbb{P}_n(\mathbb{C})$ una variedad proyectiva equidimensional de dimensión compleja m. Sea $\varepsilon > 0$ un número real positivo. Entonces,*

$$\frac{\nu_m[V_\varepsilon]}{\vartheta_n} \leq 2\deg(V)\left(\frac{e\ n\ \varepsilon}{n-m}\right)^{2(n-m)}.$$

donde $\deg(V)$ es el grado de V y e es la base del logaritmo neperiano.

Podemos aplicar directamente este resultado a obtener una cota para la distribución de probabilidad del condicionamiento usual en el caso de matrices cuadradas (caso estudiado por Smale, Demmel o Edelman en [117, 39, 42]), obteniendo:

$$\mathrm{Prob}[\{A \in \mathcal{M}_n(\mathbb{C}) : \kappa_D(A) > \varepsilon^{-1}\}] \leq 2e^2n^5\varepsilon^2,$$

donde Prob significa probabilidad. Esta acotación es, obviamente, menos precisa que la igualdad obtenida por Edelman, sin embargo proviene de una técnica más general. También demostraremos que las constantes que aparecen en la parte derecha de la desigualdad del Teorema 2.1.1 son esencialmente óptimas.
De todos modos, ni los resultados de Smale, Renegar, Demmel o Edelman ni el Teorema 2.1.1 se pueden utilizar para estimar la probabilidad de distribución del número de condicionamiento singular (o la cantidad de la ecuación (2) de la Introducción). Para alcanzar ese objetivo necesitamos acotar el volumen de la intersección de un tubo extrínseco con otra variedad. El siguiente resultado es el más relevante de este capítulo y responde a este requerimiento.

Teorema 2.1.2 *Sean $V, V' \subseteq \mathbb{P}_n(\mathbb{C})$ dos variedades proyectivas equidimensionales de dimensiones respectivas $m > m' \geq 1$. Sea $\varepsilon > 0$ un número*

positivo. Con las notaciones anteriores, tenemos:

$$\frac{\nu_m[V'_\varepsilon \cap V]}{\nu_m[V]} \leq 2\deg(V')\left(\frac{en}{n-m'}\right)^{2(n-m')}\left[e\frac{n-m'}{m-m'}\ \varepsilon\right]^{2(m-m')},$$

donde $\deg(V')$ *es el grado de* V'*.*

La constante dependiente de n, m, m' que aparece en el Teorema 2.1.2 es esencialmente igual al cuadrado del coeficiente multinomial

$$\frac{n!}{(m')!(n-m)!(m-m')!}.$$

El siguiente resultado, que da nombre al capítulo, se sigue (casi) inmediatamente del Teorema 2.1.2, o de su versión más técnica, el Teorema 2.3.6 que expondremos más adelante.

Teorema 2.1.3 *Con las notaciones que acabamos de introducir,*

$$\frac{vol[\{A \in \Sigma^{n-1}_{\mathcal{M}} : \kappa_D^{(n-1)}(A) > \varepsilon^{-1}\}]}{vol[\Sigma^{n-1}_{\mathcal{M}}]} \leq (n^{10/3}\varepsilon)^6.$$

Además, la esperanza de $\kappa_D^{(n-1)}$ *en el espacio de matrices singulares satisface la siguiente desigualdad:*

$$E_{\Sigma^{n-1}_{\mathcal{M}}}[\kappa_D^{(n-1)}] \leq 2n^{10/3}.$$

2.2. Teoría de la Intersección y Geometría Integral en el espacio proyectivo complejo

En el capítulo preliminar (Sección 1.3) de esta memoria hemos recordado algunas definiciones y propiedades de las variedades algebraicas en el espacio proyectivo complejo. A continuación expondremos algunos resultados que relacionan estos elementos de Teoría de la Intersección con resultados clásicos de Geometría Integral, así como algunas consecuencias de esta combinación. Nuestro punto de partida es la llamada Fórmula de Poincaré, en su versión proyectiva compleja. Ésta fórmula puede encontrarse en el artículo de Ralph Howard [75]. De hecho, la versión que utilizamos es ligeramente más general que la que aparece en el artículo de Howard.

Teorema 2.2.1 *Sean* M *y* M' *dos subvariedades diferenciables complejas de* $\mathbb{P}_n(\mathbb{C})$*, de dimensiones respectivas* $m, p \in \mathbb{N}$*. Sea* $f : M \longrightarrow [0, \infty]$ *una función integrable, y supongamos que* $m + p \geq n$*. Entonces, tenemos:*

$$\nu_m[M'] \int_M f\ dM = \frac{\vartheta_p\ \vartheta_m}{\vartheta_{p+m-n}} \int_{U \in \mathcal{U}_{n+1}} \int_{x \in UM' \cap M} f(x)\ d(UM' \cap M)\ d\mathcal{U}_{n+1}.$$

El Teorema de Convergencia Monótona permite probar este resultado a partir del casi idéntico encontrado en [75, pp. 13-18]. También se puede obtener directamente a partir de la Fórmula de la Co–área (Teorema 1.1.13). La correcta interpretación de este resultado pasa por el hecho de que, con las notaciones del Teorema 2.2.1, para casi todas las matrices unitarias $U \in \mathcal{U}_{n-1}$, el conjunto $UM' \cap M$ es o bien vacío o bien una variedad diferenciable de dimensión $p+m-n$. No hemos encontrado una demostración apropiada de este hecho, por lo que la incluimos aquí. Necesitaremos un lema previo.

Lema 2.2.2 *Sea $x \in \mathbb{C}^{n+1} \setminus \{0\}$ un punto. Sea $S^{\|x\|_2}(\mathbb{C}^{n+1})$ la esfera de centro 0 y radio $\|x\|_2$ en $\mathbb{C}^{n+1} \equiv \mathbb{R}^{2n+2}$. Entonces, la siguiente aplicación es una submersión (esto es, el conjunto de valores críticos es vacío):*

$$\begin{array}{rccc} \psi: & \mathcal{U}_{n+1} & \longrightarrow & S^{\|x\|_2}(\mathbb{C}^{n+1}) \\ & U & \mapsto & Ux. \end{array}$$

Demostración.– Obsérvese que ψ es sobreyectiva. Sean pues dos puntos $z_1, z_2 \in S^{\|x\|_2}(\mathbb{C}^{n+1})$, y sean $U_1, U_2 \in \mathcal{U}_{n+1}$ dos matrices unitarias tales que

$$\psi(U_1) := U_1 x = z_1, \quad \psi(U_2) := U_2 x = z_2.$$

Sea $U' := U_2 U_1^{-1} \in \mathcal{U}_{n+1}$ la matriz unitaria tal que $U'U_1 = U_2$. Entonces, $Uz_1 = z_2$ y el siguiente es un diagrama conmutativo:

$$\begin{array}{ccc} \mathcal{U}_{n+1} & \xrightarrow{\psi} & S^{\|x\|_2}(\mathbb{C}^{n+1}) \\ U' \downarrow & & \downarrow Iso_{U'} \\ \mathcal{U}_{n+1} & \xrightarrow{\psi} & S^{\|x\|_2}(\mathbb{C}^{n+1}) \end{array}$$

donde $U'(U) = U'_L(U) = U'U$ es la multiplicación a izquierda definida por U' y $Iso_{U'}$ es la isometría definida por U' en $S^{2n+1}(\|x\|_2)$ ($Iso_{U'}(v) = U'v \; \forall v \in S^{2n+1}(\|x\|_2)$). Se tiene, por tanto, la siguiente igualdad entre aplicaciones diferenciales

$$d_{z_1}(Iso_{U'}) d_{U_1}\psi = d_{U_1}(Iso_{U'} \circ \psi) = d_{U_1}(\psi \circ U') = d_{U_2}\psi d_{U_1} U'.$$

Como $d_{z_1}(Iso_{U'})$ y $d_{U_1}U'$ son isomorfismos lineales, deducimos que $d_{U_1}\psi$ es sobreyectiva si y sólo si $d_{U_2}\psi$ es sobreyectiva. Esto es, z_1 es un valor regular de ψ si y sólo si z_2 también lo es. Por otro lado el Teorema 1.1.4 nos asegura que el conjunto de valores regulares de ψ es un conjunto denso en $S^{2n+1}(\|x\|_2)$, lo que acaba la demostración del Lema.

■

Lema 2.2.3 *Sean M y M' dos subvariedades diferenciables complejas de $\mathbb{P}_n(\mathbb{C})$, de dimensiones respectivas $m, p \in \mathbb{N}$. Entonces, existe un conjunto $W \subseteq \mathcal{U}_{n+1}$ que depende sólo de M y M' tal que su complementario tiene medida nula en $\mathcal{U}_{n+1}$ y que las propiedades siguientes se satisfacen:*

1. *Si $m+p<n$, para toda $U \in W$, $M \cap UM' = \emptyset$.*

2. *Si $m+p \geq n$, para toda $U \in W$, $M \cap UM'$ es vacío o una subvariedad diferenciable de $\mathbb{P}_n(\mathbb{C})$ de dimensión compleja $m+p-n$.*

Demostración.– Denotemos por $\pi : \mathbb{C}^{n+1} \longrightarrow \mathbb{P}_n(\mathbb{C})$ la proyección natural de $\mathbb{C}^{n+1}$ en $\mathbb{P}_n(\mathbb{C})$ que lleva cada punto afín $\underline{x}$ a su clase proyectiva $x = \pi(\underline{x})$. Sean $\widetilde{M}, \widetilde{M'} \subset \mathbb{C}^{n+1} \setminus \{0\}$ los conos respectivos sobre M y M'. Esto es,

$$\widetilde{M} := \pi^{-1}(M), \quad \widetilde{M'} := \pi^{-1}(M').$$

Obsérvese que $\widetilde{M}$ y $\widetilde{M'}$ son subvariedades diferenciables complejas de $\mathbb{C}^{n+1}$ de dimensiones complejas:

$$\begin{aligned} \dim(\widetilde{M}) &= \dim(M)+1, \\ \dim(\widetilde{M'}) &= \dim(M')+1. \end{aligned}$$

Consideremos la siguiente aplicación entre variedades diferenciables reales:

$$\begin{array}{rccc} \varphi: & \mathcal{U}_{n+1} \times \widetilde{M'} \times \widetilde{M} & \longrightarrow & \mathbb{C}^{n+1} \\ & (U, \underline{y}, \underline{x}) & \mapsto & U\underline{y} - \underline{x}. \end{array}$$

Demostraremos que φ es transversal a $\{0\} \in \mathbb{C}^{n+1}$. Esto es, que $0 \in \mathbb{C}^{n+1}$ no es un valor crítico de φ. Sea $\varphi^{-1}(\{0\})$ la fibra en $\{0\}$. Queremos probar que todo punto $P := (U, \underline{y}, \underline{x}) \in \varphi^{-1}(\{0\})$ es un punto regular de φ. Esto es, que la diferencial $d_P\varphi : T_U\mathcal{U}_{n+1} \times T_{\underline{y}}\widetilde{M'} \times T_{\underline{x}}\widetilde{M} \longrightarrow T_0\mathbb{C}^{n+1}$ es sobreyectiva. Como $\widetilde{M'}$ es un cono, se tiene que $\underline{y} \in T_{\underline{y}}\widetilde{M'}$. Por tanto,

$$d_P\varphi(0, \underline{y}, 0) = U\underline{y} = \underline{x} \in T_0\mathbb{C}^{n+1},$$

Veamos ahora que los elementos del espacio ortogonal a $\underline{x}$ también están en la imagen de $d_P\varphi$. En primer lugar, $U\underline{y} = \underline{x}$ implica que $\|\underline{y}\|_2 = \|\underline{x}\|_2$. Sea $\varphi_{\underline{y},\underline{x}}$ la restricción de φ a $\mathcal{U}_{n+1} \times \{\underline{y}\} \times \{\underline{x}\}$, y consideremos la aplicación

$$\begin{array}{rccc} \psi_{\underline{y},\underline{x}}: & \mathcal{U}_{n+1} & \longrightarrow & S^{\|y\|_2}(\mathbb{C}^{n+1}) \\ & U & \mapsto & U\underline{y}. \end{array}$$

Sea $S^{\|y\|_2}_{-x}(\mathbb{C}^{n+1}) = \{\underline{z} \in \mathbb{C}^{n+1} : \|\underline{z} + \underline{x}\|_2 = \|\underline{y}\|_2\}$ la esfera centrada en $-\underline{x}$ de radio $\|\underline{y}\|_2$, y sea $t_{\underline{x}}$ la traslación

$$\begin{array}{rccc} t_{\underline{x}}: & S^{\|y\|_2}_{-x}(\mathbb{C}^{n+1}) & \longrightarrow & S^{\|x\|_2}(\mathbb{C}^{n+1}) \\ & v & \mapsto & v + \underline{x}. \end{array}$$

Entonces, se tiene que $\psi_{\underline{y},\underline{x}} = t_{\underline{x}} \circ \varphi_{\underline{y},\underline{x}}$.

Por el Lema 2.2.2 sabemos que $\psi_{\underline{y},\underline{x}}$ no tiene valores críticos y, por tanto, $\varphi_{\underline{y},\underline{x}}$ tampoco tiene valores críticos. En particular, se cumple que

$$T_0 S_{-x}^{\|y\|_2}(\mathbb{C}^{n+1}) \subseteq Im(d_U \varphi_{\underline{y},\underline{x}}) \subseteq Im(d_P \varphi).$$

Ahora, $T_0 S_{-x}^{\|y\|_2}(\mathbb{C}^{n+1})$ es exactamente el espacio ortogonal a $\underline{x}$ en 0.
Concluimos que $d_P\varphi$ es sobreyectiva y todo $P \in \varphi^{-1}(\{0\})$ es un punto regular de φ. Por tanto, el Corolario 1.1.6 nos asegura que $\varphi^{-1}(\{0\})$ es una variedad diferenciable real y el Teorema 1.1.7 garantiza la existencia de un subconjunto $W \subseteq \mathcal{U}_{n+1}$ cuyo complementario tiene medida nula y tal que 0 es un valor regular de la aplicación

$$\begin{array}{rccc} \varphi_U: & \widetilde{M'} \times \widetilde{M} & \longrightarrow & \mathbb{C}^{n+1} \\ & (\underline{y},\underline{x}) & \mapsto & U\underline{y} - \underline{x}. \end{array}$$

En particular, para toda matriz $U \in W$ la fibra $\varphi_U^{-1}(\{0\})$ es una variedad diferenciable (posiblemente vacía) cuya dimensión compleja satisface:

$$\dim(\varphi_U^{-1}(\{0\})) = \dim(\widetilde{M'}) + \dim(\widetilde{M}) - \operatorname{codim}_{\mathbb{C}^{n+1}}(\{0\}) = m + p - n + 1, \tag{2.1}$$

para toda matriz $U \in W$.
Por otro lado, sea $U \in W$ una matriz unitaria del conjunto que hemos obtenido. Sea $M \cap UM' \subset \mathbb{P}_n(\mathbb{C})$ ese conjunto proyectivo y sea $\widetilde{M \cap UM'}$ el cono sobre $M \cap UM'$. Esto es,

$$\widetilde{M \cap UM'} := \pi^{-1}(M \cap UM').$$

Obsérvese que la siguiente aplicación es un isomorfismo entre $\varphi_U^{-1}(\{0\})$ y $\widetilde{M \cap UM'}$:

$$\begin{array}{rccc} \pi_2: & \varphi_U^{-1}(\{0\}) & \longrightarrow & \widetilde{M \cap UM'} \\ & (\underline{y},\underline{x}) & \mapsto & \underline{x}. \end{array}$$

El inverso de π_2 viene dado por la igualdad

$$\pi_2^{-1}(\underline{x}) = (U^{-1}\underline{x}, \underline{x}).$$

Por tanto, para toda matriz $U \in W$, $\widetilde{M \cap UM'}$ es vacío o una subvariedad diferenciable compleja de $\mathbb{C}^{n+1}$ de dimensión compleja $m+p-n+1$. Como $\widetilde{M \cap UM'}$ es el cono sobre $M \cap UM'$, concluimos que para todo $U \in W$, $M \cap UM'$ es vacío o una subvariedad diferenciable compleja de $\mathbb{P}_n(\mathbb{C})$ de dimensión compleja $m+p-n$. Finalmente, el lema se sigue de que $M \cap UM' = \emptyset$ si y sólo si $\dim(M \cap UM') = m + p - n < 0$. ∎

Nota 2.2.4 *El Lema 2.2.3 permite una interpretación más correcta del Teorema 2.2.1, como hemos indicado. En efecto, la integración en el espacio de matrices unitarias $\mathcal{U}_{n+1}$ que se propone en el Teorema 2.2.1 se realiza en realidad en el conjunto W que existe por el Lema 2.2.3. En este conjunto, podemos asegurar que la intersección $UM' \cap M$ es una variedad diferenciable compleja de la dimensión apropiada.*

Corolario 2.2.5 *Sea $f : \mathbb{P}_n(\mathbb{C}) \longrightarrow \mathbb{R}$ una función integrable. Sea $z \in \mathbb{P}_n(\mathbb{C})$ un punto cualquiera del espacio proyectivo complejo. Entonces, tenemos:*

$$\int_{x \in \mathbb{P}_n(\mathbb{C})} f(x)\ d\mathbb{P}_n(\mathbb{C}) = \vartheta_n \int_{U \in \mathcal{U}_{n+1}} f(Uz)\ d\mathcal{U}_{n+1}.$$

Demostración.– Basta aplicar el Teorema 2.2.1 en el caso de que $M = \mathbb{P}_n(\mathbb{C})$, $M' = \{z\}$. ∎

El resultado que demostramos a continuación es una generalización de la Fórmula de Poincaré (Teorema 2.2.1) que resultará de gran utilidad en el estudio de las propiedades de las variedades proyectivas.

Corolario 2.2.6 *Sean $V, V' \subseteq \mathbb{P}_n(\mathbb{C})$ dos variedades proyectivas equidimensionales de dimensiones respectivas m y p. Supongamos que $m + p - n \geq 0$. Sean $A \subset V$, $A' \subset V'$ dos conjuntos abiertos (para la topología inducida por la de $\mathbb{P}_n(\mathbb{C})$) de V y V'. Entonces, para casi toda matriz $U \in \mathcal{U}_{n+1}$, $V \cap UV'$ es una variedad proyectiva de dimensión $m+p-n$. Además, tenemos:*

$$\nu_m[A]\ \nu_p[A'] = \frac{\vartheta_p\ \vartheta_m}{\vartheta_{p+m-n}} \int_{U \in \mathcal{U}_{n+1}} \nu_{m+p-n}[A \cap UA']\ d\mathcal{U}_{n+1}.$$

Demostración.– Sea $W_1 \subseteq \mathcal{U}_{n+1}$ el conjunto asociado a $Reg(V)$ y $Reg(V')$ cuyo complementario tiene medida nula proporcionado por el Lema 2.2.3. Esto es, para toda matrix $U \in W_1$, $Reg(V) \cap URreg(V')$ es una variedad diferenciable proyectiva (posiblemente vacía) de dimensión compleja $m + p - n$. Por otro lado, $V \setminus Reg(V)$ puede ser descrita como una unión finita de variedades diferenciables disjuntas de dimensión a lo más $m - 1$. De modo similar, $V' \setminus Reg(V')$ puede también ser descrita mediante una unión finita de variedades diferenciables de dimensión a lo más $p-1$. Por tanto, de nuevo por el Lema 2.2.3, existe un subconjunto $W_2 \subseteq \mathcal{U}_{n+1}$ cuyo complementario tiene medida nula y tal que

$$U(V' \setminus Reg(V')) \cap Reg(V),\ URreg(V') \cap (V \setminus Reg(V)) \text{ y}$$

$$U(V' \setminus Reg(V')) \cap (V \setminus Reg(V))$$

son uniones finitas disjuntas de variedades diferenciables de dimensión a lo más $m + p - n - 1$. Entonces, para toda matriz $U \in W = W_1 \cap W_2$ las siguientes propiedades se satisfacen:

- $V \cap UV'$ es una variedad proyectiva.
- $V \cap UV'$ es una unión finita disjunta de variedades diferenciables de dimensión a lo más $m + p - n$.
- $Reg(V) \cap URreg(V')$ es una variedad diferenciable de dimensión compleja $m + p - n$.
- $(V \cap UV') \setminus (Reg(V) \cap URreg(V'))$ es un conjunto constructible de dimensión a lo más $m + p - n - 1$.

Por lo tanto, $V \cap UV'$ es una variedad proyectiva de dimensión $m + p - n$. Ahora, existen conjuntos abiertos $T, T' \subseteq \mathbb{P}_n(\mathbb{C})$ tales que $A = V \cap T, A' = V' \cap T'$. Por tanto,

$$A \cap UA' = (T \cap UT') \cap (V \cap UV').$$

Luego para toda matriz unitaria $U \in W$, $A \cap UA'$ es un subconjunto abierto de $V \cap UV'$y se tiene que

$$\nu_{m+p-n}[A \cap UA'] = \nu_{m+p-n}[(A \cap UA') \cap Reg(V \cap UV')] =$$

$$= \nu_{m+p-n}[(A \cap Reg(V)) \cap (UA' \cap Reg(UA'))].$$

Además, tenemos:

$$\nu_m[A] = \nu_m[A \cap Reg(V)], \qquad \nu_m[A'] = \nu_m[A' \cap Reg(V')].$$

El corolario se sigue inmediatamente del Teorema 2.2.1, aplicado a las variedades diferenciables $A \cap Reg(V)$ and $A' \cap Reg(V')$. ∎

El siguiente corolario relaciona el grado de una variedad proyectiva y su volumen. Éste es un resultado clásico de Teoría de la Intersección, al menos en el caso de variedades lisas (i.e. variedades sin puntos singulares). Por ejemplo, una demostración para ese caso puede encontrarse en [96, Teor. 5.22]. Nosotros proponemos una demostración alternativa basada en el uso de la Geometría Integral, que es independiente del hecho de que la variedad sea o no lisa.

Corolario 2.2.7 *Sea $V \subseteq \mathbb{P}_n(\mathbb{C})$ una variedad proyectiva equi–dimensional de dimensión m. Entonces, tenemos:*

$$\nu_m[V] = \vartheta_m \deg(V).$$

Demostración.– Sea $M := Reg(V)$ la variedad diferenciable formada por los puntos simples de V. Sea $L^{n-m} \subseteq \mathbb{P}_n(\mathbb{C})$ un subespacio lineal de $P_n(\mathbb{C})$ de dimensión $n - m$. Por la prueba de Lema 2.2.3, existe un conjunto cuyo complementario tiene medida nula $W \subseteq \mathcal{U}_{n+1}$ tal que para toda matriz

$U \in W$, UL^{n-m} y M son transversales en cualquier cero común. Esto es, $UL^{n-m} \cap M$ es una variedad diferenciable cero–dimensional y para todo punto $x \in UL^{n-m} \cap M$, los espacios tangentes T_xUL^{n-m} y T_xM son transversales. Por el Teorema [96, Teor. 5.16],concluimos que para toda matriz $U \in W$, $\sharp(UL^{n-m} \cap M) = \sharp(UL^{n-m} \cap V) = \deg(V)$.
Del Corolario 2.2.6 concluimos:

$$\nu_m[V]\ \nu_{n-m}[L^{n-m}] = \vartheta_{n-m}\vartheta_m \int_{U\in\mathcal{U}_{n+1}} \sharp(V \cap UL^{n-m})\ d\mathcal{U}_{n+1}.$$

Por tanto,

$$\nu_m[V]\ \vartheta_{n-m} = \vartheta_{n-m}\vartheta_m \deg(V),$$

y el corolario queda demostrado. ∎

Corolario 2.2.8 *Sea $V \subseteq \mathbb{P}_n(\mathbb{C})$ una variedad proyectiva equi–dimensional de dimensión m. Sea $A \subseteq V$ un subconjunto abierto de V y $0 \leq \varepsilon \leq 1$ un número positivo. Entonces, tenemos:*

$$\nu_m[A]\vartheta_n\varepsilon^{2n} = \int_{x\in\mathbf{P}_n(\mathbb{C})} \nu_m[B_{\mathbf{P}}(x,\varepsilon)\cap A]\ d\mathbb{P}_n(\mathbb{C}).$$

Demostración.– El Corolario 2.2.6 aplicado a los conjuntos A y $B_{\mathbf{P}}(e_0,\varepsilon)$ implica:

$$\nu_m[A]\nu_n[B_{\mathbf{P}}(e_0,\varepsilon)] = \frac{\vartheta_n\ \vartheta_m}{\vartheta_m} \int_{U\in\mathcal{U}_{n+1}} \nu_m[UB_{\mathbf{P}}(e_0,\varepsilon)\cap A]\ d\mathcal{U}_{n+1}.$$

Ahora, por el Corolario 2.2.5 se tiene que

$$\int_{U\in\mathcal{U}_{n+1}} \nu_m[UB_{\mathbf{P}}(e_0,\varepsilon)\cap A]\ d\mathcal{U}_{n+1} = \frac{1}{\vartheta_n}\int_{x\in\mathbf{P}_n(\mathbb{C})} \nu_m[B_{\mathbf{P}}(x,\varepsilon)\cap A]\ d\mathbb{P}_n(\mathbb{C}).$$

Por tanto, hemos obtenido que:

$$\nu_m[A]\nu_n[B_{\mathbf{P}}(e_0,\varepsilon)] = \int_{x\in\mathbf{P}_n(\mathbb{C})} \nu_m[B_{\mathbf{P}}(x,\varepsilon)\cap A]\ d\mathbb{P}_n(\mathbb{C}).$$

La siguiente igualdad (véase Sección 1.2) termina la demostración:

$$\nu_n[B_{\mathbf{P}}(e_0,\varepsilon)] = \vartheta_n\varepsilon^{2n}.$$

∎

El siguiente corolario puede verse como una igualdad de Bézout en media:

Corolario 2.2.9 *Sean $V, V' \subseteq \mathbb{P}_n(\mathbb{C})$ dos variedades proyectivas equi–dimensionales de dimensiones respectivas m y p. Supongamos que $m+p \geq n$. Entonces, para casi toda matriz $U \in \mathcal{U}_{n+1}$, $V \cap UV'$ es una variedad proyectiva equidimensional de dimensión $m+p-n$ y se tiene la siguiente igualdad:*

$$\deg(V)\deg(V') = \int_{\mathcal{U}_{n+1}} \deg(V \cap UV') \; d\mathcal{U}_{n+1}.$$

Demostración.– Aplicamos el Corolario 2.2.6 a V y V'. Entonces, usamos el Corolario 2.2.7 para reemplazar $\nu_m[V]$ por $\vartheta_m \deg(V)$, y hacemos lo mismo para V' y $UV' \cap V$. ∎

Nota 2.2.10 *En [108], un resultado similar al del Corolario 2.2.9 es anunciado sin demostración. Si combinamos la desigualdad clásica de Bézout (véase [66]) y el Corolario 2.2.9, obtenemos:*

$$\nu_{\mathcal{U}_{n+1}}[U \in \mathcal{U}_{n+1} : \deg(V \cap UV') \neq \deg(V)\deg(V')] = 0,$$

para V, V' variedades proyectivas equi–dimensionales de dimensiones m, p con $m+p \geq n$.

2.3. Volumen de tubos en el espacio proyectivo complejo

2.3.1. Algunos resultados de Geometría Proyectiva

Para demostrar las acotaciones del volumen de tubos y sus consecuencias utilizaremos los resultados de la Sección 2.2 y algunas herramientas de Teoría de la Integración Geométrica, como la Fórmula de la Co–área (Teorema 1.1.13).

Lema 2.3.1 *Sea $\{(\mathbb{A}_i^n, \varphi_i) : 0 \leq i \leq n\}$ el atlas de $\mathbb{P}_n(\mathbb{C})$ dado por las cartas afines. Es decir,*

$$\begin{array}{rccc} \varphi_i : & \mathbb{C}^n & \longrightarrow & \mathbb{A}_i^n := \{x \in \mathbb{P}_n(\mathbb{C}) : x_i \neq 0\} \subseteq \mathbb{P}_n(\mathbb{C}) \\ & (z_1, \ldots, z_n) & \mapsto & (z_1 : \cdots : z_i : 1 : z_{i+1} : \cdots : z_n). \end{array}$$

Entonces, para todo punto $\underline{z} \in \mathbb{C}^n$ tenemos las siguientes propiedades:

1. *Para cada vector tangente $v \in T_{\underline{z}}\mathbb{C}^n$, $\|v\|_{T_{\underline{z}}\mathbb{C}^n} = 1$,*

$$\frac{1}{1+\|z\|_2^2} \leq \|d_{\underline{z}}\varphi_i(v)\|_{T_{\varphi_i(\underline{z})}\mathbb{P}_n(\mathbb{C})} \leq \frac{1}{(1+\|z\|_2^2)^{1/2}}.$$

2. *El jacobiano normal de φ_i (tal como ha sido definido en el capítulo de preliminares, Sección 1.1.2) satisface:*

$$NJ_{\underline{z}}\ \varphi_i = \frac{1}{(1+\|\underline{z}\|_2^2)^{n+1}}.$$

3. *Para cada subvariedad diferenciable compleja $M \subseteq \mathbb{C}^n$ de dimensión $m \geq 1$, y para cada punto $\underline{z} \in M$, el jacobiano normal de la restricción $\varphi_i\mid_M : M \longrightarrow \varphi_i(M)$ satisface:*

$$\frac{1}{(1+\|\underline{z}\|_2^2)^{m+1}} \leq NJ_{\underline{z}}(\varphi_i\mid_M) \leq \frac{1}{(1+\|\underline{z}\|_2^2)^{m}}.$$

Demostración.– Primero, observamos que basta con hacer la prueba para el caso $i = 0$. Denotamos pues $\varphi := \varphi_0$. Esto es,

$$\begin{array}{rccc} \varphi := \varphi_0 : & \mathbb{C}^n & \longrightarrow & \mathbb{A}_0^n := \{x \in \mathbb{P}_n(\mathbb{C}) : x_0 \neq 0\} \subseteq \mathbb{P}_n(\mathbb{C}) \\ & (z_1, \ldots, z_n) & \mapsto & (1 : z_1 : \cdots : z_n). \end{array}$$

Sea $0 \in \mathbb{C}^n$ el origen y $e_0 = \varphi(0) = (1 : 0 : \cdots : 0)$ su imagen. Por definición de la estructura Riemanniana de $\mathbb{P}_n(\mathbb{C})$, φ es una isometría en $0 \in \mathbb{C}^n$. Esto es, la diferencial

$$d_0\varphi : T_0\mathbb{C}^n \longrightarrow T_{e_0}\mathbb{P}_n(\mathbb{C})$$

es una isometría lineal y, por lo tanto, $NJ_0\varphi = 1$. Sea $\underline{z} \in \mathbb{C}^n$ un punto cualquiera, $\underline{z} = (z_1, \ldots, z_n)$. Sea $U \in \mathcal{U}_{n+1}$, $U = (u_{ij})_{i,j=0\ldots n}$ una matriz unitaria tal que $U\varphi(\underline{z}) = e_0$. Esto es,

$$U \begin{pmatrix} 1 \\ \underline{z}^t \end{pmatrix} = \begin{pmatrix} (1+\|\underline{z}\|_2^2)^{1/2} \\ 0 \end{pmatrix}. \tag{2.2}$$

Sean $U_0, \ldots, U_n$ las filas de U. Obsérvese que podemos elegir U_0 de forma que

$$U_0 = \frac{1}{(1+\|\underline{z}\|_2^2)^{1/2}}\ (1, \overline{z_1}, \ldots, \overline{z_n}),$$

donde $\overline{\cdot}$ denota simplemente conjugación compleja. Además, $U_1, \ldots, U_n$ son vectores (complejos) ortogonales a U_0. Por otro lado, $U : \mathbb{P}_n(\mathbb{C}) \longrightarrow \mathbb{P}_n(\mathbb{C})$ es también una isometría en cualquier punto proyectivo y, por tanto, $NJ_{\varphi(\underline{z})}U = 1$.Finalmente, sea $\phi : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ la aplicación dada por:

$$\phi := \varphi^{-1} \circ U \circ \varphi.$$

Obsérvese que $\phi(z) = 0$ y $\varphi \circ \phi = U \circ \varphi$. Por tanto por la Proposición 1.1.11, tenemos la siguiente igualdad de jacobianos normales:

$$NJ_0\varphi\ NJ_{\underline{z}}\phi = NJ_{\varphi(\underline{z})}U\ NJ_{\underline{z}}\varphi.$$

Concluimos que $NJ_{\underline{z}}\phi = NJ_{\underline{z}}\varphi$.
Además, para cada vector tangente $v \in T_{\underline{z}}\mathbb{C}^n$, tenemos

$$d_{\underline{z}}\phi(v) = \frac{1}{(1+\|\underline{z}\|_2^2)^{1/2}}\left(U_1\begin{pmatrix}0\\v^t\end{pmatrix},\ldots,U_n\begin{pmatrix}0\\v^t\end{pmatrix}\right),$$

donde v^t es el traspuesto del vector v. Sean $v, w \in T_{\underline{z}}\mathbb{C}^n$ dos vectores tangentes. Entonces, tenemos

$$\langle d_{\underline{z}}\phi(v), d_{\underline{z}}\phi(w)\rangle_{T_0\mathbb{C}^n} = \frac{1}{1+\|\underline{z}\|_2^2}\sum_{i=1}^{n} U_i\begin{pmatrix}0\\v^t\end{pmatrix}\overline{U_i\begin{pmatrix}0\\w^t\end{pmatrix}}.$$

Por tanto,

$$\langle d_{\underline{z}}\phi(v), d_{\underline{z}}\phi(w)\rangle_{T_0\mathbb{C}^n} = \frac{1}{1+\|\underline{z}\|_2^2}\left[\langle v, w\rangle_{\mathbb{C}^n} - U_0\begin{pmatrix}0\\v^t\end{pmatrix}\overline{U_0\begin{pmatrix}0\\w^t\end{pmatrix}}\right].$$

Supongamos ahora que $\langle v, \underline{z}\rangle_{\mathbb{C}^n} = 0$. Entonces, se tiene que

$$U_0\begin{pmatrix}0\\v^t\end{pmatrix} = \frac{1}{(1+\|\underline{z}\|_2^2)^{1/2}}\langle v, \underline{z}\rangle_{\mathbb{C}^n} = 0.$$

Concluimos que para cada vector $v \in T_{\underline{z}}\mathbb{C}^n$ tal que $\langle v, \underline{z}\rangle_{\mathbb{C}^n} = 0$, y para cada vector $w \in T_{\underline{z}}\mathbb{C}^n$,

$$\langle d_{\underline{z}}\phi(v), d_{\underline{z}}\phi(w)\rangle_{T_0\mathbb{C}^n} = \frac{1}{1+\|\underline{z}\|_2^2}\langle v, w\rangle_{\mathbb{C}^n}.$$

Sea ahora $\{b_1,\ldots,b_n\}$ una base ortonormal de $T_{\underline{z}}\mathbb{C}^n$ tal que $b_n = \frac{1}{\|\underline{z}\|_2}\underline{z}$. Esto implica que $\langle b_i, \underline{z}\rangle_{T_{\underline{z}}\mathbb{C}^n} = 0$ para $i = 1\ldots n-1$. Entonces, tenemos:

$$\langle d_{\underline{z}}\phi(b_i), d_{\underline{z}}\phi(b_j)\rangle_{T_0\mathbb{C}^n} = \frac{1}{1+\|\underline{z}\|_2^2}\langle b_i, b_j\rangle_{\mathbb{C}^n} = 0 \quad i \neq j.$$

Además, para cada i, $1 \leq i \leq n-1$,

$$\langle d_{\underline{z}}\phi(b_i), d_{\underline{z}}\phi(b_i)\rangle_{T_0\mathbb{C}^n} = \frac{1}{1+\|\underline{z}\|_2^2}.$$

Para $i = n$, se tiene que

$$\langle d_{\underline{z}}\phi(b_n), d_{\underline{z}}\phi(b_n)\rangle_{T_0\mathbb{C}^n} = \frac{1}{1+\|\underline{z}\|_2^2}\left[1 - \frac{1}{\|\underline{z}\|_2^2}U_0\begin{pmatrix}0\\\underline{z}^t\end{pmatrix}\overline{U_0\begin{pmatrix}0\\\underline{z}^t\end{pmatrix}}\right].$$

Ahora, obsérvese que:

$$U_0\begin{pmatrix}0\\\underline{z}^t\end{pmatrix}\overline{U_0\begin{pmatrix}0\\\underline{z}^t\end{pmatrix}} = \left[U_0\begin{pmatrix}1\\\underline{z}^t\end{pmatrix} - u_{00}\right]\left[\overline{U_0\begin{pmatrix}1\\\underline{z}^t\end{pmatrix} - u_{00}}\right] = \frac{\|\underline{z}\|_2^4}{1+\|\underline{z}\|_2^2}.$$

Concluimos por tanto que

$$\|d_{\underline{z}}\phi(b_n)\|^2_{T_0\mathbb{C}^n} = \frac{1}{1+\|\underline{z}\|_2^2}\left[1 - \frac{\|\underline{z}\|_2^2}{1+\|\underline{z}\|_2^2}\right] = \frac{1}{(1+\|\underline{z}\|_2^2)^2}.$$

Obtenemos de inmediato el item ii), puesto que

$$NJ_{\underline{z}}\varphi = NJ_{\underline{z}}\phi = \prod_{i=1}^{n} \|d_{\underline{z}}\phi(b_i)\|^2_{T_0\mathbb{C}^n} = \frac{1}{(1+\|\underline{z}\|_2^2)^{n+1}}.$$

Ahora, sea $v = T_{\underline{z}}\mathbb{C}^n$, $v = \sum_{i=1}^n \lambda_i b_i$, $\sum_{i=1}^n |\lambda_i|^2 = 1$. Entonces,

$$\|d_{\underline{z}}\phi(v)\|^2_{T_0\mathbb{C}^n} = \sum_{i=1}^{n} |\lambda_i|^2 \|d_{\underline{z}}\phi(b_i)\|^2_{T_0\mathbb{C}^n} =$$

$$\frac{1}{1+\|\underline{z}\|_2^2}\left(\sum_{i=1}^{n-1}|\lambda_i|^2 + |\lambda_n|^2\frac{1}{1+\|\underline{z}\|_2^2}\right),$$

por lo que

$$\frac{1}{1+\|\underline{z}\|_2^2} \leq \|d_{\underline{z}}\phi(v)\|_{T_0\mathbb{C}^n} \leq \frac{1}{(1+\|\underline{z}\|_2^2)^{1/2}}. \tag{2.3}$$

Ahora, dado que $\phi = \varphi^{-1}\circ U \circ \varphi$, tenemos

$$d_0\varphi \; d_{\underline{z}}\phi(v) = d_{\varphi(\underline{z})}U \; d_{\underline{z}}\varphi(v),$$

donde $d_0\varphi$ y $d_{\varphi(\underline{z})}U$ son isometrías lineales. Concluimos que

$$\|d_{\underline{z}}\phi(v)\|_{T_0\mathbb{C}^n} = \|d_{\underline{z}}\varphi(v)\|_{T_{\varphi(\underline{z})}\mathbf{P}_n(\mathbb{C})},$$

y el item i) se sigue de las desigualdades (2.3).
Denotemos por $\{b'_1,\ldots,b'_n\}$ las imágenes por $d_{\underline{z}}\varphi$ de la base $\{b_1,\ldots,b_n\}$. Esto es,

$$b'_i = d_{\underline{z}}\varphi(b_i), \quad i = 1\ldots n.$$

Entonces, hemos demostrado que $\{b'_1,\ldots,b'_n\}$ es ortogonal. En efecto,

$$\langle b'_j, b'_i\rangle_{T_{\varphi(\underline{z})}\mathbf{P}_n(\mathbb{C})} = \langle d_{\underline{z}}\phi(b_i), d_{\underline{z}}\phi(b_j)\rangle_{T_0\mathbb{C}^n} = 0, \quad i \neq j.$$

Además,

$$\|b'_i\|_{T_{\varphi(\underline{z})}\mathbf{P}_n(\mathbb{C})} = \frac{1}{(1+\|\underline{z}\|_2^2)^{1/2}}, \quad i = 1\ldots n-1,$$

y

$$\|b'_n\|_{T_{\varphi(\underline{z})}\mathbf{P}_n(\mathbb{C})} = \frac{1}{1+\|\underline{z}\|_2^2}.$$

Sea $M \subseteq \mathbb{C}^n$ una subvariedad diferenciable de dimensión compleja m, y sea $\underline{z} \in M$ un punto. Sabemos que $T_{\underline{z}}M$ es un subespacio complejo de dimensión

m de $T_{\underline{z}}\mathbb{C}^n$, que hereda un producto interior de $T_{\underline{z}}\mathbb{C}^n$. Entonces, la siguiente expresión define un subespacio lineal de $\mathbb{C}^n \equiv T_{\underline{z}}\mathbb{C}^n$ de dimensión compleja al menos $m+n-1-n = m-1$:

$$W := T_{\underline{z}}M \cap \langle\{b_1, \ldots, b_{n-1}\}\rangle,$$

donde $\langle\{b_1, \ldots, b_{n-1}\}\rangle$ es el subespacio complejo de $\mathbb{C}^n$ generado por esos vectores. Entonces, podemos encontrar una base ortonormal $\{c_1, \ldots, c_m\}$ de $T_{\underline{z}}M$ tal que $c_1, \ldots, c_{m-1} \in W$. Por lo tanto, para todo $i = 1 \ldots m-1$ tenemos

$$\|d_{\underline{z}}(\varphi\mid_M)(c_i)\|_{T_{\varphi(\underline{z})}\mathbf{P}_n(\mathbb{C})} = \frac{1}{(1+\|\underline{z}\|_2^2)^{1/2}},$$

el número real $\|d_{\underline{z}}(\varphi\mid_M)(c_m)\|_2$ está acotado por la ecuación (2.3). Sin pérdida de generalidad podemos asumir que $c_i = b_i$, para $1 \leq i \leq m-1$. Entonces, $d_{\underline{z}}(\varphi\mid_M (c_m))$ está en el subespacio complejo $\langle b'_m, \ldots, b'_n\rangle$ y es ortogonal al espacio complejo generado por $\{d_{\underline{z}}(\varphi\mid_M)(c_i) : 1 \leq i \leq m-1\}$. En particular, hemos visto que la familia de vectores $\{d_{\underline{z}}(\varphi\mid_M)(c_1), \ldots, d_{\underline{z}}(\varphi\mid_M)(c_m)\}$ es ortogonal. Por lo tanto, el jacobiano normal satisface:

$$NJ_{\underline{z}}(\varphi\mid_M) = \prod_{i=1}^{m} \|d_{\underline{z}}(\varphi\mid_M)(c_i)\|^2_{T_{\varphi(\underline{z})}\mathbf{P}_n(\mathbb{C})},$$

y obtenemos el item *iii*). ∎

A continuación presentamos un resultado que establece una cota inferior para el volumen de la intersección de una variedad proyectiva equidimensional con una bola de radio creciente centrada en un punto de la variedad. Un resultado similar para el caso afín existe como consecuencia de teoremas muy fuertes de geometría compleja e integración geométrica. La versión de este resultado para el caso afín, que será usado durante la prueba, se debe probablemente a Wirtinger, aunque las primeras referencias que conocemos son debidas a Lelong y Federer, véanse [86, 44] (recomendamos al lector el libro de Stolzenberg [123], mucho más accesible para los no–expertos en Teoría de la Medida).

Teorema 2.3.2 *Sea $V \subseteq \mathbb{P}_n(\mathbb{C})$ una variedad proyectiva equidimensional de dimensión $m \geq 1$. Sea $x \in V$ un punto y sea $0 < \varepsilon \leq 1$ un número positivo. Entonces, tenemos:*

$$\nu_m[V \cap B_{\mathbf{P}}(x, \varepsilon)] \geq \vartheta_m \varepsilon^{2m}(1-\varepsilon^2).$$

En particular, para todo $\varepsilon > 0$ tal que $\varepsilon \leq \frac{\sqrt{2}}{2}$, se tiene que

$$\nu_m[V \cap B_{\mathbf{P}}(e_0, \varepsilon)] \geq \frac{1}{2}\vartheta_m \varepsilon^{2m}.$$

Demostración.– Sean $\mathbb{A}_0^n$ y φ_0 como en el Lema 2.3.1. Sin pérdida de generalidad podemos suponer que

$$x = e_0 = (1 : 0 : \cdots : 0) \in V \cap \mathbb{A}_0^n \neq \emptyset.$$

Nótese que la siguiente igualdad se satisface para todo valor de ε, $0 < \varepsilon < 1$:

$$\varphi_0^{-1}(B_{\mathbb{P}}(e_0, \varepsilon)) = B_{\mathbb{C}^n}(0, \frac{\varepsilon}{\sqrt{1-\varepsilon^2}}). \tag{2.4}$$

En efecto, se tiene que:

$$d_{\mathbb{P}}(e_0, \varphi_0(\underline{z})) = \sqrt{1 - \frac{|\langle e_0, (1, \underline{z})\rangle^2_{\mathbb{C}^{n+1}}|}{\|(1, \underline{z})\|^2_{\mathbb{C}^{n+1}}}} = \frac{\|\underline{z}\|_{\mathbb{C}^n}}{\sqrt{1 + \|\underline{z}\|^2_{\mathbb{C}^n}}} = \frac{d_{\mathbb{C}^n}(0, \underline{z})}{\sqrt{1 + d_{\mathbb{C}^n}(0, \underline{z})^2}},$$

y por lo tanto,

$$d_{\mathbb{C}^n}(0, \underline{z}) = \frac{d_{\mathbb{P}}(e_0, \varphi_0(\underline{z}))}{\sqrt{1 - d_{\mathbb{P}}(e_0, \varphi_0(\underline{z}))^2}},$$

lo que demuestra la igualdad (2.4).
Sea $W = Reg(V) \subseteq \mathbb{P}_n(\mathbb{C})$ la variedad diferenciable proyectiva de dimensión compleja m definida como el conjunto de los puntos regulares de V. Sea $\overline{W} = \varphi_0^{-1}(W)$ la imagen inversa de W mediante φ_0. Entonces, $\overline{W}$ es la variedad diferenciable afín de dimensión compleja m que consiste en el conjunto de los puntos regulares de $\overline{V} = \varphi_0^{-1}(V)$. Por el Lema 2.3.1 se satisface la siguiente igualdad:

$$NJ_{\underline{z}}(\varphi_0 \mid_{\overline{W}}) \geq \frac{1}{(1 + \|\underline{z}\|^2_{\mathbb{C}^n})^{m+1}}.$$

Ahora, por el Teorema 1.1.13 sabemos que:

$$\nu_m[W \cap B_{\mathbb{P}}(e_0, \varepsilon)] = \int_{\underline{z} \in \overline{W} \cap B_{\mathbb{C}^n}(0, \frac{\varepsilon}{\sqrt{1-\varepsilon^2}})} NJ_{\underline{z}}(\varphi_0 \mid_{\overline{W}}) \, d\overline{W} \geq$$

$$\int_{\underline{z} \in \overline{W} \cap B_{\mathbb{C}^n}(0, \frac{\varepsilon}{\sqrt{1-\varepsilon^2}})} \frac{1}{\left(1 + \|\underline{z}\|^2_{\mathbb{C}^n}\right)^{m+1}} \, d\overline{W} \geq$$

$$\frac{1}{\left(1 + \frac{\varepsilon^2}{1-\varepsilon^2}\right)^{m+1}} \mathscr{H}^{2m}\left[\overline{W} \cap B_{\mathbb{C}^n}(0, \frac{\varepsilon}{\sqrt{1-\varepsilon^2}})\right] =$$

$$= \mathscr{H}^{2m}\left[\overline{W} \cap B_{\mathbb{C}^n}(0, \frac{\varepsilon}{\sqrt{1-\varepsilon^2}})\right](1-\varepsilon^2)^{m+1},$$

donde $\mathscr{H}^{2m}$ denota la medida de Hausdorff $2m-$dimensional. Como $\overline{V} \backslash \overline{W} = \varphi_0^{-1}(V \setminus W)$ está contenida es una variedad algebraica afín de dimensión compleja a lo más $m-1$, tenemos:

$$\mathscr{H}^{2m}\left[\overline{W} \cap B_{\mathbb{C}^n}(0, \frac{\varepsilon}{\sqrt{1-\varepsilon^2}})\right] = \mathscr{H}^{2m}\left[\overline{V} \cap B_{\mathbb{C}^n}(0, \frac{\varepsilon}{\sqrt{1-\varepsilon^2}})\right].$$

Por otro lado, el Teorema (B) de [123] implica que:

$$\mathscr{H}^{2m}\left[\overline{V}\cap B_{\mathbb{C}^n}(0,\frac{\varepsilon}{\sqrt{1-\varepsilon^2}})\right]\geq \mathscr{H}^{2m}[B_{\mathbb{C}^m}(0,1)]\left(\frac{\varepsilon}{\sqrt{1-\varepsilon^2}}\right)^{2m}.$$

Finalmente, observamos que

$$\mathscr{H}^{2m}[B_{\mathbb{C}^m}(0,1)]=\frac{\pi^m}{m!}=\vartheta_m.$$

Por tanto, concluimos la desigualdad

$$\nu_m[V\cap B_{\mathbb{P}}(e_0,\varepsilon)]\geq \vartheta_m\varepsilon^{2m}(1-\varepsilon^2).$$

■

El siguiente resultado se sigue de forma inmediata a partir del Teorema 2.3.2:

Corolario 2.3.3 *Sea $V\subseteq \mathbb{P}_n(\mathbb{C})$ una variedad proyectiva (posiblemente no equi–dimensional), y sea m el máximo de las dimensiones de sus componentes irreducibles. Sea $x\in V$ un punto y sea $0<\varepsilon\leq 1$ un número real. Entonces, tenemos:*

$$\nu_m[V\cap B_{\mathbb{P}}(x,\varepsilon)]\geq \mathcal{C}(V,x)\ \vartheta_m\varepsilon^{2m}(1-\varepsilon^2),$$

donde $\mathcal{C}(V,x)$ es el número de componentes irreducibles de V de dimensión exactamente igual a m que contienen x.

Corolario 2.3.4 *Sea $V\subseteq \mathbb{P}_n(\mathbb{C})$ una variedad proyectiva equidimensional de dimensión m. Sean $0<\varepsilon<\varepsilon_1$ dos números reales tales que $\varepsilon<1$. Supongamos además que $\varepsilon_1-\varepsilon\leq\frac{\sqrt{2}}{2}$. Entonces, para todo punto $z\in V_\varepsilon$, se satisface la siguiente desigualdad:*

$$\frac{\nu_m[B_{\mathbb{P}}(z,\varepsilon_1)\cap V]}{\vartheta_m}\geq\frac{1}{2}(\varepsilon_1-\varepsilon)^{2m}.$$

Demostración.– Como $z\in V_\varepsilon$, existe un punto $y\in V$ tal que $d_{\mathbb{P}}(z,y)<\varepsilon$. Por la desigualdad triangular,

$$B_{\mathbb{P}}(z,\varepsilon_1)\supseteq B_{\mathbb{P}}(y,\varepsilon_1-\varepsilon).$$

Por el Teorema 2.3.2 tenemos:

$$\frac{\nu_m[B_{\mathbb{P}}(z,\varepsilon_1)\cap V]}{\vartheta_m}\geq\frac{\nu_m[B_{\mathbb{P}}(y,\varepsilon_1-\varepsilon)\cap V]}{\vartheta_m}\geq\frac{1}{2}(\varepsilon_1-\varepsilon)^{2m}.$$

■

El siguiente resultado nos proporciona una cota más ajustada para el caso de que V sea una variedad lineal.

Lema 2.3.5 *Sea $V \subseteq \mathbb{P}_n(\mathbb{C})$ un subespacio proyectivo de dimensión m. Sean $0 < \varepsilon < \varepsilon_1 \leq 1$ dos números positivos. Entonces, para todo punto $z \in V_\varepsilon$, tenemos:*

$$\frac{\nu_m[B_{\mathbf{P}}(z, \varepsilon_1) \cap V]}{\vartheta_m} \geq (\varepsilon_1^2 - \varepsilon^2)^m \frac{1 - \varepsilon_1^2}{(1 - \varepsilon^2)^m}.$$

Demostración.– Sean $\mathbb{A}_0^n = \mathbb{P}_n(\mathbb{C}) \setminus \{x_0 = 0\}$ y $\varphi = \varphi_0$ como en el Lema 2.3.1. Sin pérdida de generalidad podemos suponer que $z = e_0 := (1 : 0 : \cdots : 0) \in V_\varepsilon$. Sea $z' \in V$ un punto tal que

$$d_{\mathbf{P}}(e_0, V) = d_{\mathbf{P}}(e_0, z') = d < \varepsilon.$$

También podemos suponer que $\varepsilon_1 < 1$, esto es que $z' \in \mathbb{A}_0^n \cap V$. Como en la demostración del Teorema 2.3.2 tenemos

$$d_{\mathbb{C}^n}(0, \varphi^{-1}(z')) = \frac{d}{\sqrt{1 - d^2}} \leq \frac{\varepsilon}{\sqrt{1 - \varepsilon^2}}.$$

Además,

$$\nu_m[V \cap B_{\mathbf{P}}(e_0, \varepsilon_1)] \geq \mathscr{H}^{2m}\left[\varphi^{-1}(V) \cap B_{\mathbb{C}^n}\left(0, \frac{\varepsilon_1}{\sqrt{1 - \varepsilon_1^2}}\right)\right](1 - \varepsilon_1^2)^{m+1},$$

donde $\mathscr{H}^{2m}$ denota de nuevo la medida de Hausdorff $2m-$dimensional. Ahora, observamos que $\varphi^{-1}(V) \subseteq \mathbb{C}^n$ es un subespacio lineal afín . Por lo tanto,

$$\|\varphi^{-1}(z')\|_2 = \frac{d}{\sqrt{1 - d^2}} = d_{\mathbb{C}^n}(0, \varphi^{-1}(V)).$$

Además, $\varphi^{-1}(z')$ es ortogonal al espacio vectorial de direcciones de $\varphi^{-1}(V)$. Esto es, para todo $x \in \varphi^{-1}(V)$, $x - \varphi^{-1}(z')$ y $\varphi^{-1}(z')$ son ortogonales. Por tanto, para todo $x \in \varphi^{-1}(V)$,

$$\|x\|_2^2 = \|x - \varphi^{-1}(z')\|_2^2 + \|\varphi^{-1}(z')\|_2^2.$$

Esto implica de modo inmediato que

$$\varphi^{-1}(V) \cap B_{\mathbb{C}^n}\left(\varphi^{-1}(z'), \left(\frac{\varepsilon_1^2}{1 - \varepsilon_1^2} - \frac{d^2}{1 - d^2}\right)^{1/2}\right) \subseteq$$

$$\varphi^{-1}(V) \cap B_{\mathbb{C}^n}\left(0, \frac{\varepsilon_1}{\sqrt{1 - \varepsilon_1^2}}\right).$$

Ahora, dado que $\varphi^{-1}(V)$ es un subespacio afín de dimensión compleja m, podemos calcular

$$\mathscr{H}^{2m}\left[\varphi^{-1}(V) \cap B_{\mathbb{C}^n}\left(\varphi^{-1}(z'), \left(\frac{{\varepsilon_1}^2}{1 - {\varepsilon_1}^2} - \frac{d^2}{1 - d^2}\right)^{1/2}\right)\right] =$$

$$\mathscr{H}^{2m}[B_{\mathbb{C}^m}(0,1)]\left(\frac{{\varepsilon_1}^2}{1-{\varepsilon_1}^2}-\frac{d^2}{1-d^2}\right)^m \geq$$
$$\mathscr{H}^{2m}[B_{\mathbb{C}^m}(0,1)]\left(\frac{{\varepsilon_1}^2}{1-{\varepsilon_1}^2}-\frac{\varepsilon^2}{1-\varepsilon^2}\right)^m =$$
$$\mathscr{H}^{2m}[B_{\mathbb{C}^m}(0,1)]\left(\frac{{\varepsilon_1}^2-\varepsilon^2}{(1-{\varepsilon_1}^2)(1-\varepsilon^2)}\right)^m.$$

Concluimos que

$$\nu_m[V\cap B_{\mathbb{P}}(e_0,\varepsilon_1)] \geq (1-{\varepsilon_1}^2)^{m+1}\mathscr{H}^{2m}[B_{\mathbb{C}^m}(0,1)]\left(\frac{{\varepsilon_1}^2-\varepsilon^2}{(1-{\varepsilon_1}^2)(1-\varepsilon^2)}\right)^m.$$

Ahora, $\mathscr{H}^{2m}[B_{\mathbb{C}^m}(0,1)] = \vartheta_m$. Esto termina la demostración del lema. ■

2.3.2. La cota para el volumen del tubo

A continuación exponemos el resultado más importante de esta sección, que es la versión técnica del Teorema 2.1.1. Para cada par de números naturales $1 \leq m < n$, sea $C(n,m)$ la constante definida como sigue:

$$C(n,m) := 2\frac{n^{2n}}{m^{2m}(n-m)^{2(n-m)}} \leq 2\left(\frac{e\ n}{n-m}\right)^{2(n-m)}.$$

Teorema 2.3.6 *Sea $V \subset \mathbb{P}_n(\mathbb{C})$ una variedad proyectiva (posiblemente singular) equi–dimensional de dimensión $m < n$. Sea $0 < \varepsilon \leq 1$ un número real. Entonces, las siguientes desigualdades se satisfacen:*

$$\varepsilon^{2(n-m)} \leq \frac{\nu_n[V_\varepsilon]}{\vartheta_n} \leq C(n,m)\deg(V)\varepsilon^{2(n-m)}.$$

En particular,

$$\frac{\nu_n[V_\varepsilon]}{\vartheta_n} \leq 2\deg(V)\left(\frac{e\ n\ \varepsilon}{n-m}\right)^{2(n-m)}.$$

Demostración.– Fijemos $L \subseteq \mathbb{P}_n(\mathbb{C})$ un subespacio lineal proyectivo cualquiera de dimensión $n-m$. Por el Corolario 2.2.6 tenemos

$$\nu_n[V_\varepsilon] = \frac{\vartheta_n}{\vartheta_{n-m}}\int_{U\in\mathcal{U}_{n+1}} \nu_{n-m}[V_\varepsilon \cap UL]\ d\mathcal{U}_{n+1}.$$

Como V y UL son variedades proyectivas de dimensiones respectivas m y $n-m$, un resultado básico de Teoría de la Intersección (véase por ejemplo [109, 65]) garantiza que

$$V \cap UL \neq \emptyset \quad \forall\ U \in \mathcal{U}_{n+1}.$$

Además, para todo punto $z \in V \cap UL$ se tiene que:

$$\nu_{n-m}[V_\varepsilon \cap UL] \geq \nu_{n-m}[B_{\mathbf{P}}(z,\varepsilon) \cap UL] = \vartheta_{n-m}\varepsilon^{2(n-m)},$$

de donde se sigue la cota inferior del teorema.
Para la cota superior, observamos que si $\varepsilon > 0$ satisface

$$\frac{\sqrt{2}}{2}\frac{n-m}{m} \leq \varepsilon \leq 1,$$

entonces tenemos que

$$C(n,m)\deg(V)\varepsilon^{2(n-m)} \geq 1.$$

En efecto, basta comprobar que la siguiente función es siempre mayor que 1 en el intervalo $[1, n-1]$:

$$f(x) := 2\frac{n^{2n}}{x^{2x}(n-x)^{2n-2x}}\left(\frac{\sqrt{2}}{2}\frac{n-x}{x}\right)^{2n-2x} = 2\left(\frac{n}{x}\right)^{2n}\frac{1}{2^{n-x}}.$$

Ahora, $f'(x) \leq 0$ es siempre negativa, y como consecuencia $f(x) \geq f(n-1) > 1$. La cota superior del Teorema se sigue inmediatamente para ese caso. Supongamos por tanto que $0 < \varepsilon < \text{mín}\{1, \frac{\sqrt{2}}{2}\frac{n-m}{m}\}$. Sea $\varepsilon_1 > 0$ otro número positivo, $0 < \varepsilon < \varepsilon_1$. Consideramos la cantidad

$$\varphi_V(\varepsilon_1,\varepsilon) = \inf_{z\in V_\varepsilon}(\nu_m[B_{\mathbf{P}}(z,\varepsilon_1) \cap V]).$$

Demostraremos más adelante que $\varphi_V(\varepsilon_1,\varepsilon) > 0$. Por lo tanto, tenemos que:

$$\nu_n[V_\varepsilon] = \int_{V_\varepsilon} 1 \; d\mathbb{P}_n(\mathbb{C}) \leq \int_{z\in V_\varepsilon} \frac{\nu_m[B_{\mathbf{P}}(z,\varepsilon_1)\cap V]}{\varphi_V(\varepsilon_1,\varepsilon)} \; d\mathbb{P}_n(\mathbb{C}) \leq$$

$$\frac{1}{\varphi_V(\varepsilon_1,\varepsilon)}\int_{z\in\mathbb{P}_n(\mathbb{C})} \nu_m[B_{\mathbf{P}}(z,\varepsilon_1)\cap V] \; d\mathbb{P}_n(\mathbb{C}).$$

Por el Corolario 2.2.8, deducimos:

$$\nu_n[V_\varepsilon] \leq \frac{\vartheta_n}{\varphi_V(\varepsilon_1,\varepsilon)}\nu_m[V]\varepsilon_1^{2n}.$$

Ahora, por el Corolario 2.2.7, $\nu_m[V] = \vartheta_m \deg(V)$. Por tanto:

$$\nu_n[V_\varepsilon] \leq \frac{\vartheta_n}{\varphi_V(\varepsilon_1,\varepsilon)}\vartheta_m\deg(V)\varepsilon_1^{2n}, \tag{2.5}$$

Por el Corolario 2.3.4, siempre que tengamos $\varepsilon_1 - \varepsilon \leq \frac{\sqrt{2}}{2}$ y $z \in V_\varepsilon$ se tendrá la siguiente desigualdad:

$$\frac{\nu_m[B_{\mathbf{P}}(z,\varepsilon_1)\cap V]}{\vartheta_m} \geq \frac{1}{2}(\varepsilon_1 - \varepsilon)^{2m}.$$

Por lo tanto, siempre que $\varepsilon_1 - \varepsilon \leq \frac{\sqrt{2}}{2}$, tendremos que

$$\varphi_V(\varepsilon_1, \varepsilon) \geq \vartheta_m \frac{1}{2}(\varepsilon_1 - \varepsilon)^{2m}.$$

Finalmente, elegimos $\varepsilon_1 = \frac{n}{n-m}\varepsilon$. Observamos que

$$\varepsilon_1 - \varepsilon = \frac{m}{n-m}\varepsilon < \frac{m}{n-m}\frac{\sqrt{2}}{2}\frac{n-m}{m} = \frac{\sqrt{2}}{2}.$$

Por la desigualdad (2.5), concluimos que

$$\nu_n[V_\varepsilon] \leq \frac{\vartheta_n \vartheta_m \deg(V) \left(\frac{n}{n-m}\varepsilon\right)^{2n}}{\vartheta_m \frac{1}{2} \left(\frac{m}{n-m}\right)^{2m} \varepsilon^{2m}} = \vartheta_n C(n,m) \deg(V) \varepsilon^{2(n-m)},$$

como queríamos. La última afirmación del teorema se sigue de la siguiente desigualdad.

$$C(n,m) = 2\left(1 + \frac{n-m}{m}\right)^{2m} \left(\frac{n}{n-m}\right)^{2(n-m)} \leq 2\left(\frac{en}{n-m}\right)^{2(n-m)}.$$

■

La constante $C(n,m)$ que aparece en el Teorema 2.3.6 también satisface la siguiente desigualdad:

$$C(n,m) = C(n,n-m) \leq 2\left(\frac{e\ n}{m}\right)^{2m}.$$

Podemos dar también otra acotación, como consecuencia de versiones finas de las desigualdades de Stirling-Gautschi (véase por ejemplo [120]):

$$2\sqrt{\pi}\frac{\sqrt{m}\sqrt{n-m}}{\sqrt{n}}\binom{n}{m} < C(n,m)^{1/2} < 2e^{1/6}\sqrt{\pi}\frac{\sqrt{m}\sqrt{n-m}}{\sqrt{n}}\binom{n}{m}.$$

La cota superior obtenida en el teorema 2.3.6 es esencialmente óptima, al menos en el caso de que $V \subseteq \mathbb{P}_n(\mathbb{C})$ es un subespacio lineal de dimensión m. Este hecho viene expresado por el siguiente resultado.

Teorema 2.3.7 *Sea $V \subset \mathbb{P}_n(\mathbb{C})$ un subespacio lineal de dimensión $1 \leq m < n$. Sea $0 < \varepsilon$ un número real positivo tal que*

$$\varepsilon \leq \left(\frac{n-m}{2n}\right)^{1/2}.$$

Entonces, se tiene:

$$\binom{n}{m}\varepsilon^{2(n-m)}(1-\varepsilon^2)^m \leq \frac{\nu_n[V_\varepsilon]}{\vartheta_n} \leq 6\sqrt{m}\binom{n}{m}\varepsilon^{2(n-m)}(1-\varepsilon^2)^m.$$

Demostración.– La cota inferior se obtiene del artículo de Alfred Gray [60]. En efecto, [60, Cor. 1.3] implica

$$\frac{\nu_n[V_\varepsilon]}{\vartheta_n} \geq \binom{n}{m}\varepsilon^{2(n-m)}(1-\varepsilon^2)^m.$$

Para la cota superior, seguimos esencialmente los mismos pasos que en la prueba del Teorema 2.3.7, sustituyendo el Corolario 2.3.4 por el Lema 2.3.5 a la hora de estimar $\varphi_V(\varepsilon_1, \varepsilon)$. Esto es, dados dos números positivos $0 < \varepsilon < \varepsilon_1 < 1$ definimos la función

$$\varphi_V(\varepsilon_1, \varepsilon) = \inf_{z \in V_\varepsilon} (\nu_m[B_{\mathbf{P}}(z, \varepsilon_1) \cap V]).$$

Como en la demostración del Teorema 2.3.6, concluimos:

$$\nu_n[V_\varepsilon] \leq \frac{\vartheta_n}{\varphi_V(\varepsilon_1, \varepsilon)}\vartheta_m \varepsilon_1^{2n},$$

puesto que en este caso $\deg(V) = 1$. Además, por el Lema 2.3.5 tenemos que

$$\varphi_V(\varepsilon_1, \varepsilon) \geq \vartheta_m(\varepsilon_1^2 - \varepsilon^2)^m \frac{1-\varepsilon_1^2}{(1-\varepsilon^2)^m}.$$

Por lo tanto,

$$\nu_n[V_\varepsilon] \leq \vartheta_n \frac{\varepsilon_1^{2n}}{(\varepsilon_1^2 - \varepsilon^2)^m} \frac{(1-\varepsilon^2)^m}{1-\varepsilon_1^2}.$$

Supongamos que $\varepsilon \leq \left(\frac{n-m}{2n}\right)^{1/2}$. Entonces, elegimos

$$\varepsilon_1 := \left(\frac{n}{n-m}\right)^{1/2} \varepsilon \leq \frac{\sqrt{2}}{2} < 1$$

y concluimos que:

$$\nu_n[V_\varepsilon] \leq \vartheta_n \frac{n^n}{m^m(n-m)^{n-m}} \varepsilon^{2(n-m)}(1-\varepsilon^2)^m \frac{n-m}{n-m-n\varepsilon^2} \leq$$

$$\leq 2\vartheta_n \frac{n^n}{m^m(n-m)^{n-m}} \varepsilon^{2(n-m)}(1-\varepsilon^2)^m.$$

La siguiente estimación de [120] termina la prueba:

$$\frac{n^n}{m^m(n-m)^{n-m}} < e^{1/6}\sqrt{2\pi}\frac{\sqrt{m}\sqrt{n-m}}{\sqrt{n}}\binom{n}{m} < 3\sqrt{m}\binom{n}{m}.$$

■

El siguiente resultado se sigue del Teorema 2.3.7.

Corolario 2.3.8 *Sea $V \subseteq \mathbb{P}_n(\mathbb{C})$ una variedad proyectiva equi–dimensional, con $n > 1$, $\dim(V) = m$, y sea $z \in \mathbb{P}_n(\mathbb{C})$ un punto cualquiera. Sea $0 < \varepsilon \leq 1$ un número real positivo, tal que*

$$\varepsilon < \left(\frac{m}{2n}\right)^{1/2}.$$

Entonces, se tiene la siguiente desigualdad para todo $1 \leq m \leq n-1$:

$$\frac{\nu_m[V \cap B_{\mathbf{P}}(z,\varepsilon)]}{\nu_m[V]} \leq 6\sqrt{n-m}\binom{n}{m}\varepsilon^{2m}(1-\varepsilon^2)^{n-m}.$$

Demostración.– Sea L un subespacio lineal cualquiera de $\mathbb{P}_n(\mathbb{C})$ de dimensión $n-m$. Por el Corolario 2.2.6

$$\nu_m[V \cap B_{\mathbf{P}}(z,\varepsilon)] = \vartheta_m \int_{U \in \mathcal{U}_{n+1}} \sharp(UL \cap V \cap B_{\mathbf{P}}(z,\varepsilon))\ d\mathcal{U}_{n+1}.$$

Por lo tanto, se tiene:

$$\begin{gathered}\nu_m[V \cap B_{\mathbf{P}}(z,\varepsilon)] \leq \\ \deg(V)\vartheta_m\nu_{\mathcal{U}_{n+1}}[U \in \mathcal{U}_{n+1} : UL \cap V \cap B_{\mathbf{P}}(z,\varepsilon) \neq \emptyset] \leq \\ \leq \nu_m[V]\nu_{\mathcal{U}_{n+1}}[U \in \mathcal{U}_{n+1} : UL \cap B_{\mathbf{P}}(z,\varepsilon) \neq \emptyset] = \\ = \nu_m[V]\nu_{\mathcal{U}_{n+1}}[U \in \mathcal{U}_{n+1} : L \cap U^* B_{\mathbf{P}}(z,\varepsilon) \neq \emptyset],\end{gathered}$$

donde U^* es la matriz conjugada traspuesta de U. La aplicación $A \longmapsto A^*$ define una isometría en $\mathcal{U}_{n+1}$. Por tanto, tenemos:

$$\begin{gathered}\nu_{\mathcal{U}_{n+1}}[U \in \mathcal{U}_{n+1} : L \cap U^* B_{\mathbf{P}}(z,\varepsilon) \neq \emptyset] = \\ \nu_{\mathcal{U}_{n+1}}[U \in \mathcal{U}_{n+1} : L \cap B_{\mathbf{P}}(Uz,\varepsilon) \neq \emptyset],\end{gathered}$$

dado que $UB_{\mathbf{P}}(z,\varepsilon) = B_{\mathbf{P}}(Uz,\varepsilon)$. Sea $L_\varepsilon \subseteq \mathbb{P}_n(\mathbb{C})$ el tubo de radio ε alrededor del subespacio L y sea $U(z,L,\varepsilon) \subseteq \mathcal{U}_{n+1}$ el conjunto dado por la siguiente identidad:

$$U(z,L,\varepsilon) := \{U \in \mathcal{U}_{n+1} : L \cap B_{\mathbf{P}}(Uz,\varepsilon) \neq \emptyset\} = \{U \in \mathcal{U}_{n+1} : Uz \in L_\varepsilon\}.$$

Hemos demostrado que:

$$\nu_m[V \cap B_{\mathbf{P}}(z,\varepsilon)] \leq \nu_m[V] \int_{\mathcal{U}_{n+1}} \chi_{U(z,L,\varepsilon)}\ d\mathcal{U}_{n+1}.$$

Ahora, el Corolario 2.2.5 implica:

$$\nu_m[V \cap B_{\mathbf{P}}(z,\varepsilon)] \leq \frac{\nu_m[V]}{\vartheta_n} \int_{\mathbf{P}_n(\mathbb{C})} \chi_{L_\varepsilon}\ d\mathbb{P}_n(\mathbb{C}) = \frac{\nu_n[L_\varepsilon]}{\vartheta_n}\nu_m[V].$$

El Teorema 2.3.7 nos permite concluir que:

$$\frac{\nu_m[V \cap B_{\mathbf{P}}(z,\varepsilon)]}{\nu_m[V]} \leq 6\sqrt{n-m}\binom{n}{m}\varepsilon^{2m}(1-\varepsilon^2)^{n-m}.$$

■

2.4. Tubos extrínsecos

En esta sección demostraremos el Teorema 2.1.2. Para cada terna de naturales positivos $1 \leq m' < m < n$, sea $C(n, m, m') \in \mathbb{Q}$ la constante definida como

$$C(n, m, m') := \frac{1}{2} C(n, m') C(n - m', n - m),$$

donde $C(a, b)$ es la constante del Teorema 2.3.6. El siguiente resultado es la versión técnica del Teorema 2.1.2.

Teorema 2.4.1 *Sean $V, V' \subseteq \mathbb{P}_n(\mathbb{C})$ dos variedades proyectivas equi–dimensionales, de dimensiones respectivas $m > m' \geq 1$, con $n > m$. Sea $0 < \varepsilon \leq 1$ un número real positivo. Entonces, tenemos la siguiente desigualdad:*

$$\frac{\nu_m[V'_\varepsilon \cap V]}{\nu_m[V]} \leq C(n, m, m') \deg(V') \varepsilon^{2(m-m')}. \tag{2.6}$$

Además, si $V' \subseteq V$, también se tiene que:

$$\frac{\nu_m[V'_\varepsilon \cap V]}{\vartheta_m} \geq \varepsilon^{2(m-m')}(1 - \varepsilon^2). \tag{2.7}$$

Demostración.– Para demostrar la desigualdad (2.6), consideramos dos casos. En primer lugar, si $\frac{m-m'}{n-m'} \leq \varepsilon \leq 1$, entonces la cantidad en el lado derecho de la desigualdad (2.6) es mayor que 1 y hemos terminado la demostración.
Buscamos por lo tanto una cota superior para el caso de que $\varepsilon < \frac{m-m'}{n-m'} < 1$. Sea $\varepsilon_1 > 0$ un número real positivo tal que $\varepsilon_1 + \varepsilon < 1$. Entonces, para cada punto $z \in \mathbb{P}_n(\mathbb{C})$ tenemos que:

$$d_{\mathbf{P}}(z, V') \geq \varepsilon_1 + \varepsilon \Longrightarrow B_{\mathbf{P}}(z, \varepsilon_1) \cap {V'}_\varepsilon = \emptyset. \tag{2.8}$$

Sea $L_0 \subseteq \mathbb{P}_n(\mathbb{C})$ un subespacio lineal fijo de dimensión $n-m$ tal que $e_0 \in L_0$. Por el Corolario 2.2.6, se tiene que

$$\nu_m[{V'}_\varepsilon \cap V] \nu_{n-m}[B_{\mathbf{P}}(e_0, \varepsilon_1) \cap L_0] = \vartheta_m \vartheta_{n-m}$$

$$\int_{U \in \mathcal{U}_{n+1}} \sharp[B_{\mathbf{P}}(Ue_0, \varepsilon_1) \cap {V'}_\varepsilon \cap V \cap UL_0] \, d\mathcal{U}_{n+1}.$$

Ahora, observamos que

$$\sharp[B_{\mathbf{P}}(Ue_0, \varepsilon_1) \cap {V'}_\varepsilon \cap V \cap UL_0] \leq \sharp[V \cap UL_0] \leq \deg(V).$$

Por otro lado, la expresión (2.8) implica que si $Ue_0 \notin V'_{\varepsilon_1 + \varepsilon}$, entonces

$$\sharp[B_{\mathbf{P}}(Ue_0, \varepsilon_1) \cap {V'}_\varepsilon \cap V \cap UL_0] = 0.$$

Sea $A_1 \subseteq \mathcal{U}_{n+1}$ el conjunto dado por

$$A_1 := \{U \in \mathcal{U}_{n+1} : Ue_0 \in V'_{\varepsilon_1+\varepsilon}\}.$$

Concluimos que

$$\nu_m[{V'}_\varepsilon \cap V]\nu_{n-m}[B_{\mathbb{P}}(e_0, \varepsilon_1) \cap L_0] \leq \vartheta_m \vartheta_{n-m} \int_{A_1} \deg(V)\ d\mathcal{U}_{n+1}.$$

Por el Corolario 2.2.5 sabemos que

$$\int_{A_1} \deg(V)\ d\mathcal{U}_{n+1} = \frac{1}{\vartheta_n} \int_{z \in V'_{\varepsilon_1+\varepsilon}} \deg(V)\ d\mathbb{P}_n(\mathbb{C}) = \frac{\deg(V)\nu_n[V'_{\varepsilon_1+\varepsilon}]}{\vartheta_n}.$$

Por tanto, tenemos

$$\nu_m[{V'}_\varepsilon \cap V]\nu_{n-m}[B_{\mathbb{P}}(e_0, \varepsilon_1) \cap L_0] \leq \frac{\vartheta_m \vartheta_{n-m}}{\vartheta_n} \deg(V)\nu_n[V'_{\varepsilon_1+\varepsilon}].$$

Además, tenemos la siguiente igualdad:

$$\nu_{n-m}[B_{\mathbb{P}}(e_0, \varepsilon_1) \cap L_0] = \vartheta_{n-m} {\varepsilon_1}^{2(n-m)}.$$

Por el Corolario 2.2.7, sabemos que $\vartheta_m \deg(V) = \nu_m[V]$. Por tanto, hemos demostrado que

$$\nu_m[{V'}_\varepsilon \cap V]\varepsilon_1^{2(n-m)} \leq \frac{\nu_m[V]}{\vartheta_n}\nu_n[V'_{\varepsilon_1+\varepsilon}].$$

Por el Teorema 2.3.6, deducimos que

$$\frac{\nu_m[{V'}_\varepsilon \cap V]}{\nu_m[V]} \leq C(n, m')\deg(V')\frac{(\varepsilon_1 + \varepsilon)^{2(n-m')}}{{\varepsilon_1}^{2(n-m)}}.$$

Eligiendo $\varepsilon_1 = \frac{n-m}{m-m'}\varepsilon$ (que satisface $\varepsilon_1 + \varepsilon < 1$), tenemos:

$$\frac{\nu_m[{V'}_\varepsilon \cap V]}{\nu_m[V]} \leq C(n, m, m')\deg(V')\varepsilon^{2(m-m')},$$

y la desigualdad (2.6) queda probada.
Demostremos ahora la desigualdad (2.7). Sea $L \subseteq \mathbb{P}_n(\mathbb{C})$ un subespacio lineal de dimensión $n - m'$. Por el corolario 2.2.6 tenemos que

$$\nu_m[V'_\varepsilon \cap V] = \frac{\vartheta_m}{\vartheta_{m-m'}} \int_{U \in \mathcal{U}_{n+1}} \nu_{m-m'}[V'_\varepsilon \cap V \cap UL]\ d\mathcal{U}_{n+1}.$$

Ahora, observamos que $V' \cap UL \neq \emptyset$ para toda matriz $U \in \mathcal{U}_{n+1}$, y si $z \in V' \cap UL$ (lo que implica que también $z \in V$) se tiene que:

$$\nu_{m-m'}[V'_\varepsilon \cap V \cap UL] \geq \nu_{m-m'}[B_{\mathbb{P}}(z, \varepsilon) \cap (V \cap UL)].$$

Por el Lema 2.2.3, existe un conjunto cuyo complementario tiene medida nula, $W \subseteq \mathcal{U}_{n+1}$, tal que para toda matriz $U \in W$, $V \cap UL$ es una variedad proyectiva de dimensión compleja $m-m'$. Por tanto, el Teorema 2.3.2 implica que:

$$\nu_{m-m'}[B_{\mathbb{P}}(z,\varepsilon) \cap (V \cap UL)] \geq \vartheta_{m-m'}\varepsilon^{2(m-m')}(1-\varepsilon^2),$$

y la desigualdad queda demostrada. ∎

2.5. Distribución de probabilidad del condicionamiento en el Álgebra Lineal

El objetivo principal de esta sección es aplicar los anteriores resultados de Geometría Integral al estudio de la distribución de probabilidad de las diversas variantes del número de condicionamiento, así como al análisis de la esperanza del mismo cuando se mide en diferentes espacios. Así, obtendremos resultados que controlan con cierta precisión el condicionamiento esperable en el conjunto de matrices singulares (el Teorema 2.1.3 de la Introducción de este capítulo); también podremos analizar el comportamiento del condicionamiento en otras clases de matrices (matrices simétricas, rectangulares, por bloques), y dar resultados muy generales sobre cómo se comportan estas cantidades cuando se hace variar el rango de las matrices. Varios de los resultados aquí obtenidos serán utilizados con posterioridad en esta memoria. Todos los resultados relativos a volúmenes y probabilidades de esta sección se obtienen de manera inmediata a partir de los teoremas generales 2.3.6 y 2.4.1 demostrados en las secciones anteriores.

Introducimos de nuevo brevemente las notaciones básicas para facilitar la lectura. Sean $2 \leq n_1 \leq n_2$ dos números naturales y sea $\mathcal{M}_{n_1\times n_2}(\mathbb{C})$ el espacio vectorial formado por las matrices $n_1 \times n_2$ con coeficientes complejos. Sea $\mathbb{P}(\mathcal{M}_{n_1\times n_2}(\mathbb{C}))$ el espacio proyectivo asociado. Podemos identificar de forma natural $\mathcal{M}_{n_1\times n_2}(\mathbb{C})$ con $\mathbb{C}^{n_1 n_2}$, luego también tenemos la identificación natural

$$\mathbb{P}(\mathcal{M}_{n_1\times n_2}(\mathbb{C})) \equiv \mathbb{P}(\mathbb{C}^{n_1 n_2}) = \mathbb{P}_{n_1 n_2 - 1}(\mathbb{C}).$$

Por tanto, podemos considerar $\mathbb{P}(\mathcal{M}_{n_1\times n_2}(\mathbb{C}))$ como una variedad Riemanniana compleja de dimensión $n_1 n_2 - 1$ con una medida de volumen natural (como se ha indicado en la Sección 1.1) que denotaremos $\nu_{n_1 n_2 - 1}$. Asimismo, como viene siendo habitual, denotaremos por $\vartheta_{n_1 n_2 - 1}$ el volumen del espacio proyectivo complejo $\mathbb{P}(\mathcal{M}_{n_1\times n_2}(\mathbb{C}))$ (véase la identidad (1.2)). Recordemos que para cada número natural $1 \leq r \leq n_1$, hemos denotado por $\Sigma^r_{\mathcal{M}}$ la variedad proyectiva formada por todas las matrices de rango a lo más r. Esto es,

$$\Sigma^r_{\mathcal{M}} := \{A \in \mathbb{P}(\mathcal{M}_{n_1\times n_2}(\mathbb{C})) : rank(A) \leq r\}.$$

Nótese que no incluimos en la notación de $\Sigma^r_{\mathcal{M}}$ referencia alguna a los valores n_1 y n_2 por evitar el uso de expresiones engorrosas, sin embargo hay que tener presente que sólo podremos definir $\Sigma^r_{\mathcal{M}}$ cuando n_1 y n_2 estén fijados. Como hemos visto en el Corolario 1.5.9, se pueden estimar fácilmente la dimensión y el grado de la variedad $\Sigma^r_{\mathcal{M}}$. En efecto, se tiene:

$$\dim(\Sigma^r_{\mathcal{M}}) = r(n_1+n_2) - r^2 - 1, \quad \deg(\Sigma^r_{\mathcal{M}}) \leq (r+1)^{(n_1-r)(n_2-r)}.$$

En la Sección 1.5.1 del capítulo de preliminares hemos definido el condicionamiento generalizado de una matriz $A \in \mathcal{M}_{n_1\times n_2}(\mathbb{C})$ (véase definición 1.5.1). Además por el Teorema 1.5.11 sabemos que para toda matriz proyectiva $A \in \mathbb{P}(\mathcal{M}_{n_1\times n_2}(\mathbb{C}))$ y para todo número natural $2 \leq r \leq n_1$, se tiene:

$$d_{\mathbb{P}}(A, \Sigma^{r-1}_{\mathcal{M}}) = \frac{1}{\kappa^{(r)}_D(A)}.$$

Esto permite transformar estudios de probabilidad de condicionamiento en estudios de volúmenes de tubos.

2.5.1. Condicionamiento de las matrices de corango dado.

Los siguientes dos resultados describen la distribución de probabilidad del número de condicionamiento generalizado en todos los casos posibles.

Corolario 2.5.1 *Con las notaciones anteriores, sea $r \in \mathbb{N}$ un número natural, $2 \leq r \leq n_1$. Para todo número real $\varepsilon > 0$, la probabilidad que una matriz elegida al azar $A \in \mathbb{P}(\mathcal{M}_{n_1\times n_2}(\mathbb{C}))$ verifique $\kappa^{(r)}_D(A) > \frac{1}{\varepsilon}$ es menor o igual que:*

$$2\left[\frac{e\ (n_1n_2-1)\sqrt{r}}{(n_1-r+1)(n_2-r+1)}\ \varepsilon\right]^{2(n_1-r+1)(n_2-r+1)}.$$

Demostración.– Por el Teorema 1.5.11, tenemos que:

$$\frac{\nu_{n_1n_2-1}[A \in \mathbb{P}(\mathcal{M}_{n_1\times n_2}(\mathbb{C})) : \kappa^{(r)}_D(A) > \frac{1}{\varepsilon}]}{\vartheta_{n_1n_2-1}} =$$

$$= \frac{\nu_{n_1n_2-1}[A \in \mathbb{P}(\mathcal{M}_{n_1\times n_2}(\mathbb{C})) : d_{\mathbb{P}}(A, \Sigma^{r-1}_{\mathcal{M}}) < \varepsilon]}{\vartheta_{n_1n_2-1}}.$$

Ahora, el Teorema 2.3.6 proporciona de inmediato una cota para esa cantidad, dado que conocemos la codimensión y el grado de $\Sigma^r_{\mathcal{M}}$. ∎

El Corolario 2.5.1 estudia la distribución de probabilidad del condicionamiento generalizado en el caso de que nuestro espacio de probabilidad (o espacio de inputs) sea el espacio total de matrices proyectivas. Podemos también demostrar un resultado de distribución de probabilidad en el caso de

que el espacio de inputs sean las matrices de rango acotado por una cantidad cualquiera. Éste es el objetivo del siguiente resultado, que, aunque similar al Corolario 2.5.1, es de una naturaleza muy distinta, y las técnicas que requiere son las de volúmenes de tubos intersecados con variedades proyectivas, como han sido estudiados en la Sección 2.4.

Corolario 2.5.2 *Con las notaciones anteriores, sea $r \in \mathbb{N}$ un número natural, $2 \leq r \leq n_2$. Para todo número real $\varepsilon > 0$, tenemos:*

$$\frac{\nu_{r(n_1+n_2)-r^2-1}[A \in \Sigma^r_{\mathcal{M}} : \kappa_D^{(r)}(A) > \frac{1}{\varepsilon}]}{\nu_{r(n_1+n_2)-r^2-1}[\Sigma^r_{\mathcal{M}}]} \leq$$

$$\deg(\Sigma^{r-1}_{\mathcal{M}}) D(n_1, n_2, r) \varepsilon^{2(n_1+n_2-2r+1)},$$

donde

$$D(n_1, n_2, r) := C(n_1 n_2 - 1, r(n_1+n_2) - r^2 - 1, (r-1)(n_1+n_2) - (r-1)^2 - 1)$$

y $C(a, b, c)$ es la constante del Teorema 2.4.1 para cada terna de números naturales $a > b > c \in \mathbb{N}$.

Demostración.– Por el Teorema 1.5.11, tenemos:

$$\frac{\nu_{r(n_1+n_2)-r^2-1}[A \in \Sigma^r_{\mathcal{M}} : \kappa_D^{(r)}(A) > \frac{1}{\varepsilon}]}{\nu_{r(n_1+n_2)-r^2-1}[\Sigma^r_{\mathcal{M}}]} =$$

$$= \frac{\nu_{r(n_1+n_2)-r^2-1}[A \in \Sigma^r_{\mathcal{M}} : d_{\mathbf{P}}(A, \Sigma^{r-1}_{\mathcal{M}}) < \varepsilon]}{\nu_{r(n_1+n_2)-r^2-1}[\Sigma^r_{\mathcal{M}}]}.$$

En esta ocasión, el Teorema 2.4.1 proporciona la cota del enunciado. De nuevo, usamos que conocemos el valor de la dimensión de las variedades $\Sigma^r_{\mathcal{M}}$ y $\Sigma^{r-1}_{\mathcal{M}}$. ∎

Nota 2.5.3 *De hecho, aún queda un caso por analizar, que es la distribución de probabilidad del número de condicionamiento $\kappa_D^{(r')}$ en $\Sigma^r_{\mathcal{M}}$, cuando $r' < r$. Con la misma técnica que hemos utilizado, se obtiene también una acotación para ese caso.*

Para valores concretos de n_1, n_2, r las estimaciones que obtenemos tienen expresiones mucho más sencillas y manejables que en el caso general. Como un ejemplo de aplicación, demostramos una versión ligeramente más general del Teorema 2.1.3, que lo engloba como un caso particular. Nos resultará de utilidad el siguiente lema técnico que es consecuencia de un resultado bien conocido de Teoría de Probabilidad.

Lema 2.5.4 *Sea X una variable aleatoria real positiva tal que*

$$P[X > t] \leq ct^{-\alpha}, \quad \forall t \geq 1,$$

para algunas constantes positivas $c > 1, \alpha > 1$. Entonces, la esperanza de X satisface la siguiente desigualdad:

$$E[X] \leq c^{\frac{1}{\alpha}} \ \frac{\alpha}{\alpha - 1}.$$

El mismo resultado es cierto si exigimos

$$P[X > t] < ct^{-\alpha}, \quad \forall t \geq 0,$$

y permitimos la hipótesis menos restrictiva $c > 0$.

Demostración.– Utilizamos la siguiente igualdad, que es consecuencia del Teorema de Fubini y puede encontrarse en cualquier libro introductorio de Teoría de la Probabilidad.

$$E[X] = \int_0^\infty P[X > t] \ dt.$$

Entonces, observamos que para todo número positivo $s > 1$,

$$E[X] = \int_0^\infty P[X > t] \ dt \leq s + c \int_s^\infty t^{-\alpha} \ dt = s + c\frac{s^{1-\alpha}}{\alpha - 1}.$$

Tomando $s := c^{\frac{1}{\alpha}}$, el lema queda demostrado. La última afirmación se demuestra del mismo modo. ∎

Corolario 2.5.5 *Con las notaciones anteriores, tenemos la siguiente desigualdad:*

$$\frac{\nu_{\dim(\Sigma_{\mathcal{M}}^{n_1-1})}[A \in \Sigma_{\mathcal{M}}^{n_1-1} : \kappa_D^{(n_1-1)}(A) > \frac{1}{\varepsilon}]}{\nu_{\dim(\Sigma_{\mathcal{M}}^{n_1-1})}[\Sigma_{\mathcal{M}}^{n_1-1}]} \leq (e \ n_1^3 n_2^2 \varepsilon)^{2(n_2 - n_1 + 3)}. \tag{2.9}$$

Además, en el caso $n_1 = n_2 = n$, tenemos:

$$\frac{\nu_{\dim(\Sigma_{\mathcal{M}}^{n-1})}[A \in \Sigma_{\mathcal{M}}^{n-1} : \kappa_D^{(n-1)}(A) > \frac{1}{\varepsilon}]}{\nu_{\dim(\Sigma_{\mathcal{M}}^{n-1})}[\Sigma_{\mathcal{M}}^{n-1}]} \leq \frac{7}{10} n^{20} \varepsilon^6, \tag{2.10}$$

y la esperanza de $\kappa_D^{(n-1)}$ en el espacio de matrices singulares satisface la siguiente desigualdad:

$$E_{\Sigma_{\mathcal{M}}^{n-1}}[\kappa_D^{(n-1)}] \leq 2n^{10/3}.$$

Demostración.– Observamos el resultado del Corolario 2.5.2 en el caso de que $r = n_1 - 1$. La constante del Corolario 2.5.2 se traduce en:

$$C(n_1n_2 - 1, n_1n_2 - n_1 + n_2 - 2, n_1n_2 - 2n_2 + 2n_1 - 5) \leq$$

$$\leq 2\left(\frac{e\ n_1n_2}{2n_2 - 2n_1 + 4}\right)^{4n_2-4n_1+8}(2e)^{2n_2-2n_1+6} \leq \left(\frac{2e^3}{16}n_1^2n_2^2\right)^{2n_2-2n_1+6} <$$

$$< \left(e\ n_1^2n_2^2\right)^{2n_2-2n_1+6}.$$

Además, el grado de $\Sigma_{\mathcal{M}}^{n_1-2}$ es conocido por la Proposición 1.5.8:

$$\deg(\Sigma_{\mathcal{M}}^{n_1-2}) = \binom{n_2}{n_1-2}\binom{n_2+1}{n_1-1}\frac{1}{n_2-n_1+3} \leq n_1^{2(n_2-n_1+2)}.$$

Queda demostrada la desigualdad (2.9) del corolario. Para la desigualdad (2.10), observamos que en el caso de que $n_1 = n_2 = n$ sean iguales, podemos acotar de forma más precisa la constante:

$$C(n^2 - 1, n^2 - 2, n^2 - 5)\deg(\Sigma_{\mathcal{M}}^{n-2}) \leq 2\frac{e^8}{3^6}\frac{n^{20}}{12} \leq \frac{7}{10}n^{20}.$$

Finalmente, para obtener el resultado de la esperanza, aplicamos el Lema 2.5.4 a la estimación de la desigualdad (2.10). ∎

También podemos obtener resultados de esperanza para el caso más general tratado en el Corolario 2.5.1. Ése es el objetivo del siguiente enunciado.

Corolario 2.5.6 *Sean $n_2 \geq n_1 \geq 2$ dos naturales positivos. Entonces, el valor esperable del condicionamiento $\kappa_D^{(n_1)}$ satisface la siguiente desigualdad:*

$$\mathrm{E}_{\mathbf{P}(\mathcal{M}_{n_1\times n_2}(\mathbb{C}))}[\kappa_D^{(n_1)}] \leq 2^{\frac{1}{2(n_2-n_1+1)}}e\frac{n_1^{3/2}n_2}{n_2-n_1+1/2}.$$

Además, si $n_1 = 1$ se tiene:

$$\mathrm{E}_{\mathbf{P}(\mathcal{M}_{1\times n_2}(\mathbb{C}))}[\kappa_D^{(1)}] = 1.$$

Demostración.– El caso $n_1 = 1$ es trivial a partir de la definición. Aceptemos pues que $n_1 \geq 2$. Entonces, por el Corolario 2.5.1, para todo número real $t > 0$ se tiene:

$$Prob[A \in \mathbb{P}(\mathcal{M}_{n_1\times n_2}(\mathbb{C})) : \kappa_D^{(n_1)} > t] \leq 2\left[\frac{en_1^{3/2}n_2}{n_2-n_1+1}\frac{1}{t}\right]^{2(n_2-n_1+1)}.$$

Por el Lema 2.5.4, esto implica que:

$$\mathrm{E}_{\mathbf{P}(\mathcal{M}_{n_1\times n_2}(\mathbb{C}))}[\kappa_D^{(n_1)}] \leq 2^{\frac{1}{2(n_2-n_1+1)}}\frac{en_1^{3/2}n_2}{(n_2-n_1+1)}\ \frac{2(n_2-n_1+1)}{2(n_2-n_1+1)-1},$$

de donde se sigue el corolario. ∎

2.5.2. Condicionamiento de matrices simétricas y por bloques.

A continuación demostramos dos resultados más que muestran cómo las cotas generales obtenidas en el Teorema 2.3.6 se aplican a otros subespacios del espacio general de matrices.

Corolario 2.5.7 *Sea $n \geq 2$ un número natural, y sea $\mathcal{SIM}_n(\mathbb{C}) \subseteq \mathcal{M}_n(\mathbb{C})$ el conjunto de matrices simétricas de lado n. Entonces, $\mathbb{P}(\mathcal{SIM}_n(\mathbb{C}))$ es un espacio proyectivo complejo de dimensión $\frac{n(n+1)}{2} - 1$. Además, tenemos:*

$$\frac{\nu_{\frac{n^2+n}{2}-1}[A \in \mathbb{P}(\mathcal{SIM}_n(\mathbb{C})) : \kappa_D^{(n)}(A) > \frac{1}{\varepsilon}]}{\nu_{\frac{n^2+n}{2}-1}[\mathbb{P}(\mathcal{SIM}_n(\mathbb{C}))]} \leq 2\left[e\left(\frac{n^2+n}{2} - 1\right)\sqrt{n}\ \varepsilon\right]^2.$$

Demostración.– Por el Teorema 1.5.11, sabemos que

$$\frac{\nu_{\frac{n^2+n}{2}-1}[A \in \mathbb{P}(\mathcal{SIM}_n(\mathbb{C})) : \kappa_D^{(n)}(A) > \frac{1}{\varepsilon}]}{\nu_{\frac{n^2+n}{2}-1}[\mathbb{P}(\mathcal{SIM}_n(\mathbb{C}))]} =$$

$$= \frac{\nu_{\frac{n^2+n}{2}-1}[A \in \mathbb{P}(\mathcal{SIM}_n(\mathbb{C})) : d_{\mathbb{P}}(A, \Sigma_{\mathcal{M}}^{n-1}) < \varepsilon]}{\nu_{\frac{n^2+n}{2}-1}[\mathbb{P}(\mathcal{SIM}_n(\mathbb{C}))]}.$$

Sin embargo, esta propiedad no basta para demostrar el corolario. Necesitamos la siguiente igualdad que será demostrada a continuación:

$$\frac{\nu_{\frac{n^2+n}{2}-1}[A \in \mathbb{P}(\mathcal{SIM}_n(\mathbb{C})) : d_{\mathbb{P}}(A, \Sigma_{\mathcal{M}}^{n-1}) < \varepsilon]}{\nu_{\frac{n^2+n}{2}-1}[\mathbb{P}(\mathcal{SIM}_n(\mathbb{C}))]} =$$

$$\frac{\nu_{\frac{n^2+n}{2}-1}[A \in \mathbb{P}(\mathcal{SIM}_n(\mathbb{C})) : d_{\mathbb{P}}(A, \Sigma_{\mathcal{M}}^{n-1} \cap \mathbb{P}(\mathcal{SIM}_n(\mathbb{C}))) < \varepsilon]}{\nu_{\frac{n^2+n}{2}-1}[\mathbb{P}(\mathcal{SIM}_n(\mathbb{C}))]}. \quad (2.11)$$

Demostremos pues esta igualdad. En primer lugar, observamos que basta con demostrarla para el conjunto de matrices simétricas que tienen todos sus valores singulares distintos y no nulos. En efecto, el complementario de este conjunto es un conjunto de medida nula en $\mathbb{P}(\mathcal{SIM}_n(\mathbb{C}))$ y no afecta a las estimaciones de volumen.
Sea pues $A \in \mathbb{P}(\mathcal{SIM}_n(\mathbb{C}))$ una matriz simétrica. Supongamos que todos los valores singulares de A, $\sigma_1, \ldots, \sigma_n$, son distintos y no nulos. Sea $A = UDV^*$ una SVD de A. Sea $D' = Diag(\sigma_1, \ldots, \sigma_{n-1}, 0)$ la matriz que se obtiene de reemplazar el último elemento de la diagonal de D por 0. Como hemos visto en la prueba del Teorema 1.5.11, se tiene la siguiente igualdad:

$$d_{\mathbb{P}}(A, \Sigma_{\mathcal{M}}^{n-1}) = d_{\mathbb{P}}(A, UD'V^*).$$

Por lo tanto, para demostrar la igualdad (2.11) debemos comprobar que $UD'V^* \in \mathbb{P}(\mathcal{SIM}_n(\mathbb{C}))$. Como A es simétrica,sabemos que:

$$UDV^* = (V^*)^t DU^t,$$

luego tenemos dos SVDs distintas de A. Dado que todos los valores singulares de A son distintos y no nulos, es bien conocido que esto implica:

$$U = (V^*)^t \lambda_1, \quad V^* = \lambda_2 U^t,$$

donde $\lambda_1, \lambda_2 \in \mathcal{M}_n(\mathbb{C})$ son dos matrices diagonales cuyas entradas son números complejos de módulo 1, y además $\lambda_1\lambda_2 = Id_n$. Concluimos que:

$$UD'V^* = (V^*)^t \lambda_1 D' \lambda_2 U^t = (V^*)^t D' U^t.$$

Por tanto, $UD'V^* \in \mathbb{P}(\mathcal{SIM}_n(\mathbb{C}))$ y la igualdad (2.11) queda demostrada. El Teorema 2.3.6 acota superiormente el término en la derecha de la ecuación (2.11), puesto que $\Sigma_{\mathcal{M}}^{n-1} \cap \mathbb{P}(\mathcal{SIM}_n(\mathbb{C}))$ es una variedad proyectiva de $\mathbb{P}(\mathcal{SIM}_n(\mathbb{C}))$ de codimensión 1 y grado acotado por la desigualdad de Bézout:

$$\deg(\Sigma_{\mathcal{M}}^{n-1} \cap \mathbb{P}(\mathcal{SIM}_n(\mathbb{C})) \leq \deg(\Sigma_{\mathcal{M}}^{n-1}) = n.$$

■

Corolario 2.5.8 *Sea $\mathcal{B}_{ij}(\mathbb{C}) \subseteq \mathcal{M}_{n_1 \times n_2}(\mathbb{C})$ el conjunto de matrices A de la forma:*

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix},$$

donde $A_1 \in \mathcal{M}_{i\times j}(\mathbb{C})$,$A_2 \in \mathcal{M}_{(n_1-i)\times(n_2-j)}(\mathbb{C})$. Esto es,

$$\mathcal{B}_{ij}(\mathbb{C}) \equiv \mathcal{M}_{ij}(\mathbb{C}) \oplus \mathcal{M}_{(n_1-i)\times(n_2-j)}(\mathbb{C})$$

puede identificarse con la suma directa de $\mathcal{M}_{ij}(\mathbb{C})$ y $\mathcal{M}_{(n_1-i)\times(n_2-j)}(\mathbb{C})$. Entonces, $\mathbb{P}(\mathcal{B}_{ij}(\mathbb{C}))$ es un subespacio proyectivo complejo de dimensión $ij + (n_1 - i)(n_2 - j) - 1$, y se tiene la siguiente desigualdad:

$$\frac{\nu_{ij+(n_1-i)(n_2-j)-1}[A \in \mathbb{P}(\mathcal{B}_{ij}(\mathbb{C})) : \kappa_D^{(n_1)}(A) > \frac{1}{\varepsilon}]}{\nu_{ij+(n_1-i)(n_2-j)-1}[\mathbb{P}(\mathcal{B}_{ij}(\mathbb{C}))]} \leq$$

$$\leq 2\left[\frac{e(ij + (n_1 - i)(n_2 - j) - 1)}{n_2 - n_1 + 1}\sqrt{n_1}\ \varepsilon\right]^{2(n_2-n_1+1)}.$$

Demostración.– Por el Teorema 1.5.11, se satisface la siguiente igualdad:

$$\frac{\nu_{ij+(n_1-i)(n_2-j)-1}[A \in \mathbb{P}(\mathcal{B}_{ij}(\mathbb{C})) : \kappa_D^{(n_1)}(A) > \frac{1}{\varepsilon}]}{\nu_{ij+(n_1-i)(n_2-j)-1}[\mathbb{P}(\mathcal{B}_{ij}(\mathbb{C}))]} =$$

$$= \frac{\nu_{ij+(n_1-i)(n_2-j)-1}[A \in \mathbb{P}(\mathcal{B}_{ij}(\mathbb{C})) : d_{\mathbf{P}}(A, \Sigma_{\mathcal{M}}^{n_1-1}) < \varepsilon]}{\nu_{ij+(n_1-i)(n_2-j)-1}[\mathbb{P}(\mathcal{B}_{ij}(\mathbb{C}))]}.$$

De nuevo, esto no basta para demostrar el corolario. Necesitamos la siguiente expresión:

$$\frac{\nu_{ij+(n_1-i)(n_2-j)-1}[A \in \mathbb{P}(\mathcal{B}_{ij}(\mathbb{C})) : d_{\mathbf{P}}(A, \Sigma_{\mathcal{M}}^{n_1-1}) < \varepsilon]}{\nu_{ij+(n_1-i)(n_2-j)-1}[\mathbb{P}(\mathcal{B}_{ij}(\mathbb{C}))]} =$$

$$\frac{\nu_{ij+(n_1-i)(n_2-j)-1}[A \in \mathbb{P}(\mathcal{B}_{ij}(\mathbb{C})) : d_{\mathbf{P}}(A, \Sigma_{\mathcal{M}}^{n_1-1} \cap \mathbb{P}(\mathcal{B}_{ij}(\mathbb{C}))) < \varepsilon]}{\nu_{ij+(n_1-i)(n_2-j)-1}[\mathbb{P}(\mathcal{B}_{ij}(\mathbb{C}))]}. \quad (2.12)$$

En efecto, sea $A \in \mathbb{P}(\mathcal{B}_{ij}(\mathbb{C}))$. Sea $A' \in \Sigma_{\mathcal{M}}^{n_1-1}$ una matriz singular tal que $d_{\mathbf{P}}(A, \Sigma_{\mathcal{M}}^{n_1-1}) = d_{\mathbf{P}}(A, A')$. Por la expresión de A' (véase el Teorema 1.5.11) es obvio que $A' \in \mathbb{P}(\mathcal{B}_{ij}(\mathbb{C}))$ y la igualdad (2.12) queda demostrada. Ahora, el Teorema 2.3.6 proporciona una acotación superior para el término en la derecha de la ecuación (2.12), puesto que $\Sigma_{\mathcal{M}}^{n_1-1} \cap \mathbb{P}(\mathcal{B}_{ij}(\mathbb{C}))$ es una variedad proyectiva de $\mathbb{P}(\mathcal{B}_{ij}(\mathbb{C}))$ de codimensión $n_2 - n_1 + 1$ y grado acotado por la desigualdad de Bézout:

$$\deg(\Sigma_{\mathcal{M}}^{n_1-1} \cap \mathbb{P}(\mathcal{B}_{ij}(\mathbb{C}))) \leq \deg(\Sigma_{\mathcal{M}}^{n_1-1}) \leq n_1^{n_2-n_1+1}.$$

Obtenemos por tanto la desigualdad:

$$\frac{\nu_{ij+(n_1-i)(n_2-j)-1}[A \in \mathbb{P}(\mathcal{B}_{ij}(\mathbb{C})) : d_{\mathbf{P}}(A, Sigma_{\mathcal{M}}^{n_1-1} \cap \mathbb{P}(\mathcal{B}_{ij}(\mathbb{C}))) < \varepsilon]}{\nu_{ij+(n_1-i)(n_2-j)-1}[\mathbb{P}(\mathcal{B}_{ij}(\mathbb{C}))]} \leq$$

$$2\left[\frac{e(ij+(n_1-i)(n_2-j)-1)}{n_2-n_1+1}\sqrt{n_1}\,\varepsilon\right]^{2(n_2-n_1+1)}.$$

■

2.5.3. Un resultado general.

El lector puede observar que los corolarios 2.5.7 y 2.5.8 son casos particulares de un resultado general que podemos escribir como sigue.

Teorema 2.5.9 *Sea r un número natural, tal que $2 \leq r \leq n_1$. Sea $\mathcal{C} \subseteq \mathbb{P}(\mathcal{M}_{n_1\times n_2}(\mathbb{C}))$ una variedad proyectiva equi–dimensional de dimensión m. Supongamos que existe una variedad proyectiva $\mathcal{C}' \subseteq \mathbb{P}(\mathcal{M}_{n_1\times n_2}(\mathbb{C}))$ equi–dimensional de dimensión $m(r) < m$ tal que para toda matriz proyectiva $A \in \mathcal{C}$ se satisface la siguiente igualdad:*

$$d_{\mathbf{P}}(A, \Sigma_{\mathcal{M}}^{r-1}) = d_{\mathbf{P}}(A, \mathcal{C}').$$

Entonces, también se satisface la desigualdad:

$$\frac{\nu_m[\{A \in \mathcal{C} : \kappa_D^{(r)} > \varepsilon^{-1}\}]}{\nu_m[\mathcal{C}]} \leq C(n_1n_2-1, m, m(r)) \deg(\mathcal{C}')\varepsilon^{2(m-m(r))},$$

donde $C(n_1n_2-1, m, m(r))$ es la constante del Teorema 2.4.1.

Demostración.– Por el Teorema 1.5.11, sabemos que:

$$\frac{\nu_m[\{A \in \mathcal{C} : \kappa_D^{(r)} > \varepsilon^{-1}\}]}{\nu_m[\mathcal{C}]} = \frac{\nu_m[\{A \in \mathcal{C} : d_{\mathbf{P}}(A, \Sigma_{\mathcal{M}}^{r-1}) < \varepsilon\}]}{\nu_m[\mathcal{C}]}.$$

Por lo tanto,

$$\frac{\nu_m[\{A \in \mathcal{C} : \kappa_D^{(r)} > \varepsilon^{-1}\}]}{\nu_m[\mathcal{C}]} = \frac{\nu_m[\{A \in \mathcal{C} : d_{\mathbf{P}}(A, \mathcal{C}') < \varepsilon\}]}{\nu_m[\mathcal{C}]},$$

y el resultado se sigue de Teorema 2.4.1. ■

Capítulo 3

El Condicionamiento Generalizado de Sistemas de Ecuaciones Polinomiales: Un Análisis de Probabilidad

En este capítulo discutiremos una generalización $\mu_{\text{norm}}^{(r)}$ del número de condicionamiento no–lineal μ_{norm} de M. Shub y S. Smale. Nuestra generalización contempla los casos singulares y los casos sub–determinados (es decir, cuando $V(f)$ define una intersección completa de dimensión positiva). Las principales propiedades que satisface $\mu_{\text{norm}}^{(r)}$ son discutidas: Tanto la estabilidad –global y promedio– de variedades bajo pequeñas perturbaciones, como la convergencia del operador de Newton en el caso sub–determinado.
Analizaremos la distribución de probabilidad y las esperanzas de estos condicionamientos generalizados. Para ello, introduciremos una generalización de las técnicas introducidas por M. Shub y S. Smale para la integración de funciones unitariamente invariantes sobre variedades de incidencia generalizadas.

3.1. Introducción y conceptos básicos

Durante las páginas que siguen discutiremos ampliamente sobre el valor en promedio de los distintos números de condicionamiento $\mu_{\text{norm}}^{(r)}$ de la Definición 1.6.2. Ya hemos indicado la capacidad de estos números para controlar la convergencia del operador de Newton proyectivo cero–dimensional (véase la Nota 1.7.3). En los Capítulos 4 y 5 utilizaremos en toda su potencia un resultado debido a Shub & Smale que relaciona el condicionamiento μ_{norm} con la complejidad de los algoritmos de homotopía para la búsqueda de ceros aproximados proyectivos y afines, en el caso cero–dimensional (véase

por ejemplo la Proposición 4.2.6). Dejando de lado, por el momento, esa importante propiedad, expondremos a continuación algunos otros resultados que muestran la gran versatilidad del número de condicionamiento $\mu^{(r)}_{\mathrm{norm}}$. En algunos casos, las propiedades interesantes aparecen al restringirnos a ciertos valores de n, m, r. Dividimos esta tarea en varias subsecciones.

3.1.1. Teorema del Número de Condicionamiento No–Lineal.

El resultado que enunciamos a continuación es una generalización del resultado para el caso cero–dimensional debido a Shub & Smale que puede leerse por ejemplo en [14, Teor. 3, pág 234]. Se trata de una igualdad que permite relacionar, al igual que hemos hecho en el caso lineal, el condicionamiento de un problema y el inverso de la distancia al conjunto de los problemas mal condicionados; en esta ocasión, el condicionamiento no depende sólo del sistema f (como ocurría en el caso lineal), sino que también depende de la solución $\zeta \in V(f)$ en que nos centremos. El conjunto de problemas mal condicionados estará relacionado con la solución particular. Mediremos por tanto la distancia en la "fibra", esto es, en el conjunto de sistemas que se anulan en ζ.
Para cada valor positivo de $r \in \mathbb{N}$, $1 \leq r \leq m$, hemos definido el conjunto $\Sigma^r_{(d)} \subseteq \mathbb{P}(\mathcal{H}^m_{(d)})$ como el conjunto de sistemas que tienen alguna solución ζ con $rank(d_\zeta f) \leq r$. En la Sección 3.4 demostraremos que $\Sigma^r_{(d)}$ es una variedad proyectiva, y estimaremos su dimensión y su grado. Sin embargo, ahora estamos interesados en "localizar" $\Sigma^r_{(d)}$ en cada solución particular ζ. Por ello para cada $\zeta \in \mathbb{P}_n(\mathbb{C})$, definimos el conjunto

$$\Sigma^r_{(d)}(\zeta) := \{f \in \mathbb{P}(\mathcal{H}^m_{(d)}) : \zeta \in V(f), rank(d_\zeta f) \leq r\}.$$

Presentamos el siguiente resultado, que será demostrado en la Sección 3.2. Este tipo de resultados, que relacionan el condicionamiento de un problema con la distancia a un conjunto de problemas "singulares", puede encontrarse en otros muchos tipos de problemas (véase por ejemplo [103, 30]).

Teorema 3.1.1 *Con las notaciones anteriores, para todo punto $(f, \zeta) \in W$ se tiene:*

$$\mu^{(r)}_{\mathrm{norm}}(f, \zeta) = \frac{1}{d_{\mathbb{P}}(f, \Sigma^{r-1}_{(d)}(\zeta))},$$

donde $d_{\mathbb{P}}$ es la distancia proyectiva en $\mathbb{P}(\mathcal{H}^m_{(d)})$.

Por lo tanto, el número de condicionamiento generalizado $\mu^{(r)}_{\mathrm{norm}}(f, \zeta)$ es exactamente la distancia de f al conjunto de sistemas g que tienen a ζ como solución de "rango" a lo más $r - 1$, esto es, tales que $rank(d_\zeta g) \leq r - 1$.

3.1.2. El control de la convergencia del operador de Newton en el caso de dimensión positiva

En la Nota 1.7.3, hemos indicado la capacidad del número de condicionamiento para controlar la convergencia del operador de Newton proyectivo cero–dimensional. A continuación mostraremos un resultado que permite extender esta conclusión al caso afín de dimensión positiva. En la Sección 1.8, hemos introducido la cantidad $\gamma(f,\zeta)$, definida para $f \in \mathcal{H}^m_{(d)}$, $\zeta \in V_{\mathbb{C}^n}(f)$. Entonces, para cada $f \in \mathcal{H}^m_{(d)}$ hemos definido la cantidad del caso peor $\gamma_{\mathrm{worst}}(f) := \sup_{\zeta \in V_{\mathbb{C}^n}(f)} \gamma(f,\zeta)$, y hemos demostrado que existe una constante universal $u_0 > 0$ tal que todos los puntos afines en un entorno de $V_{\mathbb{C}^n}(f)$ de radio $u_0 \gamma_{\mathrm{worst}}(f)^{-1}$ son ceros aproximados afines de f (véase Corolario 1.8.2). Además, hemos descrito una función $\varphi_0 : \mathbb{C}^n \longrightarrow \mathbb{P}_n(\mathbb{C})$ que relaciona las soluciones afines de un sistema con su conjunto de soluciones proyectivas (véase la fórmula (1.5) y los comentarios que la siguen). Parece por tanto razonable considerar, para un sistema polinomial $f \in \mathbb{P}(\mathcal{H}^m_{(d)})$, el peor de los condicionamientos de sus soluciones proyectivas, definido como sigue:

$$\mu^{(r)}_{\mathrm{worst}}(f) := \max_{\zeta \in V(f)} \mu^{(r)}_{\mathrm{norm}}(f,\zeta). \tag{3.1}$$

Tendría sentido definir esto para todo $r \geq 2$, pero nuestro interés se centra ahora en el caso de que r sea igual al número de ecuaciones m, esto es, el caso de $\mu^{(m)}_{\mathrm{norm}}$. En efecto, el siguiente resultado relaciona la cantidad γ_{worst} con el número de condicionamiento $\mu^{(m)}_{\mathrm{worst}}$.

Proposición 3.1.2 *Sea $f \in \mathcal{H}^m_{(d)}$ un sistema de ecuaciones polinomiales, y sea $\zeta \in V_{\mathbb{C}^n}(f)$ una solución de f. Entonces, tenemos la siguiente desigualdad:*

$$\gamma(f,\zeta) \leq \frac{d^{3/2}}{2} \mu^{(m)}_{\mathrm{norm}}(f, \varphi_0(\zeta)).$$

Además, se tiene que

$$\gamma_{\mathrm{worst}}(f) \leq \frac{d^{3/2}}{2} \mu^{(m)}_{\mathrm{worst}}(f).$$

Por lo tanto, si queremos controlar el radio de convergencia del operador de Newton afín en dimensión positiva para un sistema $f \in \mathcal{H}^m_{(d)}$, basta con controlar su condicionamiento $\mu^{(m)}_{\mathrm{worst}}(f)$.

3.1.3. Separación y estabilidad del conjunto de soluciones

Englobamos en esta subsección dos resultados que vienen de la mano. El primero de ellos, debido a J. P. Dedieu (cf. [31]), se aplica en el caso cero–dimensional: En él se utiliza el condicionamiento $\mu^{(n)}_{\mathrm{worst}}$ para establecer una

cota inferior de la distancia entre dos soluciones proyectivas distintas de un sistema de ecuaciones. Esto es, se tiene el siguiente importante (y elegante) resultado:

Teorema 3.1.3 (Separación, Dedieu [31]) *Sea $f \in \mathbb{P}(\mathcal{H}^n_{(d)})$ un sistema de ecuaciones homogéneas cero–dimensional. Entonces, se tiene la siguiente desigualdad:*

$$\min_{x,y \in V(f), x \neq y} d_{\mathbb{P}}(x,y) \geq \frac{1}{d^{3/2}\mu^{(n)}_{\text{worst}}(f)}.$$

Por lo tanto, si tenemos un control sobre el condicionamiento de un sistema, podemos dar una estimación sobre el diámetro que debe tener un conjunto para poseer, a lo sumo, una solución proyectiva del mismo.
Finalmente, nos ocupamos ahora de la estabilidad del conjunto de soluciones, que fue estudiada por Shub & Smale para el caso cero–dimensional (véase por ejemplo [112]), y por Degót en el caso de dimensión positiva (véase [37]). En realidad, dicha estabilidad puede verse como consecuencia de que el número $\mu^{(m)}_{\text{norm}}(f,\zeta)$ controla (módulo algunos elementos técnicos) la norma como aplicación lineal de la inversa generalizada de la diferencial $d_{(f,\zeta)}p_1$, donde $p_1 : W \longrightarrow \mathbb{P}(\mathcal{H}^m_{(d)})$ es la proyección canónica sobre la primera componente. Resumimos a continuación el caso de dimensión positiva, que engloba al caso cero–dimensional.

Teorema 3.1.4 (Estabilidad, Shub & Smale, Degót) *Sea $f \in \mathbb{P}(\mathcal{H}^m_{(d)})$ un sistema de ecuaciones, y sea $x \in V(f)$ una solución regular de f. Entonces, existe un entorno $U_{f,x}$ de f en $\mathbb{P}(\mathcal{H}^m_{(d)})$ tal que para todo $g \in U_{f,x}$ se satisface:*

$$d_{\mathbb{P}}(x, V(g)) \leq 2\mu^{(m)}_{\text{norm}}(f,x) d_{\mathbb{P}}(f,g).$$

En otras palabras, $\mu^{(m)}_{\text{norm}}(f,x)$ controla, en cierto sentido, cómo cambia la variedad solución cerca de x cuando perturbamos ligeramente el sistema f. Otra forma de interpretar el teorema anterior es como sigue: Para un sistema $g \in \mathbb{P}(\mathcal{H}^m_{(d)})$ suficientemente cercano a f, se tiene que la variedad solución $V(f)$ está contenida en un tubo de radio $2\mu^{(m)}_{\text{worst}}(f) d_{\mathbb{P}}(f,g)$ en torno a la variedad solución $V(g)$.
A la vista de este resultado, si para un sistema polinomial $f \in \mathbb{P}(\mathcal{H}^m_{(d)})$ consideramos el peor de los condicionamientos de sus soluciones $\mu^{(m)}_{\text{worst}}$ de la ecuación (3.1), concluimos que ese número controla la estabilidad global, o estabilidad en el caso peor, del conjunto de soluciones de f. Si queremos garantizar que *todo el conjunto de soluciones de f sea estable*, debemos exigir que $\mu^{(m)}_{\text{worst}}(f)$ sea pequeño.
Sin embargo, el condicionamiento en el caso peor no siempre se corresponde con la realidad computacional. En efecto, el conjunto de soluciones de un sistema sub–determinado es en general una variedad proyectiva de dimensión positiva, o incluso en el caso cero–dimensional el número de puntos

que lo componen es genéricamente igual al número de Bézout, que es una cantidad exponencial en el número de incógnitas n. Difícilmente podremos por lo tanto abordar la cuestión de calcular (o aproximar) todas las soluciones de un sistema de ecuaciones. Por este motivo, puede ser mucho más interesante definir el condicionamiento de un sistema no como el peor de los condicionamientos posibles, sino como el caso más probable: Si garantizamos que *la mayor parte de las soluciones de un sistema de ecuaciones es estable*, tendremos una herramienta para estimar, probabilísticamente, la precisión de nuestros cálculos.
Tiene por tanto sentido definir la siguiente cantidad, para todo $f \in \mathbb{P}(\mathcal{H}^m_{(d)})$:

$$\mu^{(m)}_{\mathrm{av}}(f) := \mathrm{E}_{V(f)}[\mu^{(m)}_{\mathrm{norm}}(f, \cdot)], \tag{3.2}$$

donde E significa esperanza. Vemos pues que $\mu^{(m)}_{\mathrm{av}}(f)$ es la esperanza del condicionamiento en el conjunto de soluciones de f. En otras palabras, si $\mu^{(m)}_{\mathrm{av}}(f)$ es pequeño, *la mayor parte del conjunto de soluciones de f es estable.*

3.1.4. Algunas estimaciones de probabilidad

Una vez ha quedado clara la importancia del número de condicionamiento generalizado, aparece de modo natural la pregunta de cómo de estiman o calculan esos números. Lamentablemente, en el caso de sistemas de ecuaciones polinomiales, estimar $\mu^{(m)}_{\mathrm{worst}}$ o $\mu^{(m)}_{\mathrm{av}}$ para un sistema dado es una labor muy complicada: Requiere del conocimiento previo del conjunto de soluciones, que en general es el objetivo a calcular, no un dato que conozcamos de antemano. Por ello, el diseño de procedimientos que resuelvan de modo eficiente los problemas a los que nos enfrentamos requiere algún tipo de conocimiento teórico que determine *a priori* el valor esperado para el número de condicionamiento de nuestro problema.
Nos preguntamos por tanto, de modo natural, por la distribución de probabilidad de $\mu^{(r)}_{\mathrm{norm}}$ y temas relacionados. Una primera observación es que en el Teorema del Condicionamiento de la Sección 3.1.1 la distancia que aparece relacionada al condicionamiento $\mu^{(r)}_{\mathrm{norm}}$ no es la distancia proyectiva en $\mathbb{P}(\mathcal{H}^m_{(d)})$, sino una distancia "en la fibra" muy particular. Una pregunta inicial es cómo se distribuye la distancia proyectiva, antes de analizar el caso más complicado de la distancia en la fibra. A continuación exponemos una respuesta parcial a este problema. El motivo de que sea parcial es que nos restringimos al caso cero–dimensional. Nuestro espacio de medida es por tanto, de momento, $\mathbb{P}(\mathcal{H}_{(d)}) = \mathbb{P}(\mathcal{H}^n_{(d)})$. El siguiente corolario será demostrado en la Sección 3.4 (véase también [13]).

Corolario 3.1.5 *Sea* $1 \leq r \leq n-1$ *un número natural y sea* $\frac{1}{\mathrm{dist_r}} : \mathbb{P}(\mathcal{H}_{(d)}) \longrightarrow [0,\infty]$ *la función definida como*

$$\left(\frac{1}{\mathrm{dist_r}}\right)(f) := \frac{1}{d_{\mathbb{P}}(f, \Sigma^r_{(d)})}.$$

Entonces, la esperanza en $\mathbb{P}(\mathcal{H}_{(d)})$ *de la función* $\frac{1}{\mathrm{dist_r}}$ *satisface la siguiente desigualdad:*

$$E\left[\frac{1}{\mathrm{dist_r}}\right] \leq 8N(r+1)^3 \; d\left[\prod_{i=1}^{n}(d_i+1)\right]^{\frac{1}{2(n-r)^2}}.$$

Este resultado (y su versión probabilística, véase Teorema 3.4.1)tiene interés propio: Nos proporciona una cota de la distancia esperable de un sistema de ecuaciones a los sistemas con soluciones singulares de corango dado. De otra manera, se trata de un resultado de probabilidad que controla la distribución de los sistemas en función de la clase de singularidades que presenten. Se trata por tanto, en última instancia, de un resultado de distribución de las singularidades según su clasificación, y de los puntos cercanos a las singularidades de distintos tipos.
La demostración del Corolario 3.1.5, aunque contenga algunos elementos técnicos algo complejos, no requiere de un análisis demasiado específico si utilizamos las herramientas del Capítulo 2. En efecto, al tratarse de la distancia proyectiva en $\mathbb{P}(\mathcal{H}_{(d)})$, nos enfrentamos simplemente a la tarea de estimar los volúmenes de ciertos tubos, lo que se puede hacer, salvando ciertas complicaciones técnicas, con los teoremas generales del Capítulo 2.
Sin embargo, el análisis de los números de condicionamiento $\mu^{(r)}_{\mathrm{norm}}$ será bastante más complicado. Para estimar convenientemente sus distribuciones de probabilidad, tendremos que generalizar ampliamente las técnicas de integración en la variedad de incidencia generalizada W (véase la identidad (1.6)) desarrolladas por Shub & Smale en su artículo [113]. También tendremos que utilizar los resultados de comportamiento en media del condicionamiento lineal de sistemas obtenidos en el Capítulo 2, y resolver no pocos problemas técnicos de considerable dificultad. Un resultado destacable que demostraremos es el siguiente, que generaliza las técnicas de integración en la variedad de incidencia W del caso cero–dimensional (debidas a Shub & Smale, véase [113]) al caso de dimensión positiva. Se trata de una igualdad integral que permite transformar integrales en el espacio total de sistemas en integrales en el conjunto V_{e_0} de sistemas que se anulan en e_0.

Teorema 3.1.6 *Sea* $\phi : W \longrightarrow \mathbb{R}$ *una función integrable, tal que para toda matriz unitaria* $U \in \mathcal{U}_{n+1}$*, se satisface que:*

$$\phi(f,\zeta) = \phi(f \circ U, U^{-1}\zeta).$$

Entonces, se tiene la siguiente igualdad:

$$\int_{f\in\mathbb{P}(\mathcal{H}^m_{(d)})}\int_{\zeta\in V(f)}\phi(f,\zeta)\ dV(f)\ d\mathbb{P}(\mathcal{H}^m_{(d)}) =$$

$$\vartheta_n\int_{f\in V_{e_0}}\phi(f,e_0)\det(T_{e_0}f(T_{e_0}f)^*)\ dV_{e_0}.$$

Como consecuencia de estas reflexiones obtendremos algunos resultados que arrojan cierta luz a las preguntas que nos hemos hecho en las últimas secciones. Comenzamos con el siguiente teorema (véase el Teorema 3.6.1 para una versión más precisa), que estima la distribución de probabilidad del condicionamiento generalizado $\mu^{(r)}_{\mathrm{norm}}$ en su forma más general, esto es, abarcando a la vez el caso singular (esto es, caso $r < m$) y el caso sub–determinado (esto es, $m < n$).

Teorema 3.1.7 *Sea $(d) = (d_1, \ldots, d_m)$ tal que $d_i > 1$ para algún i, $1 \leq i \leq m$. Sea $\varepsilon > 0$ un número real. Entonces, tenemos:*

$$\frac{1}{\nu_{\mathbb{P}(\mathcal{H}^m_{(d)})}[\mathbb{P}(\mathcal{H}^m_{(d)})]}\int_{f\in\mathbb{P}(\mathcal{H}^m_{(d)})}\nu_{n-m}[\zeta\in V(f):\mu^{(r)}_{\mathrm{norm}}(f,\zeta)>\varepsilon^{-1}]\ d\mathbb{P}(\mathcal{H}^m_{(d)}) \leq$$

$$2\pi e^{1/3}\vartheta_{n-m}\mathcal{D}\left(\sqrt{Nmr(n+1)}\ \varepsilon\right)^{2(m-r+1)(n-r+2)}.$$

A continuación, un resultado que acota la esperanza del condicionamiento $\mu^{(m)}_{\mathrm{av}}$.

Teorema 3.1.8 *Sea $(d) = (d_1, \ldots, d_m)$ tal que $d_i > 1$ para algún i, $1 \leq i \leq m$. Entonces, el valor esperado del número de condicionamiento $\mu^{(m)}_{\mathrm{av}}$ de la ecuación (3.2) satisface la siguiente desigualdad:*

$$\mathrm{E}_{\mathbb{P}(\mathcal{H}^m_{(d)})}[\mu^{(m)}_{\mathrm{av}}] \leq 3m\sqrt{nN}.$$

En el caso $m = 1$, podemos incluso obtener una igualdad (véase el Teorema 3.7.1).

En cuanto al condicionamiento del caso peor $\mu^{(m)}_{\mathrm{worst}}$, también obtenemos una cota, aunque mucho peor que la que acabamos de exponer:

Teorema 3.1.9 *Sea $(d) = (d_1, \ldots, d_m)$ tal que $d_i > 1$ para algún i, $1 \leq i \leq m$. Entonces, el valor esperado del número de condicionamiento $\mu^{(m)}_{\mathrm{worst}}$ de la ecuación (3.1) satisface la siguiente desigualdad:*

$$\mathrm{E}_{\mathbb{P}(\mathcal{H}^m_{(d)})}[\mu^{(m)}_{\mathrm{worst}}] \leq \frac{\mathcal{D}^{1/4}}{d^{3/2}}[10N^{1/2}mn^{1/2}d^{3/2}]^{\frac{n-m+2}{2}}.$$

Podemos escribir la conclusión de estos dos resultados como sigue: Por el Teorema 3.1.8, el valor esperado del condicionamiento $\mu_{\mathrm{av}}^{(m)}(f)$ es muy pequeño (parecido a la raíz cuadrada del tamaño del input $N+1$). Por lo tanto, para un sistema de ecuaciones elegido al azar, *podemos esperar que la mayor parte del conjunto de soluciones sea extremadamente estable*. Sin embargo, si pedimos que la totalidad del conjunto de soluciones sea estable, nos encontramos con que la situación puede ser totalmente distinta: Como se ve en el Teorema 3.1.9, solamente podemos obtener una cota exponencial para el valor esperable de $\mu_{\mathrm{worst}}^{(m)}$. Puede suceder por lo tanto que la mayoría de los sistemas de ecuaciones tengan algunas soluciones muy inestables (sobre todo si el valor de $n-m$ es grande), aunque sepamos que la mayoría de esas soluciones son muy estables.
Combinando este último resultado con la tesis del Corolario 1.8.2 y la Proposición 3.1.2, obtendremos un interesante teorema que proporciona una cota inferior para el radio de convergencia del operador de Newton en el caso de dimensión positiva.

Teorema 3.1.10 *Con las hipótesis y notaciones del Teorema 3.1.9, se tiene:*

1. $\gamma_{\mathrm{worst}}(f) < +\infty$ *para casi todo* $\mathcal{H}_{(d)}^m$. *En otras palabras,*

$$Prob[f \in \mathcal{H}_{(d)}^m : \gamma_{\mathrm{worst}}(f) < +\infty] = 1.$$

2. *La esperanza de* γ_{worst} *satisface la siguiente desigualdad:*

$$\mathrm{E}_{\mathcal{H}_{(d)}^m}[\gamma_{\mathrm{worst}}] \leq \frac{\mathcal{D}^{1/4}}{2}[10m\sqrt{nN}d^{3/2}]^{\frac{n-m+2}{2}},$$

3. *La esperanza del radio de convergencia del operador de Newton,* $\frac{u_0}{\gamma_{\mathrm{worst}}}$, *satisface la siguiente desigualdad:*

$$\mathrm{E}_{\mathcal{H}_{(d)}^m}\left[\frac{u_0}{\gamma_{\mathrm{worst}}}\right] \geq \frac{2u_0}{\mathcal{D}^{1/4}[10m\sqrt{nN}d^{3/2}]^{\frac{n-m+2}{2}}},$$

Este resultado significa lo siguiente: Para casi todas las variedades intersección completa $V \subseteq \mathbb{C}^n$, existe un tubo V_R de radio $R > 0$ tal que todos los puntos de V_R son ceros aproximados afines de V (en el sentido de la Sección 3.3). Además, podemos proporcionar una cota inferior para el valor esperable del radio R.

3.2. Demostración del Teorema del Número de Condicionamiento

Comenzamos demostrando el Teorema 3.1.1. Utilizaremos las notaciones ya introducidas en las secciones 1.4 y 1.6. Sea $(f,\zeta) \in W$ un punto de la

variedad de incidencia. Sea $U \in \mathcal{U}_{n+1}$ una matriz unitaria tal que $Ue_0 = \zeta$. Recordemos que $\mu^{(r)}_{\text{norm}}$ es unitario invariante (identidad (1.11)). Además, como U define una isometría en $\mathbb{P}(\mathcal{H}^m_{(d)})$ que deja fijo $\Sigma^{r-1}_{(d)}$ y transforma V_ζ en V_{e_0}, tenemos que

$$\frac{1}{d_{\mathbb{P}}(f, \Sigma^{r-1}_{(d)}(\zeta))} = \frac{1}{d_{\mathbb{P}}(f \circ U, \Sigma^{r-1}_{(d)}(e_0))}.$$

Por lo tanto, basta con demostrar el resultado para el caso de que $\zeta = e_0$. Sea pues $f \in V_{e_0}$ un sistema. Escribimos $f = h \oplus h_L \in L^{\perp}_{e_0} \oplus L_{e_0}$, donde hemos elegido un representante afín tal que $\|f\|_\Delta = 1$. Por definición, se tiene que

$$\mu^{(r)}_{\text{norm}}(f, e_0) = \frac{1}{\sqrt{\sigma_r^2 + \cdots + \sigma_m^2}},$$

donde $\sigma_1 \geq \cdots \geq \sigma_m$ son los valores singulares de

$$\Delta(d)^{-1/2} T_{e_0} f = \Delta(d)^{-1/2} T_{e_0} h_L.$$

Por el Teorema 1.5.10, concluimos que $\mu^{(r)}_{\text{norm}}(f, e_0)$ es igual al inverso de la distancia de Frobenius de $\Delta(d)^{-1/2} T_{e_0} h_L$ al conjunto de matrices en $\mathcal{M}_{m \times n}(\mathbb{C})$ de rango menor o igual que $r-1$. Por lo tanto, existe una matriz $M \in \mathcal{M}_{m \times n}(\mathbb{C})$ tal que

$$\|\Delta(d)^{-1/2} T_{e_0} h_L - M\|_F = \frac{1}{\mu^{(r)}_{\text{norm}}(f, e_0)}, \qquad rank(M) \leq r-1.$$

Como ψ_{e_0} es una isometría, existe un elemento $g_L := \psi^{-1}_{e_0}(M) \in L_{e_0}$ tal que

$$\|h_L - g_L\|_\Delta = \frac{1}{\mu^{(r)}_{\text{norm}}(f, e_0)}, \qquad rank(d_{e_0} g_L) \leq r-1.$$

Además, a partir de la descripción explícita de $\Delta(d)^{-1/2} T_{e_0} h_L$ proporcionada por la demostración del Teorema 1.5.10 (véase [122], por ejemplo), tenemos:

$$\langle h_L, g_L \rangle_\Delta = \langle \Delta(d)^{-1/2} T_{e_0} h_L, \Delta(d)^{-1/2} T_{e_0} g_L \rangle_F =$$

$$= \sigma_1^2 + \ldots + \sigma_{r-1}^2 = \langle \Delta(d)^{-1/2} T_{e_0} g_L, \Delta(d)^{-1/2} T_{e_0} g_L \rangle_F = \langle g_L, g_L \rangle_\Delta.$$

Por lo tanto, se tiene que el elemento $g = h + g_L$ satisface:

$$rank(d_{e_0} g) = rank(d_{e_0} g_L) \leq r-1, \Rightarrow g \in \Sigma^{r-1}_{(d)}(e_0),$$

y además

$$d_{\mathbb{P}}(g, f) = d_{\mathbb{P}}(h + g_L, h + h_L) = \sqrt{1 - \frac{|\langle h + g_L, h + h_L \rangle_\Delta|^2}{\langle h + g_L, h + g_L \rangle_\Delta}} =$$

$$\sqrt{1-\frac{|\langle h,h\rangle_\Delta+\langle g_L,h_L\rangle_\Delta|^2}{\langle h,h\rangle_\Delta+\langle g_L,g_L\rangle_\Delta}}=$$

$$\sqrt{1-\|h\|_\Delta^2-(\sigma_1^2+\ldots+\sigma_{r-1}^2)}=\sqrt{\|h_L\|_\Delta^2-(\sigma_1^2+\ldots+\sigma_{r-1}^2)}=$$

$$\sqrt{\sigma_r^2+\ldots+\sigma_m^2}=\frac{1}{\mu^{(r)}_{\mathrm{norm}}(f,e_0)}.$$

Queda por tanto demostrado que $d_{\mathbb{P}}(f,\Sigma^{r-1}_{(d)}(e_0))\leq\frac{1}{\mu^{(r)}_{\mathrm{norm}}(f,e_0)}$.

Para demostrar la otra desigualdad, sea $g\in\Sigma^{r-1}_{(d)}(e_0)$ otro sistema, $g=h'\oplus g_L\in L^\perp_{e_0}\oplus L_{e_0}$. Elegimos un representante de g tal que se satisfacen las condiciones:

$$\|g_L\|_\Delta^2=\|h_L\|_\Delta^2-(\sigma_r^2+\cdots+\sigma_m^2),\qquad\langle g_L,h_L\rangle_\Delta\in\mathbb{R}^{0,+}.$$

Se tiene que la matriz $\Delta(d)^{-1/2}T_{e_0}g=\Delta(d)^{-1/2}T_{e_0}g_L$ tiene rango menor o igual que $r-1$. Por tanto, de nuevo por el Teorema 1.5.10, sabemos que

$$\sigma_r^2+\cdots+\sigma_m^2\leq\|\Delta(d)^{-1/2}T_{e_0}g_L-\Delta(d)^{-1/2}T_{e_0}h_L\|_F^2=$$

$$\langle\Delta(d)^{-1/2}(T_{e_0}g_L-T_{e_0}h_L),\Delta(d)^{-1/2}(T_{e_0}g_L-T_{e_0}h_L)\rangle_F=$$

$$\langle g_L-h_L,g_L-h_L\rangle_\Delta=\|g_L\|_\Delta^2+\|h_L\|_\Delta^2-2\langle g_L,h_L\rangle_\Delta=$$

$$\|h_L\|_\Delta^2-(\sigma_r^2+\cdots+\sigma_m^2)+\|h_L\|_\Delta^2-2\langle g_L,h_L\rangle_\Delta.$$

Concluimos la siguiente desigualdad,

$$\langle g_L,h_L\rangle_\Delta\leq\|h_L\|_\Delta^2-(\sigma_r^2+\cdots+\sigma_m^2)=\|g_L\|_\Delta^2$$

Por lo tanto, se tiene:

$$d_{\mathbb{P}}(g,f)=d_{\mathbb{P}}(h'+g_L,h+h_L)=\sqrt{1-\frac{|\langle h'+g_L,h+h_L\rangle_\Delta|^2}{\langle h'+g_L,h'+g_L\rangle_\Delta}}=$$

$$\sqrt{1-\frac{|\langle h',h\rangle_\Delta+\langle g_L,h_L\rangle_\Delta|^2}{\|h'\|_\Delta^2+\|g_L\|_\Delta^2}}\geq$$

$$\sqrt{1-\frac{|\langle h',h\rangle_\Delta|^2+\langle g_L,h_L\rangle_\Delta^2+2|\langle h',h\rangle_\Delta|\langle g_L,h_L\rangle_\Delta}{\|h'\|_\Delta^2+\|g_L\|_\Delta^2}}\geq$$

$$\sqrt{1-\frac{\|h'\|_\Delta^2\|h\|_\Delta^2+\langle g_L,h_L\rangle_\Delta^2+2\|h'\|_\Delta\|h\|_\Delta\langle g_L,h_L\rangle_\Delta}{\|h'\|_\Delta^2+\|g_L\|_\Delta^2}}=$$

$$\sqrt{1-\frac{(\|h'\|_\Delta\|h\|_\Delta+\langle g_L,h_L\rangle_\Delta)^2}{\|h'\|_\Delta^2+\|g_L\|_\Delta^2}}\geq$$

$$\sqrt{1-\frac{(\|h'\|_\Delta\|h\|_\Delta+\|g_L\|_\Delta^2)^2}{\|h'\|_\Delta^2+\|g_L\|_\Delta^2}}.$$

Consideramos la función real

$$t\mapsto\frac{(t\|h\|_\Delta+\|g_L\|_\Delta^2)^2}{t^2+\|g_L\|_\Delta^2}.$$

Algunos cálculos elementales confirman que el máximo de esa función se alcanza en $t=\|h\|_\Delta$. Por lo tanto,

$$\frac{(\|h'\|_\Delta\|h\|_\Delta+\|g_L\|_\Delta^2)^2}{\|h'\|_\Delta^2+\|g_L\|_\Delta^2}\leq\frac{(\|h\|_\Delta^2+\|g_L\|_\Delta^2)^2}{\|h\|_\Delta^2+\|g_L\|_\Delta^2}=\|h\|_\Delta^2+\|g_L\|_\Delta^2=$$

$$\|h\|_\Delta^2+\|h_L\|_\Delta^2-(\sigma_r^2+\cdots+\sigma_m^2)=1-(\sigma_r^2+\cdots+\sigma_m^2).$$

Deducimos que

$$d_{\mathbf{P}}(g,f)\geq\sqrt{1-(1-(\sigma_r^2+\cdots+\sigma_m^2))}=\sqrt{\sigma_r^2+\cdots+\sigma_m^2}=\frac{1}{\mu_{\mathrm{norm}}^{(r)}(f,e_0)},$$

lo que termina la demostración del resultado. ∎

3.3. El método de Newton en dimensión positiva

En esta Sección demostraremos la Proposición 3.1.2, que relaciona el condicionamiento de un problema con el comportamiento del operador de Newton asociado, en el caso de dimensión positiva.

Para un sistema de ecuaciones $f\in\mathcal{H}_{(d)}^m$ y para una solución afín $\zeta\in V_{\mathbb{C}^n}(f)$, hemos definido la cantidad $\gamma(f,\zeta)$ que controla el comportamiento local del operador de Newton afín cerca de ζ (véase Teorema 1.8.1). Además, para una solución proyectiva $\zeta'\in V(f)$, hemos definido el número de condicionamiento $\mu_{\mathrm{norm}}^{(m)}(f,\zeta')$. A continuación, consideramos un tercer número que resultará de utilidad en los lemas siguientes. Para $f\in\mathcal{H}_{(d)}^m$, $\zeta\in V_{\mathbb{C}^n}(f)$ regular, definimos:

$$\mu_{\mathrm{afin}}^{(m)}(f,\zeta):=\|f\|_\Delta\|(\partial_\zeta f)^\dagger Diag(d_i^{1/2}\|(1,\zeta)\|^{d_i-1})\|_2,$$

donde estamos denotando por $(1,\zeta)\in\mathbb{C}^{n+1}$ el punto afín formado por las coordenadas de ζ precedidas de un 1. Si $\zeta\in V_{\mathbb{C}^n}(f)$ es solución singular de f, definimos simplemente $\mu_{\mathrm{afin}}^{(m)}(f,\zeta):=+\infty$.

Obsérvese que en general las cantidades $\mu_{\mathrm{afin}}^{(m)}(f,\zeta)$ y $\mu_{\mathrm{norm}}^{(m)}(f,\varphi_0(\zeta))$ pueden no coincidir. A continuación hacemos un cierto esfuerzo por relacionar estos conceptos. Comenzamos con resultado esencial debido a Shub & Smale, conocido como "the higher derivative estimate" (véase por ejemplo [14, Lema 12, pág. 269], [112]).

Lema 3.3.1 (Shub & Smale) *Sea $f \in \mathcal{H}^m_{(d)}$ un sistema, y sea $x \in \mathbb{C}^{n+1}$ un punto cualquiera, tal que $\|x\|_2 = 1$. Entonces, se tiene la siguiente desigualdad:*

$$\left(\frac{\|\Delta(d)^{-1/2} d_x^{(k)} f\|_2}{\|f\|_\Delta k!}\right)^{1/(k-1)} \leq \frac{d^{3/2}}{2}. \tag{3.3}$$

El siguiente resultado, que se sigue fácilmente del Lema 3.3.1, relaciona γ con $\mu^{(m)}_{\text{afin}}$.

Lema 3.3.2 *Sea $f \in \mathcal{H}^m_{(d)}$ y sea $\zeta \in V_{\mathbb{C}^n}(f)$ una solución de f. Entonces,*

$$\|(1,\zeta)\|_2 \gamma(f,\zeta) \leq \frac{d^{3/2}}{2} \mu^{(m)}_{\text{afin}}(f,\zeta).$$

Demostración.– Basta comprobar el resultado para el caso de que ζ sea una solución regular de f. Primero, observamos que

$$\|(1,\zeta)\|_2 \gamma(f,\zeta) = \|(1,\zeta)\|_2 \sup_{k\geq 2} \left\| (\partial_\zeta f)^\dagger \frac{\partial^{(k)}_\zeta f}{k!} \right\|_2^{\frac{1}{k-1}} =$$

$$\sup_{k\geq 2} \left\| (\partial_\zeta f)^\dagger Diag(d_i^{1/2} \|(1,\zeta)\|_2^{d_i - 1}) Diag(d_i^{-1/2} \|(1,\zeta)\|_2^{k-d_i}) \frac{\partial^{(k)}_\zeta f}{k!} \right\|_2^{\frac{1}{k-1}} \leq$$

$$\sup_{k\geq 2} \mu^{(m)}_{\text{afin}}(f,\zeta)^{\frac{1}{k-1}} \sup_{k\geq 2} \left\| Diag(\|(1,\zeta)\|_2^{k-d_i} d_i^{-1/2}) \frac{\partial^{(k)}_\zeta f}{\|f\|_\Delta k!} \right\|_2^{\frac{1}{k-1}}.$$

Por la identidad (1.7), deducimos la siguiente desigualdad:

$$\sup_{k\geq 2} \left\| Diag(\|(1,\zeta)\|_2^{k-d_i} d_i^{-1/2}) \frac{\partial^{(k)}_\zeta f}{\|f\|_\Delta k!} \right\|_2^{\frac{1}{k-1}} \leq$$

$$\sup_{k\geq 2} \left\| Diag(\|(1,\zeta)\|_2^{k-d_i} d_i^{-1/2}) \frac{d^{(k)}_{(1,\zeta)} f}{\|f\|_\Delta k!} \right\|_2^{\frac{1}{k-1}} =$$

$$\sup_{k\geq 2} \left\| \Delta(d)^{-1/2} \frac{d^{(k)}_{\frac{(1,\zeta)}{\|(1,\zeta)\|_2}} f}{\|f\|_\Delta k!} \right\|_2^{\frac{1}{k-1}}.$$

Por el Lema 3.3.1 deducimos que esta última cantidad es a lo sumo $\frac{d^{3/2}}{2}$, y el lema queda demostrado. ∎

A continuación demostramos un resultado que relaciona el número de condicionamiento $\mu^{(m)}_{\text{norm}}$ con su equivalente afín $\mu^{(m)}_{\text{afin}}$.

Lema 3.3.3 *Sea $f \in \mathcal{H}^m_{(d)}$ un sistema, $\zeta \in \mathbb{C}^n$ una solución de f. Entonces, se tiene la siguiente desigualdad:*

$$\mu^{(m)}_{\text{afin}}(f,\zeta) \leq \|(1,\zeta)\|_2 \mu^{(m)}_{\text{norm}}(f,\varphi_0(\zeta)).$$

Demostración.–
De nuevo, basta con demostrar el resultado en el caso de que $\zeta \in V_{\mathbb{C}^n}(f)$ sea una solución afín regular de f, lo que implica que $\varphi_0(\zeta) \in V(f)$ es una solución proyectiva regular de f. Además, $(1,\zeta) \in \mathbb{C}^{n+1}$ es un representante afín del punto proyectivo $\varphi_0(\zeta)$. Por la identidad (1.7) y la definición de $\mu^{(m)}_{\text{afin}}$, podemos escribir

$$\mu^{(m)}_{\text{afin}}(f,\zeta) = \|f\|_\Delta \|(d_{(1,\zeta)} f \mid_{e_0^\perp})^\dagger Diag(d_i^{1/2}\|(1,\zeta)\|_2^{d_i-1})\|_2.$$

Ahora, obsérvese que

$$\|(d_{(1,\zeta)} f \mid_{e_0^\perp})^\dagger Diag(d_i^{1/2}\|(1,\zeta)\|_2^{d_i-1})\|_2 =$$

$$\|(d_{(1,\zeta)} f \mid_{e_0^\perp})^\dagger (d_{(1,\zeta)} f \mid_{(1,\zeta)^\perp})(d_{(1,\zeta)} f \mid_{(1,\zeta)^\perp})^\dagger Diag(d_i^{1/2}\|(1,\zeta)\|_2^{d_i-1})\|_2 \leq$$

$$\|(d_{(1,\zeta)} f \mid_{e_0^\perp})^\dagger (d_{(1,\zeta)} f \mid_{(1,\zeta)^\perp})\|_2 \|(d_{(1,\zeta)} f \mid_{(1,\zeta)^\perp})^\dagger Diag(d_i^{1/2}\|(1,\zeta)\|_2^{d_i-1})\|_2$$

Por la identidad (1.10), concluimos:

$$\mu^{(m)}_{\text{afin}}(f,\zeta) \leq \mu^{(m)}_{\text{norm}}(f,\varphi_0(\zeta))\ \|(d_{(1,\zeta)} f \mid_{e_0^\perp})^\dagger (d_{(1,\zeta)} f \mid_{(1,\zeta)^\perp})\|_2.$$

Por lo tanto, basta con demostrar que para una solución cualquiera $(1,\zeta)$ de f, se tiene que

$$\|(d_{(1,\zeta)} f \mid_{e_0^\perp})^\dagger (d_{(1,\zeta)} f \mid_{(1,\zeta)^\perp})\|_2 \leq \|(1,\zeta)\|_2.$$

Comprobemos esta última desigualdad. En efecto, sea $w \in (1,\zeta)^\perp$un vector. Si $w \in e_0^\perp$, entonces

$$\|(d_{(1,\zeta)} f \mid_{e_0^\perp})^\dagger (d_{(1,\zeta)} f \mid_{(1,\zeta)^\perp})(v)\|_2 =$$

$$\|(d_{(1,\zeta)} f \mid_{e_0^\perp})^\dagger (d_{(1,\zeta)} f \mid_{e_0^\perp})(v)\|_2 \leq \|v\|_2,$$

por las propiedades elementales de la inversa generalizada (véase por ejemplo [34]). Supongamos ahora que $v \in (1,\zeta)^\perp \cap ((1,\zeta)^\perp \cap e_0^\perp)^\perp$, que es un subespacio complejo de dimensión a lo más 1. Entonces, $v = t(-\|\zeta\|_2^2, \zeta) \in \mathbb{C}^{n+1}$ para algún $t \in \mathbb{C}$. Además, podemos escribir v como

$$v = w - t\|\zeta\|_2^2(1,\zeta), \quad w \in e_0^\perp.$$

Entonces,

$$\frac{1}{\|v\|_2}\|(d_{(1,\zeta)} f \mid_{e_0^\perp})^\dagger (d_{(1,\zeta)} f \mid_{(1,\zeta)^\perp})(v)\|_2 =$$

$$\frac{1}{\|v\|_2}\|(d_{(1,\zeta)}f\mid_{e_0^\perp})^\dagger(d_{(1,\zeta)}f)(v)\|_2 =$$

$$\frac{1}{\|v\|_2}\|(d_{(1,\zeta)}f\mid_{e_0^\perp})^\dagger(d_{(1,\zeta)}f)(w)\|_2 \leq \frac{\|w\|_2}{\|v\|_2}.$$

Por último,

$$\frac{\|w\|_2}{\|v\|_2} = \frac{\|t(-\|\zeta\|_2^2,\zeta)+t\|\zeta\|_2^2(1,\zeta)\|_2}{\|t(-\|\zeta\|_2^2,\zeta)\|_2} = \|(1,\zeta)\|_2.$$

Esto acaba la demostración del lema. ∎

3.3.1. Demostración de la Proposición 3.1.2

La segunda afirmación de la proposición se sigue de la primera de forma inmediata. Demostremos pues la primera de las afirmaciones. Por el Lema 3.3.2,

$$\gamma(f,\zeta) \leq \frac{1}{\|(1,\zeta)\|_2}\frac{d^{3/2}}{2}\mu_{\text{afin}}^{(m)}(f,\zeta).$$

Por el Lema 3.3.3, esta última cantidad es a lo sumo

$$\frac{1}{\|(1,\zeta)\|_2}\frac{d^{3/2}}{2}\|(1,\zeta)\|_2\mu_{\text{norm}}^{(m)}(f,\varphi_0(\zeta)),$$

como queríamos. ∎

3.4. La distancia a las variedades singulares de corango dado

Comenzamos ahora con la estimación probabilística y en media de las cantidades indicadas en la introducción de este capítulo. El primer caso del que nos ocupamos es el de las variedades de corango dado. Esto es, demostraremos el Corolario 3.1.5. Dicho corolario se obtendrá como consecuencia el siguiente teorema de distribución de probabilidad.

Teorema 3.4.1 *Sea $\varepsilon > 0$ un número real, y sea $1 \leq r \leq n-1$ un número natural. Sea P_ε^r la probabilidad de que un sistema $f \in \mathbb{P}(\mathcal{H}_{(d)}^n)$ elegido al azar esté a distancia menor que ε del conjunto $\Sigma_{(d)}^r$. Entonces, P_ε^r es a lo sumo*

$$2\prod_{i=1}^{n}(d_i+1)\binom{n+1}{r}\binom{n}{r}\left(\frac{e\ N(r+1)\ d\ \varepsilon}{(n-r)^2}\right)^{2(n-r)^2}.$$

Demostración.– Primero, demostraremos que $\Sigma^r_{(d)} \subseteq \mathbb{P}(\mathcal{H}^n_{(d)})$ es una variedad proyectiva compleja de codimensión al menos $(n-r)^2$ y grado a lo más

$$\prod_{i=1}^{n}(d_i+1)\binom{n+1}{r}\binom{n}{r}((r+1)d)^{2(n-r)^2}.$$

Introduzcamos algunas notaciones nuevas.

$$W^r := \{(f,\zeta) \in \mathbb{P}(\mathcal{H}^n_{(d)}) \times \mathbb{P}_n(\mathbb{C}) : f(\zeta) = 0,\ rank(d_\zeta f) \leq r\},$$

$$W^r_0 := \{f \in \mathbb{P}(\mathcal{H}^n_{(d)}) : f(e_0) = 0,\ rank(d_{e_0} f) \leq r\},$$

$$S^r := \{M \in \mathcal{M}_n(\mathbb{C}) : rank(M) \leq r\}.$$

Por el Teorema del Ideal Principal de Krull (véase por ejemplo [85, Cor. 3.8]), se tiene que:

$$\dim(W^r_0) \geq \dim(W^r) - n. \tag{3.4}$$

Por otro lado, observemos que el conjunto $L^\perp_{e_0}$ definido en la Sección 1.4 es un subespacio lineal complejo de $\mathcal{H}^n_{(d)}$ de dimensión (compleja) $N+1-n-n^2$. Por otro lado, tenemos que $V_{e_0} = \mathbb{P}(\mathcal{M}_n(\mathbb{C}) \times L^\perp_{e_0})$. Además, tenemos la siguiente igualdad:

$$W^r_0 = \mathbb{P}(S^r \times L^\perp_{e_0}).$$

Como se demuestra en [47, 18], el conjunto S^r es una variedad algebraica irreducible de $\mathcal{M}_n(\mathbb{C})$ de dimensión compleja $n^2-(n-r)^2$. Por tanto, W^r_0 es también una variedad algebraica irreducible de V_{e_0} de dimensión compleja $N-n-(n-r)^2$. Deducimos que W^r es una variedad algebraica de dimensión compleja a lo más

$$\dim(W^r) \leq \dim(W^r_0) + n = N - (n-r)^2.$$

Sea $p_1 : \mathbb{P}(\mathcal{H}^n_{(d)}) \times \mathbb{P}_n(\mathbb{C}) \longrightarrow \mathbb{P}(\mathcal{H}^n_{(d)})$ la proyección en la primera componente. Por el Teorema Fundamental de Teoría de la Eliminación (véase por ejemplo [109]), $\Sigma^r_{(d)} = p_1(W^r)$ es una variedad algebraica, y su dimensión es a lo más $N-(n-r)^2$.
Para obtener la cota en el grado, sea $\widetilde{W}^r \subseteq \mathcal{H}^n_{(d)} \times \mathbb{C}^{n+1}$ el conjunto definido como sigue:

$$\widetilde{W}^r := \{(f,\zeta) \in \mathcal{H}^n_{(d)} \times \mathbb{C}^{n+1} : f(\zeta) = 0, rank(d_\zeta f) = 0\},$$

y sea $\widetilde{\pi}_1 : \mathcal{H}^n_{(d)} \times \mathbb{C}^{n+1} \longrightarrow \mathcal{H}^n_{(d)}$ la proyección ortogonal. Obsérvese que

$$\Sigma^r_{(d)} = \mathbb{P}(\widetilde{\pi}_1(\widetilde{W}^r)).$$

Por lo tanto, $\deg(\Sigma^r_{(d)})$ está acotado superiormente por $\deg(\widetilde{W}^r)$. Además, para $(f,\zeta) \in \mathcal{H}^n_{(d)} \times \mathbb{C}^{n+1}$, el hecho de que $rank(d_\zeta f) = r$ es equivalente a la satisfacción del siguiente sistema de igualdades y desigualdades algebraicas:

$$\bigvee_{\substack{1\le i_1\le\cdots\le i_r\le n\\ 0\le j_1\le\cdots\le j_r\le n}}\left[det(M^{j_1,\ldots,j_r}_{i_1,\ldots,i_r})\neq 0,\ \bigwedge_{\substack{k_1\neq i_1,\ldots,i_r\\ k_2\neq j_1,\ldots,j_r}}\det(M^{j_1,\ldots,j_r,k_2}_{i_1,\ldots,i_r,k_1})=0\right],$$

donde $M^{j_1,\ldots,j_r,k_2}_{i_1,\ldots,i_r,k_1}$ es el menor $(r+1)\times(r+1)$ de $d_\zeta f$ obtenido de las filas $i_1,\ldots,i_r,k_1$ y las columnas $j_1,\ldots,j_r,k_2$. Por el Teorema de Bézout (como en [66]), tenemos que

$$\deg(\widetilde{W}^r)=\deg(\widetilde{W}^r\setminus\widetilde{W}^{r-1})\le$$

$$\prod_{i=1}^n(d_i+1)\sum_{\substack{1\le i_1\le\cdots\le i_r\le n\\ 0\le j_1\le\cdots\le j_r\le n}}\prod_{\substack{k_1\neq i_1,\ldots,i_r\\ k_2\neq j_1,\ldots,j_r}}(d_{i_1}+\cdots+d_{i_r}+d_{k_1})\le$$

$$\prod_{i=1}^n(d_i+1)\binom{n+1}{r}\binom{n}{r}((r+1)d)^{(n-r+1)(n-r)}\le$$

$$\prod_{i=1}^n(d_i+1)\binom{n+1}{r}\binom{n}{r}((r+1)d)^{2(n-r)^2},$$

como queríamos.
Hemos denotado por $(\mathbb{P}(\mathcal{H}^n_{(d)}),can)$ el espacio proyectivo con la estructura Riemanniana canónica. Por el Lema 1.4.1, sabemos que Δ^{-1} define una isometría lineal entre $(\mathbb{P}(\mathcal{H}^n_{(d)}),can)$ y $\mathbb{P}(\mathcal{H}^n_{(d)})$ con la estructura Riemanniana de Kostlan. Por lo tanto, Δ^{-1} conserva dimensiones y grados, no sólo volumen. Por el Teorema 2.3.6, deducimos que:

$$P^r_\varepsilon\le 2\deg(\Sigma^r_{(d)})\left(\frac{e\ N\ \varepsilon}{codim(\Sigma^r_{(d)})}\right)^{2codim(\Sigma^r_{(d)})}.$$

Utilizando las cotas para $\deg(\Sigma^r_{(d)})$ y $codim(\Sigma^r_{(d)})$ obtenidas arriba, deducimos la cota del teorema. ∎

Demostración del Corolario 3.1.5. Por el Teorema 3.4.1, sabemos que para todo real positivo t,

$$P\left[\frac{1}{\mathrm{dist_r}}>t\right]\le 2\prod_{i=1}^n(d_i+1)\binom{n+1}{r}\binom{n}{r}\left(\frac{e\ N(r+1)\ d}{t(n-r)^2}\right)^{2(n-r)^2}.$$

Podemos acotar groseramente esta cantidad por

$$2\prod_{i=1}^n(d_i+1)\left(\frac{e\ N(r+1)^3\ d}{(n-r)^2}\right)^{2(n-r)^2}t^{-2(n-r)^2}.$$

Por lo tanto, por el Lema 2.5.4, se tiene que

$$E\left[\frac{1}{\mathrm{dist_r}}\right] \leq \frac{e\ N(r+1)^3\ d}{(n-r)^2}\frac{2(n-r)^2}{2(n-r)^2-1}\left(2\prod_{i=1}^{n}(d_i+1)\right)^{\frac{1}{2(n-r)^2}}.$$

Ahora, como $(n-r)^2 \geq 1$, esta última expresión vale a lo más

$$8N(r+1)^3\ d\left[\prod_{i=1}^{n}(d_i+1)\right]^{\frac{1}{2(n-r)^2}},$$

como queríamos demostrar. ∎

3.5. Integración en la variedad de incidencia generalizada

En esta sección desarrollaremos ampliamente algunas herramientas técnicas de gran aplicabilidad que nos permitirán demostrar los teoremas 3.1.8, 3.1.9 y 3.1.10 de la introducción de este capítulo.
Para todo número real $\varepsilon > 0$ y todo número natural r, $1 \leq r \leq m$, denotamos por $\chi_\varepsilon^{(r)}$ la función característica del conjunto

$$\{(f,\zeta) : \zeta \in V(f), \mu_{\mathrm{norm}}^{(r)}(f,\zeta) > \varepsilon^{-1}\}.$$

Sean $x \in \mathbb{P}_n(\mathbb{C})$ y $f \in \mathbb{P}(\mathcal{H}_{(d)}^m)$ dos elementos. Sean $\underline{x}$ y $\underline{f}$ representantes afines respectivos de x y f, tales que $\|x\|_2 = \|f\|_\Delta = 1$. Consideremos los espacios ortogonales $\underline{x}^\perp$ y $\underline{f}^\perp$. En el Capítulo 1, identidades (1.1) y (1.3), hemos definido aplicaciones $\varphi_{\underline{x}}$ y $\varphi_{\underline{f}}$ como sigue:

$$\begin{array}{rccc} \varphi_{\underline{x}}: & x^\perp & \longrightarrow & \mathbb{P}_n(\mathbb{C}) \setminus x^\perp, \\ & y & \mapsto & \underline{x}+y \end{array}$$

$$\begin{array}{rccc} \varphi_{\underline{f}}: & f^\perp & \longrightarrow & \mathbb{P}(\mathcal{H}_{(d)}^m) \setminus f^\perp, \\ & g & \mapsto & \underline{f}+g, \end{array}$$

de las cuales sabemos que son difeomorfismos y además isometrías en 0 (esto es, $d_0\varphi_{\underline{x}}$ y $d_0\varphi_{\underline{f}}$ son isometrías). El siguiente resultado (véase [14, pág. 193]) resume algunas de las propiedades de la variedad de incidencia. Incluimos la demostración por completitud de esta memoria, dado que se trata de un resultado esencial en nuestro esquema.

Proposición 3.5.1 *La variedad de incidencia W es una variedad diferenciable de dimensión compleja $N+n-m$. Además, sea $(f,\zeta) \in W$ un punto cualquiera, y sean $\underline{f}, \underline{\zeta}$ representantes afines respectivos de f, ζ tales que*

$\|\underline{f}\|_\Delta = \|\underline{\zeta}\|_2 = 1$. *Entonces, el espacio tangente* $T_{(f,\zeta)}W \subseteq f^\perp \times \zeta^\perp$ *puede identificarse con el espacio vectorial definido a continuación:*

$$T_{(f,\zeta)}W \equiv \{(g,x) \in f^\perp \times \zeta^\perp : g(\underline{\zeta}) + (d_{\underline{\zeta}}\underline{f})x = 0\}.$$

La identificación viene dada por la isometría

$$d_{(0,0)}(\varphi_{\underline{f}} \times \varphi_{\underline{\zeta}}).$$

Demostración.– Sea $(f,\zeta) \in W$ un punto, y sean $\underline{f}, \underline{\zeta}$ representantes afines respectivos de f, ζ, tales que $\|\underline{f}\|_\Delta = \|\underline{\zeta}\|_2 = 1$. Entonces, la siguiente aplicación es un difeomorfismo sobre su imagen:

$$\begin{array}{rccc} \varphi_{\underline{f}} \times \varphi_{\underline{\zeta}} : & f^\perp \times \zeta^\perp & \longrightarrow & \mathbb{P}(\mathcal{H}^m_{(d)}) \times \mathbb{P}_n(\mathbb{C}) \\ & (g,x) & \mapsto & (\underline{f}+g, \underline{\zeta}+x). \end{array}$$

Sea $\widehat{W} := (\varphi_{\underline{f}} \times \varphi_{\underline{\zeta}})^{-1}(W)$ la anti–imagen de W mediante este difeomorfismo. Es decir,

$$\widehat{W} := \{(g,x) \in f^\perp \times \zeta^\perp : (\underline{f}+g)(\underline{\zeta}+x) = 0\}.$$

Entonces, $\widehat{W}$ también es la anti–imagen de 0 mediante la aplicación diferenciable

$$\begin{array}{rccc} \varphi : & f^\perp \times \zeta^\perp & \longrightarrow & \mathbb{C}^m \\ & (g,x) & \mapsto & (\underline{f}+g)(\underline{\zeta}+x). \end{array}$$

Ahora, se comprueba trivialmente que 0 es un valor regular de esta aplicación. Esto es, para todo punto $(g,x) \in \widehat{W}$, $d_{(g,x)}\varphi$ es sobreyectiva. Por lo tanto, $\widehat{W}$ es una variedad diferenciable compleja de dimensión $N+n-m$. Deducimos que W es también una variedad diferenciable compleja de dimensión $N+n-m$.
Además, un vector $(g,x) \in f^\perp \times \zeta^\perp$ pertenece a $T_{(0,0)}\widehat{W}$ si y sólo si

$$g(\underline{\zeta}) + d_{\underline{\zeta}}\underline{f}(x) = 0,$$

lo que termina la demostración. ∎

Como hemos indicado en la Sección 1.6 del Capítulo 1, para cada matriz unitaria $U \in \mathcal{U}_{n+1}$, se definen isometrías en $\mathbb{P}(\mathcal{H}^m_{(d)})$ y $\mathbb{P}_n(\mathbb{C})$, como sigue:

$$\begin{array}{ccccccc} \mathbb{P}(\mathcal{H}^m_{(d)}) & \longrightarrow & \mathbb{P}(\mathcal{H}^m_{(d)}), & \quad \mathbb{P}_n(\mathbb{C}) & \longrightarrow & \mathbb{P}_n(\mathbb{C}) \\ f & \mapsto & f \circ U & \quad x & \mapsto & U^{-1}x. \end{array}$$

Por tanto, observamos que queda también definida una isometría en W, de la forma

$$\begin{array}{ccc} W & \longrightarrow & W, \\ (f,\zeta) & \mapsto & (f \circ U, U^{-1}\zeta) \end{array}$$

Consideremos ahora las dos proyecciones canónicas

$$p_1 : W \longrightarrow \mathbb{P}(\mathcal{H}^m_{(d)}), \quad p_2 : W \longrightarrow \mathbb{P}_n(\mathbb{C}).$$

Podemos identificar de manera obvia $p_1^{-1}(f)$ y $V(f)$ (lo que demuestra que para casi todo punto $f \in \mathbb{P}(\mathcal{H}^m_{(d)})$, el conjunto de soluciones $V(f)$ es una variedad diferenciable de dimensión compleja $n-m$). De la misma manera, identificamos $p_2^{-1}(x)$ y V_x. A partir de ahora, no distinguiremos entre estos conceptos. El siguiente teorema generaliza ampliamente un resultado de Shub & Smale (véase por ejemplo [14, Prop. 2, pág. 244], [113]).

Teorema 3.5.2 *Sea $\phi : W \longrightarrow \mathbb{R}$ una función integrable, tal que para toda matriz unitaria $U \in \mathcal{U}_{n+1}$, se satisface que:*

$$\phi(f, \zeta) = \phi(f \circ U, U^{-1}\zeta).$$

Sea $\mathcal{J}$ la integral definida a continuación:

$$\mathcal{J} := \int_{(f,\zeta)\in W} \phi(f,\zeta) NJ_{(f,\zeta)}p_1 \; dW.$$

Entonces, las siguientes igualdades satisfacen:

$$\mathcal{J} = \int_{f\in\mathbb{P}(\mathcal{H}^m_{(d)})} \int_{\zeta\in V(f)} \phi(f,\zeta) \; dV(f) \; d\mathbb{P}(\mathcal{H}^m_{(d)}),$$

$$\mathcal{J} = \vartheta_n \int_{f\in V_{e_0}} \phi(f, e_0) \frac{NJ_{(f,e_0)}p_1}{NJ_{(f,e_0)}p_2} \; dV_{e_0}.$$

Demostración.– La primera de las dos igualdades resulta del Teorema 1.1.13 aplicado a p_1. Para la segunda, también el Teorema 1.1.13 nos asegura que

$$\mathcal{J} = \int_{x\in\mathbb{P}_n(\mathbb{C})} \int_{f\in V_x} \phi(f,x) \frac{NJ_{(f,x)}p_1}{NJ_{(f,x)}p_2} \; dV_x \; d\mathbb{P}_n(\mathbb{C}).$$

Ahora, sea $x \in \mathbb{P}_n(\mathbb{C})$ un punto proyectivo cualquiera y sea $U \in \mathcal{U}_{n+1}$ una matriz unitaria tal que $Ue_0 = x$. Entonces, la aplicación que manda f a $f \circ U$ es una isometría de V_x en V_{e_0}. Por lo tanto,

$$\int_{f\in V_x} \phi(f,x) \frac{NJ_{(f,x)}p_1}{NJ_{(f,x)}p_2} \; dV_x = \int_{f\in V_{e_0}} \phi(f\circ U^{-1}, Ue_0) \frac{NJ_{(f\circ U^{-1},Ue_0)}p_1}{NJ_{(f\circ U^{-1},Ue_0)}p_2} \; dV_{e_0}.$$

Ahora, $\phi(f \circ U^{-1}, Ue_0) = \phi(f, e_0)$. Además, las aplicaciones definidas a continuación son isometrías:

$$\begin{array}{ccc} W & \longrightarrow & W, \\ (g,z) & \mapsto & (g\circ U^{-1}, Uz) \end{array} \qquad \begin{array}{ccc} \mathbb{P}_n(\mathbb{C}) & \longrightarrow & \mathbb{P}_n(\mathbb{C}) \\ z & \mapsto & Uz \end{array}$$

Por tanto, estamos en las condiciones del Corolario 1.1.12. Deducimos que

$$NJ_{(f\circ U^{-1},Ue_0)}p_2 = NJ_{(f,e_0)}p_2.$$

Un argumento simétrico con la aplicación $\mathbb{P}(\mathcal{H}^m_{(d)}) \longrightarrow \mathbb{P}(\mathcal{H}^m_{(d)})$ dada como $f \mapsto f \circ U^{-1}$ demuestra que

$$NJ_{(f\circ U^{-1},Ue_0)}p_1 = NJ_{(f,e_0)}p_1,$$

y el teorema queda demostrado. ∎

A continuación calculamos el valor de los jacobianos normales que aparecen en el Teorema 3.5.2. De nuevo, el resultado que sigue generaliza los cálculos realizados por Shub & Smale en [113] (véase también [14, págs. 241–243]).

Lema 3.5.3 *Sea $f \in V_{e_0}$ un sistema tal que $rank(T_{e_0}f) = m$. Entonces, se tiene:*

$$NJ_{(f,e_0)}p_1 = \frac{1}{\det(Id_m + ((T_{e_0}f)^\dagger)^*(T_{e_0}f)^\dagger)},$$

$$NJ_{(f,e_0)}p_2 = \frac{1}{\det(Id_m + (T_{e_0}f)(T_{e_0}f)^*)}.$$

Demostración.– Como hemos visto en la Proposición 3.5.1,

$$T_{(f,e_0)}W \equiv \{(g,x) \in f^\perp \times e_0^\perp : g(e_0) + (T_{e_0}f)x = 0\},$$

donde hemos elegido un representante de norma 1 de f. Sea K_1 el núcleo de la matriz diferencial de p_1 en (f,e_0). Esto es, $K_1 := Ker(d_{(f,e_0)}p_1)$. Entonces, $K_1 = \{(0,x) : x \in Ker(T_{e_0}f)\}$, y

$$NJ_{(f,e_0)}p_1 = NJ_{(0,0)}((d_{(f,e_0)}p_1) \mid_{K_1^\perp}) = \frac{1}{NJ_{(0,0)}(((d_{(f,e_0)}p_1) \mid_{K_1^\perp})^{-1})} =$$

$$= \frac{1}{NJ_{(0,0)}((d_{(f,e_0)}p_1)^\dagger)}.$$

Sea β una base ortonormal de $f^\perp$ tal que los primeros m elementos de la base son los sistemas

$$\beta_1 := [X_0^{d_1}, 0, \ldots, 0],$$
$$\vdots$$
$$\beta_m := [0, \ldots, 0, X_0^{d_m}].$$

Observamos que las primeras m coordenadas de cualquier sistema $g := [g_1, \ldots, g_m] \in \mathcal{H}^m_{(d)}$ en esta bases coinciden exactamente con

$$g(e_0) = (g_1(e_0), \ldots, g_m(e_0)).$$

Además, tenemos las siguientes propiedades:

- $(d_{(f,e_0)}p_1)^\dagger(\beta_i) = (\beta_i, x_i),\ x_i := -(T_{e_0}f)^\dagger(e_i)$, para $1 \leq i \leq m$.
- $(d_{(f,e_0)}p_1)^\dagger(v) = (v, 0)$, para $v \in \beta, v \notin \{\beta_1, \ldots, \beta_m\}$.

Por lo tanto,

$$NJ_{(0,0)}((d_{(f,e_0)}p_1)^\dagger) = \det(Id_m + ((T_{e_0}f)^\dagger)^*(T_{e_0}f)^\dagger).$$

En cuanto a p_2, obsérvese de nuevo que

$$NJ_{(f,e_0)}p_2 = \frac{1}{NJ_{(0,0)}((d_{(f,e_0)}p_2)^\dagger)}.$$

Ahora, tenemos la siguiente igualdad

$$Ker(d_{(f,e_0)}p_2)^\perp = \{(g,0) : g(e_0) = 0\}^\perp = \langle \beta_1, \ldots, \beta_m \rangle \times e_0^\perp,$$

donde $\langle \beta_1, \ldots, \beta_m \rangle$ denota simplemente el subespacio generado por esos vectores. Por lo tanto,

$$(d_{(f,e_0)}p_2)^\dagger(e_i) = (g_i, e_i), \quad 1 \leq i \leq n.$$

donde las primeras m coordenadas de g_i en la base β forman el vector $-(T_{e_0}f)e_i$, y el resto de coordenadas son nulas. Concluimos que

$$NJ_{(0,0)}((d_{(f,e_0)}p_2)^\dagger) = \det(Id_n + (T_{e_0}f)^*(T_{e_0}f)).$$

Finalmente, sea $T_{e_0}f = U(D\ 0)V^*$ una descomposición en valores singulares de $T_{e_0}f$. Entonces, tenemos

$$\det(Id_n + (T_{e_0}f)^*(T_{e_0}f)) = \det\left(Id_n + V\begin{pmatrix} D^2 & 0 \\ 0 & 0 \end{pmatrix}V^*\right) = \det(Id_m + D^2),$$

y además

$$\det(Id_m + (T_{e_0}f)(T_{e_0}f)^*) = \det\left(Id_m + UD^2U^*\right) = \det(Id_m + D^2),$$

luego deducimos que

$$\det(Id_n + (T_{e_0}f)^*(T_{e_0}f)) = \det(Id_m + (T_{e_0}f)(T_{e_0}f)^*),$$

lo que termina la demostración del Lema. ∎

3.5.1. Demostración del Teorema 3.1.6

El Teorema 3.1.6 es consecuencia inmediata del Teorema 3.5.2 y del siguiente lema técnico.

Lema 3.5.4 *Sea $f \in V_{e_0}$ un sistema tal que $rank(T_{e_0} f) = m$. Con las notaciones anteriores, tenemos:*

$$\frac{NJ_{(f,e_0)}p_1}{NJ_{(f,e_0)}p_2} = \det((T_{e_0}f)(T_{e_0}f)^*).$$

Demostración.– Por el Lema 3.5.3,

$$\frac{NJ_{(f,e_0)}p_1}{NJ_{(f,e_0)}p_2} = \frac{\det(Id_m + BB^*)}{\det(Id_m + (B^\dagger)^* B^\dagger)},$$

donde $B := T_{e_0}f \in \mathcal{M}_{m\times n}(\mathbb{C})$. Entonces,

$$\frac{1}{\det(BB^*)}\frac{NJ_{(f,e_0)}p_1}{NJ_{(f,e_0)}p_2} = \frac{\det(Id_m + BB^*)}{\det(BB^* + BB^*(B^\dagger)^* B^\dagger)}.$$

Ahora, observamos que $BB^*(B^\dagger)^* B^\dagger = B(B^\dagger B)^* B^\dagger$. Además, por las propiedades elementales de la pseudo inversa de Moore-Penrose, $B^\dagger B \in \mathcal{M}_n(\mathbb{C})$ es autoadjunta. Esto es,

$$B^\dagger B = (B^\dagger B)^*.$$

Por último, B es sobreyectiva, luego $BB^\dagger = Id_m$. Por lo tanto,

$$\det(BB^* + BB^*(B^\dagger)^* B^\dagger) = \det(BB^* + BB^\dagger BB^\dagger) = \det(BB^* + Id_m),$$

y el lema queda demostrado. ∎

3.5.2. Algunas consecuencias

Como consecuencia casi directa del Teorema 3.1.6 obtenemos algunos resultados que pasamos a exponer a continuación.

Lema 3.5.5 *Sea $\Phi : [0, +\infty] \longrightarrow [0, +\infty]$ una función integrable. Entonces, se tiene la siguiente igualdad:*

$$\int_{f\in\mathbb{P}(\mathcal{H}^m_{(d)})} \int_{\zeta\in V(f)} \Phi\left(\mu^{(r)}_{\mathrm{norm}}(f,\zeta)\right)\ dV(f)\ d\mathbb{P}(\mathcal{H}^m_{(d)}) =$$

$$\vartheta_n \int_{f\in V_{e_0}} \Phi(\mu^{(r)}_{\mathrm{norm}}(f,e_0)) \det((T_{e_0}f)(T_{e_0}f)^*)\ dV_{e_0}.$$

Demostración.– Sabemos que para todo elemento $(f, \zeta) \in W$ y para toda matriz unitaria $U \in \mathcal{U}_{n+1}(\mathbb{C})$, se tiene que

$$\mu^{(r)}_{\text{norm}}(f, \zeta) = \mu^{(r)}_{\text{norm}}(f \circ U, U^{-1}\zeta).$$

(véase la identidad (1.11)). Por lo tanto, también se satisface la siguiente igualdad:

$$\Phi(\mu^{(r)}_{\text{norm}}(f, \zeta)) = \Phi(\mu^{(r)}_{\text{norm}}(f \circ U, U^{-1}\zeta)).$$

El lema se sigue del Teorema 3.1.6, aplicado a $\phi := \Phi \circ \mu^{(r)}_{\text{norm}}$. ∎

Este último lema adquiere una forma muy particular cuando los grados d_i, $1 \leq i \leq n$ son todos iguales a 1. En efecto, presentamos el siguiente resultado,

Lema 3.5.6 *Sea $\Phi : [0, +\infty] \longrightarrow [0, +\infty]$ una función integrable. Entonces, se tiene la siguiente igualdad:*

$$\vartheta_{n-m} \int_{M \in \mathbb{P}(\mathcal{M}_{m\times(n+1)}(\mathbb{C}))} \Phi\left(\kappa^{(r)}_D(M)\right)\, d\mathbb{P}(\mathcal{M}_{m\times(n+1)}(\mathbb{C})) =$$

$$\vartheta_n \int_{M \in \mathbb{P}(\mathcal{M}_{m\times n}(\mathbb{C}))} \Phi(\kappa^{(r)}_D(M)) \det(MM^*)\, d\mathbb{P}(\mathcal{M}_{m\times n}(\mathbb{C})),$$

donde el representante de M en la última integral se elige de modo que $\|M\|_F = 1$.

Demostración.– Consideremos el Lema 3.5.5 en el caso de que $(d) := (1, \ldots, 1) \in \mathbb{N}^n$. Entonces, $\mathbb{P}(\mathcal{H}^m_{(d)})$ se convierte en $\mathbb{P}(\mathcal{M}_{m\times(n+1)}(\mathbb{C}))$, y el número de condicionamiento $\mu^{(r)}_{\text{norm}}(M, \zeta)$, donde $\zeta \neq 0$ está en el núcleo de M, coincide con $\kappa^{(r)}_D(M)$. Por lo tanto, $\mu^{(r)}_{\text{norm}}(M, \zeta)$ no depende de la solución ζ elegida, y el Lema 3.5.5 se escribe como sigue.

$$\int_{M \in \mathbb{P}(\mathcal{M}_{m\times(n+1)}(\mathbb{C}))} \Phi\left(\kappa^m(M)\right) \nu_{n-m}[Ker(M)]\, d\mathbb{P}(\mathcal{M}_{m\times(n+1)}(\mathbb{C})) =$$

$$\vartheta_n \int_{M \in V_{e_0}} \Phi(\kappa^{(r)}_D(M) \det((T_{e_0}M)(T_{e_0}M)^*))\, dV_{e_0}.$$

Ahora, en el caso lineal tenemos que

$$V_{e_0} = \{M \in \mathbb{P}(\mathcal{M}_{m\times(n+1)}(\mathbb{C})) : Me_0 = 0\},$$

es un subespacio lineal de $\mathbb{P}(\mathcal{M}_{m\times(n+1)}(\mathbb{C}))$ que se identifica del modo obvio con $\mathbb{P}(\mathcal{M}_{m\times n}(\mathbb{C}))$. En efecto, una matriz pertenece a V_{e_0} si y solamente si su primera columna es igual a cero. Además, bajo esta identificación, el condicionamiento $\kappa^{(r)}_D$ –definido independientemente en $\mathbb{P}(\mathcal{M}_{m\times(n+1)}(\mathbb{C}))$ y en $\mathbb{P}(\mathcal{M}_{m\times n}(\mathbb{C}))$– no cambia. Finalmente, sea $M \in \mathbb{P}(\mathcal{M}_{m\times n}(\mathbb{C}))$ una

matriz cualquiera, y sea $(0\ M) \in \mathbb{P}(\mathcal{M}_{m\times n+1}(\mathbb{C}))$ la matriz obtenida al añadir a M una columna de ceros. Entonces, se tiene que

$$T_{e_0}(0\ M) = (0\ M)\mid_{e_0^\perp} = M,$$

para algún representante fijado tal que $\|M\|_F = 1$. El lema se sigue, dado que $\nu_{n-m}[Ker(M)] = \vartheta_{n-m}$ para casi toda matriz $M \in \mathbb{P}(\mathcal{M}_{m\times(n+1)}(\mathbb{C}))$. ■

A continuación escribimos un útil resultado técnico que nos permite relacionar la integración de ciertas funciones en el espacio de sistemas de ecuaciones con la integración en el espacio de matrices asociado. El esquema de prueba es muy similar al que se sigue en la demostración del teorema principal de Shub & Smale en [113], aunque nuestro enunciado es mucho más general.

Teorema 3.5.7 *Sea $(d) = (d_1, \ldots, d_m)$ tal que $d_i > 1$ para algún i, $1 \leq i \leq m$. Sea $\Phi : [0, +\infty] \longrightarrow [0, +\infty]$ una función integrable. Entonces, se tiene la siguiente igualdad:*

$$\int_{f\in\mathbb{P}(\mathcal{H}^m_{(d)})} \int_{\zeta\in V(f)} \Phi(\mu^{(r)}_{\mathrm{norm}}(f,\zeta))\ dV(f)\ d\mathbb{P}(\mathcal{H}^m_{(d)}) =$$

$$2\pi\vartheta_{N-m-nm}\vartheta_{n-m}\mathcal{D}\int_0^1 (1-t^2)^{N-m-nm}t^{2nm+2m-1}\mathcal{G}_t(\Phi)\ dt,$$

donde

$$\mathcal{G}_t(\Phi) := \int_{M\in\mathbb{P}(\mathcal{M}_{m\times(n+1)}(\mathbb{C}))} \Phi\left(\frac{\kappa^{(r)}_D(M)}{t}\right)\ d\mathbb{P}(\mathcal{M}_{m\times(n+1)}(\mathbb{C})).$$

Demostración.– Sea $S^1(V_{e_0})$ la esfera de radio 1 en V_{e_0}. Esto es,

$$S^1(V_{e_0}) := \{f \in \mathcal{H}^m_{(d)} : \|f\|_\Delta = 1, f(e_0) = 0\}.$$

Sea $\hat{\pi} : S^1(V_{e_0}) \longrightarrow L_{e_0}$ la proyección ortogonal. Observamos que $S^1(V_{e_0})$ es una esfera de dimensión $2N - 2m + 1$, y para cada $f \in L_{e_0}, \|f\|_\Delta < 1$, el conjunto $\hat{\pi}^{-1}(f)$ es una esfera de dimensión real $2N-2m+1-2nm$ y radio $(1 - \|f\|^2_\Delta)^{1/2}$. Por lo tanto, el volumen $(2N - 2m + 1 - 2nm)$–dimensional de $\hat{\pi}^{-1}(f)$ es

$$\nu_{\hat{\pi}^{-1}(f)}[\hat{\pi}^{-1}(f)] = (1 - \|f\|^2_\Delta)^{N-m-nm+1/2}2\pi\vartheta_{N-m-nm}. \tag{3.5}$$

Además, se comprueba fácilmente que

$$NJ_f\hat{\pi} = (1 - \|\hat{\pi}(f)\|^2_\Delta)^{1/2}.$$

Denotamos por $I(\Phi)$ la integral en el espacio de sistemas polinomiales que queremos calcular. Esto es,

$$I(\Phi) := \int_{f\in\mathbb{P}(\mathcal{H}^m_{(d)})} \int_{\zeta\in V(f)} \Phi(\mu^{(r)}_{\mathrm{norm}}(f,\zeta))\ dV(f)\ d\mathbb{P}(\mathcal{H}^m_{(d)}).$$

Por el Lema 3.5.5, se tiene que

$$I(\Phi) = \vartheta_n \int_{f\in V_{e_0}} \Phi(\mu^{(r)}_{\mathrm{norm}}(f,e_0)) \det((T_{e_0}f)(T_{e_0}f)^*)\ dV_{e_0}.$$

Ahora, por la estructura Riemanniana del espacio proyectivo complejo, tenemos que:

$$\vartheta_n \int_{f\in V_{e_0}} \Phi(\mu^{(r)}_{\mathrm{norm}}(f,e_0)) \det((T_{e_0}f)(T_{e_0}f)^*)\ dV_{e_0} =$$

$$\frac{\vartheta_n}{2\pi} \int_{f\in S^1(V_{e_0})} \Phi(\mu^{(r)}_{\mathrm{norm}}(f,e_0)) \det((T_{e_0}f)(T_{e_0}f)^*)\ dS^1(V_{e_0}).$$

Por el Teorema 1.1.13 (Fórmula de la Co–área), esta última expresión es igual a:

$$\frac{\vartheta_n}{2\pi} \int_{f\in L_{e_0}} \int_{g\in\hat{\pi}^{-1}(f)} \Phi(\mu^{(m)}_{\mathrm{norm}}(g,e_0)) \frac{\det((T_{e_0}g)(T_{e_0}g)^*)}{(1-\|f\|^2_\Delta)^{1/2}}\ d\hat{\pi}^{-1}(f)\ dL_{e_0}.$$

Ahora, si $g\in\hat{\pi}^{-1}(f)$, entonces se tiene que

$$\mu^{(r)}_{\mathrm{norm}}(g,e_0) = \frac{\kappa^{(r)}_D(\psi_{e_0}(f))}{\|\psi_{e_0}(f)\|_F},\ T_{e_0}g = T_{e_0}f.$$

Deducimos pues que $I(\Phi)$ es igual a

$$\frac{\vartheta_n}{2\pi} \int_{\substack{f\in L_{e_0}\\ \|f\|_\Delta\le 1}} \nu_{\hat{\pi}^{-1}(f)}[\hat{\pi}^{-1}(f)]\Phi\left(\frac{\kappa^{(r)}_D(\psi_{e_0}(f))}{\|\psi_{e_0}(f)\|_F}\right) \frac{\det((T_{e_0}f)(T_{e_0}f)^*)}{(1-\|f\|^2_\Delta)^{1/2}}\ dL_{e_0},$$

Por la ecuación (3.5), deducimos que $I(\Phi)$ es igual a $\vartheta_n\vartheta_{N-m-nm}$ multiplicado por

$$\int_{\substack{f\in L_{e_0}\\ \|f\|_\Delta\le 1}} (1-\|f\|^2_\Delta)^{N-m-nm}\Phi\left(\frac{\kappa^{(r)}_D(\psi_{e_0}(f))}{\|\psi_{e_0}(f)\|_F}\right) \det((T_{e_0}f)(T_{e_0}f)^*)\ dL_{e_0}.$$

Denotemos $\alpha := N-m-nm$. Entonces, de nuevo el Teorema 1.1.13 aplicado a ψ_{e_0} muestra que

$$\int_{\substack{f\in L_{e_0}\\ \|f\|_\Delta\le 1}} (1-\|f\|^2_\Delta)^{\alpha}\Phi\left(\frac{\kappa^{(r)}_D(\psi_{e_0}(f))}{\|\psi_{e_0}(f)\|_F}\right) \det((T_{e_0}f)(T_{e_0}f)^*)\ dL_{e_0} =$$

$$\det(\Delta(d)) \int_{\substack{M\in\mathcal{M}_{m\times n}(\mathbb{C})\\ \|M\|_F\leq 1}} (1-\|M\|_F^2)^\alpha \Phi\left(\frac{\kappa_D^{(r)}(M)}{\|M\|_F}\right) \det(MM^*)\ d\mathcal{M}_{m\times n}(\mathbb{C}) =$$

$$\mathcal{D} \int_{\substack{M\in\mathcal{M}_{m\times n}(\mathbb{C})\\ \|M\|_F\leq 1}} (1-\|M\|_F^2)^\alpha \Phi\left(\frac{\kappa_D^{(r)}(M)}{\|M\|_F}\right) \det(MM^*)\ d\mathcal{M}_{m\times n}(\mathbb{C}).$$

En coordenadas polares, esta última expresión es igual a

$$\mathcal{D} \int_0^1 (1-t^2)^\alpha \int_{\|M\|_F=t} \Phi\left(\frac{\kappa_D^{(r)}(M)}{t}\right) \det(MM^*)\ dS^t(\mathcal{M}_{m\times n}(\mathbb{C}))\ dt =$$

$$\mathcal{D} \int_0^1 (1-t^2)^\alpha t^\beta \int_{\|M\|_F=1} \Phi\left(\frac{\kappa_D^{(r)}(M)}{t}\right) \det(MM^*)\ dS^1(\mathcal{M}_{m\times n}(\mathbb{C}))\ dt,$$

donde hemos denotado $\beta := 2mn+2m-1$. Ahora, para cada elección posible de $t\in[0,1]$, se tiene que

$$\int_{\|M\|_F=1} \Phi\left(\frac{\kappa_D^{(r)}(M)}{t}\right) \det(MM^*)\ dS^1(\mathcal{M}_{m\times n}(\mathbb{C})) =$$

$$= 2\pi \int_{M\in\mathbb{P}(\mathcal{M}_{m\times n}(\mathbb{C}))} \Phi\left(\frac{\kappa_D^{(r)}(M)}{t}\right) \det(MM^*)\ d\mathbb{P}(\mathcal{M}_{m\times n}(\mathbb{C})),$$

donde el representante de M en la última fórmula se elige de modo que $\|M\|_F = 1$. Consideremos la función real

$$\begin{array}{rccc} \Phi_t: & [0,+\infty] & \longrightarrow & [0,+\infty] \\ & s & \mapsto & \Phi\left(\frac{s}{t}\right) \end{array}$$

Entonces, podemos escribir

$$2\pi \int_{M\in\mathbb{P}(\mathcal{M}_{m\times n}(\mathbb{C}))} \Phi\left(\frac{\kappa_D^{(r)}(M)}{t}\right) \det(MM^*)\ d\mathbb{P}(\mathcal{M}_{m\times n}(\mathbb{C})) =$$

$$2\pi \int_{M\in\mathbb{P}(\mathcal{M}_{m\times n}(\mathbb{C}))} \Phi_t(\kappa_D^{(r)}(M)) \det(MM^*)\ d\mathbb{P}(\mathcal{M}_{m\times n}(\mathbb{C})),$$

y por el Lema 3.5.6, esta última expresión es igual a

$$2\pi \frac{\vartheta_{n-m}}{\vartheta_n} \int_{M\in\mathbb{P}(\mathcal{M}_{m\times (n+1)}(\mathbb{C}))} \Phi_t(\kappa_D^{(r)}(M))\ d\mathbb{P}(\mathcal{M}_{m\times (n+1)}(\mathbb{C})).$$

Hemos demostrado por lo tanto que $I(\Phi)$ es igual a $2\pi\vartheta_{n-m}\vartheta_\alpha\mathcal{D}$ multiplicado por

$$\int_0^1 (1-t^2)^\alpha t^\beta \int_{M\in\mathbb{P}(\mathcal{M}_{m\times (n+1)}(\mathbb{C}))} \Phi_t(\kappa_D^{(r)}(M))\ d\mathbb{P}(\mathcal{M}_{m\times (n+1)}(\mathbb{C})),$$

y el teorema queda demostrado. ∎

Demostramos a continuación un resultado ya conocido, si bien nuestra prueba es particularmente elegante a partir del Teorema 3.5.7.

Corolario 3.5.8 *Para todo sistema de ecuaciones $f \in \mathbb{P}(\mathcal{H}^m_{(d)})$, excepto para un conjunto de medida nula, se tiene que el conjunto de soluciones $V(f)$ es una variedad proyectiva lisa de dimensión $n-m$, y su volumen es:*

$$\nu_{n-m}[V(f)] = \vartheta_{n-m}\mathcal{D}.$$

Demostración.– La primera parte del corolario ha sido ya indicada con anterioridad. En efecto, una de las posibles formas de demostrarla es utilizando el Teorema 1.1.4 aplicado a la proyección p_1: Deducimos que para casi todo sistema de ecuaciones $f \in \mathbb{P}(\mathcal{H}^m_{(d)})$, se tiene que f es un valor regular de p_1. Por el Corolario 1.1.6, concluimos que su anti–imagen por p_1 es una variedad diferenciable compleja de dimensión $n-m$. Ahora, hemos visto que $p_1^{-1}(f)$ es difeomorfa a $V(f)$, luego $V(f)$ es una variedad diferenciable (además de algebraica) de dimensión $n-m$. Nos ocupamos ahora de la igualdad para el volumen de esa variedad (módulo un conjunto de medida cero). Si aplicamos el Teorema 3.5.7 a la función constante $\Phi \equiv 1$, obtenemos que

$$\int_{f\in\mathbb{P}(\mathcal{H}^m_{(d)})} \nu_{n-m}[V(f)]\; d\mathbb{P}(\mathcal{H}^m_{(d)}) = 2\pi\vartheta_{N-m-nm}\vartheta_{n-m}\vartheta_{nm+m-1}\mathcal{D} \times$$

$$\times \int_0^1 (1-t^2)^{N-m-nm} t^{2nm+2m-1}\; dt.$$

El valor de esta última integral es bien conocido:

$$\frac{1}{2}\frac{\Gamma(nm+m)\Gamma(N-m-nm+1)}{\Gamma(N+1)},$$

Usando que para todo número natural $k \in \mathbb{N}$ se tiene que

$$\vartheta_k = \frac{\pi^k}{\Gamma(k+1)},$$

concluimos que

$$\frac{1}{\nu_{\mathbb{P}(\mathcal{H}^m_{(d)})}[\mathbb{P}(\mathcal{H}^m_{(d)})]} \int_{f\in\mathbb{P}(\mathcal{H}^m_{(d)})} \nu_{n-m}[V(f)]\; d\mathbb{P}(\mathcal{H}^m_{(d)}) = \vartheta_{n-m}\mathcal{D}.$$

Por otro lado, para todo sistema $f \in \mathbb{P}(\mathcal{H}^m_{(d)})$ tal que $V(f)$ es una variedad proyectiva de dimensión $n-m$ (condición que acabamos de ver que es se cumple para casi todo f), el Corolario 2.2.7 implica que

$$\nu_{n-m}[V(f)] = \vartheta_{n-m}\deg(V(f)), \tag{3.6}$$

donde $\deg(V)$ es el grado geométrico de V. Concluimos que

$$\frac{1}{\nu_{\mathbb{P}(\mathcal{H}_{(d)}^m)}[\mathbb{P}(\mathcal{H}_{(d)}^m)]}\int_{f\in\mathbb{P}(\mathcal{H}_{(d)}^m)}\deg(V(f))\;d\mathbb{P}(\mathcal{H}_{(d)}^m)=\mathcal{D}.$$

Por otro lado, la desigualdad de Bézout garantiza que

$$\deg(V(f))\leq\mathcal{D},\qquad\forall f\in\mathbb{P}(\mathcal{H}_{(d)}^m).$$

Deducimos que $\deg(V(f))=\mathcal{D}$ para casi todo $f\in\mathbb{P}(\mathcal{H}_{(d)}^m)$, y el corolario se sigue de la ecuación (3.6). ∎

3.6. Distribuciones de probabilidad de los números de condicionamiento generalizados

En esta sección aplicaremos el Teorema 3.5.7 y los resultados obtenidos en el caso lineal (Capítulo 2) para estudiar la distribución de probabilidad del condicionamiento generalizado $\mu_{\mathrm{norm}}^{(r)}$. Esto es, demostraremos el siguiente resultado técnico a partir del cual se deducirán las estimaciones de probabilidad y las cotas para esperanzas de los teoremas de la Sección 3.1.4.

Teorema 3.6.1 *Sea $(d)=(d_1,\ldots,d_m)$ tal que $d_i>1$ para algún i, $1\leq i\leq m$. Sea $\varepsilon>0$ un número real. Supongamos que $r\geq 2$. Entonces, tenemos:*

$$\frac{1}{\nu_{\mathbb{P}(\mathcal{H}_{(d)}^m)}[\mathbb{P}(\mathcal{H}_{(d)}^m)]}\int_{f\in\mathbb{P}(\mathcal{H}_{(d)}^m)}\nu_{n-m}[\zeta\in V(f):\mu_{\mathrm{norm}}^{(r)}(f,\zeta)>\varepsilon^{-1}]\;d\mathbb{P}(\mathcal{H}_{(d)}^m)\leq$$

$$2\pi e^{1/3}\vartheta_{n-m}\mathcal{D}\left(\sqrt{Nmr(n+1)}\;\varepsilon\right)^{2(m-r+1)(n-r+2)}.$$

Además, si $r=1$, tenemos:

$$\frac{1}{\nu_{\mathbb{P}(\mathcal{H}_{(d)}^m)}[\mathbb{P}(\mathcal{H}_{(d)}^m)]}\int_{f\in\mathbb{P}(\mathcal{H}_{(d)}^m)}\nu_{n-m}[\zeta\in V(f):\mu_{\mathrm{norm}}^{(1)}(f,\zeta)>\varepsilon^{-1}]\;d\mathbb{P}(\mathcal{H}_{(d)}^m)\leq$$

$$\vartheta_{n-m}\mathcal{D}(\sqrt{N}\varepsilon)^{2nm+2m}.$$

Demostración.– Aplicamos el Teorema 3.5.7 a la función $\Phi_\varepsilon:[0,+\infty]\longrightarrow[0,+\infty]$ definida como

$$\Phi_\varepsilon(s):=\begin{cases}1 & si\ s>\varepsilon^{-1}\\ 0 & en\ otro\ caso.\end{cases}$$

Esto es, Φ_ε es la función característica del intervalo $(\varepsilon^{-1},+\infty)$. Obtenemos que

$$\frac{1}{\nu_{\mathbb{P}(\mathcal{H}_{(d)}^m)}[\mathbb{P}(\mathcal{H}_{(d)}^m)]}\int_{f\in\mathbb{P}(\mathcal{H}_{(d)}^m)}\nu_{n-m}[\zeta\in V(f)]:\mu_{\mathrm{norm}}^{(r)}(f,\zeta)>\varepsilon^{-1}]\;d\mathbb{P}(\mathcal{H}_{(d)}^m)$$

$$= 2\pi\vartheta_{N-m-nm}\vartheta_{n-m}\mathcal{D}\int_0^1 (1-t^2)^{N-m-nm}t^{2nm+2m-1}\mathcal{G}_t(\varepsilon)\ dt,$$

donde

$$\mathcal{G}_t(\varepsilon) := \nu_{nm+m-1}[M \in \mathbb{P}(\mathcal{M}_{m\times(n+1)}(\mathbb{C})) : \kappa_D^{(r)}(M) > \varepsilon^{-1}t].$$

Primero, supongamos que $r \geq 2$. Por el Lema 2.5.1, conocemos una cota superior para la cantidad $\mathcal{G}_t(\varepsilon)$. De todas formas, preferimos utilizar la cota proporcionada por el Teorema 2.3.6 sin simplificaciones. Esto es, si denotamos

$$a := nm + m - 1, \quad b := (m - r + 1)(n - r + 2),$$

entonces, tenemos que

$$\mathcal{G}_t(\varepsilon) \leq \vartheta_a C(a,b)\left(\frac{\sqrt{r}\varepsilon}{t}\right)^{2b},$$

donde $C(a,b)$ es la constante del Teorema 2.3.6.
Ahora, tenemos la siguiente acotación (expuesta después del Teorema 2.3.6):

$$C(a,b) \leq 4e^{1/3}\pi\frac{b(a-b)}{a}\binom{a}{b}^2 = 4e^{1/3}\pi\frac{b(a-b)}{a}\frac{\Gamma(a+1)}{\Gamma(b+1)\Gamma(a-b+1)}\binom{a}{b}.$$

Hemos demostrado pues que

$$\frac{1}{\nu_{\mathbb{P}(\mathcal{H}_{(d)}^m)}[\mathbb{P}(\mathcal{H}_{(d)}^m)]}\int_{f\in\mathbb{P}(\mathcal{H}_{(d)}^m)} \nu_{n-m}[\zeta \in V(f)] : \mu_{\text{norm}}^{(r)}(f,\zeta) > \varepsilon^{-1}]\ d\mathbb{P}(\mathcal{H}_{(d)}^m) \leq$$

$$\frac{2}{\nu_{\mathbb{P}(\mathcal{H}_{(d)}^m)}[\mathbb{P}(\mathcal{H}_{(d)}^m)]}\pi\vartheta_{N-a-1}\vartheta_{n-m}\vartheta_a\mathcal{D}\ \times$$

$$4e^{1/3}\pi\frac{\Gamma(a+1)(a-b)}{\Gamma(b+1)\Gamma(a-b+1)}\binom{a-1}{b-1}(\sqrt{r}\varepsilon)^{2b}\ \times$$

$$\int_0^1 (1-t^2)^{N-a-1}t^{2a-2b+1}\ dt.$$

Esta última integral tiene un valor conocido:

$$\frac{1}{2}\frac{\Gamma(N-a)\Gamma(a-b+1)}{\Gamma(N-b+1)}.$$

Sabiendo que para todo número natural $k \geq 0$ se tiene que

$$\vartheta_k = \frac{\pi^k}{\Gamma(k+1)},$$

obtenemos que

$$\frac{1}{\nu_{\mathbb{P}(\mathcal{H}_{(d)}^m)}[\mathbb{P}(\mathcal{H}_{(d)}^m)]}\int_{f\in\mathbb{P}(\mathcal{H}_{(d)}^m)} \nu_{n-m}[\zeta \in V(f)] : \mu_{\text{norm}}^{(r)}(f,\zeta) > \varepsilon^{-1}]\ d\mathbb{P}(\mathcal{H}_{(d)}^m) \leq$$

$$4\pi e^{1/3}\mathcal{D}\vartheta_{n-m}(a-b)\binom{N}{b}\binom{a-1}{b-1}(\sqrt{r}\varepsilon)^{2b} \leq$$

$$4\pi e^{1/3}\mathcal{D}\vartheta_{n-m}(a-b)\frac{N^b}{b!}\frac{a^b}{a(b-1)!}(\sqrt{r}\varepsilon)^{2b} \leq$$

$$2\pi e^{1/3}\mathcal{D}\vartheta_{n-m}(\sqrt{Nar}\varepsilon)^{2b},$$

puesto que $b \geq 2$, $a \geq 5$. Queda demostrado pues el caso de que $r \geq 2$. Supongamos ahora que $r = 1$. Para toda matriz $A \in \mathbb{P}(\mathcal{M}_{m\times(n+1)})$ se tiene que $\kappa_D^{(1)}(A) := 1$. Por tanto, tenemos que:

$$\mathcal{G}_t(\varepsilon) = \nu_{nm+m-1}[\{M \in \mathbb{P}(\mathcal{M}_{m\times(n+1)}(\mathbb{C})) : \kappa_D^{(1)}(M) > \varepsilon^{-1}t\}] =$$

$$\begin{cases} 0 & \text{si } \varepsilon^{-1}t \geq 1 \\ \vartheta_{nm+m-1} & \text{en otro caso .} \end{cases}$$

Por tanto, tenemos que

$$\frac{1}{\nu_{\mathbb{P}(\mathcal{H}_{(d)}^m)}[\mathbb{P}(\mathcal{H}_{(d)}^m)]}\int_{f\in\mathbb{P}(\mathcal{H}_{(d)}^m)} \nu_{n-m}[\zeta \in V(f)] : \mu_{\text{norm}}^{(1)}(f,\zeta) > \varepsilon^{-1}]\; d\mathbb{P}(\mathcal{H}_{(d)}^m)$$

$$= \frac{2\pi\vartheta_{N-m-nm}\vartheta_{n-m}\vartheta_{nm+m-1}\mathcal{D}}{\nu_{\mathbb{P}(\mathcal{H}_{(d)}^m)}[\mathbb{P}(\mathcal{H}_{(d)}^m)]}\int_0^\varepsilon (1-t^2)^{N-m-nm}t^{2nm+2m-1}\; dt \leq$$

$$\frac{2\pi\vartheta_{N-m-nm}\vartheta_{n-m}\vartheta_{nm+m-1}\mathcal{D}}{\nu_{\mathbb{P}(\mathcal{H}_{(d)}^m)}[\mathbb{P}(\mathcal{H}_{(d)}^m)]}\frac{\varepsilon^{2nm+2m}}{2nm+2m} =$$

$$\vartheta_{n-m}\mathcal{D}\binom{N}{nm+m}\varepsilon^{2nm+2m} \leq \vartheta_{n-m}\mathcal{D}(\sqrt{N}\varepsilon)^{2nm+2m},$$

como queríamos. ∎

3.7. Estabilidad en media del conjunto de soluciones

Como hemos indicado en la introducción de este capítulo, la estabilidad del conjunto solución de un sistema de ecuaciones f puede ser medida de dos modos diferentes (ambos con la forma de un número de condicionamiento):

$$\mu_{\text{worst}}^{(m)}(f) := \max_{\zeta\in V(f)} \mu_{\text{norm}}^{(m)}(f,\zeta), \quad \textit{o bien} \quad \mu_{\text{av}}^{(m)}(f) := \mathrm{E}_{\zeta\in V(f)}[\mu_{\text{norm}}^{(m)}(f,\cdot)].$$

Estudiamos a continuación el valor esperable de ambos condicionamientos. Las conclusiones de nuestro estudio son muy significativas, y pueden leerse en la introducción de este capítulo. Durante esta sección, suponemos que $(d) = (d_1, \ldots, d_m)$ es tal que $d_i \geq 2$ para algún i, $1 \leq i \leq m$.

3.7.1. El valor esperable de $\mu_{\mathrm{av}}^{(m)}$

En esta sección demostramos el Teorema 3.1.8. Escribamos la versión técnica de este enunciado:

Teorema 3.7.1 *Sea $m \geq 2$. Entonces, el valor esperable del número de condicionamiento $\mu_{\mathrm{av}}^{(m)}$ satisface:*

$$\mathrm{E}_{\mathbb{P}(\mathcal{H}_{(d)}^m)}[\mu_{\mathrm{av}}^{(m)}] \leq 3m\sqrt{nN}.$$

Además, si $m = 1$, tenemos la igualdad:

$$\mathrm{E}_{\mathbb{P}(\mathcal{H}_{(d)}^m)}[\mu_{\mathrm{av}}^{(1)}] = \frac{\Gamma(N+1)\Gamma(n+1/2)}{\Gamma(N+1/2)\Gamma(n+1)} \sim \sqrt{\frac{N}{n}}.$$

Demostración.– La cantidad $\mathrm{E}_{\mathbb{P}(\mathcal{H}_{(d)}^m)}[\mu_{\mathrm{av}}^m]$ que queremos estimar es la esperanza en el espacio de sistemas $\mathbb{P}(\mathcal{H}_{(d)}^m)$ de la función

$$\mu_{\mathrm{av}}^{(m)}(f) = \frac{1}{\nu_{n-m}[V(f)]} \int_{\zeta \in V(f)} \mu_{\mathrm{norm}}^{(m)}(f,\zeta)\ dV(f).$$

Por tanto, se trata de estimar la siguiente cantidad:

$$\frac{1}{\nu_{\mathbb{P}(\mathcal{H}_{(d)}^m)}[\mathbb{P}(\mathcal{H}_{(d)}^m)]} \int_{f \in \mathbb{P}(\mathcal{H}_{(d)}^m)} \frac{1}{\nu_{n-m}[V(f)]} \int_{\zeta \in V(f)} \mu_{\mathrm{norm}}^{(m)}(f,\zeta)\ dV(f)\ d\mathbb{P}(\mathcal{H}_{(d)}^m).$$

Definimos la siguiente cantidad auxiliar,

$$\mathcal{K}_{(d)} := \int_{f \in \mathbb{P}(\mathcal{H}_{(d)}^m)} \int_{\zeta \in V(f)} \mu_{\mathrm{norm}}^{(m)}(f,\zeta)\ dV(f)\ d\mathbb{P}(\mathcal{H}_{(d)}^m).$$

El Corolario 3.5.8 garantiza que:

$$\mathrm{E}_{\mathbb{P}(\mathcal{H}_{(d)}^m)}[\mu_{\mathrm{av}}^{(m)}] = \frac{\mathcal{K}_{(d)}}{\nu_{\mathbb{P}(\mathcal{H}_{(d)}^m)}[\mathbb{P}(\mathcal{H}_{(d)}^m)]\vartheta_{n-m}\mathcal{D}}.$$

Calculemos pues una estimación para $\mathcal{K}_{(d)}$. Por el Teorema 3.5.7, sabemos que

$$\mathcal{K}_{(d)} = 2\pi\vartheta_{N-m-nm}\vartheta_{n-m}\mathcal{D}\int_0^1 (1-t^2)^{N-m-nm}t^{2nm+2m-2}\ dt$$

multiplicado por

$$\int_{M \in \mathbb{P}(\mathcal{M}_{m\times(n+1)}(\mathbb{C}))} \kappa_D^{(m)}(M)\ d\mathbb{P}(\mathcal{M}_{m\times(n+1)}(\mathbb{C})).$$

Ahora, se tiene que

$$\int_0^1 (1-t^2)^{N-m-nm} t^{2nm+2m-2}\, dt =$$

$$\frac{1}{2}\frac{\Gamma(N-m-nm+1)\Gamma(nm+m-1/2)}{\Gamma(N+1/2)}.$$

Por lo tanto, se tiene la siguiente igualdad:

$$\mathcal{K}_{(d)} = \vartheta_{n-m}\mathcal{D}\pi^N \frac{\Gamma(nm+m-1/2)}{\Gamma(N+1/2)\Gamma(nm+m)} \mathrm{E}_{\mathbf{P}(\mathcal{M}_{m\times(n+1)}(\mathbb{C}))}[\kappa_D^{(m)}].$$

Deducimos que

$$\mathrm{E}_{\mathbf{P}(\mathcal{H}^m_{(d)})}[\mu_{\mathrm{av}}^{(m)}] = \frac{\Gamma(N+1)\Gamma(nm+m-1/2)}{\Gamma(N+1/2)\Gamma(nm+m)} \mathrm{E}_{\mathbf{P}(\mathcal{M}_{m\times(n+1)}(\mathbb{C}))}[\kappa_D^{(m)}].$$

El caso $m = 1$ del teorema se sigue del Corolario 2.5.6, que afirma que

$$\mathrm{E}_{\mathbf{P}(\mathcal{M}_{1\times(n+1)}(\mathbb{C}))}[\kappa_D^{(1)}] = 1.$$

Para el caso $m \geq 2$, también por el Corolario 2.5.6 tenemos que

$$\mathrm{E}_{\mathbf{P}(\mathcal{H}^m_{(d)})}[\mu_{\mathrm{av}}^{(m)}] \leq \frac{\Gamma(N+1)\Gamma(nm+m-1/2)}{\Gamma(N+1/2)\Gamma(nm+m)} 2^{1/4} e \frac{m^{3/2}(n+1)}{n-m+3/2}.$$

Las desigualdades de Gautschi (véase [43, Th. 3] para cotas muy finas) garantizan que, para $x > 0$,

$$\sqrt{x+1/4} \leq \frac{\Gamma(x+1)}{\Gamma(x+1/2)} \leq \sqrt{x+1/\pi}.$$

Por lo tanto, tenemos que

$$\mathrm{E}_{\mathbf{P}(\mathcal{H}^m_{(d)})}[\mu_{\mathrm{av}}^{(m)}] \leq 2^{1/4} e \sqrt{N+1/\pi} \frac{m^{3/2}(n+1)}{(n-m+3/2)\sqrt{nm+m-3/4}}.$$

Algunos cálculos elementales muestran que esta cantidad está acotada superiormente por

$$3m\sqrt{nN},$$

para toda elección posible de $n \geq m \geq 2$. En efecto, tenemos que

$$N \geq (n+1)m - 1 > nm.$$

Por lo tanto, deducimos que

$$\frac{1}{3m\sqrt{nN}}\ 2^{1/4} e \sqrt{N+1/\pi} \frac{m^{3/2}(n+1)}{(n-m+3/2)\sqrt{nm+m-3/4}} \leq$$

$$\frac{2^{1+1/4}e}{9}\sqrt{1+\frac{1}{\pi N}}\sqrt{1+\frac{1}{n}}\sqrt{\frac{mn+m}{nm+m-3/4}} \leq$$

$$\frac{2^{1+1/4}e}{9}\sqrt{1+\frac{1}{4\pi}}\sqrt{1+\frac{1}{2}}\sqrt{\frac{6}{6-3/4}} < 1.$$

Finalmente, hemos demostrado la desigualdad

$$\mathrm{E}_{\mathbf{P}\,(\mathcal{H}^m_{(d)})}[\mu^{(m)}_{\mathrm{av}}] \leq 3m\sqrt{nN},$$

que era el objetivo del teorema. ∎

3.7.2. El valor esperable de $\mu^{(m)}_{\mathrm{worst}}$

En esta sección demostraremos el Teorema 3.1.9. La demostración de dicho enunciado requerirá una serie de resultados previos. Comenzamos con el siguiente

Lema 3.7.2 *Sean $f \in \mathcal{H}^m_{(d)}$, $\zeta \in \mathbb{C}^{n+1}$ tales que $f(\zeta) = 0$, $rank(T_\zeta f) = m$. Entonces, para todo vector $v \in \mathbb{C}^m$, se tiene la siguiente igualdad:*

$$(d_\zeta f)^\dagger v = ((d_\zeta f)\mid_{\zeta^\perp})^\dagger v.$$

Demostración.– Para un operador lineal sobreyectivo entre dos espacios de Hilbert $L : E_1 \longrightarrow E_2$, se tiene que

$$L^\dagger = i \circ (L_{(KerL)^\perp})^{-1},$$

donde i es la inclusión en E_1. Ahora, si f es un sistema de ecuaciones y ζ es una solución de f, se tiene que $d_\zeta f(\zeta) = 0$. Por lo tanto,

$$(d_\zeta f)^\dagger = i \circ ((d_\zeta f)\mid_{(Ker(d_\zeta f))^\perp})^{-1},$$

y

$$((d_\zeta f)\mid_{\zeta^\perp})^\dagger = i \circ ((d_\zeta f)\mid_{(Ker(d_\zeta f))^\perp})^{-1},$$

donde i es la inclusión en $\mathbb{C}^{n+1}$. Esto termina la demostración del lema. ∎

Sea $(f, \zeta) \in W$ un punto de la variedad de incidencia. En la ecuación (1.12) hemos definido la cantidad $\gamma_0(f, \zeta)$, que permite controlar la convergencia del operador de Newton proyectivo (véase Teorema 1.7.2). Además, la definición de $\gamma_0(f, \zeta)$ es independiente de los representantes de f y ζ usados para calcularlo.
El siguiente resultado se demuestra con facilidad a partir del Lema 3.3.1.

Lema 3.7.3 *Sea $(f,\zeta) \in W$ un punto en la variedad de incidencia. Entonces, se tiene la siguiente desigualdad:*

$$\gamma_0(f,\zeta) \leq \frac{d^{3/2}}{2}\mu^{(m)}_{\text{norm}}(f,\zeta).$$

Demostración.– Consideramos dos representantes fijados de f y ζ tales que $\|f\|_\Delta = \|\zeta\|_2 = 1$. Podemos escribir

$$\gamma_0(f,\zeta) = \max_{k>1} \left\| (T_\zeta f)^\dagger \Delta(d)^{1/2} \Delta(d)^{-1/2} \frac{d^{(k)}_\zeta f}{k!} \right\|_2^{1/(k-1)} \leq$$

$$\max_{k>1} \|(T_\zeta f)^\dagger \Delta(d)^{1/2}\|_2^{1/(k-1)} \left(\frac{\|\Delta(d)^{-1/2} d^{(k)}_\zeta f\|_2}{k!} \right)^{1/(k-1)}.$$

Por el Lema 3.3.1, obtenemos que

$$\gamma_0(f,\zeta) \leq \frac{d^{3/2}}{2} \max_{k>1} \|(T_\zeta f)^\dagger \Delta(d)^{1/2}\|_2^{1/(k-1)} =$$

$$\frac{d^{3/2}}{2} \|(T_\zeta f)^\dagger \Delta(d)^{1/2}\|_2 = \frac{d^{3/2}}{2}\mu^{(m)}_{\text{norm}}(f,\zeta),$$

como queríamos. La última igualdad de este razonamiento se sigue de la identidad (1.10) ∎

El siguiente resultado se sigue de los argumentos de Shub & Smale y Dedieu en [112, 34].

Proposición 3.7.4 *Sea $f \in \mathbb{P}(\mathcal{H}^m_{(d)})$, $\zeta \in V(f)$ tales que $\mu^{(m)}_{\text{norm}}(f,\zeta) < \infty$. Sea $\zeta' \in V(f)$ otra solución de f, tal que*

$$u := d_{\mathbb{P}}(\zeta',\zeta)\mu^{(m)}_{\text{norm}}(f,\zeta)\frac{\sqrt{2}d^{3/2}}{2} < 1 - \sqrt{2}/2.$$

Entonces, se tiene la siguiente desigualdad:

$$\mu^{(m)}_{\text{norm}}(f,\zeta') \leq \frac{(1-u)^2}{2u^2 - 4u + 1}\mu^{(m)}_{\text{norm}}(f,\zeta).$$

Demostración.– Denotamos por $\underline{f}, \underline{\zeta}, \underline{\zeta}'$ algunos representantes fijados de f, ζ, ζ', tales que $\|\underline{f}\|_\Delta = \|\underline{\zeta}\|_2 = \|\underline{\zeta}'\|_2 = 1$. Además, podemos elegirlos de tal modo que

$$\langle \underline{\zeta}, \underline{\zeta}' \rangle_2 \in \mathbb{R}^{0,+}.$$

Observamos que $T_{\underline{\zeta}}\underline{f}(T_{\underline{\zeta}}\underline{f})^\dagger$ es la identidad en $\underline{\zeta}^\perp$. Por lo tanto,

$$\mu^{(m)}_{\text{norm}}(\underline{f},\zeta') = \|(T_{\underline{\zeta}'}\underline{f})^\dagger \Delta(d)^{1/2}\|_2 \leq$$

$$\|(T_{\underline{\zeta}'}\underline{f})^\dagger T_{\underline{\zeta}}\underline{f}\|_2 \|(T_{\underline{\zeta}}\underline{f})^\dagger \Delta(d)^{1/2}\|_2 = \|(T_{\underline{\zeta}'}\underline{f})^\dagger T_{\underline{\zeta}}\underline{f}\|_2 \mu^{(m)}_{\text{norm}}(\underline{f},\underline{\zeta}).$$

Concluimos que basta con demostrar que en las condiciones del lema, se tiene la siguiente desigualdad:

$$\|(T_{\underline{\zeta}'}\underline{f})^\dagger T_{\underline{\zeta}}\underline{f}\|_2 \leq \frac{(1-u)^2}{2u^2-4u+1}.$$

Ahora, por el Lema 3.7.2,

$$\|(T_{\underline{\zeta}'}\underline{f})^\dagger T_{\underline{\zeta}}\underline{f}\|_2 = \|((d_{\underline{\zeta}'}\underline{f})\mid_{(\underline{\zeta}')^\perp})^\dagger d_{\underline{\zeta}}\underline{f}\mid_{(\underline{\zeta})^\perp}\|_2 = \|(d_{\underline{\zeta}'}\underline{f})^\dagger d_{\underline{\zeta}}\underline{f}\|_2.$$

Sea $\gamma(\underline{f},\underline{\zeta})$ la siguiente cantidad, que es muy similar a la introducida en la Sección 1.8:

$$\gamma(\underline{f},\underline{\zeta}) := \max_{k>1} \left\| (d_{\underline{\zeta}}\underline{f})^\dagger \frac{d^{(k)}_{\underline{\zeta}}\underline{f}}{k!} \right\|_2^{1/(k-1)},$$

si $d_{\underline{\zeta}}\underline{f}$ es sobreyectiva. Por el Lema 3.7.2, tenemos que

$$\gamma(\underline{f},\underline{\zeta}) = \gamma_0(f,\zeta).$$

Por lo tanto, por el Lema 3.7.3 deducimos:

$$\gamma(\underline{f},\underline{\zeta}) \leq \mu^{(m)}_{\text{norm}}(f,\zeta)\frac{d^{3/2}}{2}.$$

Por otro lado, se tiene la siguiente cadena de desigualdades:

$$\|\underline{\zeta}-\underline{\zeta}'\|_2 = \sqrt{2}(1-\langle\underline{\zeta},\underline{\zeta}'\rangle_2)^{1/2} = \sqrt{2}(1-\sqrt{1-d_{\mathbf{P}}(\zeta,\zeta')^2})^{1/2} \leq$$

$$\sqrt{2}d_{\mathbf{P}}(\zeta,\zeta').$$

Deducimos que

$$\|\underline{\zeta}-\underline{\zeta}'\|_2\gamma(\underline{f},\underline{\zeta}) \leq \sqrt{2}d_{\mathbf{P}}(\zeta,\zeta')\mu^{(m)}_{\text{norm}}(f,\zeta)\frac{d^{3/2}}{2} = u.$$

Finalmente, por [116, pg. 20] o [34, Lem. 127] sabemos que esto implica que

$$\|(d_{\underline{\zeta}'}\underline{f})^\dagger d_{\underline{\zeta}}\underline{f}\|_2 \leq \frac{(1-u)^2}{2u^2-4u+1},$$

y el lema queda demostrado.

■

Corolario 3.7.5 *Sean $\varepsilon > 0, s > 1$ dos números reales positivos. Sean $f \in \mathbb{P}(\mathcal{H}^m_{(d)})$, $\zeta' \in V(f)$ tales que $\varepsilon^{-1} < \mu^{(m)}_{\text{norm}}(f, \zeta') < +\infty$. Sea $\zeta \in V(f)$ otra solución de f, tal que*

$$d_{\mathbb{P}}(\zeta', \zeta) \leq \frac{\sqrt{2}\varepsilon}{d^{3/2}} s \left(1 - \sqrt{\frac{s}{2s-1}}\right).$$

Entonces, tenemos la siguiente desigualdad:

$$\mu^{(m)}_{\text{norm}}(f, \zeta) > \frac{1}{s\varepsilon}.$$

Demostración.– Procedemos por reducción al absurdo. Supongamos que

$$\mu^{(m)}_{\text{norm}}(f, \zeta) \leq \frac{1}{s\varepsilon}.$$

Entonces, se tiene:

$$u := d_{\mathbb{P}}(\zeta', \zeta)\mu^{(m)}_{\text{norm}}(f, \zeta)\frac{\sqrt{2}d^{3/2}}{2} \leq \frac{\sqrt{2}\varepsilon}{d^{3/2}} s \left(1 - \sqrt{\frac{s}{2s-1}}\right) \frac{1}{s\varepsilon} \frac{\sqrt{2}d^{3/2}}{2} =$$

$$\left(1 - \sqrt{\frac{s}{2s-1}}\right) < 1 - \frac{\sqrt{2}}{2}$$

Por lo tanto, por la Proposición 3.7.4, tenemos que

$$\mu^{(m)}_{\text{norm}}(f, \zeta') \leq \frac{(1-u)^2}{2u^2 - 4u + 1}\mu^{(m)}_{\text{norm}}(f, \zeta) \leq$$

$$\frac{\left(1 - \left(1 - \sqrt{\frac{s}{2s-1}}\right)\right)^2}{2\left(1 - \sqrt{\frac{s}{2s-1}}\right)^2 - 4\left(1 - \sqrt{\frac{s}{2s-1}}\right) + 1}\mu^{(m)}_{\text{norm}}(f, \zeta) \leq$$

$$\frac{\frac{s}{2s-1}}{\frac{1}{2s-1}} \frac{1}{s\varepsilon} = \frac{1}{\varepsilon},$$

lo que es falso por hipótesis. ∎

El siguiente resultado es una cota superior para la distribución de probabilidad de $\mu^{(m)}_{\text{worst}}$ en $\mathbb{P}(\mathcal{H}^m_{(d)})$.

Teorema 3.7.6 *Sea $0 < \varepsilon < d^{3/2}$ un número positivo cualquiera, y supongamos que $m < n$. Entonces, para un sistema elegido al azar $f \in \mathbb{P}(\mathcal{H}^m_{(d)})$, la probabilidad de que $\mu^{(m)}_{\text{worst}}(f) > \varepsilon^{-1}$ es a lo sumo*

$$2\mathcal{D}\left[10N^{1/2}mn^{1/2}d^{3/2}\right]^{2(n-m)} [6N^{1/2}mn^{1/2}\ \varepsilon]^4.$$

Demostración.– Sea $T_\varepsilon \subseteq \mathbb{P}(\mathcal{H}^m_{(d)})$ el conjunto definido como sigue:

$$T_\varepsilon := \{f \in \mathbb{P}(\mathcal{H}^m_{(d)}) : \exists \zeta \in V(f), \mu^{(m)}_{\text{norm}}(f,\zeta) > \varepsilon^{-1}\}.$$

La probabilidad que queremos estimar es igual a

$$\frac{\nu_{\mathbb{P}(\mathcal{H}^m_{(d)})}[T_\varepsilon]}{\nu_{\mathbb{P}(\mathcal{H}^m_{(d)})}[\mathbb{P}(\mathcal{H}^m_{(d)})]} = \frac{1}{\nu_{\mathbb{P}(\mathcal{H}^m_{(d)})}[\mathbb{P}(\mathcal{H}^m_{(d)})]} \int_{f\in T_\varepsilon} 1 d\mathbb{P}(\mathcal{H}^m_{(d)}).$$

Para todo real positivo $s > 1$, definimos la siguiente cantidad:

$$MIN_{\varepsilon,s} := \min_{f\in T_\varepsilon} \nu_{n-m}[\zeta \in V(f) : \mu^{(m)}_{\text{norm}}(f,\zeta) > 1/(s\varepsilon)].$$

Demostraremos que $MIN_{\varepsilon,s}$ es un número positivo para $s > 1$. Por lo tanto, se tiene que

$$\frac{\nu_{\mathbb{P}(\mathcal{H}^m_{(d)})}[T_\varepsilon]}{\nu_{\mathbb{P}(\mathcal{H}^m_{(d)})}[\mathbb{P}(\mathcal{H}^m_{(d)})]} \leq$$

$$\frac{1}{\nu_{\mathbb{P}(\mathcal{H}^m_{(d)})}[\mathbb{P}(\mathcal{H}^m_{(d)})] MIN_{\varepsilon,s}} \int_{f\in T_\varepsilon} \nu_{n-m}[\zeta \in V(f) : \mu^{(m)}_{\text{norm}}(f,\zeta) > 1/(s\varepsilon)] d\mathbb{P}(\mathcal{H}^m_{(d)}) \leq$$

$$\frac{1}{\nu_{\mathbb{P}(\mathcal{H}^m_{(d)})}[\mathbb{P}(\mathcal{H}^m_{(d)})] MIN_{\varepsilon,s}} \int_{f\in \mathbb{P}(\mathcal{H}^m_{(d)})} \nu_{n-m}[\zeta \in V(f) : \mu^{(m)}_{\text{norm}}(f,\zeta) > 1/(s\varepsilon)] d\mathbb{P}(\mathcal{H}^m_{(d)}).$$

Por el Teorema 3.6.1, tenemos que

$$\frac{1}{\nu_{\mathbb{P}(\mathcal{H}^m_{(d)})}[\mathbb{P}(\mathcal{H}^m_{(d)})]} \int_{f\in \mathbb{P}(\mathcal{H}^m_{(d)})} \nu_{n-m}[\zeta \in V(f) : \mu^{(m)}_{\text{norm}}(f,\zeta) > 1/(s\varepsilon)] d\mathbb{P}(\mathcal{H}^m_{(d)}) \leq$$

$$2\pi e^{1/3} \vartheta_{n-m} \mathcal{D}[sN^{1/2} m(n+1)^{1/2}\ \varepsilon]^{2(n-m+2)} \leq$$
$$\vartheta_{n-m} \mathcal{D}[seN^{1/2} mn^{1/2}\ \varepsilon]^{2(n-m+2)}.$$

Concluimos la siguiente desigualdad:

$$\frac{\nu_{\mathbb{P}(\mathcal{H}^m_{(d)})}[T_\varepsilon]}{\nu_{\mathbb{P}(\mathcal{H}^m_{(d)})}[\mathbb{P}(\mathcal{H}^m_{(d)})]} \leq \frac{\vartheta_{n-m} \mathcal{D}[seN^{1/2} mn^{1/2}\ \varepsilon]^{2(n-m+2)}}{MIN_{\varepsilon,s}},$$

para todo real positivo $s > 1$. Ahora, podemos dar una cota inferior para $MIN_{\varepsilon,s}$. En efecto, sea $f \in T_\varepsilon$ un sistema, y sea $\zeta' \in V(f)$ una solución tal que $\mu^{(m)}_{\text{norm}}(f,\zeta') > \varepsilon^{-1}$. Podemos asumir que todo punto de $V(f)$ es regular, pues el conjunto de sistemas que no satisfacen esa condición tiene medida cero en $\mathbb{P}(\mathcal{H}^m_{(d)})$ y no afecta para propósitos de integración. Entonces, por el Corolario 3.7.5, tenemos:

$$\nu_{n-m}[\zeta \in V(f) : \mu^{(m)}_{\text{norm}}(f,\zeta) > 1/(s\varepsilon)] \geq$$

$$\nu_{n-m}\left[V(f)\cap B_{\mathbb{P}}\left(\zeta',\frac{\sqrt{2}\varepsilon}{d^{3/2}}s\left(1-\sqrt{\frac{s}{2s-1}}\right)\right)\right],$$

donde $B_{\mathbb{P}}(x,\lambda)$ es la bola en $\mathbb{P}_n(\mathbb{C})$ centrada en x de radio λ, para la distancia proyectiva $d_{\mathbb{P}}$. Además, $V(f)$ es una variedad proyectiva de dimensión $n-m$. Por el Teorema 2.3.2 podemos dar una cota inferior para esta cantidad:

$$\nu_{n-m}\left[V(f)\cap B_{\mathbb{P}}\left(\zeta',\frac{\sqrt{2}\varepsilon}{d^{3/2}}s\left(1-\sqrt{\frac{s}{2s-1}}\right)\right)\right]\geq$$

$$\frac{1}{2}\vartheta_{n-m}\left(\frac{\sqrt{2}\varepsilon}{d^{3/2}}s\left(1-\sqrt{\frac{s}{2s-1}}\right)\right)^{2(n-m)},$$

siempre que se satisfaga la siguiente desigualdad:

$$\frac{\sqrt{2}\varepsilon}{d^{3/2}}s\left(1-\sqrt{\frac{s}{2s-1}}\right)\leq\frac{\sqrt{2}}{2}. \tag{3.7}$$

Deducimos que en ese caso,

$$MIN_{\varepsilon,s}\geq\frac{1}{2}\vartheta_{n-m}\left(\frac{\sqrt{2}\varepsilon}{d^{3/2}}s\left(1-\sqrt{\frac{s}{2s-1}}\right)\right)^{2(n-m)}.$$

Finalmente, esto implica la siguiente desigualdad:

$$\frac{\nu_{\mathbb{P}(\mathcal{H}^m_{(d)})}[T_\varepsilon]}{\nu_{\mathbb{P}(\mathcal{H}^m_{(d)})}[\mathbb{P}(\mathcal{H}^m_{(d)})]}\leq 2\mathcal{D}\left[\frac{e}{\sqrt{2}}N^{1/2}mn^{1/2}d^{3/2}\right]^{2(n-m)}\frac{s^4[eN^{1/2}mn^{1/2}\ \varepsilon]^4}{\left(1-\sqrt{\frac{s}{2s-1}}\right)^{2(n-m)}},$$

que se satisface para todo número $s>1$, siempre que se tenga la condición (3.7). Sea $s:=\frac{6}{e}>1$ ese número. Entonces, tenemos que

$$\frac{\nu_{\mathbb{P}(\mathcal{H}^m_{(d)})}[T_\varepsilon]}{\nu_{\mathbb{P}(\mathcal{H}^m_{(d)})}[\mathbb{P}(\mathcal{H}^m_{(d)})]}\leq 2\mathcal{D}\left[\frac{e}{\sqrt{2}}N^{1/2}mn^{1/2}d^{3/2}\right]^{2(n-m)}\frac{[6N^{1/2}mn^{1/2}\ \varepsilon]^4}{\left(1-\sqrt{\frac{6/e}{12/e-1}}\right)^{2(n-m)}},$$

y el teorema se sigue del hecho de que

$$\frac{e}{\sqrt{2}\left(1-\sqrt{\frac{6/e}{12/e-1}}\right)}\leq 10.$$

Hemos impuesto la condición (3.7). Algunos cálculos elementales muestran que basta con

$$\varepsilon\leq d^{3/2}.$$

■

Finalmente, obtenemos la cota para la esperanza buscada, demostrando el Teorema 3.1.9.

Demostración del Teorema 3.1.9 Primero, supongamos que $n > m$. Sea $t > 1/d^{3/2}$ un número positivo. Por el Teorema 3.7.6 tenemos que

$$Prob[f \in \mathbb{P}(\mathcal{H}^m_{(d)}) : \mu^{(m)}_{\text{worst}}(f) > t] = Prob[f \in \mathbb{P}(\mathcal{H}^m_{(d)}) : \mu^{(m)}_{\text{worst}}(f) > \frac{1}{1/t}] \leq$$

$$2\mathcal{D}\left[10N^{1/2}mn^{1/2}d^{3/2}\right]^{2(n-m)}[6N^{1/2}mn^{1/2}]^4\frac{1}{t^4}.$$

Por el Lema 2.5.4, concluimos que:

$$\mathrm{E}_{\mathbb{P}(\mathcal{H}^m_{(d)})}[\mu^{(m)}_{\text{worst}}] \leq \frac{4}{3}(2\mathcal{D})^{1/4}\left[10N^{1/2}mn^{1/2}d^{3/2}\right]^{\frac{n-m}{2}} 6N^{1/2}mn^{1/2}.$$

Ahora, observamos que

$$\frac{4}{3}2^{1/4}6 \leq 10,$$

y el teorema se sigue en el caso que $m < n$.
Finalmente, supongamos que $m = n$ (el caso estudiado por Shub & Smale en [113]). En este caso, tenemos que

$$Prob[f \in \mathbb{P}(\mathcal{H}^m_{(d)}) : \mu^{(m)}_{\text{worst}}(f) > t] = Prob[f \in \mathbb{P}(\mathcal{H}^m_{(d)}) : \mu^{(m)}_{\text{worst}}(f) > \frac{1}{1/t}] \leq$$

$$\frac{1}{\nu_{\mathbb{P}(\mathcal{H}^m_{(d)})}[\mathbb{P}(\mathcal{H}^m_{(d)})]}\int_{f\in\mathbb{P}(\mathcal{H}^m_{(d)})} \sharp[\zeta \in V(f) : \mu^{(m)}_{\text{norm}}(f,\zeta) > \frac{1}{1/t}]\; d\mathbb{P}(\mathcal{H}^m_{(d)}).$$

Por el Teorema 3.6.1, esta última cantidad es a lo sumo

$$\vartheta_0\mathcal{D}\left(en\sqrt{nN}\;\frac{1}{t}\right)^4 = \mathcal{D}\left(en\sqrt{nN}\;\frac{1}{t}\right)^4.$$

Por el Lema 2.5.4, esto implica que

$$\mathrm{E}_{\mathbb{P}(\mathcal{H}^m_{(d)})}[\mu^{(m)}_{\text{worst}}] \leq \frac{4}{3}\mathcal{D}^{1/4}en\sqrt{nN}.$$

En particular, el teorema es cierto pues $\frac{4}{3}e \leq 10$. ∎

3.8. El valor esperable del radio de convergencia del operador de Newton en dimensión positiva

Exponemos a continuación la demostración del Teorema 3.1.10. El primero de los items es consecuencia del segundo, pues si la función $\gamma_{\text{worst}} \geq 0$ alcanzase el infinito en un conjunto de medida no nula necesariamente la

esperanza sería infinita. Nos centramos pues en los dos últimos items del teorema. Como hemos indicado en la Sección 1.4, ecuación (1.4), tenemos:

$$\mathrm{E}_{\mathcal{H}^m_{(d)}}[\gamma_{\mathrm{worst}}] \equiv \mathrm{E}_{\mathbb{P}(\mathcal{H}^m_{(d)})}[\gamma_{\mathrm{worst}}].$$

Por la Proposición 3.1.2, se tiene que

$$\gamma_{\mathrm{worst}} \leq \frac{d^{3/2}}{2}\mu^m_{\mathrm{worst}}(f).$$

Deducimos que

$$\mathrm{E}_{\mathcal{H}^m_{(d)}}[\gamma_{\mathrm{worst}}] \leq \frac{d^{3/2}}{2}\mathrm{E}_{\mathbb{P}(\mathcal{H}^m_{(d)})}[\mu^m_{\mathrm{worst}}],$$

y el segundo item se sigue del Teorema 3.1.9. El tercer item se sigue de forma inmediata de la desigualdad de Jensen, esto es,

$$\mathrm{E}_X\left[\frac{1}{\phi}\right] \geq \frac{1}{\mathrm{E}_X[\phi]},$$

donde X es un espacio de probabilidad y ϕ una variable aleatoria. Como, en nuestro caso, por el Teorema 1.8.2, el radio de convergencia es $\frac{u_0}{\gamma_{\mathrm{worst}}(f)}$, hemos terminado la demostración. ■

Capítulo 4

Una Solución Probabilista al Problema 17 de Smale

4.1. Introducción.

Durante la primera mitad de los años 90, M. Shub y S. Smale sentaron las bases para una nueva concepción del análisis numérico. Se centraron en el problema de la resolución numérica de sistemas de ecuaciones polinomiales en la serie de artículos [112, 113, 114, 115, 116]. El trabajo realizado por Shub & Smale quedó, no obstante, incompleto. En efecto, el problema esencial de encontrar una solución aproximada de un sistema de ecuaciones no pudo ser resuelto en su totalidad, ya que si bien se abría la puerta por primera vez a la existencia de un algoritmo que realizase esa tarea, se trataba en todo caso de una demostración no–constructiva que no permitía trasladar a la práctica los resultados obtenidos. En vista de ello, Stephen Smale propuso en su lista de problemas matemáticos para el siglo XXI [119] superar esa barrera y encontrar un algoritmo descrito explícitamente que funcionase, en media, en tiempo polinomial en el tamaño del input.
En el presente capítulo completamos parte del programa iniciado en la serie de trabajos [112, 113, 114, 115, 116]. Los resultados que aquí exponemos están esencialmente contenidos en el artículo [11]. Como en [115], el espacio de inputs es el espacio de los sistemas de ecuaciones homogéneas cero–dimensionales $\mathcal{H}_{(d)} := \mathcal{H}^n_{(d)}$, con la estructura de Kostlan definida en la Sección 1.4. Como hemos venido haciendo en capítulos anteriores, denotaremos por $d := \max\{d_i\}$ el máximo de los grados de los polinomios, y por $N+1$ la dimensión compleja de $\mathcal{H}_{(d)}$ como espacio vectorial.
El problema 17 de Smale es enunciado, en sus propias palabras, como sigue:

Can a zero of n complex polynomial equations in n unknowns be found approximately, on the average, in polynomial time with a uniform algorithm?

En otras palabras,

Existe un algoritmo que calcule ceros aproximados de sistemas de ecuaciones polinomiales en tiempo polinomial en promedio?

Smale pide por tanto un resultado de naturaleza probabilista. La palabra "uniform" referida al término algoritmo, se refiere simplemente que el algoritmo encontrado debe ser descrito explícitamente, no mediante argumentos no–constructivos como en [115].
El siguiente teorema es el resultado principal de este capítulo. Supone una respuesta probabilística positiva al Problema 17 de Smale.

Teorema 4.1.1 *Existe un algoritmo probabilístico, con probabilidad de fracaso acotada, verificando las siguientes propiedades:*

- *El tiempo de ejecución es polinomial en n, N, d.*
- *Para cada input $f \in \mathcal{H}_{(d)}$, el algoritmo calcula o bien un cero aproximado proyectivo de algún cero (exacto) de f, o bien* failure.
- *Sea $\mathcal{F}$ el conjunto de los inputs $f \in \mathcal{H}_{(d)}$, tales que el algoritmo calcula un cero aproximado proyectivo de f. Entonces, la probabilidad de que un input al azar f esté en $\mathcal{F}$ es mayor que*

$$1 - \frac{1}{N}.$$

4.1.1. Resultados principales.

Queremos por tanto un algoritmo con la siguiente estructura:

Input: Un sistema de ecuaciones homogéneas $f \in \mathcal{H}_{(d)}$.

Output: Un cero aproximado proyectivo $z \in \mathbb{P}_n(\mathbb{C})$ de f asociado con algún cero exacto $\zeta \in V(f)$ (en el sentido de la Sección 1.7.1).

Esta clase de algoritmos no se diseñan, en principio, para resolver todos los sistemas $f \in \mathcal{H}_{(d)}$, sino una larga clase de ellos. En principio, por ejemplo, el procedimiento que propondremos no pretende resolver sistemas singulares.
A lo largo de este capítulo, denotaremos simplemente por $\Sigma_{(d)} \subseteq \mathcal{H}_{(d)}$ (o indistintamente $\Sigma_{(d)} \subseteq \mathbb{P}(\mathcal{H}_{(d)})$) la clase de los sistemas f tales que $V(f)$ contiene una solución singular, que ha sido denotada en ocasiones anteriores por $\Sigma_{(d)}^{n-1}$. Como acabamos de indicar, las páginas que siguen no pretenden resolver sistemas en $\Sigma_{(d)}$.
El esquema básico que seguiremos es el algoritmo de deformación homotópica de Newton como ha sido descrito en [116, 115]: Dados dos sistemas $f, g \in \mathcal{H}_{(d)} \setminus \Sigma_{(d)}$, consideramos el "segmento" de sistemas entre f y g,

$$\Gamma := \{f_t := (1-t)g + tf, t \in [0,1]\}. \tag{4.1}$$

Si $\Gamma \cap \Sigma_{(d)} = \emptyset$, entonces existen curvas reales formadas por pares de solución y sistema asociadas con ese segmento:

$$C_i(\Gamma) := \{(f_t, \zeta_t) : \zeta_t \in V(f_t), t \in [0,1]\}, \qquad 1 \leq i \leq \mathcal{D} := \prod_{i=1}^{n} d_i.$$

El operador de Newton puede ser utilizado para construir una poligonal que sigue muy de cerca esas curvas de soluciones $C_i(\Gamma)$ en la variedad de incidencia. El procedimiento así descrito trata de producir un cero aproximado proyectivo z_1 de f (esto es, en $t = 1$), a partir de un cero aproximado o exacto z_0 de g (esto es, en $t = 0$). Podemos formalizar estas ideas como sigue.

Definición 4.1.2 *Un esquema de deformación homotópica de Newton (que abreviaremos como* NHD*) con par inicial* $(g, z_0) \in \mathcal{H}_{(d)} \times \mathbb{P}_n(\mathbb{C})$ *y función de recursos* $\varphi : \mathcal{H}_{(d)} \times \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ *es un esquema algorítmico basado en la siguiente estrategia:*

Input: $f \in \mathcal{H}_{(d)}$, $\varepsilon \in \mathbb{R}^+$.

- *Realizar* $\varphi(f, \varepsilon)$ *"pasos de homotopía" siguiendo el segmento* $(1-t)g + tf$, $t \in [0,1]$, *comenzando en* (g, z_0), *donde* z_0 *es un cero aproximado proyectivo de* g *asociado con algún cero* $\zeta_0 \in V(g)$.

Output:

O bien failure*, o bien*
un cero aproximado proyectivo $z_1 \in \mathbb{P}_n(\mathbb{C})$ *de* f.

Un algoritmo basado en NHD es un algoritmo que construye una poligonal con $\varphi(f, \varepsilon)$ vértices, de modo que el vértice inicial es (g, z_0) y el vértice final es el punto (f, z_1) para algún $z_1 \in \mathbb{P}_n(\mathbb{C})$, siendo z_1 el output del algoritmo. La poligonal es construida por "pasos de homotopía" que van de un vértice al siguiente. Por lo tanto, $\varphi(f, \varepsilon)$ es el número de pasos de homotopía realizados por el algoritmo. Hay distintos modos de realizar esos pasos de homotopía, entre los cuales se encuentra el operador proyectivo de Newton, como ha sido descrito en [110, 112, 88].
El número real ε se utiliza normalmente para controlar el número de pasos (a través de la función $\varphi(f, \varepsilon)$) y la probabilidad de fracaso del algoritmo (esto es, la probabilidad de que un input $f \in \mathcal{H}_{(d)}$ no sea resuelto en $\varphi(f, \varepsilon)$ pasos con par inicial (g, z_0)).
Una buena elección del par inicial (g, z_0) debe garantizar que el número de pasos $\varphi(f, \varepsilon)$ será razonablemente pequeño con alta probabilidad. Estableceremos este concepto en la siguiente definición, inspirada en el paradigma de algoritmo numérico que hemos esbozado en la Introducción de esta memoria. Las construcciones que se hacen a continuación pueden hacerse para

cualquier función polinomial $p \in \mathbb{R}[T_1, T_2, T_3, T_4]$. Para evitar arrastrar notaciones engorrosas, a partir de ahora consideraremos ese polinomio p fijado, como sigue:

$$\forall\ n, N, d \geq 1, \varepsilon > 0,\ \ p\ (n, N, d, \varepsilon^{-1}) := 10^8 n^5 N^3 d^4 \varepsilon^{-2}. \tag{4.2}$$

Definición 4.1.3 *Sea $\varepsilon > 0$ un número real positivo. Decimos que un par inicial $(g, z_0) \in \mathcal{H}_{(d)} \times \mathbb{P}_n(\mathbb{C})$ es ε–eficiente para* NHD *si el esquema* NHD *con par inicial (g, z_0) y función de recursos*

$$\varphi(f, \varepsilon) := 10^8 n^5 N^3 d^4 \varepsilon^{-2},\ \ \forall\ f \in \mathbb{S}_\Delta,\ \varepsilon > 0$$

satisface la siguiente propiedad:

$$\mathrm{Prob}_{f \in \mathbb{S}_\Delta}[f \in \mathbb{S}_\Delta :\ \text{NHD } \textit{encuentra un cero aproximado proyectivo de } f\,]$$

$$\geq 1 - \varepsilon.$$

El resultado principal del artículo de Shub & Smale [115] demuestra que para cada $\varepsilon > 0$, existe un par $(g_\varepsilon, \zeta_\varepsilon) \in \mathcal{H}_{(d)} \times \mathbb{P}_n(\mathbb{C})$ que es ε–eficiente (véase el Teorema 0.0.1). Este resultado supone un avance espectacular en la historia de la eficiencia de resolución de sistemas polinomiales. Conduce al siguiente procedimiento basado en NHD:

Input: $f \in \mathcal{H}_{(d)}$, $\varepsilon \in \mathbb{R}^+$.

- *Computar $(g_\varepsilon, \zeta_\varepsilon)$ (el par ε–eficiente cuya existencia garantiza [115]).*
- *Realizar un número polinomial (en $\varepsilon^{-1}, n, N, d$) de pasos de homotopía, siguiendo el segmento $(1-t)g + tf$, $t \in [0, 1]$, empezando en $(g_\varepsilon, \zeta_\varepsilon)$.*

Output:

O bien *failure*, o bien
un cero aproximado proyectivo $z \in \mathbb{P}_n(\mathbb{C})$ de f.

Este procedimiento parece dar la respuesta deseada, puesto que puede calcular ceros aproximados proyectivos de la mayor parte de sistemas de ecuaciones (con probabilidad $1 - \varepsilon$). Sin embargo, tiene tres puntos débiles. Primero, los autores de [115] demuestran la existencia de algún par $(g_\varepsilon, \zeta_\varepsilon)$, pero sin obtener detalle alguno sobre cómo construirlo. Evidentemente, sin un modo de calcular este par inicial, el esquema que acabamos de escribir no puede ser llevado a la práctica. De hecho, no podemos garantizar que sea un algoritmo, pues ni siquiera sabemos si g_ε es calculable o si ζ_ε es computable. La falta de pares ε–eficientes conducen a la conjetura de Shub & Smale (como se escribe en [115]) y al Problema 17 de Smale.

Un segundo aspecto que desearíamos mejorar es la dependencia del punto inicial $(g_\varepsilon, \zeta_\varepsilon)$ del valor ε. En tercer lugar, observamos que no bastaría con calcular el sistema inicial g_ε, sino que nos hace falta conocer la solución particular ζ_ε, que debe ser un cero aproximado proyectivo de g_ε. De hecho, Shub & Smale en [115] asumían que ζ_ε es un cero exacto de g_ε, lo que añade el problema de calcularlo, pues esto no es garantizado ni siquiera a partir del conocimiento de un cero aproximado proyectivo.
Por lo tanto, cualquier algoritmo basado en esta NHD requiere la resolución "a priori" de dos tareas de considerable dificultad: Calcular un sistema de ecuaciones g_ε tal que alguno de sus ceros ζ_ε proporciona un par ε–eficiente, y resolver el sistema g_ε para calcular la solución "exacta" ζ_ε, sin que quede claro de qué naturaleza es esa solución.
El cálculo de soluciones "exactas" debe ser descartado para obtener un algoritmo razonable. Procedemos por lo tanto del modo opuesto al que acabamos de sugerir: Elegimos a priori un punto complejo $\zeta_\varepsilon \in \mathbb{P}_n(\mathbb{C})$, para el que deberíamos conocer la existencia de un sistema g_ε tal que $(g_\varepsilon, \zeta_\varepsilon)$ es un par ε–eficiente. La existencia de ese sistema es un hecho que se sigue fácilmente de los argumentos en [115]. Pero, una vez más, esa demostración no–constructiva no permite garantizar la existencia de un algoritmo ni tampoco describirlo explícitamente.
En este capítulo mostraremos una solución a estos problemas. El método que proponemos es probabilístico y no asume ningún conocimiento a priori, y es capaz de resolver la mayor parte de los sistemas de ecuaciones polinomiales. Comenzamos con la siguiente noción.

Definición 4.1.4 *Una clase $\mathcal{G} \subseteq \mathcal{H}_{(d)} \times \mathbb{P}_n(\mathbb{C})$ se llama conjunto questor para pares eficientes si para todo número real $\varepsilon > 0$, la probabilidad de que un par elegido al azar $(g, \zeta) \in \mathcal{G}$ sea ε–eficiente es mayor o igual que*

$$1 - \varepsilon.$$

Entonces, demostraremos el siguiente resultado.

Teorema 4.1.5 *Para toda lista de grados $(d) = (d_1, \ldots, d_n)$, existe un conjunto questor para pares eficientes, $\mathcal{G}_{(d)}$, que resuelve la mayor parte de los sistemas en $\mathcal{H}_{(d)}$ en tiempo polinomial en el tamaño del input N.*

La existencia de un conjunto questor $\mathcal{G}_{(d)} \subseteq \mathcal{H}_{(d)} \times \mathbb{P}_n(\mathbb{C})$ como el que acabamos de describir proporciona otra variación, de naturaleza probabilista, de los algoritmos basados en NHD. En efecto, fijemos un conjunto questor $\mathcal{G}_{(d)}$ (que no depende del $\varepsilon > 0$ que elijamos). Entonces, tenemos el siguiente esquema basado en NHD.

Input: $f \in \mathcal{H}_{(d)}$, $\varepsilon \in \mathbb{R}^+$.

- *Elegir al azar* $(g, \zeta) \in \mathcal{G}_{(d)}$.
- *Realizar un número polinomial (en* $\varepsilon^{-1}, n, N, d$*) de pasos de homotopía siguiendo el segmento* $(1-t)g + tf$*,* $t \in [0,1]$*, empezando en* (g, ζ).

Output:

o bien failure*, o bien*
un cero aproximado proyectivo $z \in \mathbb{P}_n(\mathbb{C})$ *de* f.

De todas maneras, un resultado existencial como el Teorema 4.1.5 no proporciona una solución al problema principal de [115], pues seguiríamos siendo incapaces de diseñar un algoritmo concreto. Por ello, mostraremos una clase tratable algorítmicamente $\mathcal{G}_{(d)}$, y demostraremos que es un conjunto questor para pares eficientes. La exposición concreta de la clase que proponemos requiere algunas notaciones previas que pasamos a describir a continuación.

4.1.2. Descripción explícita del conjunto questor

Recuperamos las notaciones y conceptos de la Sección 1.4 (nos restringimos, naturalmente, al caso cero–dimensional, esto es, $m = n$). Entonces, sea Y el siguiente conjunto compacto, definido como el producto de bolas afines cerradas:

$$Y := [0,1] \times B^1(L_{e_0}^{\perp}) \times B^1(\mathcal{M}_{n\times(n+1)}(\mathbb{C})) \subseteq \mathbb{R} \times \mathbb{C}^{N+1},$$

donde $B^1(L_{e_0}^{\perp})$ es la bola cerrada de radio 1 en $L_{e_0}^{\perp}$ con respecto a la métrica canónica (no la métrica heredada de Kostlan), y $B^1(\mathcal{M}_{n\times(n+1)}(\mathbb{C}))$ es la bola cerrada de radio 1 en el espacio de matrices complejas $n\times(n+1)$, con respecto a la métrica estándar de Frobenius (véase Sección 1.2). Consideramos Y equipada con la métrica producto, y la estructura Riemanniana asociada. Sea $\tau \in \mathbb{R}$ el número real definido como sigue:

$$\tau := \sqrt{\frac{n^2+n}{N}}.$$

Ahora, fijemos una aplicación cualquiera $\Omega : \mathcal{M}_{n\times(n+1)}(\mathbb{C}) \longrightarrow \mathcal{U}_{n+1}$ tal que para cada matriz de rango maximal $M \in \mathcal{M}_{n\times(n+1)}(\mathbb{C})$, $\Omega(M)$ es una matriz unitaria con

$$M\Omega(M)e_0 = 0.$$

En otras palabras, $\Omega(M)$ transforma e_0 en un vector de módulo 1 del núcleo de M. Los enunciados que siguen son independientes del modo en que diseñemos la aplicación Ω (hay muchas formas de diseñarla, con técnicas

de Álgebra Lineal elemental). Definimos la aplicación $G_{(d)} : Y \longrightarrow V_{e_0}$ como sigue: Para cada punto $(t, h, M) \in Y$, $G_{(d)}(t, h, M) \in V_{e_0}$ es igual a

$$\left(1 - \tau^2 t^{\frac{1}{n^2+n}}\right)^{1/2} \frac{\Delta^{-1} h}{\|h\|_2} + \tau t^{\frac{1}{2n^2+2n}} \psi_{e_0}^{-1} \left(T_{e_0}\left(\frac{M}{\|M\|_F}\Omega\left(\frac{M}{\|M\|_F}\right)\right)\right).$$

Observamos que $G_{(d)}$ no está definida en el caso de que M sea de rango no maximal. No obstante, este detalle carece de importancia pues estamos tratando con probabilidades y el conjunto de matrices no regulares de $\mathcal{M}_{n\times(n+1)}(\mathbb{C})$ tiene medida nula.
Finalmente, denotamos por $\mathcal{G}_{(d)}$ la clase definida por la siguiente igualdad:

$$\mathcal{G}_{(d)} := Im(G_{(d)}) \times \{e_0\} \subseteq \mathcal{H}_{(d)} \times \mathbb{P}_n(\mathbb{C}). \tag{4.3}$$

Obsérvese que $\mathcal{G}_{(d)}$ está contenida en la variedad de incidencia, y que todos los sistemas en $Im(G_{(d)})$ comparten un cero común e_0. Esto es, están todos resueltos por construcción. Consideramos $\mathcal{G}_{(d)}$ equipada con la distribución de probabilidad heredada de Y via $G_{(d)}$. Esto es, para elegir un punto al azar en $\mathcal{G}_{(d)}$, elegimos un punto al azar en $y \in Y$, y calculamos $(G_{(d)}(y), e_0) \in \mathcal{G}_{(d)}$.

Teorema 4.1.6 *Con las notaciones anteriores, la clase de sistemas $\mathcal{G}_{(d)}$ es un conjunto questor para pares eficientes en $\mathcal{H}_{(d)}$. Esto es, para todo número real positivo $\varepsilon > 0$, la probabilidad de que un par elegido al azar $(g, e_0) \in \mathcal{G}_{(d)}$ sea ε–eficiente es mayor o igual que*

$$1 - \varepsilon.$$

Obsérvese que para esos pares ε–eficientes $(g, e_0) \in \mathcal{G}_{(d)}$, la probabilidad de que un input elegido al azar $f \in \mathcal{H}_{(d)}$ sea resuelto por NHD con par inicial (g, e_0) realizando $O(n^5 N^3 d^3 \varepsilon^{-2})$ pasos de homotopía es al menos $1 - \varepsilon$. Por tanto, y como es usual, la existencia de un conjunto questor constructible implica inmediatamente la existencia de un algoritmo. Por ello, el Teorema 4.1.1 es consecuencia inmediata del Teorema 4.1.6. Como un ejemplo de aplicación, sabiendo que el número de operaciones que requiere un paso de Newton es del orden de $O(nN \log_2 d)$, tenemos el siguiente corolario.

Corolario 4.1.7 *Existe un algoritmo probabilista que resuelve los sistemas homogéneos de grado 3 (esto es, sistemas en $\mathcal{H}_{(3)}$)), realizando un número de operaciones aritméticas del orden de*

$$O(n^{22}\varepsilon^{-2}),$$

con probabilidad mayor o igual que $1 - \varepsilon$.
Tomando por ejemplo $\varepsilon = \frac{1}{n^2}$ podemos resolver casi todos los sistemas de ecuaciones de grado 3 en tiempo $O(n^{26})$. La probabilidad de éxito es al menos de $1 - \frac{1}{n^2}$.

Podemos dar una versión más fuerte del Teorema 4.1.6. En efecto, obsérvese que dicho resultado puede interpretarse como sigue: Para todo $\varepsilon > 0$, el conjunto

$$\mathcal{E}_\varepsilon := \{(g, e_0) \in \mathcal{G}_{(d)} : (g, e_0) \text{ es } \varepsilon - eficiente\} \subseteq \mathcal{G}_{(d)}$$

satisface:

$$\text{Prob}_{(g,e_0)\in\mathcal{G}_{(d)}}[(g, e_0) \in \mathcal{E}_\varepsilon] \geq 1 - \varepsilon.$$

Una pregunta natural es si existe algún par $(g, e_0) \in \mathcal{G}_{(d)}$ que sea ε–eficiente *para todo* $\varepsilon > 0$. De hecho, podemos demostrar que no solo es así, sino que es muy probable encontrar dicho par (g, e_0). Esto es, tenemos el siguiente resultado.

Teorema 4.1.8 *Con las notaciones anteriores, se* $\mathcal{E} \subseteq \mathcal{G}_{(d)}$ *la clase definida como sigue,*

$$\mathcal{E} := \bigcap_{0<\varepsilon<1/2} \mathcal{E}_\varepsilon.$$

Esto es, $\mathcal{E}$ *es la clase de los pares iniciales* (g, e_0) *que son* ε*–eficientes para todo* ε*,* $0 < \varepsilon < 1/2$*. Entonces, se satisface la siguiente desigualdad:*

$$\text{Prob}_{(g,e_0)\in\mathcal{G}_{(d)}}\left[(g, e_0) \in \mathcal{E}\right] \geq \frac{3}{4}.$$

Nota 4.1.9 *Este resultado no es consecuencia inmediata del Teorema 4.1.6, pues no podemos garantizar a priori ningún resultado de jerarquía entre las clases* $\mathcal{E}_\varepsilon$ *al hacer variar* ε*. En efecto, para* $\varepsilon, \varepsilon' > 0$*, los hechos de que un par sea* ε*–eficiente o* ε'*–eficiente no tienen por qué estar relacionados a priori.*

El Teorema 4.1.8 significa que podemos encontrar, probabilísticamente, un par inicial (g, e_0) que es ε–eficiente *para todo* $\varepsilon > 0$. Esto es, un par inicial que nos permite calcular un cero aproximado proyectivo de la mayoría de los sistemas de ecuaciones (probabilidad de éxito $1 - \varepsilon$) en un tiempo polinomial en la talla del input ($O(n^5N^3d^3\varepsilon^{-2})$ pasos de homotopía), con la ventaja adicional de que podemos cambiar el valor de ε en función de nuestros recursos, sin necesidad de cambiar de par inicial (g, e_0).
El Teorema 4.1.6 y sus consecuencias representan un pequeño paso adelante en la teoría de Shub & Smale para la resolución de sistemas de ecuaciones. Simplemente muestra la existencia de un algoritmo "uniforme", probabilístico, que computa información sobre las soluciones de la gran mayoría de sistemas de ecuaciones en tiempo polinomial en el tamaño del input.

4.2. Notaciones y resultados previos

En las secciones es que siguen utilizaremos las mismas notaciones que en las secciones 1.4, 1.6 y el Capítulo 3, con la salvedad de que ahora trabajaremos

en la esfera en el espacio de sistemas en vez de en el espacio proyectivo asociado. La estructura Riemanniana considerada en $\mathcal{H}_{(d)}$ es la misma estructura unitaria invariante que en el capítulo 3. Viene por tanto dada por la matriz de Kostlan Δ definida en la Sección 1.6 del Capítulo 1. No obstante, en ocasiones utilizaremos también la estructura Riemanniana dada por la métrica canónica usual. Representaremos por $(\mathcal{H}_{(d)}, can)$ el espacio con la métrica canónica. Como hemos visto en el Lema 1.4.1, los espacios $(\mathcal{H}_{(d)}, can)$ y $\mathcal{H}_{(d)}$ son isométricos mediante la aplicación Δ. Supondremos a lo largo de todo el capítulo que al menos uno de los grados d_i que aparecen en la expresión de (d) es mayor que 1.
Para simplificar la notación, durante este capítulo denotaremos respectivamente por $\mathbb{S}$ y $\mathbb{S}_\Delta$ las esferas en $(\mathcal{H}_{(d)}, can)$ y $\mathcal{H}_{(d)}$. Naturalmente, $\mathbb{S}$ se considera con la estructura canónica y $\mathbb{S}_\Delta$ con la estructura unitaria invariante heredada de la de $\mathcal{H}_{(d)}$ (definida por la matriz de Kostlan Δ).
También utilizaremos las nociones relacionadas con la variedad de incidencia, de un modo similar a como lo hemos hecho en el Capítulo 3. No obstante, los conceptos ahora serán sobre la esfera, en vez del proyectivo complejo. Por ello volvemos a introducir algunas notaciones, para evitar posibles confusiones.
Sea $f \in \mathcal{H}_{(d)}$ un sistema de ecuaciones y $\zeta \in V(f)$ una solución cualquiera de ese sistema. Denotaremos por $T_\zeta f$ la restricción de la aplicación diferencial $d_\zeta f$ al subespacio tangente $T_\zeta \mathbb{P}_n(\mathbb{C}) = \zeta^\perp \subseteq \mathbb{C}^{n+1}$, consistente en los elementos de $\mathbb{C}^{n+1}$ ortogonales a la recta compleja $\zeta \in \mathbb{P}_n(\mathbb{C})$. En el caso de que $\zeta = e_0$, identificaremos $T_{e_0} f$ con su matriz en base natural $\{e_1, \ldots, e_n\}$ (como en la identidad (1.8)).
Para todo punto $\zeta \in \mathbb{P}_n(\mathbb{C})$, denotamos por $\widetilde{V_\zeta} \subseteq \mathcal{H}_{(d)}$ el subespacio vectorial dado por todos los sistemas de ecuaciones en $\mathcal{H}_{(d)}$ que se anulan en ζ. En otras palabras,

$$\widetilde{V_\zeta} := \{f \in \mathcal{H}_{(d)} \ : \ f(\zeta) = 0 \in \mathbb{C}^n\}.$$

Obsérvese que $\widetilde{V_\zeta}$ es un subespacio vectorial de $\mathcal{H}_{(d)}$ de codimensión compleja n. Consideramos de nuevo la variedad de incidencia $W \subseteq \mathbb{S}_\Delta \times \mathbb{P}_n(\mathbb{C})$, aunque ahora los sistemas pertenecen a la esfera $\mathbb{S}_\Delta$ en vez del proyectivo $\mathbb{P}(\mathcal{H}_{(d)})$. Esto es, W viene dada mediante la siguiente igualdad:

$$W := \{(f, \zeta) \in \mathbb{S}_\Delta \times \mathbb{P}_n(\mathbb{C}) \ : \ \zeta \in V(f)\}.$$

También consideraremos las proyecciones canónicas:

$$p_1 : W \longrightarrow \mathbb{S}_\Delta, \quad p_1(f, \zeta) := f, \forall (f, \zeta) \in V,$$

y

$$p_2 : W \longrightarrow \mathbb{P}_n(\mathbb{C}), \quad p_2(f, \zeta) := \zeta, \forall (f, \zeta) \in V.$$

Podemos identificar de manera obvia $p_1^{-1}(f) \equiv V(f)$ y $p_2^{-1}(\zeta) \equiv \widetilde{V_\zeta} \cap \mathbb{S}_\Delta = S^1_\Delta(\widetilde{V_\zeta})$. Desde ahora, denotaremos $V_\zeta := p_2^{-1}(\zeta)$.

En la Proposición 3.5.1 hemos visto las propiedades de la variedad de incidencia en el caso proyectivo. Estas propiedades se conservan cuando consideramos que los sistemas están en la esfera. Esto es, tenemos el siguiente resultado.

Proposición 4.2.1 (Shub & Smale) *La variedad de incidencia W es una subvariedad diferenciable conexa de la variedad producto $\mathbb{S}_\Delta \times \mathbb{P}_n(\mathbb{C})$, de codimensión real $2n$. Además, las fibras V_ζ son subvariedades diferenciables de W de codimensión real $2n$ en V.*

Denotaremos por $\Sigma'_{(d)} \subseteq W$ el conjunto de puntos críticos de p_1, esto es, $\Sigma'_{(d)} := \{(f,\zeta) \in W \ : \ T_\zeta f \notin GL(n,\mathbb{C})\}$. También denotamos por $\Sigma_{(d)} := p_1(\Sigma'_{(d)})$ el conjunto de valores críticos de p_1. Como se indica en [115], los dos resultados que escribimos a continuación se siguen del Teorema de la Función Inversa. Incluimos una demostración topológica del segundo de ellos por no estar explícitamente incluida en [115] y por completitud de esta memoria.

Proposición 4.2.2 (Shub & Smale) *Sea $g \in \mathbb{S}_\Delta$ un punto, y sea $L_\Delta \subseteq \mathbb{S}_\Delta$ un círculo máximo en $\mathbb{S}_\Delta$, tal que $g \in L_\Delta$. Supongamos que $L_\Delta \cap \Sigma_{(d)} = \emptyset$. Entonces, $p_1^{-1}(L_\Delta) \setminus p_1^{-1}(\{-g\})$ consiste en $\mathcal{D}$ arcos abiertos en W. Sea $\zeta \in V(g)$ una solución de g. Denotamos por $ARC_{L_\Delta,g,\zeta}$ el arco abierto en $p_1^{-1}(L_\Delta) \setminus p_1^{-1}(\{-g\})$ que contiene al punto (g,ζ).*

Proposición 4.2.3 *Sea $(f,g,\zeta) \in \mathbb{S}_\Delta \times W$ un punto tal que $f \notin \{g,-g\}$. Sea $\mathfrak{L}_\Delta(f,g,\zeta)$ la componente conexa en $p_1^{-1}(\mathfrak{L}_\Delta(f,g)) \setminus p_1^{-1}(\{-g\}) \subseteq W$ que contiene el punto (g,ζ). Si $\mathfrak{L}_\Delta(f,g,\zeta)$ no corta a $\Sigma'_{(d)}$, entonces es una curva diferenciable. Esto es, en las condiciones de la Proposición 4.2.2 tendríamos:*

$$\mathfrak{L}_\Delta(f,g,\zeta) := ARC_{\mathfrak{L}_\Delta(f,g),g,\zeta}.$$

Además, para cada $h \in \mathfrak{L}_\Delta(f,g) \setminus \{-g\}$, existe una única solución $\zeta' \in \mathbb{P}_n(\mathbb{C})$ de h tal que $(h,\zeta') \in \mathfrak{L}_\Delta(f,g,\zeta)$.

Demostración.– Consideramos la restricción

$$\tilde{p} := p_1 \mid_{\mathfrak{L}_\Delta(f,g,\zeta)} \longrightarrow \mathfrak{L}_\Delta(f,g) \setminus \{-g\}.$$

Como $\mathfrak{L}_\Delta(f,g,\zeta)$ no corta a $\Sigma'_{(d)}$, tenemos que $\tilde{p}$ es una aplicación regular. Por el Corolario 1.1.6, $\mathfrak{L}_\Delta(f,g,\zeta)$ es una variedad diferenciable de dimensión 1, esto es, una curva diferenciable. Deducimos también que para todo $h \in Im(\tilde{p})$, $\tilde{p}^{-1}(h)$ es un conjunto discreto (y, dado que h es un sistema de ecuaciones polinomiales, un conjunto finito). Demostramos ahora que $\tilde{p}$ es sobreyectiva.
Supongamos que no es así. Entonces, existe un sistema $h' \in \mathfrak{L}_\Delta(f,g) \setminus \{-g\}$ tal que $\tilde{p}^{-1}(h') = \emptyset$. Sea $\gamma : [0,1] \longrightarrow \mathfrak{L}_\Delta(f,g) \setminus \{-g\}$ una curva tal que $\gamma(0) = g$ y $\gamma(1) = h'$. Sea $t_0 \in (0,1)$ el ínfimo de los valores de t tales

que $\tilde{p}^{-1}(\gamma(t)) = \emptyset$. Sea $\zeta_n \in \mathbb{P}_n(\mathbb{C}) \in \tilde{p}^{-1}\left(\gamma\left(t_0 - \frac{1}{n}\right)\right)$ una solución de $\gamma\left(t_0 - \frac{1}{n}\right)$, para todo $n \in \mathbb{N}$ (suficientemente grande). Entonces, la sucesión $(\zeta_n)_{n\in\mathbb{N}}$ tiene una subsucesión convergente por ser $\mathbb{P}_n(\mathbb{C})$ compacto; llamemos $y \in \mathbb{P}_n(\mathbb{C})$ al límite de esa subsucesión. Por continuidad de las soluciones, y es solución de $\gamma(t_0)$. Las componentes conexas de un espacio topológico son siempre cerradas, por lo que deducimos que $y \in \tilde{p}^{-1}(\gamma(t_0))$. Ahora, y es por tanto una solución regular de h, con lo que p_1 y $\tilde{p}$ son regulares en (h, y). Por el Teorema de la Función Inversa, existe un entorno de h que es difeomorfo, vía $\tilde{p}^{-1}$, con un entorno de (h, y) en $\mathfrak{L}_\Delta(f, g, \zeta)$. En particular, existe un entorno de h tal que todo punto de él tiene una anti–imagen en $\mathfrak{L}_\Delta(f, g, \zeta)$, lo que contradice la hipótesis.
Queda pues demostrado que $\tilde{p}$ es sobreyectiva. Por otro lado, de nuevo debido a que $\tilde{p}$ es regular, al Teorema de la Función Inversa y a que $\tilde{p}^{-1}(h)$ es finito para todo $h \in \mathfrak{L}_\Delta(f, g) \setminus \{-g\}$, deducimos que $\tilde{p}$ es una aplicación recubridora, esto es, que para todo $h \in \mathfrak{L}_\Delta(f, g) \setminus \{-g\}$, existe un entorno U_h de h en $\mathfrak{L}_\Delta(f, g) \setminus \{-g\}$ y entornos $\{V_i\}_{1\leq i\leq k}$ de las anti–imágenes de h de forma que $\tilde{p} : V_i \longrightarrow U_h$ es un homeomorfismo para todo $1 \leq i \leq k$. Por último, es bien sabido que toda aplicación recubridora de un espacio conexo en otro simplemente conexo es en realidad un homeomorfismo (véase cualquier libro de texto de Topología). En particular, $\tilde{p}$ es biyectiva y la proposición queda demostrada. ∎

Notación 4.2.4 *A partir de ahora, para $f \neq \pm g$ y $\zeta \in V(g)$, denotamos:*

$$\mu_{\mathrm{norm}}(f, g, \zeta) := \sup_{(h,\zeta')\in\mathfrak{L}_\Delta(f,g,\zeta)} \{\mu_{\mathrm{norm}}(h, \zeta')\}.$$

4.2.1. Algunas acciones unitarias

Como hemos indicado ya (véase por ejemplo la Sección 3.5), la variedad de incidencia W es invariante mediante la acción de $\mathcal{U}_{n+1}$ en el producto $\mathbb{S}_\Delta \times \mathbb{P}_n(\mathbb{C})$. Esto es, la aplicación que sigue es una isometría para toda matriz $U \in \mathcal{U}_{n+1}$:

$$\begin{array}{rccc} U: & W & \longrightarrow & W \\ & (f, y) & \mapsto & (f \circ U^{-1}, Uy). \end{array}$$

Además, $U \in \mathcal{U}_{n+1}$ define isometrías entre las fibras de p_2. En efecto, dados dos puntos proyectivos $\zeta, \zeta' \in \mathbb{P}_n(\mathbb{C})$, y dada una matriz unitaria $U \in \mathcal{U}_{n+1}$, tal que $U\zeta = \zeta'$, la restricción

$$\begin{array}{rccc} U^{-1}: & V_\zeta & \longrightarrow & V_{\zeta'} \\ & f & \mapsto & f \circ U^{-1} \end{array}$$

es también una isometría entre esferas. Obsérvese que los siguientes diagramas son conmutativos:

$$\begin{array}{ccc} W & \xrightarrow{U} & W \\ p_1 \downarrow & & \downarrow p_1 \\ \mathbb{S}_\Delta & \xrightarrow{U^{-1}} & \mathbb{S}_\Delta \end{array} \qquad \begin{array}{ccc} W & \xrightarrow{U} & W \\ p_2 \downarrow & & \downarrow p_2 \\ \mathbb{P}_n(\mathbb{C}) & \xrightarrow{U} & \mathbb{P}_n(\mathbb{C}) \end{array}$$

Sea $f \in V_{e_0}$ un sistema. Consideramos el número siguiente:

$$DET(f, e_0) := det(T_{e_0} f (T_{e_0} f)^*),$$

donde el símbolo * denota trasposición conjugada compleja.
El siguiente resultado se sigue del Lema 3.5.4.

Proposición 4.2.5 *Con las notaciones anteriores, sea $(f, \zeta) \in W$ un punto regular de p_1. Entonces, tenemos:*

$$\frac{NJ_{(f,\zeta)} p_1}{NJ_{(f,\zeta)} p_2} = \frac{NJ_{(f \circ U, e_0)} p_1}{NJ_{(f \circ U, e_0)} p_2} = DET(f \circ U, e_0),$$

donde $U \in \mathcal{U}_{n+1}$ es una matriz unitaria cualquiera tal que $Ue_0 = \zeta$.

Demostración.– La primera de las dos igualdades es consecuencia de la invariancia unitaria, utilizando el Corolario 1.1.12 del Capítulo 1. El Lema 3.5.4 demuestra la segunda igualdad en el caso en que p_1 y p_2 son consideradas con dominio en el espacio proyectivo complejo $\mathbb{P}(\mathcal{H}_{(d)})$ en vez de en la esfera $\mathbb{S}_\Delta$. Ahora, comprobamos que este cambio no afecta al cálculo. En efecto, dado un punto $(f, e_0) \in W$, los espacios tangentes $T_f \mathbb{S}_\Delta$ y $T_f \mathbb{P}(\mathcal{H}_{(d)})$ sólo se diferencian en el vector $\sqrt{-1} f \in T_f \mathbb{S}_\Delta$. Ahora, el vector $(\sqrt{-1} f, 0) \in T_{(f,e_0)} W$ satisface:

- $\sqrt{-1} f = d_{(f,e_0)} p_1 (\sqrt{-1} f, 0)$ es ortogonal a $g = d_{(f,e_0)} p_1 (g, x)$ para cada $(g, x) \in T_{(f,e_0)} W$ tal que $\langle (g, x), (\sqrt{-1} f, 0) \rangle_{T_{(f,e_0)} W} = 0$.
- $(\sqrt{-1} f, 0) \in Ker(d_{(f,e_0)} p_2)$.

Por lo tanto, el volumen de la imagen por $d_{(f,e_0)} p_1$ o $d_{(f,e_0)} p_2$ de un cubo unidad contenido en el ortogonal al núcleo respectivo no varía, y ambos jacobianos normales $NJ_{(f,\zeta)} p_1$ y $NJ_{(f,\zeta)} p_2$ quedan igual al considerar como dominio $\mathbb{S}_\Delta$ o $\mathbb{P}(\mathcal{H}_{(d)})$. ∎

4.2.2. Homotopía y condicionamiento

Para todo par $(f, \zeta) \in W$, denotaremos por $\mu_{\text{norm}}(f, \zeta)$ el número de condicionamiento no–lineal de la definición 1.6.1. Esto es,

$$\mu_{\text{norm}}(f, \zeta) := \|(T_\zeta f)^{-1} \Delta(d)^{1/2}\|_2 = \frac{\kappa_D(\Delta(d)^{-1/2} T_\zeta f)}{||\Delta(d)^{-1/2} T_\zeta f||_F},$$

donde el representante ζ ha sido elegido de modo que $\|\zeta\|_2 = 1$. Sabemos que μ_{norm} es invariante bajo la acción unitaria del grupo $\mathcal{U}_{n+1}$. Esto es, dados $(f, \zeta) \in W$ y $U \in \mathcal{U}_{n+1}$, tenemos:

$$\mu_{\text{norm}}(f, \zeta) = \mu_{\text{norm}}(f \circ U^{-1}, U\zeta).$$

Para cada número real positivo $\varepsilon > 0$, consideramos los conjuntos $(\Sigma'_{(d)})_\varepsilon \subseteq W$ y $(\Sigma_{(d)})_\varepsilon \subseteq \mathbb{S}_\Delta$ dados como sigue,

$$(\Sigma'_{(d)})_\varepsilon := \{(f, \zeta) \in W \ : \ \mu_{\text{norm}}(f, \zeta) > \varepsilon^{-1}\},$$

$$(\Sigma_{(d)})_\varepsilon := p_1((\Sigma'_{(d)})_\varepsilon) = \{f \in S^1_\Delta(\mathcal{H}_{(d)}) \ : \ \exists \zeta \in V(f), \ \mu_{\text{norm}}(f, \zeta) > \varepsilon^{-1}\}.$$

Ambos conjuntos son invariantes bajo la acción de $\mathcal{U}_{n+1}$.
Sea $g \in \mathbb{S}_\Delta$ un sistema de ecuaciones. Para cada círculo máximo L_Δ que contiene a g y tal que $L_\Delta \cap \Sigma_{(d)} = \emptyset$, y para cada número positivo $\varepsilon > 0$, denotamos por $\tau^g_\varepsilon(L)$ el número de arcos abiertos en $p_1^{-1}(L_\Delta) \setminus p_1^{-1}(\{-g\})$ que intersecan al conjunto $(\Sigma_{(d)})'_\varepsilon$. En otras palabras,

$$\tau^g_\varepsilon(L) := \sharp\{\zeta \in V(g) : ARC_{L_\Delta, g, \zeta} \cap (\Sigma_{(d)})'_\varepsilon \neq \emptyset\}.$$

Esta definición tiene sentido gracias a la Proposición 4.2.2. Entonces, para cada real positivo $\varepsilon > 0$, y para cada círculo máximo $L_\Delta \subseteq \mathbb{S}_\Delta$ tal que $L_\Delta \cap \Sigma_{(d)} = \emptyset$, definimos

$$\tau_\varepsilon(L_\Delta) := \sup_{g \in L} \tau^g_\varepsilon(L). \tag{4.4}$$

4.2.3. La homotopía lineal.

La homotopía lineal propuesta por Shub & Smale trata de encontrar un cero aproximado proyectivo de sistemas $f \in \mathbb{S}_\Delta$ mediante el siguiente método. Primero, consideramos otro sistema $g \in \mathbb{S}_\Delta$, con un cero conocido ζ_0. Elegimos también un número natural $k \geq 1$, representando el número de pasos de homotopía que vamos a realizar. Consideramos la siguiente secuencia de sistemas:

$$h_i := \left(1 - \frac{i}{k}\right) g + \frac{i}{k} f, \quad 0 \leq i \leq k. \tag{4.5}$$

Obsérvese que $h_0 = g$, $h_k = f$. Entonces, podemos considerar la secuencia de puntos definida como sigue:

$$x_0 := \zeta_0, \quad x_i := N_{h_i}(x_{i-1}), 1 \leq i \leq k. \tag{4.6}$$

Este tipo de algoritmos se conocen como Newton Deformation Homotopy. Como hemos hecho en la introducción de este capítulo, abreviaremos esta expresión por NHD. El siguiente resultado controla el número de pasos k necesarios para garantizar la convergencia en función de la cantidad μ_{norm}. Su demostración puede encontrarse (implícitamente escrita) en [14]. Incluimos una demostración por completitud de esta memoria.

Proposición 4.2.6 (Shub & Smale) *Sean $f, g \in \mathbb{S}_\Delta$ dos sistemas, y sea $\zeta \in \mathbb{P}_n(\mathbb{C})$ una solución de g. Además, sea $k(f,g,\zeta) \in \mathbb{N}$ un número natural tal que:*

$$k(f,g,\zeta) \geq c_1 d^{3/2} \mu_{\text{norm}}(f,g,\zeta)^2,$$

donde $c_1 := \frac{28}{5(3-\sqrt{7})} > 0$ es esa constante universal. Entonces, NHD *con par inicial (g,ζ) y $k(f,g,\zeta)$ pasos de homotopía, con input f, devuelve un cero aproximado proyectivo z de f. Además, se tiene la siguiente desigualdad:*

$$d_T(z,\zeta') \leq \frac{3-\sqrt{7}}{2d^{3/2}\mu_{\text{norm}}(f,g,\zeta)},$$

donde $\zeta' \in V(f)$ es la única solución de f que está en $\mathfrak{L}_\Delta(f,g,\zeta)$.

Demostración.– Denotemos $k := k(f,g,\zeta) \in \mathbb{N}$ y para cada natural i, $0 \leq i \leq k$, sean h_i, z_i como en las identidades (4.5) y (4.6). Por el Teorema 1.7.2, basta con demostrar que

$$d_T(z_i,\zeta_i)\mu_{\text{norm}}(f,g,\zeta) \leq \frac{3-\sqrt{7}}{2d^{3/2}}, \ \forall i \in \{0,\dots k\}, \tag{4.7}$$

donde ζ_i es alguna solución (exacta) de h_i. De hecho, bastaría con poner un $d^{3/2}$ en el denominador, pero la expresión (4.7) tal y como está escrita facilita el proceso de inducción. Demostramos esta afirmación por inducción en $i \in \mathbb{N}$. El caso $i = 0$ es trivial, porque $\zeta = \zeta_0$ es una solución de $h_0 = g$. Sea ahora $i \geq 1$. Estamos asumiendo que $k < +\infty$. Por lo tanto, $\mu_{\text{norm}}(f,g,\zeta) < +\infty$ y $\mathfrak{L}_\Delta(f,g,\zeta)$ no interseca a $\Sigma'_{(d)}$, que hemos definido como

$$\Sigma'_{(d)} := \{(h,y) \in \mathbb{S}_\Delta \times \mathbb{P}_n(\mathbb{C}) : h(y) = 0, rank(d_y h) \leq n-1\}.$$

Entonces $p_1 : \mathfrak{L}_\Delta(f,g,\zeta) \longrightarrow \mathfrak{L}_\Delta(f,g)$ es una biyección (véase la Proposición 4.2.3), y su inversa asigna a cada sistema $h \in \mathfrak{L}_\Delta(f,g)$, $h \neq -g$, la única solución de h en $\mathfrak{L}_\Delta(f,g,\zeta)$. Denotamos por $\zeta_i := p_1^{-1}(h_i)$ la solución asociada a h_i. Esto es, ζ_i es la única solución de h_i que pertenece a $\mathfrak{L}_\Delta(f,g,\zeta)$. Por [14, Prop. 2, page 272], se tiene la siguiente desigualdad (que en realidad es una consecuencia inmediata del hecho de que μ_{norm} es número de condicionamiento, esto es, que controla la variación de las soluciones en función de la variación del sistema):

$$d_R(\zeta_i,\zeta_{i-1}) \leq \|h_i - h_{i-1}\|_\Delta \mu_{\text{norm}}(f,g,\zeta),$$

donde d_R es la distancia de Fubini–Study en $\mathbb{P}_n(\mathbb{C})$ y $\|h_i - h_{i-1}\|_\Delta$ es la distancia entre h_i y h_{i-1} en $\mathcal{H}_{(d)}$. Ahora, obsérvese que $f, g \in \mathbb{S}_\Delta$. Por lo tanto,

$$\|h_i - h_{i-1}\|_\Delta = \frac{1}{k}\|f-g\|_\Delta \leq \frac{2}{k} \leq \frac{2}{c_1 d^{3/2}\mu_{\text{norm}}(f,g,\zeta)^2}.$$

Concluimos que

$$d_R(\zeta_i, \zeta_{i-1}) \leq \frac{2}{c_1 d^{3/2} \mu_{\text{norm}}(f,g,\zeta)} = \frac{5(3-\sqrt{7})}{14 d^{3/2} \mu_{\text{norm}}(f,g,\zeta)}.$$

Por otro lado, la desigualdad triangular implica que

$$d_R(z_{i-1}, \zeta_i) \leq d_R(z_{i-1}, \zeta_{i-1}) + d_R(\zeta_{i-1}, \zeta_i).$$

Por lo tanto, por hipótesis de inducción,

$$d_T(z_{i-1}, \zeta_i) \leq \tan(d_R(z_{i-1}, \zeta_{i-1}) + d_R(\zeta_{i-1}, \zeta_i)) \leq$$

$$\tan\left(\frac{3-\sqrt{7}}{2d^{3/2}\mu_{\text{norm}}(f,g,\zeta)} + \frac{5(3-\sqrt{7})}{14 d^{3/2}\mu_{\text{norm}}(f,g,\zeta)}\right) =$$

$$\tan\left(\frac{6(3-\sqrt{7})}{7 d^{3/2}\mu_{\text{norm}}(f,g,\zeta)}\right) \leq \frac{3-\sqrt{7}}{d^{3/2}\mu_{\text{norm}}(f,g,\zeta)}$$

La última desigualdad se sigue del hecho de que $\frac{\tan x}{x}$ es una función creciente en $(0,1)$ y que evaluada en $x = \frac{6(3-\sqrt{7})}{7}$ es menor que $\frac{7}{6}$. Por el Teorema 1.7.2, concluimos que z_{i-1} es un cero aproximado proyectivo de h_i con cero asociado ζ_i. Por la definición de cero aproximado proyectivo (Definición 1.7.1), tenemos:

$$d_T(z_i, \zeta_i) = d_T(N_{h_i}(z_{i-1}), \zeta_i) \leq \frac{1}{2} d_T(z_{i-1}, \zeta_i) \leq \frac{3-\sqrt{7}}{2d^{3/2}\mu_{\text{norm}}(f,g,\zeta)},$$

y queda demostrada la ecuación (4.7). ∎

4.2.4. Una estimación de volumen para círculos máximos

En esta subsección demostramos la Proposición 4.2.9, que atribuimos a Shub y Smale pues, si bien no está explícitamente escrita en [115], es consecuencia bastante directa de los resultados y técnicas de ese artículo. Recordemos que hemos denotado por $\mathbb{S} := (S^1(\mathcal{H}_{(d)}), can)$ la esfera de radio 1 para la métrica usual en $\mathcal{H}_{(d)}$. Durante este capítulo, usaremos notaciones más cómodas para el volumen. Para un conjunto medible $A_1 \subseteq \mathbb{S}$, denotaremos por $\nu[A]$ el volumen de A, mientras que en el caso $A_2 \subseteq \mathbb{S}_\Delta$, denotaremos simplemente $\nu_\Delta[A_2]$. También consideramos las variedades diferenciables $\mathbb{S}\times\mathbb{S}$ y $\mathbb{S}_\Delta \times \mathbb{S}_\Delta$, con la estructura producto. Para conjuntos medibles $A_1 \subseteq \mathbb{S}\times\mathbb{S}$, $A_2 \subseteq \mathbb{S}_\Delta \times \mathbb{S}_\Delta$, también denotaremos sus volúmenes respectivos por $\nu[A_1], \nu_\Delta[A_2]$. Por el Lema 1.4.1, las aplicaciones

$$\Delta^{-1} : \mathbb{S} \longrightarrow \mathbb{S}_\Delta.$$

y

$$\Delta^{-1} \times \Delta^{-1} : \mathbb{S} \times \mathbb{S} \longrightarrow \mathbb{S}_\Delta \times \mathbb{S}_\Delta.$$

son isometrías. En particular, tenemos que

$$\nu_\Delta[\mathbb{S}_\Delta] = \nu[\mathbb{S}] = 2\frac{\pi^{N+1}}{\Gamma(N+1)}.$$

Sea $\mathcal{L}$ la variedad diferenciable consistente en los círculos máximos de $\mathbb{S}$, con la estructura Riemanniana natural, invariante por transformaciones ortogonales. Para todo conjunto medible $A \subseteq \mathcal{L}$, sea $\nu_{\mathcal{L}}[A]$ el volumen de A con respecto a esa estructura Riemanniana. Suponemos que la forma de volumen $d\mathcal{L}$ ha sido normalizada de modo que $\nu_{\mathcal{L}}[\mathcal{L}] = 1$. Sea $\mathcal{O}_{2N+2}$ el grupo de matrices ortogonales de lado $2N+2$, que actúa isométricamente en el espacio vectorial $\mathbb{C}^{N+1} \equiv \mathbb{R}^{2N+2}$. Es decir, para todo conjunto medible $A \subseteq \mathbb{S}$, se tiene que:

$$\nu[A] = \nu[OA], \quad \forall O \in \mathcal{O}_{2N+2}.$$

La siguiente aplicación es una isometría para toda matriz ortogonal $O \in \mathcal{O}_{2N+2}$:

$$\begin{array}{rccl} O: & \mathcal{L} & \longrightarrow & \mathcal{L} \\ & L & \mapsto & OL := \{Of \in \mathbb{S} : f \in L\}. \end{array}$$

Para todo elemento $L \in \mathcal{L}$, consideramos el círculo máximo $L_\Delta \subseteq \mathbb{S}_\Delta$ definido como $L_\Delta := \Delta^{-1}L = \{\Delta^{-1}f : f \in L\}$. Para cada elemento $L \in \mathcal{L}$ tal que $L_\Delta \cap \Sigma_{(d)} = \emptyset$ podemos considerar el número $\tau(\varepsilon, L)$ definido como sigue:

$$\tau(\varepsilon, L) := \tau_\varepsilon(L_\Delta),$$

donde τ_ε es como fue definido en la ecuación (4.4). Para cada sistema $f \in \mathbb{S}_\Delta \setminus \Sigma_{(d)}$, podemos considerar el número positivo $\sharp(\varepsilon, f) \in \mathbb{N}$ definido como:

$$\sharp(\varepsilon, f) := \sharp\{\zeta \in V(f) : \mu_{\text{norm}}(f, \zeta) > \varepsilon^{-1}\}.$$

También denotamos por $\mathcal{L}_\Delta$ el conjunto de todos los círculos máximos en $\mathbb{S}_\Delta$. En el Capítulo 3, hemos obtenido algunas cotas sobre la distribución del número de condicionamiento no–lineal. Sin embargo, en esta sección utilizaremos la cota más fina obtenida por Shub & Smale para el caso particular de $\mu_{\text{norm}} = \mu^{(n)}_{\text{norm}}$. El siguiente resultado se puede encontrar en [113, 14].

Teorema 4.2.7 (Shub & Smale) *Para todo número real positivo $\varepsilon > 0$, se tiene:*

$$\frac{1}{\nu_\Delta[\mathbb{S}_\Delta]} \int_{f \in \mathbb{S}_\Delta} \sharp(\varepsilon, f)\ d\mathbb{S}_\Delta \leq n^3(n+1)N^2\mathcal{D}\varepsilon^4.$$

El siguiente resultado, también debido a Shub & Smale, es consecuencia inmediata de una desigualdad que aparece en la prueba del Teorema 2 de la Sección 2 de [115].

Lema 4.2.8 (Shub & Smale) *Para todo número real positivo $\varepsilon > 0$ y para todo círculo máximo $L \in \mathcal{L}$, se tiene:*

$$\tau(\varepsilon, L) \leq \left(\frac{c\varepsilon^2}{d^{3/2}}\right)^{-1} \int_{f \in \Delta^{-1}L} \sharp(2\varepsilon, f)\ d(\Delta^{-1}L),$$

donde $c \geq 0{,}09$ es una constante universal.

Demostración.– Por definición, $\tau(\varepsilon, L) = \tau_\varepsilon(\Delta^{-1}L) = \sup_{g \in \Delta^{-1}L} \tau_\varepsilon^g(\Delta^{-1}L)$. Ahora, por [115, Teor. 2, Sec. 2] sabemos que la cantidad $\tau_\varepsilon^g(\Delta^{-1}L)$ está acotada para todo $g \in L$ por

$$\left(\frac{c\varepsilon^2}{d^{3/2}}\right)^{-1} \int_{f \in \Delta^{-1}L} \sharp(2\varepsilon, f)\ d(\Delta^{-1}L),$$

y el lema se sigue. ∎

El siguiente resultado se demuestra implícitamente en [115]. Incluimos una demostración por completitud de esta memoria.

Proposición 4.2.9 *Sea $\varepsilon > 0$ un número real positivo. Entonces, tenemos:*

$$\int_{L \in \mathcal{L}} \tau(\varepsilon, L)\ d\mathcal{L} \leq \frac{32\pi}{c} d^{3/2} n^3 (n+1) N^2 \mathcal{D} \varepsilon^2,$$

donde $c > 0$ es la constante universal del Lema 4.2.8.

Demostración.– Por el Lema 4.2.8,

$$\int_{L \in \mathcal{L}} \tau(\varepsilon, L)\ d\mathcal{L} \leq \left(\frac{c\varepsilon^2}{d^{3/2}}\right)^{-1} \int_{L \in \mathcal{L}} \int_{f \in \Delta^{-1}L} \sharp(2\varepsilon, f)\ dL_\Delta\ d\mathcal{L} =$$

$$= \left(\frac{c\varepsilon^2}{d^{3/2}}\right)^{-1} \int_{L \in \mathcal{L}} \int_{f \in L} \sharp(2\varepsilon, \Delta^{-1}f)\ dL\ d\mathcal{L}.$$

Por el famoso Teorema de Santaló de Geometría Integral en la esfera (cf. Teorema 1.1.14),

$$\int_{L \in \mathcal{L}} \int_{f \in L} \sharp(2\varepsilon, \Delta^{-1}f)\ dL\ d\mathcal{L} = 2\pi \frac{\int_{f \in \mathbb{S}} \sharp(2\varepsilon, \Delta^{-1}f)\ d\mathbb{S}}{\nu[\mathbb{S}]}.$$

Como Δ^{-1} es una isometría de $\mathbb{S}$ en $\mathbb{S}_\Delta$,

$$\frac{1}{\nu[\mathbb{S}]} \int_{f \in \mathbb{S}} \sharp(2\varepsilon, \Delta^{-1}f)\ d\mathbb{S} = \frac{1}{\nu[\mathbb{S}_\Delta]} \int_{f \in \mathbb{S}_\Delta} \sharp(2\varepsilon, f)\ d\mathbb{S}_\Delta,$$

y el Teorema 4.2.7 implica que:

$$\frac{1}{\nu[\mathbb{S}_\Delta]} \int_{f \in \mathbb{S}_\Delta} \sharp(2\varepsilon, f)\ d\mathbb{S}_\Delta \leq 16 n^3 (n+1) N^2 \mathcal{D} \varepsilon^4.$$

Entonces,

$$\int_{L\in\mathcal{L}} \tau(\varepsilon, L)\ d\mathcal{L} \leq \left(\frac{c\varepsilon^2}{d^{3/2}}\right)^{-1} 32\pi n^3(n+1)N^2\mathcal{D}\varepsilon^4,$$

y la proposición queda demostrada.

■

4.3. Una serie de reducciones geométricas

En esta sección realizaremos una serie de reducciones a partir de los resultados debidos a Shub & Smale expuestos arriba. La expresión final se utilizará en secciones siguientes para demostrar los teoremas principales de la introducción de este capítulo. Cada subsección contiene una de esas reducciones.

4.3.1. De círculos máximos a pares de sistemas

Sea $\mathfrak{D} \subseteq \mathbb{S}\times\mathbb{S}$ el conjunto definido como sigue.

$$\mathfrak{D} := \{(f,g) \in \mathbb{S}\times\mathbb{S} \ :\ [f=g] \vee [f=-g]\}.$$

Definimos la aplicación

$$\mathfrak{L} : \mathbb{S}\times\mathbb{S} \setminus \mathfrak{D} \longrightarrow \mathcal{L},$$

tal que para todo $(f,g) \in \mathbb{S}\times\mathbb{S}\setminus\mathfrak{D}$, el conjunto $\mathfrak{L}(f,g) \in \mathcal{L}$ es el único círculo máximo en $\mathbb{S}$ que contiene a f y a g. También consideramos el conjunto

$$\mathfrak{D}_\Delta := \{(f,g) \in \mathbb{S}_\Delta\times\mathbb{S}_\Delta \ :\ [f=g] \vee [f=-g]\},$$

y la aplicación

$$\mathfrak{L}_\Delta : \mathbb{S}_\Delta\times\mathbb{S}_\Delta \setminus \mathfrak{D}_\Delta \longrightarrow \mathcal{L}_\Delta,$$

tal que para todo $(f,g) \in \mathbb{S}_\Delta\times\mathbb{S}_\Delta \setminus \mathfrak{D}_\Delta$, el conjunto $\mathfrak{L}_\Delta(f,g)$ es el único círculo máximo en $\mathbb{S}_\Delta$ que contiene a f y a g. Demostramos el siguiente resultado de Geometría Integral, en el estilo de los teoremas de Santaló.

Lema 4.3.1 *Sea $\Phi : \mathcal{L} \longrightarrow \mathbb{R}$ una función integrable. Entonces, tenemos la siguiente igualdad:*

$$\frac{\int_{(f,g)\in\mathbb{S}\times\mathbb{S}} \Phi(\mathfrak{L}(f,g))\ d(\mathbb{S}\times\mathbb{S})}{\nu[\mathbb{S}]^2} = \int_{L\in\mathcal{L}} \Phi(L)\ d\mathcal{L}.$$

Demostración.– El Teorema 1.1.13 (Fórmula de la Co–área) aplicado a $\mathfrak{L} : \mathbb{S} \times \mathbb{S} \setminus \mathfrak{D} \longrightarrow \mathcal{L}$, implica que:

$$\int_{(f,g)\in\mathbb{S}\times\mathbb{S}} \Phi(\mathfrak{L}(f,g))\ d(\mathbb{S}\times\mathbb{S}) =$$

$$\int_{L\in\mathcal{L}} \Phi(L) \int_{(f,g)\in\mathfrak{L}^{-1}(L)} \frac{1}{NJ_{(f,g)}\mathfrak{L}}\ d\mathfrak{L}^{-1}(L)\ d\mathcal{L}. \tag{4.8}$$

Comprobamos que la integral interior de esta última expresión es una constante. En efecto, sean $L_1, L_2 \in \mathcal{L}$ dos círculos máximos, y sea $O \in \mathcal{O}_{2N+2}$ una matriz ortogonal tal que $OL_1 = L_2$. Consideremos la siguiente isometría:

$$\begin{array}{rccc} O\times O: & \mathbb{S}\times\mathbb{S}\setminus\mathfrak{D} & \longrightarrow & \mathbb{S}\times\mathbb{S}\setminus\mathfrak{D} \\ & (f,g) & \mapsto & (Of, Og). \end{array}$$

Entonces, $(O\times O)_{|\mathfrak{L}^{-1}(L_1)}$ es una isometría entre $\mathfrak{L}^{-1}(L_1)$ y $\mathfrak{L}^{-1}(L_2)$. De nuevo el Teorema 1.1.13 aplicado a ésta nos proporciona:

$$\int_{(f_1,g_1)\in\mathfrak{L}^{-1}(L_1)} \frac{1}{NJ_{(f_1,g_1)}\mathfrak{L}}\ d\mathfrak{L}^{-1}(L_1) =$$

$$= \int_{(f_2,g_2)\in\mathfrak{L}^{-1}(L_2)} \frac{1}{NJ_{(O^{-1}f_2,O^{-1}g_2)}\mathfrak{L}}\ d\mathfrak{L}^{-1}(L_2)$$

Ahora, sea $(f_2, g_2) \in \mathfrak{L}^{-1}(L_2)$ un par cualquiera. Sean $f_2' = O^{-1}f_2$, $g_2' = O^{-1}g_2$ las preimágenes respectivas mediante O de f_2 y g_2. Observamos que:

$$O\circ\mathfrak{L} = \mathfrak{L}\circ(O\times O), \qquad (O\times O)(f_2', g_2') = (f_2, g_2).$$

Por tanto, el Corolario 1.1.12 garantiza que:

$$NJ_{(f_2,g_2)}\mathfrak{L} = NJ_{(f_2',g_2')}\mathfrak{L} = NJ_{(O^{-1}f_2,O^{-1}g_2)}\mathfrak{L},$$

y deducimos que la integral interior en la ecuación (4.8) es constante. Sea C esa constante. La misma ecuación (4.8), aplicada a la función constante $\Phi \equiv 1$, se transforma en:

$$\int_{(f,g)\in\mathbb{S}\times\mathbb{S}} 1\ d(\mathbb{S}\times\mathbb{S}) = \int_{L\in\mathcal{L}} C\ d\mathcal{L}.$$

Deducimos por tanto que $C = \nu[\mathbb{S}]^2$, y el lema queda demostrado.

■

Proposición 4.3.2 *Sea $\Phi : \mathcal{L} \longrightarrow \mathbb{R}$ una función integrable. Entonces, tenemos:*

$$\frac{\int_{(f,g)\in\mathbb{S}_\Delta\times\mathbb{S}_\Delta} \Phi(\mathfrak{L}(\Delta f, \Delta g))\ d(\mathbb{S}_\Delta\times\mathbb{S}_\Delta)}{\nu[\mathbb{S}_\Delta]^2} = \int_{L\in\mathcal{L}} \Phi(L)\ d\mathcal{L}.$$

Demostración.– La demostración se sigue inmediatamente del Lema 4.3.1, sabiendo que $\Delta^{-1} \times \Delta^{-1}$ define una isometría entre $\mathbb{S} \times \mathbb{S}$ y $\mathbb{S}_\Delta \times \mathbb{S}_\Delta$. ■

Proposición 4.3.3 *Con las notaciones anteriores, tenemos la siguiente desigualdad:*

$$\int_{\mathbb{S}_\Delta \times \mathbb{S}_\Delta} \tau_\varepsilon(\mathfrak{L}_\Delta(f,g)) d\mathbb{S}_\Delta d\mathbb{S}_\Delta \leq \nu_\Delta[\mathbb{S}_\Delta]^2 \frac{32\pi}{c} \varepsilon^2 n^3 (n+1) N^2 \mathcal{D} d^{3/2},$$

donde τ_ε es la función definida en la Subsección 4.2.2.

Demostración.– Obsérvese que $\tau_\varepsilon(\mathfrak{L}_\Delta(f,g)) = \tau(\varepsilon, \mathfrak{L}(\Delta f, \Delta g))$, definido en la Subsección 4.2.4. Por la Proposición 4.3.2, tenemos:

$$\int_{(f,g)\in \mathbb{S}_\Delta \times \mathbb{S}_\Delta} \tau(\varepsilon, \mathfrak{L}(\Delta f, \Delta g)) d\mathbb{S}_\Delta d\mathbb{S}_\Delta = \nu_\Delta[\mathbb{S}_\Delta]^2 \int_{L\in\mathcal{L}} \tau(\varepsilon, L) d\mathcal{L}.$$

La desigualdad se sigue de la Proposición 4.2.9. ■

4.3.2. De pares de sistemas a la fibra en e_0

Tenemos el siguiente resultado, que es la versión para la esfera del Teorema 3.5.2 (enunciado para el espacio proyectivo complejo). La demostración es idéntica a la de ese resultado. En esta ocasión, utilizamos las proyecciones p_1 y p_2 con dominio en la esfera $\mathbb{S}_\Delta$.

Proposición 4.3.4 *Sea $\Phi : W \longrightarrow \mathbb{R}_+$ una función integrable, tal que para toda matriz unitaria $U \in \mathcal{U}_{n+1}$ se tiene que:*

$$\Phi(f,\zeta) = \Phi(f \circ U, U^{-1}\zeta).$$

Sea $\mathcal{J}$ la integral definida a continuación:

$$\mathcal{J} := \int_{(f,\zeta)\in W} \Phi(f,\zeta) N J_{(f,\zeta)} p_1 \; dW.$$

Entonces, se tienen las siguientes igualdades:

$$\mathcal{J} = \int_{f\in\mathbb{S}_\Delta} \sum_{\zeta\in V(f)} \Phi(f,\zeta) \; d\mathbb{S}_\Delta,$$

$$\mathcal{J} = \vartheta_n \int_{f\in V_{e_0}} \Phi(f,e_0) \frac{N J_{(f,e_0)} p_1}{N J_{(f,e_0)} p_2} \; dV_{e_0}.$$

Usaremos esta proposición de un modo similar a como usamos el Teorema 3.5.2 en el Capítulo 3. La siguiente definición juega un papel principal en nuestras demostraciones posteriores. Para cada $(g,\zeta) \in W$, sea

$$A_\varepsilon(g,\zeta) := \frac{1}{\nu_\Delta[\mathbb{S}_\Delta]} \int_{f\in\mathbb{S}_\Delta} \chi_\varepsilon(\mu_{\mathrm{norm}}(f,g,\zeta))d\mathbb{S}_\Delta,$$

donde χ_ε es la función característica de intervalo $(\varepsilon^{-1},\infty]$ y $\mu_{\mathrm{norm}}(f,g,\zeta)$ es como en la nota 4.2.4. Obsérvese que para un par dado (g,ζ), la cantidad $A_\varepsilon(g,\zeta)$ es la probabilidad de que alcancemos un sistema al azar $f \in \mathbb{S}_\Delta$ de modo que encontremos un condicionamiento $\mu_{\mathrm{norm}} > \varepsilon^{-1}$ en el camino. Por lo tanto, en virtud de la Proposición 4.2.6, si encontramos un par (g,ζ) tal que $A_\varepsilon(g,\zeta)$ sea muy pequeño, habremos encontrado un buen punto inicial para comenzar la homotopía.

Proposición 4.3.5 *Con las notaciones anteriores, se cumplen las siguientes propiedades:*

- *La cantidad $A_\varepsilon(g,\zeta)$ es unitaria invariante, esto es, para toda matriz unitaria $U \in \mathcal{U}_{n+1}$ se tiene:*

$$A_\varepsilon(g,\zeta) = A_\varepsilon(g\circ U, U^{-1}\zeta).$$

- *Se tiene la siguiente desigualdad,*

$$\vartheta_n \int_{V_{e_0}} A_\varepsilon(g,e_0)DET(g,e_0)dV_{e_0} \leq \frac{32\pi}{c}\nu_\Delta[\mathbb{S}_\Delta]\varepsilon^2 n^3(n+1)N^2\mathcal{D}d^{3/2}.$$

Demostración.– Primero, observamos que para cualquier matriz unitaria $U \in \mathcal{U}_{n+1}$, se tiene que

$$\chi_\varepsilon(\mu_{\mathrm{norm}}(f,g,\zeta)) = \chi_\varepsilon(\mu_{\mathrm{norm}}(f\circ U, g\circ U, U^{-1}\zeta)). \tag{4.9}$$

En efecto, supongamos que $\chi_\varepsilon(\mu_{\mathrm{norm}}(f,g,\zeta)) = 1$. Entonces, existe un sistema $h \in \mathfrak{L}_\Delta(f,g)$ y una solución $\zeta' \in V(h)$ tal que (h,ζ') está en el arco abierto de $p_1^{-1}(\mathfrak{L}_\Delta(f,g)) \setminus p_1^{-1}(\{-g\})$ que contiene al punto (g,ζ) y tal que

$$\mu_{\mathrm{norm}}(h,\zeta') > \varepsilon^{-1}.$$

Ahora, el punto $(h\circ U, U^{-1}\zeta')$ está en el arco abierto de $p_1^{-1}(\mathfrak{L}_\Delta(f\circ U, g\circ U)) \setminus p_1^{-1}(\{-g\circ U\})$ que contiene a $(g\circ U, U^{-1}\zeta)$, y además

$$\mu_{\mathrm{norm}}(h\circ U, U^{-1}\zeta') = \mu_{\mathrm{norm}}(h,\zeta') > \varepsilon^{-1}.$$

Por lo tanto, $\chi_\varepsilon(\mu_{\mathrm{norm}}(f\circ U, g\circ U, U^{-1}\zeta)) = 1$. Queda demostrado que

$$\chi_\varepsilon(\mu_{\mathrm{norm}}(f,g,\zeta)) \leq \chi_\varepsilon(\mu_{\mathrm{norm}}(f\circ U, g\circ U, U^{-1}\zeta)).$$

Un razonamiento simétrico demuestra la otra desigualdad, y con ella la igualdad (4.9). Por otro lado, por el Teorema 1.1.13,

$$A_\varepsilon(g,\zeta) = \frac{1}{\nu_\Delta[\mathbb{S}_\Delta]} \int_{f\in\mathbb{S}_\Delta} \chi_\varepsilon(\mu_{\mathrm{norm}}(f,g,\zeta))d\mathbb{S}_\Delta =$$

$$\frac{1}{\nu_\Delta[\mathbb{S}_\Delta]} \int_{f\in\mathbb{S}_\Delta} \chi_\varepsilon(\mu_{\mathrm{norm}}(f\circ U^{-1},g,\zeta))d\mathbb{S}_\Delta.$$

De la igualdad (4.9) deducimos que

$$A_\varepsilon(g,\zeta) = \frac{1}{\nu_\Delta[\mathbb{S}_\Delta]} \int_{f\in\mathbb{S}_\Delta} \chi_\varepsilon(\mu_{\mathrm{norm}}(f,g\circ U,U^{-1}\zeta))d\mathbb{S}_\Delta = A_\varepsilon(g\circ U,U^{-1}\zeta).$$

Esto es, A_ε es unitaria invariante. Por las proposiciones 4.3.4 y 4.2.5, esto implica que

$$\vartheta_n \int_{V_{e_0}} A_\varepsilon(g,e_0)DET(g,e_0)dV_{e_0} = \int_{g\in\mathbb{S}_\Delta} \sum_{\zeta\in V(g)} A_\varepsilon(g,\zeta)\; d\mathbb{S}_\Delta.$$

Por definición de A_ε, esta última integral es igual a

$$\frac{1}{\nu_\Delta[\mathbb{S}_\Delta]} \int_{g\in\mathbb{S}_\Delta} \sum_{\zeta\in V(g)} \int_{f\in\mathbb{S}_\Delta} \chi_\varepsilon(\mu_{\mathrm{norm}}(f,g,\zeta))d\mathbb{S}_\Delta\; d\mathbb{S}_\Delta =$$

$$\frac{1}{\nu_\Delta[\mathbb{S}_\Delta]} \int_{g\in\mathbb{S}_\Delta} \int_{f\in\mathbb{S}_\Delta} \sum_{\zeta\in V(g)} \chi_\varepsilon(\mu_{\mathrm{norm}}(f,g,\zeta))d\mathbb{S}_\Delta\; d\mathbb{S}_\Delta$$

Por otro lado, para todo par de sistemas $(f,g)\in\mathbb{S}_\Delta\times\mathbb{S}_\Delta\setminus\mathfrak{D}_\Delta$ se tiene que

$$\tau_\varepsilon(\mathfrak{L}_\Delta(f,g)) \geq \tau_\varepsilon^g(\mathfrak{L}_\Delta(f,g)) = \sum_{\zeta\in V(g)} \chi_\varepsilon(\mu_{\mathrm{norm}}(f,g,\zeta)).$$

Por lo tanto, hemos demostrado que

$$\vartheta_n \int_{V_{e_0}} A_\varepsilon(g,e_0)DET(g,e_0)dV_{e_0} \leq \frac{1}{\nu_\Delta[\mathbb{S}_\Delta]} \int_{\mathbb{S}_\Delta\times\mathbb{S}_\Delta} \tau_\varepsilon(\mathfrak{L}_\Delta(f,g))\; d\mathbb{S}_\Delta d\mathbb{S}_\Delta.$$

El resultado se sigue de la Proposición 4.3.3. ∎

4.3.3. De la fibra en e_0 al espacio de matrices cuadradas

Recuperamos las notaciones de la Sección 4.2. Consideramos la proyección ortogonal

$$\pi_{(d)} : \widetilde{V_{e_0}} \longrightarrow L_{e_0}.$$

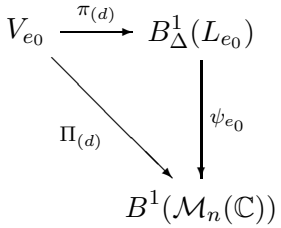
Esta proyección induce una proyección ortogonal que denotamos por el mismo símbolo,

$$\pi_{(d)} : V_{e_0} \longrightarrow B^1_\Delta(L_{e_0}),$$

donde $B^1_\Delta(L_{e_0})$ es la bola unidad en L_{e_0} con respecto a la métrica de Kostlan en L_{e_0}. También consideramos la aplicación ψ_{e_0} definida en la Sección 4.2, y la aplicación

$$\Pi_{(d)} := \psi_{e_0} \circ \pi_{(d)} : V_{e_0} \longrightarrow B^1(\mathcal{M}_n(\mathbb{C})),$$

donde $B^1(\mathcal{M}_n(\mathbb{C}))$ es la bola cerrada de radio 1 en $\mathcal{M}_n(\mathbb{C})$ con respecto a la norma de Frobenius. La situación que hemos descrito se resume en el siguiente diagrama conmutativo.



Con las notaciones de la Sección 4.2, tenemos:

$$\Pi_{(d)}(g) = \Delta(d)^{-1/2} T_{e_0} g,$$

por lo que $\Pi_{(d)}(g) = 0$ para todo $g \in L_{e_0}^\perp$.
En particular, para toda matriz $M \in \mathcal{M}_n(\mathbb{C})$ tal que $||M||_F = t \leq 1$, la fibra $\Pi^{-1}_{(d)}(M)$ puede identificarse con una esfera en $L_{e_0}^\perp$ de radio $(1-t^2)^{1/2}$. Esto es, tenemos:

$$\Pi^{-1}_{(d)}(M) \sim S^{(1-t^2)^{1/2}}_\Delta(L_{e_0}^\perp) \equiv \{(1-t^2)^{1/2} h \ : \ h \in L_{e_0}^\perp, ||h||_\Delta = 1\}.$$

Podemos calcular el volumen de esa esfera, pues conocemos su radio y su dimensión (real) es $2N - 2n - 2n^2 + 1$. Por tanto, su volumen es

$$(1 - ||M||_F^2)^{N-n^2-n+\frac{1}{2}} \nu_\Delta[S^1_\Delta(L_{e_0}^\perp)].$$

Además, observamos que para cada $g \in \widetilde{V_{e_0}}$, se tiene la siguiente igualdad:

$$DET(g, e_0) := det((T_{e_0} g)(T_{e_0} g)^*) = \mathcal{D} det((\Pi_{(d)}(g))(\Pi_{(d)}(g))^*),$$

puesto que $det(\Delta(d))^{1/2}) = \mathcal{D}^{1/2}$.
Se puede comprobar fácilmente (véase por ejemplo [14]) que el jacobiano normal de $\Pi_{(d)}$ en un punto $g \in V_{e_0}$ satisface:

$$NJ_g \Pi_{(d)} = NJ_g \pi_{(d)} = (1 - ||\pi_{(d)}(g)||^2_\Delta)^{1/2} = (1 - ||\Pi_{(d)}(g)||^2_F)^{1/2}.$$

En efecto, basta notar que $\Pi_{(d)}$ no es más que la composición de una proyección ortogonal restringida a la esfera con la isometría ψ_{e_0}. Entonces, el siguiente resultado se sigue de la Proposición 4.3.5 y del Teorema 1.1.13 aplicado a $\Pi_{(d)}$. Obsérvese que en él hacemos desaparecer el número de Bézout $\mathcal{D}$ de la cota superior.

Proposición 4.3.6 *Con las notaciones anteriores, tenemos:*

$$\frac{\vartheta_n \nu_\Delta[S^1_\Delta(L^\perp_{e_0})]}{\nu_\Delta[\mathbb{S}_\Delta]} \int_{B^1(\mathcal{M}_n(\mathbb{C}))} det(MM^*)(1-||M||_F^2)^{N-n^2-n} B_\varepsilon(M,e_0) d\mathcal{M}_n(\mathbb{C})$$

$$\leq \frac{32\pi}{c}\varepsilon^2 n^3(n+1)N^2 d^{3/2},$$

donde $\nu_\Delta[S^1_\Delta(L^\perp_{e_0})]$ *es el volumen de* $S^1_\Delta(L^\perp_{e_0})$ *como subespacio lineal de* $\mathbb{S}_\Delta$, *y*

$$B_\varepsilon(M,e_0) := \frac{1}{\nu_\Delta[S^1_\Delta(L^\perp_{e_0})]} \int_{h\in S^1_\Delta(L^\perp_{e_0})} A_\varepsilon((1-||M||_F^2)^{1/2}h + \psi^{-1}_{e_0}(M), e_0) dS^1_\Delta(L^\perp_{e_0}).$$

Demostración.– Por el Teorema 1.1.13 aplicado a $\Pi_{(d)}$ concluimos que

$$\int_{V_{e_0}} A_\varepsilon(g,e_0) DET(g,e_0) dV_{e_0} =$$

$$\int_{M\in B^1(\mathcal{M}_n(\mathbb{C}))} \int_{g\in\Pi^{-1}_{(d)}(M)} A_\varepsilon(g,e_0) \frac{DET(g,e_0)}{NJ_g\Pi_{(d)}}\, d\Pi^{-1}_{(d)}(M)\, d\mathcal{M}_n(\mathbb{C}) =$$

$$\mathcal{D} \int_{M\in B^1(\mathcal{M}_n(\mathbb{C}))} \frac{\det(MM^*)}{(1-\|M\|_F)^{1/2}} \int_{g\in\Pi^{-1}_{(d)}(M)} A_\varepsilon(g,e_0)\, d\Pi^{-1}_{(d)}(M)\, d\mathcal{M}_n(\mathbb{C}).$$

Por las observaciones hechas antes de la proposición, se tiene que

$$\int_{g\in\Pi^{-1}_{(d)}(M)} A_\varepsilon(g,e_0)\, d\Pi^{-1}_{(d)}(M) =$$

$$\int_{h\in S_\Delta^{(1-|M|_F^2)^{1/2}}(L^\perp_{e_0})} A_\varepsilon(h+\psi^{-1}_{e_0}(M), e_0) dS^1_\Delta(L^\perp_{e_0}) =$$

$$(1-||M||_F^2)^{N-n^2-n+\frac{1}{2}} \int_{h\in S^1_\Delta(L^\perp_{e_0})} A_\varepsilon((1-||M||_F^2)^{1/2}h + \psi^{-1}_{e_0}(M), e_0) dS^1_\Delta(L^\perp_{e_0}) =$$

$$(1-||M||_F^2)^{N-n^2-n+\frac{1}{2}} \nu_\Delta[S^1_\Delta(L^\perp_{e_0})] B_\varepsilon(M,e_0).$$

Concluimos que

$$\int_{V_{e_0}} A_\varepsilon(g,e_0) DET(g,e_0) dV_{e_0} =$$

$$\nu_\Delta[S^1_\Delta(L^\perp_{e_0})]\mathcal{D}\int_{M\in B^1(\mathcal{M}_n(\mathbb{C}))}(1-||M||^2_F)^{N-n^2-n}\det(MM^*)B_\varepsilon(M,e_0)\ d\mathcal{M}_n(\mathbb{C}).$$

Ahora, por la Proposición 4.3.5, sabemos que

$$\int_{V_{e_0}} A_\varepsilon(g,e_0)DET(g,e_0)dV_{e_0}\leq\frac{32\pi}{c\vartheta_n}\nu_\Delta[\mathbb{S}_\Delta]\varepsilon^2n^3(n+1)N^2\mathcal{D}d^{3/2}.$$

Queda por tanto demostrada la siguiente desigualdad,

$$\int_{M\in B^1(\mathcal{M}_n(\mathbb{C}))}(1-||M||^2_F)^{N-n^2-n}\det(MM^*)B_\varepsilon(M,e_0)\ d\mathcal{M}_n(\mathbb{C})\leq$$

$$\frac{32\pi}{c\vartheta_n\nu_\Delta[S^1_\Delta(L^\perp_{e_0})]}\nu_\Delta[\mathbb{S}_\Delta]\varepsilon^2n^3(n+1)N^2d^{3/2},$$

y con ella el resultado. ∎

Finalmente, usando coordenadas esféricas obtenemos el siguiente resultado.

Proposición 4.3.7 *Con las notaciones anteriores, se tiene que*

$$\int_0^1(1-t^2)^{N-n^2-n}t^{2n^2+2n-1}K_\varepsilon(t,e_0)dt\leq$$

$$\leq\frac{\nu_\Delta[\mathbb{S}_\Delta]}{\nu_\Delta[S^1_\Delta(L^\perp_{e_0})]\nu[S^1(\mathcal{H}_{(1)})]}\frac{32\pi}{c}\varepsilon^2n^3(n+1)N^2d^{3/2},$$

donde c es la constante del Lema 4.2.8, y

$$K_\varepsilon(t,e_0)=\frac{\vartheta_n}{\nu[S^1(\mathcal{H}_{(1)})]}\int_{S^1(\mathcal{M}_n(\mathbb{C}))}det(MM^*)B_\varepsilon(tM,e_0)dS^1(\mathcal{M}_n(\mathbb{C})).$$

Demostración.– La integral de la Proposición 4.3.6, en coordenadas esféricas se escribe como sigue,

$$\int_{B^1(\mathcal{M}_n(\mathbb{C}))}det(MM^*)(1-||M||^2_F)^{N-n^2-n}B_\varepsilon(M,e_0)\ d\mathcal{M}_n(\mathbb{C})=$$

$$\int_0^1(1-t^2)^{N-n^2-n}\int_{M\in S^t(\mathcal{M}_n(\mathbb{C}))}\det(MM^*)B_\varepsilon(M,e_0)\ dS^t(\mathcal{M}_n(\mathbb{C}))\ dt=$$

$$\int_0^1(1-t^2)^{N-n^2-n}t^{2n^2-1+2n}\int_{M\in S^1(\mathcal{M}_n(\mathbb{C}))}\det(MM^*)B_\varepsilon(tM,e_0)\ dS^1(\mathcal{M}_n(\mathbb{C}))\ dt=$$

$$\frac{\nu[S^1(\mathcal{H}_{(1)})]}{\vartheta_n}\int_0^1(1-t^2)^{N-n^2-n}t^{2n^2-1+2n}K_\varepsilon(t,e_0)\ dt.$$

Para obtener la segunda de las igualdades de esta última serie, simplemente transformamos la integral en la esfera de radio t en la integral en la esfera

de radio unidad. El factor de escala de dicha transformación (esto es, el inverso del jacobiano normal) es t^{2n^2-1}, y en t^{2n} restante viene de la igualdad $\det(tM(tM)^*) = t^{2n}\det(MM^*)$. Por la Proposición 4.3.6, tenemos:

$$\int_{B^1(\mathcal{M}_n(\mathbb{C}))} det(MM^*)(1-||M||_F^2)^{N-n^2-n}B_\varepsilon(M,e_0)d\mathcal{M}_n(\mathbb{C})$$

$$\leq \frac{\nu_\Delta[\mathbb{S}_\Delta]}{\vartheta_n\nu_\Delta[S^1_\Delta(L^\perp_{e_0})]}\frac{32\pi}{c}\varepsilon^2n^3(n+1)N^2d^{3/2}.$$

Concluimos por tanto:

$$\int_0^1 (1-t^2)^{N-n^2-n}t^{2n^2-1+2n}K_\varepsilon(t,e_0)\ dt \leq$$

$$\frac{\vartheta_n}{\nu[S^1(\mathcal{H}_{(1)})]}\frac{\nu_\Delta[\mathbb{S}_\Delta]}{\vartheta_n\nu_\Delta[S^1_\Delta(L^\perp_{e_0})]}\frac{32\pi}{c}\varepsilon^2n^3(n+1)N^2d^{3/2},$$

de donde se sigue el corolario. ∎

4.3.4. Del espacio de matrices cuadradas a los sistemas de ecuaciones lineales sub–determinados

El objetivo de esta subsección es demostrar el Corolario 4.3.10, que proporciona una definición alternativa para la cantidad $K_\varepsilon(t,e_0)$. Obsérvese que en nuestras notaciones las estructuras Riemannianas usual y de Kostlan en $\mathcal{H}_{(1)}$ coinciden. Esto es, para $M, M_1 \in \mathcal{H}_{(1)}$, tenemos el producto hermitiano

$$\langle M, M_1\rangle_F = Tr(M^*M_1).$$

Sea $W_{(1)} := \{(M,x) \in S^1(\mathcal{H}_{(1)}) \times \mathbb{P}_n(\mathbb{C}) : Mx = 0\}$ la variedad de incidencia. Para cualquier matriz $M \in \mathcal{H}_{(1)}$ de rango igual a n, consideramos el número:

$$\widetilde{B}_\varepsilon(t,M) := B_\varepsilon(tT_{e_0}(MU), e_0),$$

donde $U \in \mathcal{U}_{n+1}$ es una matriz unitaria cualquiera tal que $MUe_0 = 0$ y $T_{e_0}(MU)$ es la restricción $MU\,|_{e_0^\perp}$. En otras palabras, $T_{e_0}(MU)$ es la matriz cuadrada consistente en las últimas n columnas de MU. El lema siguiente demuestra que $\widetilde{B}_\varepsilon$ está bien definida.

Lema 4.3.8 *Sea $M \in \mathcal{H}_{(1)}$ una matriz, $rank(M) = n$, y sea $0 \leq t \leq 1$ un número real positivo. Sean $U_1, U_2 \in \mathcal{U}_{n+1}$ dos matrices unitarias tales que $MU_1e_0 = MU_2e_0 = 0$. Entonces, tenemos:*

$$B_\varepsilon(tT_{e_0}(MU_1), e_0) = B_\varepsilon(tT_{e_0}(MU_2), e_0).$$

Además, para cualquier matriz unitaria $U \in \mathcal{U}_{n+1}$ también se satisface la siguiente igualdad:

$$\widetilde{B}_\varepsilon(t,M) = \widetilde{B}_\varepsilon(t,MU).$$

Demostración.– La segunda afirmación es una consecuencia inmediata de la primera. Demostramos por tanto la primera de ellas. Observamos que

$$\frac{B_\varepsilon(tT_{e_0}(MU_1),e_0)}{B_\varepsilon(tT_{e_0}(MU_2),e_0)}=\frac{\int_{h\in\Pi_{(d)}^{-1}(tT_{e_0}(MU_1))}A_\varepsilon(h,e_0)\ d(\Pi_{(d)}^{-1}(tT_{e_0}(MU_1)))}{\int_{h\in\Pi_{(d)}^{-1}(tT_{e_0}(MU_2))}A_\varepsilon(h,e_0)\ d(\Pi_{(d)}^{-1}(tT_{e_0}(MU_2)))}. \tag{4.10}$$

Sea $U:=U_1^{-1}U_2\in\mathcal{U}_{n+1}$ una matriz unitaria tal que $U_1U=U_2$. Observamos que $Ue_0=e_0\in\mathbb{P}_n(\mathbb{C})$. Además, la siguiente aplicación es una isometría:

$$\begin{array}{ccc}\Pi_{(d)}^{-1}(tT_{e_0}(MU_1)) & \longrightarrow & \Pi_{(d)}^{-1}(tT_{e_0}(MU_2))\\ h & \mapsto & h\circ U.\end{array}$$

Por lo tanto, por el Teorema 1.1.13, la expresión en la ecuación (4.10) es igual a:

$$\frac{\int_{h\in\Pi_{(d)}^{-1}(tT_{e_0}(MU_2))}A_\varepsilon(h\circ U^{-1},e_0)\ d(\Pi_{(d)}^{-1}(tT_{e_0}(MU_2)))}{\int_{h\in\Pi_{(d)}^{-1}(tT_{e_0}(MU_2))}A_\varepsilon(h,e_0)\ d(\Pi_{(d)}^{-1}(tT_{e_0}(MU_2)))}.$$

Ahora, por la Proposición 4.3.5, $A_\varepsilon(h\circ U^{-1},e_0)=A_\varepsilon(h\circ U^{-1},Ue_0)=A_\varepsilon(h,e_0)$, y el lema queda demostrado.

■

Proposición 4.3.9 *Sean $t,\varepsilon>0$ dos números reales positivos, $0<t\leq 1$. Sea $K_\varepsilon(t,e_0)$ como en la Proposición 4.3.7. Entonces, tenemos:*

$$K_\varepsilon(t,e_0)=\frac{1}{\nu[S^1(\mathcal{H}_{(1)})]}\int_{M\in S^1(\mathcal{H}_{(1)})}\widetilde{B}_\varepsilon(t,M)\ dS^1(\mathcal{H}_{(1)}).$$

Demostración.– Para cualquier matriz $M\in\mathcal{M}_n(\mathbb{C})$, sea $(0,M)\in\mathcal{H}_{(1)}$ la matriz que se obtiene añadiendo M una primera columna de ceros. Consideramos los siguientes esquemas:

$$\begin{array}{ccc}W_{(1)} & \xrightarrow{\ i\ } & \mathcal{H}_{(1)}\times\mathbb{P}_n(\mathbb{C}),\\ & {\scriptstyle\phi_1}\searrow & \downarrow{\scriptstyle\pi_1}\\ & & \mathcal{H}_{(1)}\end{array}\qquad\qquad\begin{array}{ccc}W_{(1)} & \xrightarrow{\ i\ } & \mathcal{H}_{(1)}\times\mathbb{P}_n(\mathbb{C}),\\ & {\scriptstyle\phi_2}\searrow & \downarrow{\scriptstyle\pi_2}\\ & & \mathbb{P}_n(\mathbb{C})\end{array}$$

donde $i:W_{(1)}\longrightarrow\mathcal{H}_{(1)}\times\mathbb{P}_n(\mathbb{C})$ es la inclusión, y $\phi_1:W_{(1)}\longrightarrow\mathcal{H}_{(1)}$ y $\phi_2:W_{(1)}\longrightarrow\mathbb{P}_n(\mathbb{C})$ son las proyecciones canónicas. El valor de los jacobianos normales fue calculado en la Proposición 4.2.5, aplicada al caso particular de que $(d)=(1,\ldots,1)$:

$$\frac{NJ_{((0,M),e_0)}\phi_1}{NJ_{((0,M),e_0)}\phi_2}=det(MM^*),$$

para toda matriz regular $M \in S^1(\mathcal{M}_n(\mathbb{C}))$. Además, para toda matriz unitaria $U \in \mathcal{U}_{n+1}$, el Lema 4.3.8 nos garantiza que

$$\widetilde{B}_\varepsilon(t, M) = \widetilde{B}_\varepsilon(t, MU).$$

Por lo tanto, por la Proposición 4.3.4, tenemos:

$$\int_{M \in S^1(\mathcal{H}_{(1)})} \widetilde{B}_\varepsilon(t, M)\ dS^1(\mathcal{H}_{(1)}) =$$

$$\vartheta_n \int_{S^1(\mathcal{M}_n(\mathbb{C}))} \frac{NJ_{((0,M),e_0)}\phi_1}{NJ_{((0,M),e_0)}\phi_2} \widetilde{B}_\varepsilon(t, (0, M)) dS^1(\mathcal{M}_n(\mathbb{C})) =$$

$$\vartheta_n \int_{S^1(\mathcal{M}_n(\mathbb{C}))} \det(MM^*) \widetilde{B}_\varepsilon(t, (0, M)) dS^1(\mathcal{M}_n(\mathbb{C})),$$

Ahora, observamos que $\widetilde{B}_\varepsilon(t, (0, M)) = B_\varepsilon(tM, e_0)$ y la proposición queda demostrada.

■

Corolario 4.3.10 *Sean $\varepsilon > 0$ y $t \in (0, 1)$ dos números reales. Entonces, se tiene la siguiente igualdad:*

$$K_\varepsilon(t, e_0) = \frac{1}{\nu[S^1(\mathcal{H}_{(1)})]} \int_{M \in S^1(\mathcal{H}_{(1)})} \frac{1}{\nu_\Delta[S^1_\Delta(L^\perp_{e_0})]} \mathcal{I}(M, t)\ dS^1(\mathcal{H}_{(1)}),$$

donde $\mathcal{I}(M, t)$ es la siguiente integral:

$$\int_{h \in S^1_\Delta(L^\perp_{e_0})} A_\varepsilon \left((1 - t^2)^{1/2} h + t\psi^{-1}_{e_0} \left(T_{e_0}(M\Omega(M)) \right), e_0 \right)\ dS^1_\Delta(L^\perp_{e_0}).$$

Aquí, $\Omega(M) \in \mathcal{U}_{n+1}$ es una matriz unitaria cualquiera tal que $M\Omega(M)e_0 = 0$.

Demostración.– Por la Proposición 4.3.9 y la definición de $\tilde{B}_\varepsilon(t, M)$, se tiene:

$$K_\varepsilon(t, e_0) = \frac{1}{\nu[S^1(\mathcal{H}_{(1)})]} \int_{M \in S^1(\mathcal{H}_{(1)})} B_\varepsilon(tT_{e_0}(M\Omega(M)), e_0)\ dS^1(\mathcal{H}_{(1)}).$$

Ahora, por definición $B_\varepsilon(tT_{e_0}(M\Omega(M)), e_0)$ es igual a $\frac{1}{\nu_\Delta[S^1_\Delta(L^\perp_{e_0})]}$ multiplicado por

$$\int_{h \in S^1_\Delta(L^\perp_{e_0})} A_\varepsilon((1 - ||tT_{e_0}(M\Omega(M))||^2_F)^{1/2} h + \psi^{-1}_{e_0}(tT_{e_0}(M\Omega(M))), e_0) dS^1_\Delta(L^\perp_{e_0}).$$

Como M es un elemento de $S^1(\mathcal{H}_{(1)})$, su norma $\|M\|_F$ vale 1, y lo mismo sucede con $\|T_{e_0}(M\Omega(M))\|_F$. De aquí se sigue el corolario. ■

4.4. El último paso argumental

Recordamos ahora la construcción del conjunto questor hecha en la introducción de este capítulo. Sea Y el siguiente conjunto compacto afín

$$Y := [0,1] \times B^1(L_{e_0}^{\perp}) \times B^1(\mathcal{H}_{(1)}) \subseteq \mathbb{R} \times \mathbb{C}^{N+1},$$

donde $B^1(L_{e_0}^{\perp})$ es la bola cerrada de radio 1 en $L_{e_0}^{\perp}$ con respecto a la métrica canónica y $B^1(\mathcal{H}_{(1)})$ es la bola cerrada de radio 1 en el espacio de matrices complejas de talla $n \times (n+1)$ con respecto a la norma de Frobenius. Consideramos Y con la estructura Riemanniana producto. Sea $\tau \in \mathbb{R}^+$ el número real definido como

$$\tau := \sqrt{\frac{n^2+n}{N}}.$$

Fijemos una aplicación cualquiera $\Omega : \mathcal{H}_{(1)} \longrightarrow \mathcal{U}_{n+1}$ tal que para cada matriz $M \in \mathcal{M}_{n\times(n+1)}(\mathbb{C})$, $\Omega(M) \in \mathcal{U}_{n+1}$ es una matriz unitaria tal que $M\Omega(M)e_0 = 0$. Entonces, definimos la aplicación $G_{(d)} : Y \longrightarrow V_{e_0}$ como sigue: Para cada punto $(t,h,M) \in Y$, $G_{(d)}(t,h,M) \in V_{e_0}$ es igual a

$$\left(1-\tau^2 t^{\frac{1}{n^2+n}}\right)^{1/2} \frac{\Delta^{-1}h}{\|h\|_2} + \tau t^{\frac{1}{2n^2+2n}} \psi_{e_0}^{-1}\left(T_{e_0}\left(\frac{M}{\|M\|_F}\Omega\left(\frac{M}{\|M\|_F}\right)\right)\right).$$

Observamos que $G_{(d)}$ no está definida en los casos $M = 0$ o $h = 0$. Nuestro objetivo son las estimaciones de probabilidad, con lo que podemos simplemente omitir esos casos. En esta sección demostramos el siguiente resultado (que es una versión más precisa del Teorema 4.1.6 de la introducción de este capítulo).

Teorema 4.4.1 *Con las notaciones anteriores, para un punto elegido al azar $(t,h,M) \in Y$, la probabilidad de que*

$$g := G_{(d)}(t,h,M) \in V_{e_0}$$

satisfaga: (g, e_0) es ε–eficiente para NHD*, es al menos*

$$1-\varepsilon.$$

Además, en ese caso, la probabilidad de que un input elegido al azar $f \in \mathbb{S}_{\Delta}$ sea resuelto por NHD *con par inicial (g, e_0) realizando $6 \times 10^4 n^{7/2}(n+1)^{3/2}N^2d^3\varepsilon^{-2}$ pasos de homotopía es a lo más*

$$1-\varepsilon.$$

Para demostrar este resultado, haremos uso de algunos resultados técnicos que se exponen a continuación.

Lema 4.4.2 *Sea $f : [0,1] \longrightarrow \mathbb{R}_+$ una función real positiva integrable. Supongamos que existen dos números naturales $M > 0, p > 1$ tales que se satisface*

$$\int_0^1 (1-t^2)^M t^{p-1} f(t)dt \leq H.$$

Entonces, para todo número real $t_0 < 1$, se tiene la siguiente desigualdad:

$$(1-t_0^{2/p})^M \int_0^{t_0} f(t^{1/p})dt \leq pH.$$

Demostración.– Obsérvese que

$$p\int_0^1 (1-t^2)^M t^{p-1} f(t)dt = \int_0^1 (1-t^{2/p})^M f(t^{1/p})dt.$$

Por lo tanto,

$$(1-t_0^{2/p})^M \int_0^{t_0} f(t^{1/p})dt \leq \int_0^{t_0} (1-t^{2/p})^M f(t^{1/p})dt \leq pH.$$

■

Corolario 4.4.3 *Con las notaciones anteriores, para todo número real positivo $t_0 \in (0,1)$ tenemos la siguiente desigualdad:*

$$\left(1-t_0^{\frac{1}{n^2+n}}\right)^{N-n^2-n} \int_0^{t_0} K_\varepsilon\left(t^{\frac{1}{2n^2+2n}}, e_0\right) dt \leq$$

$$\leq \frac{\nu_\Delta[\mathbb{S}_\Delta](2n^2+2n)}{\nu_\Delta[S^1_\Delta(L^\perp_{e_0})]\nu[S^1(\mathcal{H}_{(1)})]} \frac{32\pi}{c}\varepsilon^2 n^3(n+1)N^2 d^{3/2},$$

donde K_ε es la función definida en la Proposición 4.3.7. Más aún, sea t_0 el número real definido como sigue

$$t_0 := \left(\frac{n^2+n}{N}\right)^{n^2+n}.$$

Entonces, se tiene:

$$\frac{1}{t_0}\int_0^{t_0} K_\varepsilon\left(t^{\frac{1}{2n^2+2n}}, e_0\right) dt \leq \frac{\sqrt{2\pi}e^{1/6}32\pi}{c}\varepsilon^2 n^{7/2}(n+1)^{3/2}N^2 d^{3/2},$$

donde c es la constante del Lema 4.2.8.

Demostración.– La primera desigualdad es consecuencia inmediata de la Proposición 4.3.7 y el Lema 4.4.2. En cuanto a la segunda, observamos que

$$t_0\left(1-t_0^{\frac{1}{n^2+n}}\right)^{N-n^2-n}=\frac{(N-n^2-n)^{N-n^2-n}(n^2+n)^{n^2+n}}{N^N}.$$

La desigualdad de Stirling (como se utiliza por ejemplo en [120]), nos proporciona una cota inferior para esta cantidad:

$$\left[\sqrt{2\pi}e^{1/6}\sqrt{n^2+n}\binom{N}{n^2+n}\right]^{-1}.$$

Por otro lado,

$$\frac{\nu_\Delta[\mathbb{S}_\Delta]}{\nu_\Delta[S^1_\Delta(L^\perp_{e_0})]\nu[S^1(\mathcal{H}_{(1)})]}=\frac{1}{2n^2+2n}\binom{N}{n^2+n}^{-1},$$

y obtenemos que

$$\frac{1}{t_0}\int_0^{t_0}K_\varepsilon\left(t^{\frac{1}{2n^2+2n}},e_0\right)dt\leq\sqrt{2\pi}e^{1/6}\frac{32\pi}{c}\varepsilon^2n^{7/2}(n+1)^{3/2}N^2d^{3/2},$$

de donde se sigue el corolario. ∎

Definimos ahora la función $\widetilde{A_\varepsilon}:Y\longrightarrow\mathbb{R}_+$ como sigue:

$$\widetilde{A_\varepsilon}(y):=A_\varepsilon(G_{(d)}(y),e_0).$$

El siguiente resultado es el principal resultado técnico de esta sección y muy probablemente de esta memoria. Proporciona la base a partir de la cual desarrollar los algoritmos de resolución de sistemas de ecuaciones.

Proposición 4.4.4 *Con las notaciones anteriores, tenemos:*

$$E_Y[\widetilde{A_\varepsilon}]\leq\frac{\sqrt{2\pi}e^{1/6}32\pi}{c}\varepsilon^2n^{7/2}(n+1)^{3/2}N^2d^{3/2},$$

donde $c\geq 0{,}09$ *es la constante universal del Lema 4.2.8.*

Demostración.– Sea Xel siguiente conjunto compacto afín

$$X:=[0,t_0]\times S^1_\Delta(L^\perp_{e_0})\times S^1(\mathcal{H}_{(1)})\subseteq\mathbb{R}\times\mathbb{C}^{N+1},$$

con la estructura producto. Sea $G'_{(d)}:X\longrightarrow V_{e_0}$ la aplicación definida como sigue. Para todo punto $(t,h,M)\in X$,

$$G'_{(d)}(t,h,M):=(1-t^{\frac{1}{n^2+n}})^{1/2}h+t^{\frac{1}{2n^2+2n}}\psi_{e_0}^{-1}(T_{e_0}(M\Omega(M))).$$

Por las definiciones y el Teorema de Fubini (o el Teorema 1.1.13), se tiene que $E_X[A_\varepsilon \circ G'_{(d)}]$ es igual a

$$\frac{1}{t_0 \nu[S^1(\mathcal{H}_{(1)})]\nu_\Delta[S^1_\Delta(L^\perp_{e_0})]} \int_0^{t_0} \int_{M \in S^1(\mathcal{H}_{(1)})} \mathcal{J}(M,t)\ dS^1(\mathcal{H}_{(1)})\ dt,$$

donde $\mathcal{J}(M,t)$ es la siguiente integral:

$$\int_{h \in S^1_\Delta(L^\perp_{e_0})} A_\varepsilon \left(\left(1 - t^{\frac{1}{n^2+n}}\right)^{1/2} h + t^{\frac{1}{2n^2+2n}} \psi_{e_0}^{-1} \left(T_{e_0}(M\Omega(M))\right), e_0 \right)\ dS^1_\Delta(L^\perp_{e_0}).$$

Esto es, en las notaciones del Corolario 4.3.10, se tiene que

$$\mathcal{J}(M,t) = \mathcal{I}\left(M, t^{\frac{1}{2n^2+2n}}\right).$$

Por el Corolario 4.3.10, deducimos que

$$E_X[A_\varepsilon \circ G'_{(d)}] = \frac{1}{t_0} \int_0^{t_0} K_\varepsilon \left(t^{\frac{1}{2n^2+2n}}, e_0\right)\ dt.$$

De hecho, X y $G'_{(d)}$ han sido definidos justo para que se cumpla esta igualdad. Nótese el abuso de notación en la expresión $A_\varepsilon \circ G'_{(d)}$, de un modo más correcto deberíamos decir $A_\varepsilon \circ (G'_{(d)} \times Id_{e_0})$. Por el Corolario 4.4.3, deducimos que

$$E_X[A_\varepsilon \circ G'_{(d)}] \leq \frac{300}{c} \varepsilon^2 n^{7/2} (n+1)^{3/2} N^2 d^{3/2}.$$

Ahora, sea $F : Y \longrightarrow X$ la aplicación definida como sigue:

$$F(t,h,M) := \left(t_0 t, \frac{\Delta^{-1} h}{\|h\|_2}, \frac{M}{\|M\|_F}\right),$$

donde t_0 es la constante del Corolario 4.4.3. Obsérvese que $G_{(d)} = G'_{(d)} \circ F$ (por tanto, $\widetilde{A_\varepsilon} = A_\varepsilon \circ G'_{(d)} \circ F$). Por lo tanto, el Teorema 1.1.13 aplicado a to $F : Y \longrightarrow X$ implica que

$$\int_{y \in Y} \widetilde{A_\varepsilon}(y)\ dY = \int_{x \in X} A_\varepsilon \circ G'_{(d)}(x) \int_{y \in F^{-1}(x)} \frac{1}{NJ_yF}\ dF^{-1}(x)\ dX. \quad (4.11)$$

Con un argumento muy similar al que se utiliza en la demostración del Lema 4.3.1, vamos a comprobar que la integral interior en la ecuación (4.11) es una constante. En efecto, sean $x_1 = (s_1, h_1, M_1), x_2 = (s_2, h_2, M_2) \in X$ dos puntos. Podemos suponer que $0 < s_2 \leq s_1 < 1$. Sean además $O \in \mathcal{O}_{2N+2}$ y $O' \in \mathcal{O}_{2n^2+2n}$ dos matrices ortogonales tales que se tiene:

$$\Delta^{-1} O \Delta h_1 = h_2, \quad O' M_1 = M_2,$$

donde tanto $h_i \in L_{e_0}^{\perp} \subseteq \mathcal{H}_{(d)}$ como $M_i \in \mathcal{M}_{n\times(n+1)} \equiv \mathbb{R}^{2n^2+2}$, $i = 1, 2$ son tratados como vectores reales de la dimensión apropiada. Además, O define una isometría en $(\mathcal{H}_{(d)}, can)$ y O' define una isometría en $\mathcal{M}_{n\times(n+1)}$. Observamos por tanto que la aplicación

$$\begin{array}{rccc} \varphi: & Y & \longrightarrow & Y \\ & (t,h,M) & \mapsto & (\frac{s_2}{s_1}t, Oh, O'M) \end{array}$$

es una isometría (localmente en la primera componente) y además lleva $F^{-1}(x_1)$ a $F^{-1}(x_2)$. Por tanto, por el Teorema 1.1.13, se tiene:

$$\int_{y\in F^{-1}(x_1)} \frac{1}{NJ_yF}\, dF^{-1}(x_1) = \int_{y\in F^{-1}(x_2)} \frac{1}{NJ_{\varphi^{-1}(y)}F}\, dF^{-1}(x_2).$$

Consideramos ahora la aplicación

$$\begin{array}{rccc} \varphi': & X & \longrightarrow & X \\ & (t,h,M) & \mapsto & (\frac{s_2}{s_1}t, \Delta^{-1}O\Delta h, O'M), \end{array}$$

que es también una isometría (localmente en la primera componente). Observamos que estamos en las hipótesis del Corolario 1.1.12, con lo que $NJ_yF = NJ_{\varphi^{-1}(y)}F$ y concluimos que la integral interior que aparece en el término de la derecha de la ecuación (4.11) es una constante, digamos $C \in \mathbb{R}$. Si en esa misma ecuación sustituimos el integrando $\widetilde{A_\varepsilon}$ por la función constante igual a 1 concluimos que el valor de C es exactamente

$$\frac{\nu_Y[Y]}{\nu_X[X]}.$$

Por lo tanto la ecuación (4.11) se transforma en $E_Y[\widetilde{A_\varepsilon}] = E_X[A_\varepsilon \circ G'_{(d)}]$ y la proposición queda demostrada. ∎

4.4.1. Demostración del Teorema 4.4.1

Recordemos la desigualdad de Markov que para una variable aleatoria positiva Z afirma que

$$\text{Prob}[Z \geq a] \leq \frac{E[Z]}{a}.$$

Sea $c' := \frac{10^4}{3} \geq \frac{\sqrt{2\pi}e^{1/6}32\pi}{c}$, donde c es la constante del Lema 4.2.8. Por la Proposición 4.4.4 y la desigualdad de Markov, para un input elegido al azar $(t,h,M) \in Y$, con probabilidad al menos $1 - \left(c'n^{7/2}(n+1)^{3/2}N^2d^{3/2}\right)^{1/2}\varepsilon$, tenemos que:

$$A_\varepsilon(G_{(d)}(t,h,M), e_0) \leq \left(c'n^{7/2}(n+1)^{3/2}N^2d^{3/2}\right)^{1/2}\varepsilon.$$

Ahora, este resultado se satisface para todo número positivo $\varepsilon > 0$. Por lo tanto, podemos cambiar cada aparición de ε por

$$\left(c' n^{7/2}(n+1)^{3/2} N^2 d^{3/2}\right)^{-1/2} \varepsilon.$$

Deducimos que para un elemento elegido al azar $(t, h, M) \in Y$, con probabilidad al menos $1 - \varepsilon$, se tiene que:

$$A_{(c' n^{7/2}(n+1)^{3/2} N^2 d^{3/2})^{-1/2}\varepsilon}(G_{(d)}(t, h, M), e_0) \leq \varepsilon.$$

Supongamos que $g := G_{(d)}(t, h, M)$ satisface esa fórmula. Para todo $f \in \mathbb{S}_\Delta$, sea $\tilde{k}_{f,g} \in \mathbb{R}$ el número real definido como sigue:

$$\tilde{k}_{f,g} := 18 d^{3/2} \sup_{(h,z) \in \mathfrak{L}_\Delta(f,g,e_0)} \{\mu_{norm}(h, z)\}^2.$$

Por la Proposición 4.2.6, el menor número natural $k_{f,g} \geq \tilde{k}_{f,g}$ es una cota superior para el número de pasos necesarios para la homotopía de Newton con par inicial (g, e_0).
Además, tenemos:

$$\frac{\nu_\Delta[f \in \mathbb{S}_\Delta : k_{f,g} \geq 6 \times 10^4 n^{7/2}(n+1)^{3/2} N^2 d^3 \varepsilon^{-2}]}{\nu_\Delta[\mathbb{S}_\Delta]} =$$

$$= A_{(c' n^{7/2}(n+1)^{3/2} N^2 d^{3/2})^{-1/2}\varepsilon}(g, e_0) \leq \varepsilon.$$

Por lo tanto, hemos demostrado que para un sistema elegido al azar $f \in \mathbb{S}_\Delta$, con probabilidad al menos $1 - \varepsilon$, la NHD con par inicial (g, e_0) realizando $6 \times 10^4 n^{7/2}(n+1)^{3/2} N^2 d^3 \varepsilon^{-2}$ pasos de homotopía encuentra un cero aproximado proyectivo de f. ■

4.5. Pares ε–eficientes para todo $\varepsilon > 0$

En esta sección demostraremos el Teorema 4.1.8, que es una versión fuerte del Teorema 4.1.6. Comenzamos con la siguiente

Proposición 4.5.1 *Sea $\delta \in \mathbb{R}$, $0 < \delta \leq 1$, un real positivo. Entonces, se tiene la siguiente desigualdad:*

$$\mathrm{E}_{(y,f) \in Y \times \mathbb{S}_\Delta}[\mu_{\mathrm{norm}}(f, G_{(d)}(y), e_0)^{2-\delta}] \leq \frac{c_2}{\delta} n^5 N^2 d^{3/2},$$

donde $c_2 \leq 2^{5/2} \frac{10^4}{3}$ es una constante universal.

Demostración.– Para todo $s > 0$, sea χ_s la función característica del intervalo $(s^{-1}, +\infty]$. Sea Sea $t > 0$ un real positivo. Observamos que, por el Teorema de Fubini,

$$\mathrm{Prob}_{(y,f)\in Y\times\mathbb{S}_\Delta}[\mu_{\mathrm{norm}}(f, G_{(d)}(y), e_0)^{2-\delta} > t] =$$

$$\frac{1}{\nu_Y[Y]\nu_\Delta[\mathbb{S}_\Delta]}\times$$

$$\times \int_{y\in Y}\int_{f\in\mathbb{S}_\Delta} \chi_{\frac{1}{t^{\frac{1}{2-\delta}}}}(\mu_{\mathrm{norm}}(f, G_{(d)}(y), e_0)\ d\mathbb{S}_\Delta\ dY =$$

$$\frac{1}{\nu_Y[Y]}\int_{y\in Y} A_{\frac{1}{t^{\frac{1}{2-\delta}}}}(G_{(d)}(y), e_0)\ dY = \mathrm{E}_{y\in Y}[A_{\frac{1}{t^{\frac{1}{2-\delta}}}}(G_{(d)}(y), e_0)].$$

Por la Proposición 4.4.4, deducimos que:

$$\mathrm{Prob}_{(y,f)\in Y\times\mathbb{S}_\Delta}[\mu_{\mathrm{norm}}(f, G_{(d)}(y), e_0)^{2-\delta} > t] \leq$$

$$\frac{10000}{3} n^{7/2}(n+1)^{3/2} N^2 d^{3/2} t^{-\frac{2}{2-\delta}}.$$

La proposición se sigue del Lema 2.5.4. ∎

Proposición 4.5.2 *Sea $0 < \varepsilon < \frac{1}{2}$ un número real positivo. Sea $y \in Y$, y sea $f \in \mathbb{S}_\Delta$ un sistema de ecuaciones. Denotemos $A := n^5 N^2 d^{3/2} \in \mathbb{R}$, y supongamos que se satisface la siguiente desigualdad:*

$$\mu_{\mathrm{norm}}(f, G_{(d)}(y), e_0) \leq \left(\frac{4c_2 A \log A}{\varepsilon}\right)^{\frac{1}{2-\frac{1}{\log A}}},$$

donde c_2 es la constante de la Proposición 4.5.1. Entonces, NHD *con par inicial $(G_{(d)}(y), e_0)$ y*

$$10^8 n^5 N^3 d^4 \varepsilon^{-2}$$

pasos de homotopía encuentra un cero aproximado proyectivo z de f. Mas aún, si $\zeta \in \mathbb{P}_n(\mathbb{C})$ es el cero asociado entonces se tiene que:

$$d_T(z, \zeta) \leq \frac{3-\sqrt{7}}{2d^{3/2}\mu_{\mathrm{norm}}(f, \zeta)}.$$

Demostración.– Primero, algunos cálculos elementales muestran que

$$10^8 n^5 N^3 d^4 \varepsilon^{-2} \geq 2\times 10^7 A \log A d^{3/2} \varepsilon^{-\frac{2\log A}{2\log A - 1}}$$

Por la Proposición 4.2.6, basta demostrar que

$$2\times 10^7 A \log A d^{3/2} \varepsilon^{-\frac{2\log A}{2\log A-1}} \geq c_1 d^{3/2} \left(\frac{4c_2 A\log A}{\varepsilon}\right)^{\frac{2\log A}{2\log A-1}},$$

donde $c_1 := \frac{28}{5(3-\sqrt{7})}$ es la constante universal de la Proposición 4.2.6. Esto es, basta con que se satisfaga

$$2 \times 10^7 A \log A d^{3/2} \geq c_1 d^{3/2} (4c_2 A \log A)^{\frac{2 \log A}{2 \log A - 1}} =$$

$$c_1 d^{3/2} (4c_2 A \log A)(4c_2 A \log A)^{\frac{1}{2 \log A - 1}}.$$

Hemos reducido la prueba a demostrar la siguiente desigualdad:

$$4c_1 c_2 (4c_2 A \log A)^{\frac{1}{2 \log A - 1}} \leq 2 \times 10^7. \tag{4.12}$$

Sabiendo que $A \geq 2^{7/2}$ siempre y realizando algunos cálculos elementales, concluimos:

$$4c_1 c_2 (4c_2 A \log A)^{\frac{1}{2 \log A - 1}} \leq 8 c_1 c_2 (4c_2)^{1/6} \leq 2 \times 10^7,$$

con lo que queda demostrada la igualdad (4.12) y con ella la proposición. ■

4.5.1. Demostración del Teorema 4.1.8

Basta demostrar la siguiente desigualdad:

$$\frac{\nu_Y[y : (G_{(d)}(y), e_0) \text{ es } \varepsilon - eficiente \ \ \forall\, \varepsilon, 0 < \varepsilon < 1/2]}{\nu_Y[Y]} \geq \frac{3}{4}. \tag{4.13}$$

De hecho, demostraremos la existencia de un conjunto $\mathcal{C} \subseteq Y$ tal que $\frac{\nu_Y[\mathcal{C}]}{\nu_Y[Y]} \geq \frac{3}{4}$ y tal que para todo punto $y \in Y$ y para todo real positivo $0 < \varepsilon < \frac{1}{2}$ se tiene la siguiente desigualdad:

$$\text{Prob}_{f \in \mathbb{S}_\Delta} \left[\mu_{\text{norm}}(f, G_{(d)}(y), e_0) \leq \left(\frac{4c_2 A \log A}{\varepsilon} \right)^{\frac{1}{2 - \frac{1}{\log A}}} \right] \geq 1 - \varepsilon, \tag{4.14}$$

donde c_2 es la constante de la Proposición 4.5.1, y $A > 0$ se define como en la demostración de la proposición 4.5.2. Esto es,

$$A := n^5 N^2 d^{3/2} \in \mathbb{R}.$$

Entonces, la ecuación (4.13), y con ella el Teorema 4.1.8, se sigue inmediatamente de la ecuación (4.14) y la Proposición 4.5.2, realizando algunos cálculos elementales. Demostremos por tanto la existencia de dicho conjunto $\mathcal{C}$. Por la Proposición 4.5.1,

$$\text{E}_{y \in Y}[\text{E}_{f \in \mathbb{S}_\Delta}[\mu_{\text{norm}}(f, G_{(d)}(y), e_0)^{2 - \frac{1}{\log A}}]] =$$

$$\text{E}_{(y,f) \in Y \times \mathbb{S}_\Delta}[\mu_{\text{norm}}(f, G_{(d)}(y), e_0)^{2 - \frac{1}{\log A}}] \leq c_2 A \log A.$$

Por la desigualdad de Markov, existe un conjunto $\mathcal{C} \subseteq Y$ tal que

$$\frac{\nu_Y[\mathcal{C}]}{\nu_Y[Y]} \geq \frac{3}{4},$$

y tal que para todo $y \in \mathcal{C}$, se tiene la siguiente desigualdad:

$$\mathrm{E}_{f \in \mathbb{S}_\Delta}[\mu_{\mathrm{norm}}(f, G_{(d)}(y), e_0)^{2-\frac{1}{\log A}}] \leq 4c_2 A \log A.$$

Finalmente, de nuevo por la desigualdad de Markov concluimos que para todo $y \in \mathcal{C}$ y para todo real positivo $\alpha > 0$, se tiene:

$$\mathrm{Prob}_{f \in \mathbb{S}_\Delta}\left[\mu_{\mathrm{norm}}(f, G_{(d)}(y), e_0)^{2-\frac{1}{\log A}} \geq \alpha\right] \leq \frac{4c_2 A \log A}{\alpha}.$$

La ecuación (4.14) se sigue, tomando

$$\alpha := \frac{4c_2 A \log A}{\varepsilon}.$$

■

Capítulo 5

El Problema 17 de Smale: El Caso Afín

En este capítulo desarrollamos con cierta extensión algunas de las consecuencias de los resultados del Capítulo 4. En el dicho capítulo hemos descrito por primera vez un algoritmo que calcula en tiempo polinomial un cero aproximado proyectivo de la gran mayoría de sistemas de ecuaciones polinomiales.

Una pregunta natural es, ¿hay algún método similar que nos permita encontrar ceros aproximados afines? En el presente capítulo responderemos afirmativamente a esta pregunta. Nos enfrentamos por tanto a la resolución de sistemas cero–dimensionales no–homogéneos. En este caso, buscaremos (y encontraremos) ceros aproximados afines (en el sentido de la Sección 1.7.2). Para resolver un sistema dado f nos basaremos en dos técnicas complementarias: La búsqueda de un cero aproximado proyectivo de f (usando toda la batería de resultados del Capítulo 4), y una técnica nueva de "traspaso" de soluciones proyectivas a soluciones afines. Para poder realizar dicho traspaso en tiempo polinomial en el tamaño del input, realizaremos estimaciones del tamaño medio de las soluciones de los problemas afines, resultado con interés propio.

5.1. Introducción

Durante este capítulo utilizaremos las notaciones y conceptos del Capítulo 4. En particular, utilizaremos frecuentemente el conjunto $Y \subseteq \mathbb{R} \times \mathbb{C}^{N+1}$, la aplicación $G_{(d)} : Y \longrightarrow V_{e_0}$, y el conjunto questor para pares eficientes $\mathcal{G}_{(d)}$ que hemos descrito en la introducción de dicho capítulo.

También utilizaremos las notaciones de la Sección 1.4, en lo que se refiere al espacio $\mathcal{H}_{(d)} = \mathcal{H}^n_{(d)}$. Recordemos que las soluciones proyectivas y afines de f están relacionadas como sigue: Si $\zeta := (\zeta_0 : \cdots : \zeta_n) \in V(f)$ es una

solución proyectiva de f, con $\zeta_0 \neq 0$, entonces

$$\varphi_0^{-1}(\zeta) := \left(\frac{\zeta_1}{\zeta_0}, \ldots, \frac{\zeta_n}{\zeta_0}\right) \in V_{\mathbb{C}^n}(f)$$

es una solución afín de f. Igualmente, si $(\zeta_1, \ldots, \zeta_n) \in \mathbb{C}^n$ es una solución afín de f, entonces

$$\varphi_0(\zeta) := (1 : \zeta_1 : \ldots : \zeta_n) \in V(f)$$

es una solución proyectiva de f. Una estrategia inicial para encontrar un cero aproximado afín de un sistema $f \in \mathcal{H}_{(d)}$ puede ser por lo tanto la siguiente: Primero buscamos un cero aproximado proyectivo $z \in \mathbb{P}_n(\mathbb{C})$ y luego consideramos, si existe, el punto afín $\varphi_0^{-1}(z) \in \mathbb{C}^n$. Por supuesto, $\varphi_0^{-1}(z)$ no tiene por qué ser un cero aproximado afín de f. Sin embargo, veremos una forma de calcular el punto que buscamos a partir del conocimiento de z. Este proceso requiere controlar la distribución de las soluciones proyectivas de los sistemas de ecuaciones homogéneos; el control de esa distribución es la principal aportación técnica de este capítulo. Puede escribirse como sigue,

Teorema 5.1.1 *Sea $\delta > 0$ un número real. La probabilidad de que un sistema elegido al azar $f \in \mathbb{S}_\Delta$ tenga una solución $\zeta \in V(f)$ con $\|\varphi_0^{-1}(\zeta)\|_2 \geq \delta$ es a lo sumo*

$$\frac{\mathcal{D}\sqrt{\pi n}}{\delta}.$$

Este resultado proporciona una cota probabilística para la norma de las soluciones afines de un sistema de ecuaciones polinomiales. De hecho, es posible calcular exactamente el valor esperable de dicha norma (véase el Teorema 5.3.2 y su corolario).
El procedimiento para el caso afín utiliza los argumentos para el caso homogéneo del Capítulo 4, y funciona como describimos a continuación. Sea $(g, e_0) \in \mathcal{G}_{(d)}$ un par. Entonces, se define el algoritmo NHDAFF con par inicial (g, e_0) como sigue,

ALGORITMO: NHDAFF CON PAR INICIAL (g, e_0).

Input: $f \in \mathbb{S}_\Delta$, $\varepsilon > 0$.

PRIMERA PARTE

- Aplicamos NHD con par inicial (g, e_0) y $10^8 n^5 N^3 d^4 \varepsilon^{-2}$ pasos de homotopía, con input f. Sea z el output de dicho algoritmo (que, por los resultados del Capítulo 4, es con alta probabilidad un cero aproximado proyectivo de f).

SEGUNDA PARTE

- Sea $k \in \mathbb{N}$ el número natural más pequeño tal que se satisface:

$$k \geq \log_2 \log_2(\mathcal{D}^2(10N)^8\varepsilon^{-3}).$$

- Sea $z^k = (z_0^k : \cdots : z_n^k)$ dado como $z^k := N_f^{(k)}(z)$, donde N_f es el operador de Newton proyectivo asociado a f.

- Sea $y = (y_0 : \cdots : y_n) \in \mathbb{P}_n(\mathbb{C})$ el punto proyectivo definido como sigue:

$$y_0 := z_0^k + \frac{3-\sqrt{7}}{d^{3/2}2^{2^k-1}}\|z^k\|_2 \frac{z_0^k}{|z_0^k|}, \text{ o bien } y_0 := \frac{3-\sqrt{7}}{d^{3/2}2^{2^k-1}}\|z^k\|_2 \text{ si } z_0 = 0,$$

$$y_i := z_i, 1 \leq i \leq n.$$

Output:

o bien *failure*, o bien
un cero aproximado de f, dado como $\varphi_0^{-1}(y) \in \mathbb{C}^n$.

Observamos que este algoritmo tiene un tiempo de ejecución polinomial en el tamaño del input. De hecho, el número de operaciones aritméticas que realiza es

$$\tilde{O}(n^6N^4d^4\varepsilon^{-2}).$$

Una vez el algoritmo obtiene $\varphi_0^{-1}(y) \in \mathbb{C}^n$, podemos comprobar si es en efecto un cero aproximado afín de f usando la α–teoría de Smale. Presentamos el segundo resultado principal de este capítulo. Representa una respuesta probabilística positiva al Problema 17 de Smale en el caso afín.

Teorema 5.1.2 *Existe un subconjunto $\mathcal{E}' \subseteq \mathcal{G}_{(d)}$ tal que*

$$\text{Prob}_{f \in \mathcal{G}_{(d)}}[f \in \mathcal{C}] \geq \frac{3}{4}.$$

y tal que, para todo $(g, e_0) \in \mathcal{E}'$ y para todo real positivo $0 < \varepsilon < \frac{1}{2}$ se tiene la siguiente propiedad:

$$\text{Prob}_{f \in \mathbb{S}_\Delta}[\text{NHDAFF } \textit{con par inicial } (g, e_0) \textit{ e input } (f, \varepsilon)$$

$$\textit{devuelve cero aprox. afín de } f] \geq 1 - 2\varepsilon.$$

Al igual que en el caso homogéneo, queda por tanto demostrado que, si aceptamos una pequeña probabilidad de fracaso 2ε, entonces se puede encontrar un cero aproximado afín de todo sistema de ecuaciones en un número de pasos polinomial en $n, N, d, \varepsilon^{-1}$. Además, el conjunto $\mathcal{E}'$ del Teorema 5.1.2 es un subconjunto del conjunto $\mathcal{E}$ del Teorema 4.1.8.
El siguiente corolario muestra la aplicación de este método a resolver sistemas cúbicos en tiempo polinomial en el número de incógnitas.

Corolario 5.1.3 *Sea* $(d) := (3, \ldots, 3) \in \mathbb{N}^n$. *Sea* $\mathcal{E}' \subseteq \mathcal{G}_{(d)}$ *el conjunto del Teorema 5.1.2. Para todo* $(g, e_0) \in \mathcal{E}'$, *el algoritmo* NHDAFF *con par inicial* (g, e_0) *y* $\varepsilon = \frac{1}{n^2}$ *calcula un cero aproximado afín de sistemas de ecuaciones cúbicos, con probabilidad de éxito al menos*

$$1 - \frac{2}{n^2},$$

Además, el número de operaciones aritméticas que requiere dicho algoritmo es del orden de $O(n^{26})$.

Como hemos dicho, el estudio del caso afín viene de la mano de dos técnicas distintas: Por un lado, necesitamos estimaciones cuantitativas de cómo transformar una aproximación "suficientemente buena" de un cero proyectivo de f en una aproximación "suficientemente buena" de un cero afín de f. Por otro lado, este primer estudio mostrará la necesidad de estimar de la distribución de probabilidad de las soluciones proyectivas de sistemas homogéneos. Comenzamos con la primera de las dos técnicas.

5.2. De ceros aproximados proyectivos a afines

A continuación describimos algunos resultados que permitirán realizar la siguiente tarea: A partir del conocimiento de un cero aproximado proyectivo, encontrar un cero aproximado afín. Utilizaremos frecuentemente la carta afín φ_0 definida en la igualdad (1.5). Esto es,

$$\begin{array}{rccc} \varphi_0 : & \mathbb{C}^n & \longrightarrow & \mathbb{P}_n(\mathbb{C}). \\ & (x_1, \ldots, x_n) & \mapsto & (1 : x_1 : \ldots : x_n) \end{array}$$

Comenzamos con la siguiente

Proposición 5.2.1 *Sea* $z = (z_0 : \cdots : z_n) \in \mathbb{P}_n(\mathbb{C})$ *una solución de* f, *tal que* $z_0 \neq 0$, *y sea* $y = (y_0 : \cdots : y_n) \in \mathbb{P}_n(\mathbb{C})$ *un punto proyectivo tal que* $y_0 \neq 0$. *Supongamos que*

$$\|\varphi_0^{-1}(y) - \varphi_0^{-1}(z)\|_2 \leq \frac{3 - \sqrt{7}}{d^{3/2}\mu_{\mathrm{norm}}(f, z)}.$$

Entonces, $\varphi_0^{-1}(y)$ *es un cero aproximado afín de* f, *con cero asociado* $\varphi_0^{-1}(z)$.

Demostración.– Por el Teorema 1.7.5, basta con demostrar que

$$\|\varphi_0^{-1}(y) - \varphi_0^{-1}(z)\|_2 \gamma(f, \varphi_0^{-1}(z)) \leq \frac{3 - \sqrt{7}}{2}.$$

Ahora, por la Proposición 3.1.2 sabemos que

$$\gamma(f, \varphi_0^{-1}(z)) \leq \frac{d^{3/2}}{2}\mu_{\mathrm{norm}}(f, z),$$

y el resultado queda demostrado. ∎

Lema 5.2.2 *Sean $v, w \in \mathbb{C}^n$ dos vectores complejos. Entonces, se tiene la siguiente desigualdad:*

$$\|v-w\|_2^2 \leq (1+\|v\|_2^2)(1+\|w\|_2^2) - |1+\langle v, w\rangle_2|^2.$$

Demostración.– Sean R, I respectivamente las partes real e imaginaria de $\langle v, w\rangle_2$. Por un lado, tenemos:

$$\|v-w\|_2^2 = \langle v-w, v-w\rangle_2 = \|v\|_2^2 + \|w\|_2^2 - 2R.$$

Por otro lado,

$$(1+\|v\|_2^2)(1+\|w\|_2^2) - |1+\langle v, w\rangle_2|^2 = (1+\|v\|_2^2)(1+\|w\|_2^2) - |1+R+\sqrt{-1}I|^2 =$$
$$(1+\|v\|_2^2)(1+\|w\|_2^2) - (1+R)^2 - I^2 =$$
$$1+\|v\|_2^2+\|w\|_2^2+\|v\|_2^2\ \|w\|_2^2 - 1 - 2R - (R^2+I^2) =$$
$$\|v\|_2^2+\|w\|_2^2 - 2R + \|v\|_2^2\ \|w\|_2^2 - |\langle v, w\rangle_2|^2$$

Por lo tanto,

$$(1+\|v\|_2^2)(1+\|w\|_2^2) - |1+\langle v, w\rangle_2|^2 =$$
$$\|v-w\|_2^2 + \|v\|_2^2\ \|w\|_2^2 - |\langle v, w\rangle_2|^2 \geq \|v-w\|_2^2,$$

como queríamos demostrar. ∎

Lema 5.2.3 *Sea $z := (z_0 : \cdots : z_n) \in \mathbb{P}_n(\mathbb{C})$ un punto proyectivo. Sea $\lambda \in (0,1)$ un número real positivo y sea $z^\lambda := (z_0^\lambda : \ldots : z_n^\lambda) \in \mathbb{P}_n(\mathbb{C})$ definido como sigue:*

$$\begin{array}{rcl} z_0^\lambda & = & z_0 + \lambda\|z\|_2 \frac{z_0}{|z_0|}, \quad \textit{o bien } z_0^\lambda = \lambda\|z\|_2 \textit{ si } z_0 = 0. \\ z_i^\lambda & = & z_i, \forall 1 \leq i \leq n. \end{array}$$

Entonces, se tiene la siguiente desigualdad:

$$d_T(z^\lambda, z) \leq \lambda.$$

Además, si $z_0 \neq 0$ entonces $\|\varphi_0^{-1}(z^\lambda)\|_2 \leq \|\varphi_0^{-1}(z)\|_2$.

Demostración.– Algunos cálculos elementales demuestran que:

$$d_T(z^\lambda, z) = \sqrt{\frac{\|z^\lambda\|_2^2\|z\|_2^2}{|\langle z^\lambda, z\rangle_2|^2} - 1} = \frac{\lambda\sqrt{\|z\|_2^2 - |z_0^2|}}{\lambda|z_0| + \|z\|_2} \leq \lambda,$$

y se tiene la primera desigualdad. Para la segunda, obsérvese que

$$|z_0^\lambda| = \left|z_0 + \lambda\|z\|_2\frac{z_0}{|z_0|}\right| = |z_0|\left|1 + \lambda\|z\|_2\frac{1}{|z_0|}\right| = |z_0| + \lambda\|z\|_2 \geq |z_0|.$$

Por tanto,

$$\|\varphi_0^{-1}(z^\lambda)\|_2 = \frac{1}{|z_0^\lambda|}\|(z_1, \ldots, z_n)\|_2 \leq \frac{1}{|z_0|}\|(z_1, \ldots, z_n)\|_2 = \|\varphi_0^{-1}(z)\|_2.$$

∎

Lema 5.2.4 *Sean* $z := (z_0 : \cdots : z_n), z' := (z'_0 : \cdots : z'_n) \in \mathbb{P}_n(\mathbb{C})$ *dos puntos proyectivos tales que* $z_0, z'_0 \neq 0$. *Entonces, se tiene:*

$$\|\varphi_0^{-1}(z) - \varphi_0^{-1}(z')\|_2 \leq (1 + \|\varphi_0^{-1}(z)\|_2 \ \|\varphi_0^{-1}(z')\|_2) \ d_T(z, z').$$

Demostración.– Denotamos $v := \varphi_0^{-1}(z)$, $w := \varphi_0^{-1}(z')$. Entonces, el punto $(1, v) \in \mathbb{C}^{n+1}$ es un representante afín de z y el punto $(1, w) \in \mathbb{C}^{n+1}$ es un representante afín de z'. Por tanto,

$$\frac{\|v - w\|_2^2}{d_T(z, z')^2} = \frac{\|v - w\|_2^2}{\frac{\|(1,v)\|_2^2 \|(1,w)\|_2^2}{|\langle (1,v),(1,w)\rangle_2|^2} - 1} =$$

$$\frac{\|v - w\|_2^2}{\|(1, v)\|_2^2 \|(1, w)\|_2^2 - |\langle (1, v), (1, w)\rangle_2|^2} |\langle (1, v), (1, w)\rangle_2|^2 =$$

$$\frac{\|v - w\|_2^2}{(1 + \|v\|_2^2)(1 + \|w\|_2^2) - |1 + \langle v, w\rangle_2|^2} |1 + \langle v, w\rangle_2|^2.$$

Por el Lema 5.2.2, deducimos:

$$\frac{\|v - w\|_2^2}{d_T(z, z')^2} \leq |1 + \langle v, w\rangle_2|^2.$$

Concluimos que

$$\frac{\|v - w\|_2}{d_T(z, z')} \leq |1 + \langle v, w\rangle_2| \leq 1 + |\langle v, w\rangle_2| \leq 1 + \|v\|_2 \|w\|_2,$$

y el lema queda demostrado. ∎

Lema 5.2.5 *Sea* $0 < \lambda < \frac{\sqrt{3}}{3}$ *un número real positivo. Sean* $\zeta, z, z' \in \mathbb{P}_n(\mathbb{C})$ *tres puntos tales que se tienen las siguientes propiedades:*

- $d_T(z, \zeta) \leq \lambda$.
- $\zeta_0 \neq 0$.
- *Las expresiones* $z = (z_0 : \cdots : z_n)$ *y* $z' = (z'_0 : \cdots : z'_n)$ *satisfacen:*
  1. $z'_0 = z_0 + \lambda \|z\|_2 \frac{z_0}{|z_0|}$, *o bien* $z'_0 = \lambda \|z\|_2$ *si* $z_0 = 0$.
  2. $z'_i = z_i, \ 1 \leq i \leq n$.

Entonces, $z'_0 \neq 0$ *y se tienen las siguientes desigualdades:*

$$\|\varphi_0^{-1}(z')\|_2 \leq \sqrt{1 + \|\varphi_0^{-1}(\zeta)\|_2^2}, \tag{5.1}$$

$$\|\varphi_0^{-1}(z') - \varphi_0^{-1}(\zeta)\|_2 \leq 3(2 + \|\varphi_0^{-1}(\zeta)\|_2^2)\lambda. \tag{5.2}$$

Demostración.– Comenzamos con la desigualdad (5.1). Sean r_z, r_ζ definidas como sigue:

$$r_\zeta := d_R(\zeta, e_0), \qquad r_z := d_R(z, e_0).$$

Obsérvese que se tienen las siguientes igualdades:

$$\|\varphi_0^{-1}(z)\|_2 = d_T(z, e_0) = \tan r_z, \qquad \|\varphi_0^{-1}(\zeta)\|_2 = d_T(\zeta, e_0) = \tan r_\zeta.$$

Si $r_z \leq r_\zeta$, entonces $\|\varphi_0^{-1}(z)\|_2 \leq \|\varphi_0^{-1}(\zeta)\|_2$ y por el Lema 5.2.3 tenemos:

$$\|\varphi_0^{-1}(z')\|_2 \leq \|\varphi_0^{-1}(z)\|_2 \leq \|\varphi_0^{-1}(\zeta)\|_2,$$

como queríamos. Ahora, supongamos que $r_z \geq r_\zeta$. Por la desigualdad triangular, tenemos que $r_z \leq d_R(z, \zeta) + r_\zeta$. Por lo tanto,

$$\lambda \geq d_T(z, \zeta) = \tan d_R(z, \zeta) \geq \tan(r_z - r_\zeta) \geq \cos r_\zeta - \cos r_z = \frac{|\zeta_0|}{\|\zeta\|_2} - \frac{|z_0|}{\|z\|_2}.$$

Deducimos que:

$$\|\varphi_0^{-1}(z')\|_2 = \frac{\|(z_1, \ldots, z_n)\|_2}{|z_0| + \lambda\|z\|_2} \leq \frac{\|\zeta\|_2\|(z_1, \ldots, z_n)\|_2}{|\zeta_0|\|z\|_2} \leq \frac{\|\zeta\|_2}{|\zeta_0|} =$$

$$\sqrt{1 + d_T(\zeta, e_0)^2} = \sqrt{1 + \|\varphi_0^{-1}(\zeta)\|_2^2}.$$

Demostremos ahora la desigualdad (5.2). Por el Lema 5.2.3, sabemos que $d_T(z, z') \leq \lambda$. Además,

$$\tan(d_R(z, \zeta)) = d_T(z, \zeta) \leq \lambda \leq \frac{\sqrt{3}}{3},$$

y

$$\tan(d_R(z', z)) = d_T(z, z') \leq \lambda \leq \frac{\sqrt{3}}{3}.$$

Ahora, la función tan es creciente en el intervalo $(0, \frac{\pi}{2})$. Por tanto:

$$d_T(z', \zeta) = \tan(d_R(z', \zeta)) \leq \tan(d_R(z', z) + d_R(z, \zeta)) =$$

$$\frac{\tan(d_R(z', z)) + \tan(d_R(z, \zeta))}{1 - \tan(d_R(z', z))\tan(d_R(z, \zeta))} \leq \frac{d_T(z', z) + d_T(z, \zeta)}{1 - \frac{1}{3}} \leq$$

$$\frac{3}{2}(d_T(z', z) + d_T(z, \zeta)) \leq 3\lambda.$$

Por el Lema 5.2.4, deducimos que

$$\|\varphi_0^{-1}(z') - \varphi_0^{-1}(\zeta)\|_2 \leq 3\lambda(1 + \|\varphi_0^{-1}(z')\|_2\|\varphi_0^{-1}(\zeta)\|_2),$$

y la desigualdad (5.2) se sigue de la desigualdad (5.1). ∎

El siguiente resultado proporciona un método para obtener un cero aproximado afín de f a partir de un cero aproximado proyectivo, bajo ciertas hipótesis.

Proposición 5.2.6 *Sea $f \in \mathcal{H}_{(d)}$ un sistema de ecuaciones polinomiales. Sea $\zeta \in \mathbb{P}_n(\mathbb{C})$ una solución proyectiva de f tal que $\zeta_0 \neq 0$, y sea $z := (z_0, \ldots, z_n) \in \mathbb{P}_n(\mathbb{C})$ un cero aproximado proyectivo de f con cero asociado ζ, tal que*

$$d_T(z, \zeta)\mu_{\text{norm}}(f, \zeta) \leq \frac{3 - \sqrt{7}}{d^{3/2}}.$$

Sea $z^k := N_f^{(k)}(z)$ el punto proyectivo obtenido tras k aplicaciones del operador de Newton proyectivo, donde $k \in \mathbb{N}$ es tal que

$$k \geq \log\log(6(2 + \|\varphi_0^{-1}(\zeta)\|_2^2)\mu_{\text{norm}}(f, \zeta)).$$

Finalmente, sea $y := (y_0 : \cdots : y_n) \in \mathbb{P}_n(\mathbb{C})$ el punto definido como sigue:

$$\begin{array}{lcl} y_0 & := & z_0^k + \frac{3-\sqrt{7}}{d^{3/2}2^{2^k-1}}\|z^k\|_2 \frac{z_0^k}{|z_0^k|}, \quad \acute{o} \quad y_0 := \frac{3-\sqrt{7}}{d^{3/2}2^{2^k-1}}\|z^k\|_2 \quad si \quad z_0 = 0, \\ y_i & := & z_i, 1 \leq i \leq n. \end{array}$$

Entonces, $\varphi_0^{-1}(y)$ es un cero aproximado afín de f con cero asociado $\varphi_0^{-1}(\zeta)$.

Demostración.– Por lo Proposición 5.2.1, basta con demostrar que

$$\|\varphi_0^{-1}(y) - \varphi_0^{-1}(\zeta)\|_2 \leq \frac{3 - \sqrt{7}}{d^{3/2}\mu_{\text{norm}}(f, \zeta)}.$$

Por otro lado, sabemos que

$$d_T(z^k, \zeta) \leq \frac{1}{2^{2^k-1}} d_T(z, \zeta) \leq \frac{1}{2^{2^k-1}} \frac{3 - \sqrt{7}}{d^{3/2}}.$$

Por el Lema 5.2.5, concluimos:

$$\|\varphi_0^{-1}(y) - \varphi_0^{-1}(\zeta)\|_2 \leq 3\frac{2 + \|\varphi_0^{-1}(\zeta)\|_2^2}{2^{2^k-1}} \frac{3 - \sqrt{7}}{d^{3/2}}.$$

La proposición se sigue de la cota inferior para k. ∎

5.3. El tamaño medio de las soluciones

La Proposición 5.2.6 proporciona una de las herramientas necesarias para tratar el caso afín. Esto es, para transferir ceros aproximados proyectivos a ceros aproximados afines. Sin embargo, aparece como condición indispensable para esta transferencia el conocimiento de la norma de la solución afín. En esta sección demostraremos que esa norma es, en general, suficientemente pequeña, y obtendremos el Teorema 5.1.1. Comenzamos con un lema técnico.

Lema 5.3.1 *Se tiene la siguiente igualdad:*

$$\frac{1}{\vartheta_n} \int_{x\in\mathbb{P}_n(\mathbb{C})} \|\varphi_0^{-1}(x)\|_2 \; d\mathbb{P}_n(\mathbb{C}) = \sqrt{\pi}\frac{\Gamma(n+\frac{1}{2})}{\Gamma(n)}.$$

Demostración.– Por el Lema 2.3.1, sabemos que para todo $x \in \mathbb{C}^n$,

$$NJ_x\varphi_0 = \frac{1}{(1+\|x\|_2^2)^{n+1}}.$$

Por el Teorema 1.1.13 concluimos que

$$\int_{z\in\mathbb{P}_n(\mathbb{C})} \|\varphi_0^{-1}(z)\|_2 \; d\mathbb{P}_n(\mathbb{C}) = \int_{x\in\mathbb{C}^n} \frac{\|x\|_2}{(1+\|x\|_2^2)^{n+1}} \; d\mathbb{C}^n.$$

Algunos cálculos elementales proporcionan:

$$\int_{x\in\mathbb{C}^n} \frac{\|x\|_2}{(1+\|x\|_2^2)^{n+1}} \; d\mathbb{C}^n = \frac{2\pi^n}{\Gamma(n)} \int_0^\infty \frac{t^{2n}}{(1+t^2)^{n+1}} \; dt = \frac{\pi^{n+1/2}\Gamma(n+\frac{1}{2})}{\Gamma(n)\Gamma(n+1)}.$$

Queda pues demostrado que

$$\frac{1}{\vartheta_n} \int_{x\in\mathbb{P}_n(\mathbb{C})} \|\varphi_0^{-1}(x)\|_2 d\mathbb{P}_n(\mathbb{C}) = \frac{1}{\vartheta_n}\frac{\pi^{n+1/2}\Gamma(n+\frac{1}{2})}{\Gamma(n)\Gamma(n+1)} = \sqrt{\pi}\frac{\Gamma(n+\frac{1}{2})}{\Gamma(n)}.$$

■

El siguiente resultado establece el tamaño medio de las soluciones, y es válido también en el caso de dimensión positiva. Dado un número complejo en el hiperplano del infinito, $\zeta = (0 : \zeta_1, \ldots, \zeta_n)$, denotamos $\|\varphi_0^{-1}(\zeta)\|_2 = +\infty$.

Teorema 5.3.2 *Sea $\mathcal{A}_{(d)}$ el valor esperable de la norma de las soluciones afines de un sistema de ecuaciones para la lista de grados (d). Esto es,*

$$\mathcal{A}_{(d)} := \mathrm{E}_{f\in\mathbb{S}_\Delta(\mathcal{H}^m_{(d)})}[\mathrm{E}_{\zeta\in V(f)}[\|\varphi_0^{-1}(\zeta)\|_2]],$$

donde $\mathbb{S}_\Delta(\mathcal{H}^m_{(d)})$ es la esfera en $\mathcal{H}^m_{(d)}$. Entonces, se tiene la siguiente igualdad:

$$\mathcal{A}_{(d)} = \sqrt{\pi}\frac{\Gamma(n+\frac{1}{2})}{\Gamma(n)}.$$

Demostración.– Primero, observamos que

$$\mathrm{E}_{\zeta\in V(f)}[\|\varphi_0^{-1}(\zeta)\|_2]$$

es una cantidad que depende sólo de la clase proyectiva de f. Por lo tanto, podemos escribir

$$\mathcal{A}_{(d)} = \frac{1}{\nu_{\mathbb{P}(\mathcal{H}^m_{(d)})}[\mathbb{P}(\mathcal{H}^m_{(d)})]} \int_{f\in\mathbb{P}(\mathcal{H}^m_{(d)})} \frac{1}{\nu_{n-m}[V(f)]} \int_{\zeta\in V(f)} \|\varphi_0^{-1}(\zeta)\|_2 \; d\mathbb{P}(\mathcal{H}^m_{(d)}).$$

Sea $\alpha \in \{0,1\}$ y sea $\mathcal{I}^{\alpha}_{(d)}$ la cantidad definida como sigue,

$$\mathcal{I}^{\alpha}_{(d)} := \int_{f \in \mathbb{P}(\mathcal{H}^m_{(d)})} \int_{\zeta \in V(f)} \|\varphi_0^{-1}(\zeta)\|_2^{\alpha} \, dV(f) \, d\mathbb{P}(\mathcal{H}^m_{(d)}).$$

Por el Corolario 3.5.8, sabemos que $\nu_{n-m}[V(f)] = \vartheta_{n-m}\mathcal{D}$, para casi todo sistema $f \in \mathbb{P}(\mathcal{H}^m_{(d)})$. Por lo tanto,

$$\mathcal{I}^{1}_{(d)} = \nu_{\mathbb{P}(\mathcal{H}^m_{(d)})}[\mathbb{P}(\mathcal{H}^m_{(d)})]\vartheta_{n-m}\mathcal{D}\mathcal{A}_{(d)},$$

$$\mathcal{I}^{0}_{(d)} = \nu_{\mathbb{P}(\mathcal{H}^m_{(d)})}[\mathbb{P}(\mathcal{H}^m_{(d)})]\vartheta_{n-m}\mathcal{D}. \tag{5.3}$$

Sea $W := \{(f,\zeta) \in \mathbb{P}(\mathcal{H}^m_{(d)}) \times \mathbb{P}_n(\mathbb{C}) : \zeta \in V(f)\}$ la variedad de incidencia como se ha definido en la Sección 3.5. Sean $p_1 : W \longrightarrow \mathbb{P}(\mathcal{H}^m_{(d)})$ y $p_2 : W \longrightarrow \mathbb{P}_n(\mathbb{C})$ las proyecciones en la primera y segunda coordenadas. Igual que en el demostración del Teorema 3.5.2, se tiene:

$$\mathcal{I}^{\alpha}_{(d)} = \int_{x \in \mathbb{P}_n(\mathbb{C})} \|\varphi_0^{-1}(x)\|_2^{\alpha} \int_{f \in V_x} \frac{NJ_{(f,x)}p_1}{NJ_{(f,x)}p_2} \, dV_x \, d\mathbb{P}_n(\mathbb{C}), \tag{5.4}$$

donde para cada $x \in \mathbb{P}_n(\mathbb{C})$, $V_x = p_2^{-1}(x)$ es el conjunto de sistemas que contienen a x como solución.
Sea $x \in \mathbb{P}_n(\mathbb{C})$ un punto proyectivo cualquiera, y sea $U \in \mathcal{U}_{n+1}$ una matriz unitaria tal que $Ue_0 = x$. Entonces, la aplicación de V_x a V_{e_0} que envía a cada par (f,x) sobre el par $(f \circ U, e_0)$ es una isometría. Además, como hemos demostrado en la Sección 3.5, para todo $f \in V_{e_0}$, $NJ_{(f,e_0)}p_i = NJ_{(f\circ U^{-1},x)}p_i$, $i = 1,2$. Por tanto, por el Teorema 1.1.13 se tiene que

$$\int_{f \in V_x} \frac{NJ_{(f,x)}p_1}{NJ_{(f,x)}p_2} \, dV_x = \int_{f \in V_{e_0}} \frac{NJ_{(f\circ U^{-1},x)}p_1}{NJ_{(f\circ U^{-1},x)}p_2} \, dV_{e_0} = \int_{f \in V_{e_0}} \frac{NJ_{(f,e_0)}p_1}{NJ_{(f,e_0)}p_2} \, dV_{e_0},$$

y deducimos que la integral interior en la ecuación (5.4) no depende de $x \in \mathbb{P}_n(\mathbb{C})$. Por lo tanto,

$$I^{\alpha}_{(d)} = \left(\int_{f \in V_{e_0}} \frac{NJ_{(f,e_0)}p_1}{NJ_{(f,e_0)}p_2} \, dV_{e_0} \right) \int_{x \in \mathbb{P}_n(\mathbb{C})} \|\varphi_0^{-1}(x)\|_2^{\alpha} d\mathbb{P}_n(\mathbb{C}),$$

para $\alpha = 0,1$. Primero, supongamos que $\alpha = 0$. Entonces, por la igualdad (5.3), tenemos:

$$\nu_{\mathbb{P}(\mathcal{H}^m_{(d)})}[\mathbb{P}(\mathcal{H}^m_{(d)})]\vartheta_{n-m}\mathcal{D} = I^{0}_{(d)} = \vartheta_n \int_{f \in V_{e_0}} \frac{NJ_{(f,e_0)}p_1}{NJ_{(f,e_0)}p_2} \, dV_{e_0},$$

y obtenemos que

$$\int_{f \in V_{e_0}} \frac{NJ_{(f,e_0)}p_1}{NJ_{(f,e_0)}p_2} \, dV_{e_0} = \frac{\vartheta_{n-m}}{\vartheta_n} \nu_{\mathbb{P}(\mathcal{H}^m_{(d)})}[\mathbb{P}(\mathcal{H}^m_{(d)})]\mathcal{D}.$$

Concluimos que,

$$I^1_{(d)} = \frac{\vartheta_{n-m}}{\vartheta_n} \nu_{\mathbb{P}(\mathcal{H}^m_{(d)})}[\mathbb{P}(\mathcal{H}^m_{(d)})]\mathcal{D} \int_{x\in\mathbb{P}_n(\mathbb{C})} \|\varphi_0^{-1}(x)\|_2 d\mathbb{P}_n(\mathbb{C}),$$

y por lo tanto,

$$A_{(d)} = \frac{1}{\vartheta_n} \int_{x\in\mathbb{P}_n(\mathbb{C})} \|\varphi_0^{-1}(x)\|_2 \; d\mathbb{P}_n(\mathbb{C}).$$

El teorema se sigue del Lema 5.3.1. ∎

El Teorema 5.3.2 tiene la siguiente forma en el caso cero–dimensional:

Corolario 5.3.3 *En el caso cero–dimensional (esto es, $m = n$), se tiene la siguiente igualdad:*

$$\frac{1}{\nu_\Delta[\mathbb{S}_\Delta]} \int_{f\in\mathbb{S}_\Delta} \frac{1}{\sharp V(f)} \sum_{\zeta\in V(f)} \|\varphi_0^{-1}(\zeta)\|_2 \; d\mathbb{S}_\Delta = \sqrt{\pi}\frac{\Gamma(n+\frac{1}{2})}{\Gamma(n)}.$$

5.3.1. Demostración del Teorema 5.1.1

El Corolario 3.5.8 nos asegura que $\sharp V(f) = \mathcal{D}$ para casi todo sistema $f \in \mathbb{P}(\mathcal{H}_{(d)})$, o equivalentemente, para casi todo sistema $f \in \mathbb{S}_\Delta$. Además, se tiene:

$$\mathrm{E}_{f\in\mathbb{S}_\Delta}\left[\max_{\zeta\in V(f)} \|\varphi_0^{-1}(\zeta)\|_2\right] \leq$$

$$\frac{1}{\nu_\Delta[\mathbb{S}_\Delta]} \int_{f\in\mathbb{S}_\Delta} \sum_{\zeta\in V(f)} \|\varphi_0^{-1}(\zeta)\|_2 \; d\mathbb{S}_\Delta.$$

Por el Corolario 5.3.3, esta última cantidad es igual a

$$\frac{\sqrt{\pi}\mathcal{D}\Gamma\left(n+\frac{1}{2}\right)}{\Gamma(n)}.$$

Por las desigualdades de Gautchi (véase por ejemplo [43]) deducimos:

$$\mathrm{E}_{f\in\mathbb{S}_\Delta}\left[\max_{\zeta\in V(f)} \|\varphi_0^{-1}(\zeta)\|_2\right] \leq \mathcal{D}\sqrt{\pi n}.$$

El Corolario se sigue de la desigualdad de Markov. ∎

5.4. Eficacia del algoritmo afín

En esta sección demostraremos el Teorema 5.1.2. Necesitaremos el siguiente resultado previo.

Proposición 5.4.1 *Sea $0 < \varepsilon < \frac{1}{2}$ un número positivo. Denotemos $A := n^5N^2d^{3/2}$. Sea $y \in Y$ un punto cualquiera y sea $f \in \mathcal{H}_{(d)}$ un sistema input tal que las siguientes propiedades se satisfacen:*

$$\|\varphi_0^{-1}(\zeta)\|_2 \leq \frac{\mathcal{D}\sqrt{\pi n}}{\varepsilon}, \ \ \forall\, \zeta \in V(f), \tag{5.5}$$

y

$$\mu_{\mathrm{norm}}(f, G_{(d)}(y), e_0) \leq \left(\frac{4c_2 A \log A}{\varepsilon}\right)^{\frac{1}{2-\frac{1}{\log A}}}, \tag{5.6}$$

donde c_2 es la constante de la Proposición 4.5.1. Entonces, el output de NHDAFF *con par inicial (g, e_0) e input f, ε es un cero aproximado afín de f.*

Demostración.– Por la Proposición 4.5.2 y la ecuación (5.6), la primera parte del algoritmo devuelve un cero aproximado proyectivo z de f, con un cero asociado que denotamos $\zeta \in \mathbb{P}_n(\mathbb{C})$. Además, por la Proposición 5.2.6 y las ecuaciones (5.6) y (5.5), para demostrar que el output del algoritmo es un cero aproximado afín de f con cero asociado $\varphi_0^{-1}(\zeta)$ sólo debemos comprobar que el número de veces k que NHDAFF aplica el operador de Newton proyectivo es mayor o igual que:

$$\log_2 \log_2 \left(6\left(2 + \left(\frac{\mathcal{D}\sqrt{\pi n}}{\varepsilon}\right)^2\right)\left(\frac{4c_2 A \log A}{\varepsilon}\right)^{\frac{1}{2-\frac{1}{\log A}}}\right).$$

Esto está claro por la descripción del algoritmo NHDAFF, sabiendo que

$$n^6N^2d^{3/2}\log N \leq N^8,$$

para toda elección posible de $n, (d)$. ∎

5.4.1. Demostración del Teorema 5.1.2

De hecho, demostraremos que para todo punto $y \in \mathcal{C}$ del conjunto $\mathcal{C} \subseteq Y$ de la demostración del Teorema 4.1.8, y para todo $0 < \varepsilon < \frac{1}{2}$, el par $(G_{(d)}(y), e_0)$ satisface:

$$\mathrm{Prob}_{f \in \mathbb{S}_\Delta}[\text{ se tienen las propiedades (5.6) y (5.5)}] \geq 1 - 2\varepsilon.$$

Entonces, el Teorema 5.1.2 se sigue inmediatamente de la Proposición 5.4.1. Ahora, por la demostración del Teorema 4.1.8 (véase la Sección 4.5.1), para todo $(g, e_0) \in \mathcal{C}$ y para todo $0 < \varepsilon < \frac{1}{2}$, se satisface la propiedad (5.6) con

probabilidad al menos $1 - \varepsilon$. Además, por el Teorema 5.1.1, la propiedad (5.5) se satisface con probabilidad al menos $1 - \varepsilon$. Por lo tanto, para todo $(g, e_0) \in \mathcal{C}$ y para todo $0 < \varepsilon < \frac{1}{2}$, la probabilidad de que ambas propiedades (5.6) y (5.5) se satisfagan es de al menos $1 - 2\varepsilon$, y el Teorema 5.1.2 queda demostrado. ■

Capítulo 6

Estimaciones Discretas

6.1. Introducción

El significado computacional de los estudios desarrollados en esta memoria queda mucho más claro si somos capaces de trasladarlos a resultados discretos similares. En efecto, la computación se realiza siempre en ámbitos discretos, en particular el espacio de inputs de un algoritmo será siempre una clase discreta. Este inconveniente tiene sin embargo otras ventajas. Por ejemplo, la probabilidad en un espacio discreto será siempre la probabilidad uniforme (correspondiéndose con el uso habitual de los programas y librerías de matemáticas), que es particularmente sencilla y ventajosa. Una técnica para transferir los resultados continuos al caso discreto es la Geometría de los Números de Minkowski, que fue ideada para establecer equivalencias entre el volumen de ciertos conjuntos y el número de puntos enteros que contienen. El error cometido con esas estimaciones suele recibir el nombre de *discrepancia*. Existen varias técnicas para acotar las discrepancias, un primer ejemplo es la famosa demostración de Gauss para controlar el número de puntos enteros en un círculo. Tiene especial relevancia el trabajo de Davenport [28] en que se sientan las bases para el método de demostración que aquí utilizamos. Otros autores han tratado este problema más adelante (véase por ejemplo [24, 79]). Particularmente significativos son los éxitos alcanzados recientemente en [21, 23], con aplicación tanto para el caso afín como para el proyectivo. En las secciones 6.2 y 6.3 refinaremos aún más las cotas de [23] para ambas situaciones y en las secciones posteriores aplicaremos las nuevas cotas para obtener las versiones discretas de algunos de los resultados de esta memoria. Para exponer estos resultados correctamente, recordemos muy brevemente la noción de *talla bit* de un punto proyectivo: Dado un punto proyectivo $x \in \mathbb{P}_k(\mathbb{Q})$, la *altura* $H(x)$ de x se define como sigue,

$$H(x) := \min_{y \in \mathbb{Z}^{k+1} : \pi(y) = x} \|y\|_2.$$

Esto es, $H(x)$ es el mínimo de las normas de los puntos enteros cuya clase proyectiva es x. Entonces, la talla bit de x se define como $\log_2 H(x)$ (véase la Sección 6.3 para una definición más precisa y detallada). Hay una equivalencia aproximada entre el la talla bit de un punto proyectivo y el número de bits que se requieren para representarlo en una Máquina de Turing.
Como hemos indicado, es nuestro objetivo obtener un resultado que relacione el volumen de un conjunto proyectivo con el número de puntos de talla bit dada que contiene. En esta relación jugará un papel esencial una constante geométrica asociada al conjunto en cuestión. Dado un conjunto cualquiera $\mathcal{W} \subseteq \mathbb{R}^m$, denotamos por $\beta_0(\mathcal{W})$ el número de componentes conexas de $\mathcal{W}$. Entonces, definimos la siguiente cantidad,

$$\beta(\mathcal{W}) := \sup_{L \subseteq \mathbb{R}^m} \{\beta_0(\mathcal{W} \cap L)\} \qquad \in \mathbb{Z} \cup \{\infty\},$$

donde L varía entre todas las rectas (de dimensión 1) en $\mathbb{R}^m$ paralelas a alguno de los ejes de coordenadas. En el caso $m = 1$ definimos simplemente $\beta(\mathcal{W}) := \beta_0(\mathcal{W})$. Davenport fue el primero en utilizar la constante $\beta(W)$ (o una versión similar, véase [28]).
La constante $\beta(\mathcal{W})$ que acabamos de introducir, es una característica geométrica que viene a medir la convexidad del conjunto $\mathcal{W}$ relativa a los ejes de coordenadas. Por ejemplo, si $\mathcal{W}$ es convexo, se tiene que $\beta(\mathcal{W}) = 1$.
La sección 6.3 está dedicada a demostrar el siguiente resultado (véase también [13]).

Teorema 6.1.1 *Sea $m \geq 2$ y sea $\mathcal{W} \subset \mathbb{P}_m(\mathbb{C})$ un subconjunto medible del espacio proyectivo complejo. Sea $h > 0$ un número real, tal que*

$$h = 5 + \frac{3}{2}\log_2(m+1) + \log_2(\beta(\widetilde{\mathcal{W}})) + h_1,$$

donde $\widetilde{\mathcal{W}}$ es el cono afín de $\mathcal{W}$ (origen incluido) y $h_1 \geq 0$ es algún número positivo. Sea $P \in [0,1]$ la probabilidad de que un elemento de talla bit a lo más h elegido al azar esté en $\mathcal{W}$. Entonces, se tiene que:

$$\left| P - \frac{\nu_m[\mathcal{W}]}{\vartheta_m} \right| \leq \frac{1}{2^{h_1}}.$$

Este resultado admite la siguiente lectura: Si h es "suficientemente grande" (en función de la dimensión y de la constante $\beta(\widetilde{\mathcal{W}})$), entonces los puntos de talla bit acotada por h se distribuyen muy bien, en relación al conjunto $\mathcal{W}$. En el caso de que $\mathcal{W}$ (o, más precisamente, su cono afín $\widetilde{\mathcal{W}}$), sea definible en términos de conjuntos semi–algebraicos, existen acotaciones para el valor de la constante $\beta(\widetilde{\mathcal{W}})$ (véase el Lema 6.2.4). Debido a esto, a partir del Teorema 6.1.1 se pueden obtener de manera más o menos directa las versiones discretas de muchos de los resultados de capítulos anteriores de esta memoria. Comenzamos con el siguiente resultado, que es el análisis discreto de la

distribución de probabilidad del número de condicionamiento lineal. Para todo real positivo $h > 0$, denotemos por $\mathbb{P}(\mathcal{M}_{n_1\times n_2}(\mathbb{C}))^{\mathbb{Q}[h]}$ el conjunto de las matrices proyectivas de tamaño $n_1 \times n_2$ cuya talla bit es menor o igual que h.

Corolario 6.1.2 *Con las notaciones anteriores sea $\varepsilon > 0$ un número real. Sea $h > 0$ un número real, tal que*

$$h = 5 + \frac{3}{2}\log_2(n_1 n_2) + 4(n_1 n_2 + 1)\log_2(16r + 1) + h_1,$$

La probabilidad que una matriz elegida al azar $A \in \mathbb{P}(\mathcal{M}_{n_1\times n_2}(\mathbb{C}))^{\mathbb{Q}[h]}$ verifique $\kappa_D^{(r)}(A) > \frac{1}{\varepsilon}$ es menor o igual que:

$$2\left[\frac{e\ (n_1 n_2 - 1)\sqrt{r}}{(n_1 - r + 1)(n_2 - r + 1)}\ \varepsilon\right]^{2(n_1-r+1)(n_2-r+1)} + \frac{1}{2^{h_1}}.$$

El siguiente corolario es la versión discreta del Teorema 3.4.1 (esto es, el análisis de la distancia proyectiva de un sistema a los sistemas con singularidades de corango dado). Como en ese resultado, suponemos que estamos en el caso cero–dimensional (esto es, que $m = n$). Por tanto, el espacio de sistemas es ahora $\mathbb{P}(\mathcal{H}_{(d)}) := \mathbb{P}(\mathcal{H}_{(d)}^n)$. Además, definimos $\mathbb{P}(\mathcal{H}_{(d)})^{\mathbb{Q}_\Delta[h]}$ como el conjunto de sistemas de ecuaciones cuya *talla bit representativa* es menor o igual que h. La talla bit representativa de un sistema de ecuaciones es una noción esencialmente equivalente a la talla bit usual, pero deformada de manera que se adapta a la métrica de Kostlan Δ (véase la Sección 6.4.2 para más detalles).

Corolario 6.1.3 *Sea $h > 0$ un número real, tal que*

$$h = 5 + \frac{3}{2}\log_2(N + 1) + 4(N + n + 3)\log_2(16(r + 1)d + 1) + h_1,$$

para algún número positivo h_1. Entonces, la probabilidad que un sistema elegido al azar $f \in \mathbb{P}(\mathcal{H}_{(d)})^{\mathbb{Q}_\Delta[h]}$ esté a distancia de $\Sigma_{(d)}^r$ menor que ε es menor o igual que:

$$2\prod_{i=1}^{n}(d_i + 1)\binom{n+1}{r}\binom{n}{r}\left(\frac{e\ N(r+1)\ d\ \varepsilon}{(n-r)^2}\right)^{2(n-r)^2} + \frac{1}{2^{h_1}}.$$

El siguiente corolario es la versión discreta del Teorema 3.7.6 (esto es, la distribución de probabilidad del número de condicionamiento del caso peor $\mu_{\text{worst}}^{(m)}$).

Corolario 6.1.4 *Sea $h > 0$ un número real, tal que*

$$h = 5 + \frac{3}{2}\log_2(N + 1) + 4(N + n + 3)\log_2(16md + 1) + h_1,$$

Entonces, la probabilidad que un sistema elegido al azar $f \in \mathbb{P}(\mathcal{H}_{(d)}^m)^{\mathbb{Q}_\Delta[h]}$ *verifique* $\mu_{\text{worst}}^{(m)}(f) > \frac{1}{\varepsilon}$ *es menor o igual que:*

$$2\mathcal{D}\left[10N^{1/2}mn^{1/2}d^{3/2}\right]^{2(n-m)}\left[6N^{1/2}mn^{1/2}\ \varepsilon\right]^4 + \frac{1}{2^{h_1}}.$$

Finalmente, presentamos un resultado que puede interpretarse como una versión discreta de los teoremas en el Capítulo 4. Su exposición requiere el uso de las notaciones, conceptos y resultados del Capítulo 4. Sean pues $Y, G_{(d)}, \mathcal{G}_{(d)}$ como en la introducción del Capítulo 4.
Sea $H \geq 0$ un número positivo. Sea $\mathbb{Z}^{2N+3} \subseteq \mathbb{R}^{2N+3}$ el retículo de los puntos enteros en $\mathbb{R}^{2N+3}$. Sea Y^Hel conjunto de puntos definido como sigue:

$$Y^H := Y \cap \mathbb{Z}^{2N+3}\left[\frac{1}{H}\right],$$

donde $\mathbb{Z}^{2N+3}\left[\frac{1}{H}\right]$ es el retículo dado por la ecuación

$$\mathbb{Z}^{2N+3}\left[\frac{1}{H}\right] := \left\{\frac{z}{H} : z \in \mathbb{Z}^{2N+3}\right\}.$$

Denotamos por $\mathcal{G}_{(d)}^H \subseteq \mathcal{G}_{(d)}$ el conjunto de pares definido como sigue.

$$\mathcal{G}_{(d)}^H := \left\{(G_{(d)}(y), e_0) : y \in Y^H\right\}.$$

Obsérvese que $\mathcal{G}_{(d)}^H$ es un conjunto finito. Entonces, se tiene el siguiente resultado.

Teorema 6.1.5 *Existe una constante universal* $C > 0$ *tal que para todo par de números positivos* $\varepsilon > 0, H > 0$ *tales que*

$$\log_2 H \geq Cn^2N^3 \log_2 d + 2\log_2 \varepsilon^{-1},$$

se tiene que:

$$\text{Prob}_{(g,e_0)\in\mathcal{G}_{(d)}^H}[(g, e_0)\ \textit{es}\ \varepsilon - \textit{eficiente}] \geq \varepsilon,$$

donde estamos utilizando las notaciones de la Definición 4.1.3.

Recordemos que, una vez tenemos un par ε–eficiente, el proceso descrito en la introducción del Capítulo 4 proporciona un algoritmo que resuelve la gran mayoría de sistemas de ecuaciones (probabilidad de fracaso $1/\varepsilon$). Por tanto, el Teorema 6.1.5 puede verse como una solución probabilística al Problema 17 de Smale, sólo que la elección la hacemos en un conjunto discreto $\mathcal{G}_{(d)}^H$ en lugar de un conjunto continuo $\mathcal{G}_{(d)}$ como se hizo en el Capítulo 4.

6.2. Geometría de los Números

Durante esta sección, $m \in \mathbb{N}$ será un número natural positivo, y $\mathcal{W} \subseteq \mathbb{R}^m$ un subconjunto afín medible Lebesgue. El objetivo que tenemos es establecer una cota para la discrepancia que nos permita relacionar el número de puntos enteros en $\mathcal{W}$ con su volumen. Como nuestro interés es trasladar luego estas estimaciones al caso proyectivo, tanto volúmenes como número de puntos se considerarán dentro de bolas $B_m(0,H)$ centradas en 0 de radio creciente. Esto es, sea $H > 0$ un número positivo cualquiera y sea

$$N(\mathcal{W}, H) := \sharp[\mathcal{W} \cap \mathbb{Z}^m \cap B_m(0,H)].$$

Queremos una acotación para la siguiente cantidad:

$$\left| N(\mathcal{W}, H) - \mathscr{L}^m[\mathcal{W} \cap B_m(0,H)] \right|. \tag{6.1}$$

donde $\mathscr{L}^m$ denota la medida de Lebesgue $m-$dimensional. Para todo número natural $i \geq 0$, denotamos por K_i el volumen de la bola unidad en $\mathbb{R}^i$. Es decir,

$$K_i := \frac{\pi^{i/2}}{\Gamma\left(\frac{i}{2}+1\right)},$$

donde Γ es la función Gamma.
La cota para la cantidad (6.1) dependerá de una característica geométrica de $\mathcal{W}$, que hemos definido en la introducción de este capítulo y denotado $\beta(W)$.
El siguiente resultado es una acotación más fina que la propuesta en [23].

Teorema 6.2.1 *Sea $m \geq 1$ un natural y sea $\mathcal{W} \subseteq \mathbb{R}^m$ un conjunto medible. Sea $H \geq 0$ un número real positivo. Entonces, tenemos:*

$$|N(\mathcal{W}, H) - \mathscr{L}^m[\mathcal{W} \cap B_m(0,H)]| \leq \beta(\mathcal{W}) \sum_{i=0}^{m-1} K_i \binom{m}{i} H^i.$$

Demostración.– Denotamos $C := \beta(\mathcal{W})$. Para cada número natural $n \geq 1$, denotamos por $\delta_C(n,H)$ la siguiente cantidad:

$$\delta_C(n,H) := \sup_{\substack{V \subseteq \mathbb{R}^n \\ \beta(V) \leq C}} \{|\sharp[V \cap \mathbb{Z}^n \cap B_n(0,H)] - \mathscr{L}^n[V \cap B_n(0,H)]|\},$$

Demostramos la siguiente igualdad por inducción en n:

$$\delta_C(n,H) \leq C \sum_{i=0}^{n-1} K_i \binom{n}{i} H^i, \quad \forall H, C > 0. \tag{6.2}$$

Primero, sea $n = 1$, y sea $V \subseteq \mathbb{R}$ tal que $\beta(V) \leq C$. En este caso tenemos obviamente que

$$|\sharp[V \cap \mathbb{Z} \cap B_1(0,H)] - \mathscr{L}^1[V \cap B_1(0,H)]| \leq \beta_0(V) = \beta(V) \leq C,$$

y la desigualdad (6.2) se satisface de forma trivial. Ahora, demostramos que para $n \geq 1$,

$$\delta_C(n, H) \leq CK_{n-1}H^{n-1} + \sum_{x \in \mathbb{Z} \cap [-H,H]} \delta_C(n-1, \sqrt{H^2 - \|x\|_2^2}). \qquad (6.3)$$

En efecto, sea $V \subseteq \mathbb{R}^n$ tal que $\beta(V) \leq C$. Para todo punto $x \in \mathbb{R}$, sea V_x el conjunto definido como sigue:

$$V_x := \{y \in \mathbb{R}^{n-1} : (x, y) \in V\}.$$

También, para todo punto $y \in \mathbb{R}^{n-1}$, sea V^y el siguiente conjunto.

$$V^y := \{x \in \mathbb{R} : (x, y) \in V\}.$$

Obsérvese que para toda elección posible de x o y, los conjuntos V_x, V^y satisfacen $\beta(V_x), \beta(V^y) \leq C$. Desde ahora, denotamos:

$$H_{(x)} := \sqrt{H^2 - \|x\|_2^2}, \quad H^{(y)} := \sqrt{H^2 - \|y\|_2^2}.$$

Introducimos dos cantidades auxiliares:

$$S_1 := \left| \sharp[V \cap \mathbb{Z}^n \cap B_n(0, H)] - \sum_{x \in \mathbb{Z} \cap [-H,H]} \mathscr{L}^{n-1}[V_x \cap B_{n-1}(0, H_{(x)})] \right|,$$

and

$$S_2 := \left| \sum_{x \in \mathbb{Z} \cap [-H,H]} \mathscr{L}^{n-1}[V_x \cap B_{n-1}(0, H_{(x)})] - \mathscr{L}^n[V \cap B_n(0, H)] \right|.$$

Observamos que

$$|\sharp[V \cap \mathbb{Z}^n \cap B_n(0, H)] - \mathscr{L}^n[V \cap B_n(0, H)]| \leq S_1 + S_2.$$

Acotaremos cada uno de esos dos términos por separado.
Por un lado, S_1 es el valor absoluto de

$$\sum_{x \in \mathbb{Z} \cap [-H,H]} \sharp[V_x \cap \mathbb{Z}^{n-1} \cap B_{n-1}(0, H_{(x)})] - \sum_{x \in \mathbb{Z} \cap [-H,H]} \mathscr{L}^{n-1}[V_x \cap B_{n-1}(0, H_{(x)})].$$

Por lo tanto, S_1 es a lo sumo

$$\sum_{x \in \mathbb{Z} \cap [-H,H]} |\sharp[V_x \cap \mathbb{Z}^{n-1} \cap B_{n-1}(0, H_{(x)})] - \mathscr{L}^{n-1}[V_x \cap B_{n-1}(0, H_{(x)})]| \leq$$

$$\sum_{x \in \mathbb{Z} \cap [-H,H]} \delta_C(n-1, H_{(x)}).$$

Por otro lado, S_2 iguala el valor absoluto de

$$\sum_{x\in\mathbb{Z}\cap[-H,H]}\int_{y\in\mathbb{R}^{n-1}}\chi_{V^y\cap B_1(0,H^{(y)})}(x)\ dy+$$

$$\int_{y\in\mathbb{R}^{n-1}}\int_{x\in B_1(0,H^{(y)})}\chi_{V^y}(x)\ dx\ dy\leq$$

Por lo tanto, S_2 está acotado superiormente por

$$\int_{y\in\mathbb{R}^{n-1}\cap B_{n-1}(0,H)}\left|\sharp[V^y\cap\mathbb{Z}\cap B_1(0,H^{(y)})]-\mathscr{L}^1[V^y\cap B_1(0,H^{(y)})]\right|\ dy\leq$$

$$K_{n-1}H^{n-1}\max_{y\in B_{n-1}(0,H)}\delta_C(1,H^{(y)})\leq CK_{n-1}H^{n-1},$$

de donde se sigue la desigualdad (6.3). Por la desigualdad (6.3) y la hipótesis de inducción tenemos que:

$$\frac{\delta_C(n,H)}{C}\leq K_{n-1}H^{n-1}+\sum_{i=0}^{n-2}K_i\binom{n-1}{i}\sum_{x\in\mathbb{Z}\cap[-H,H]}(H^2-\|x\|_2^2)^{i/2}.$$

Ahora, para todo número natural $i\geq 0$, tenemos que

$$\sum_{x\in\mathbb{Z}\cap[-H,H]}(H^2-\|x\|_2^2)^{i/2}\leq H^i+2\int_0^{\lfloor H\rfloor}(H^2-t^2)^{i/2}\ dt\leq$$

$$H^i+2\int_0^H(H^2-t^2)^{i/2}\ dt=H^i+B\left(\frac{1}{2},\frac{i}{2}+1\right)H^{i+1},$$

donde B es la función Beta, esto es,

$$B(a,b)=\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)},\ \ \forall a,b\in\mathbb{R}^+.$$

Por lo tanto,

$$\frac{\delta_C(n,H)}{C}\leq K_{n-1}H^{n-1}+\sum_{i=0}^{n-2}K_i\binom{n-1}{i}\left[H^i+B\left(\frac{1}{2},\frac{i}{2}+1\right)H^{i+1}\right]=$$

$$=1+\sum_{i=1}^{n-1}K_i\left[\binom{n-1}{i}+\frac{K_{i-1}}{K_i}\binom{n-1}{i-1}B\left(\frac{1}{2},\frac{i+1}{2}\right)\right]H^i.$$

Observamos que para todo valor positivo de i,

$$\frac{K_{i-1}}{K_i}=B\left(\frac{1}{2},\frac{i+1}{2}\right)^{-1}.$$

Además,

$$\binom{n-1}{i} + \binom{n-1}{i-1} = \binom{n}{i},$$

y queda demostrado que

$$\frac{\delta_C(n,H)}{C} \leq 1 + \sum_{i=1}^{n-1} K_i \binom{n}{i} H^i = \sum_{i=0}^{n-1} K_i \binom{n}{i} H^i,$$

como queríamos.

■

El siguiente corolario se sigue de forma casi inmediata del Teorema 6.2.1.

Corolario 6.2.2 *Sea $H > 0$ un número real positivo. Con las notaciones del Teorema 6.2.1 se tiene que*

- *Si $0 < H < 1$, entonces*

$$|N(\mathcal{W}, H) - \mathscr{L}^m[\mathcal{W} \cap B_m(0,H)]| \leq \text{máx}\{1, K_m\}.$$

- *Si $H \geq 1$, entonces*

$$|N(\mathcal{W}, H) - \mathscr{L}^m[\mathcal{W} \cap B_m(0,H)]| \leq \beta(\mathcal{W}) 6m H^{m-1} \left(1 + \frac{1}{H}\right)^{m-1}.$$

- *Si además $H \geq m^2$, entonces*

$$|N(\mathcal{W}, H) - \mathscr{L}^m[\mathcal{W} \cap B_m(0,H)]| \leq \beta(\mathcal{W}) 2m K_{m-1} H^{m-1},$$

Demostración.– El primer item es trivial, pues si $0 < H < 1$ entonces $\text{máx}\{N(\mathcal{W}, H), \mathscr{L}^m[\mathcal{W} \cap B_m(0,H)]\} \leq \text{máx}\{1, K_m\}$. El segundo item se sigue inmediatamente del Teorema 6.2.1, sabiendo que

$$\sup_{i \in \mathbb{N}} K_i \leq 6, \quad \binom{m}{i} \leq m \binom{m-1}{i}.$$

$$|N(\mathcal{W}, H) - \mathscr{L}^m[\mathcal{W} \cap B_m(0,H)]| \leq \beta(\mathcal{W}) 6m \sum_{i=0}^{m-1} \binom{m-1}{i} H^i,$$

pero $\sum_{i=0}^{m-1} \binom{m-1}{i} H^i = (H+1)^{m-1}$, y se sigue el segundo item. Demostremos el último item. Por el Teorema 6.2.1, basta comprobar que para todo $H \geq m^2$, se tiene que

$$\sum_{i=0}^{m-1} K_i \binom{m}{i} H^i \leq 2m K_{m-1} H^{m-1}. \tag{6.4}$$

El caso $m = 1$ es inmediato. Para el caso $m > 1$, sea $T_i := K_i\binom{m}{i}H^i$ el i–ésimo término de la suma que aparece en la ecuación (6.4). Observamos que

$$\frac{T_i}{T_{i+1}} = B\left(\frac{1}{2}, \frac{i+2}{2}\right)^{-1} \frac{i+1}{m-i}\frac{1}{H}.$$

Por las desigualdades de Gautschi (véase [43, Th. 3]), sabemos que

$$B\left(\frac{1}{2}, \frac{i+2}{2}\right)^{-1} \leq \sqrt{\frac{i+2}{2\pi}}.$$

Por tanto, si $H \geq m^2$, obtenemos que $T_i \leq \frac{T_{i+1}}{\sqrt{m}}$ y tenemos la siguiente desigualdad.

$$\sum_{i=0}^{m-1} T_i \leq T_{m-1} \sum_{i=0}^{m-1} \frac{1}{\sqrt{m}^{m-1-i}} = T_{m-1}\frac{\sqrt{m} - \frac{1}{\sqrt{m}^{m-1}}}{\sqrt{m}-1} \leq 2T_{m-1}.$$

Queda probada la ecuación (6.4) y con ella el corolario. ∎

Un caso particularmente interesante es el de los conjuntos definibles en términos semi–algebraicos. En este caso, podemos dar una cota para la constante $\beta(\mathcal{W})$ en función de características sintácticas del conjunto. Para concretar estas ideas, introduzcamos algunos conceptos previos.
Un conjunto semi–algebraico es un subconjunto $\mathcal{W}$ de un espacio afín real $\mathbb{R}^m$ que puede ser definido por una fórmula de primer orden sobre los reales, libre de cuantificadores. Sea $\mathcal{F} := \{f_1, \ldots, f_s\} \subseteq \mathbb{R}[X_1, \ldots, X_m]$ un conjunto finito de polinomios. Decimos que un conjunto semi–algebraico $\mathcal{W} \subseteq \mathbb{R}^m$ es una $\mathcal{F}$–celda si existe una lista de condiciones de signo

$$\underline{\epsilon} := (\epsilon_1, \ldots, \epsilon_s) \in \{<, =, >\}^s,$$

tal que

$$\mathcal{W} := \{x \in \mathbb{R}^{m+1} : f_i(x)\epsilon_i\ 0,\ 1 \leq i \leq s\}.$$

Un conjunto semialgebraico $\mathcal{F}$–definible es una unión finita de $\mathcal{F}$–celdas.

Definición 6.2.3 *Sean $s, d \in \mathbb{N}$ dos números naturales positivos. Decimos que un conjunto semi–algebraico $\mathcal{W} \subseteq \mathbb{R}^m$ es (s,d)–definible si existe un conjunto finito de polinomios $\mathcal{F} \subseteq \mathbb{R}[X_1, \ldots, X_m]$ tales que:*

- *$\mathcal{W}$ es $\mathcal{F}$–definible,*
- *$\sharp(\mathcal{F}) = s$,*
- *$\deg(f_i) \leq d,\ \forall f \in \mathcal{F}$.*

Decimos que un conjunto semi–algebraico $\mathcal{W} \subseteq \mathbb{R}^m$ es la M–proyección de un conjunto semi–algebraico (s,d)–definible si existe un conjunto (s,d)–definible $\mathcal{W}' \subseteq \mathbb{R}^{M+m}$ de modo que:

$$\mathcal{W} := \{x \in \mathbb{R}^m : \exists y \in \mathbb{R}^M \text{ tal que } (y,x) \in \mathcal{W}')\}.$$

Entonces, tenemos el siguiente resultado (véase [23] y las referencias en él incluidas).

Lema 6.2.4 *Si $\mathcal{W} \subseteq \mathbb{R}^m$ es un conjunto semi–algebraico (s,d)–definible, entonces*

$$\beta(\mathcal{W}) \leq sd + 1.$$

Si $\mathcal{W} \subseteq \mathbb{R}^m$ es la M–proyección de un conjunto semi–algebraico (s,d)–definible, entonces

$$\beta(\mathcal{W}) \leq (4sd+1)^{2(M+2)}.$$

Demostración.– Sea $L \subseteq \mathbb{R}^m$ una recta cualquiera, paralela a alguno de los ejes coordenados. Podemos suponer sin pérdida de generalidad que es paralela al eje de x_1. Esto es, existe $(y_2, \ldots, y_m) \in \mathbb{R}^{m-1}$ tal que L viene dada por las ecuaciones:

$$\begin{aligned} X_2 &= y_2, \\ &\cdots \\ X_m &= y_m. \end{aligned}$$

Supongamos que $\mathcal{W}$ es un conjunto semi–algebraico (s,d)–definible. Entonces, se tiene que $L \cap \mathcal{W}$ es también un conjunto semi–algebraico (s,d)–definible de dimensión 1. Por [23, Teor. 9], sabemos que el número de componentes conexas de $L \cap \mathcal{W}$ está acotado por $sd+1$. Si $\mathcal{W}$ es la M–proyección de un conjunto semi–algebraico (s,d)–definible, entonces también $L \cap \mathcal{W}$ lo es. De nuevo [23, Teor. 9] indica que el número de componentes conexas de $L \cap \mathcal{W}$ está acotado por $(4sd+1)^{2(M+2)}$, y el lema queda demostrado. ∎

6.3. El caso de los conjuntos proyectivos

Sea $\mathbb{Q}[i]$ el cuerpo de racionales de Gauss y sea $\mathbb{Z}[i] \subseteq \mathbb{Q}[i]$ el anillo de enteros de Gauss, que es un dominio de ideales principales y dominio de factorización única con unidades $S^1(1) = \{a \in \mathbb{Z}[i] : |a| = 1\} = \{1, -1, i, -i\}$.
Diremos que un conjunto $\Lambda \subseteq \mathbb{C}^{m+1}$ es un $\mathbb{Z}[i]$–retículo en $\mathbb{C}^{m+1}$ si Λ es un $\mathbb{Z}[i]-$módulo libre generado por una base β de $\mathbb{C}^{m+1}$ como espacio vectorial complejo. Esto es, si $\beta = \{v_0, \ldots, v_m\}$ es una base de $\mathbb{C}^{m+1}$ como espacio vectorial complejo, el $\mathbb{Z}[i]$–retículo que genera es el conjunto:

$$\Lambda(\beta) = \{\lambda_0 v_0 + \cdots + \lambda_m v_m : \lambda_i \in \mathbb{Z}[i], 0 \leq i \leq m\}$$

Sea $\Lambda \subseteq \mathbb{C}^{m+1}$ un $\mathbb{Z}[i]$–retículo y sea $x \in \Lambda$ un elemento distinto de cero. Sea $\langle x \rangle \subseteq \mathbb{C}^{m+1}$ el $\mathbb{Q}[i]$–espacio vectorial generado por x. Es decir, $\langle x \rangle = \{\lambda x : \lambda \in \mathbb{Q}[i]\}$. El $\mathbb{Z}[i]$–módulo $\langle x \rangle \cap \Lambda$ un submódulo libre de torsión de rango 1 de Λ. Por tanto, tiene una base formada por un solo elemento. Decimos que x es visible desde el origen en Λ si $\{x\}$ es una base del $\mathbb{Z}[i]$–módulo $\langle x \rangle \cap \Lambda$. Obsérvese que un punto distinto de cero $x \in \Lambda$ es visible desde el origen en Λ si y solo si:

$$\|x\|_2 = \text{mín}\{\|y\|_2 : y \in \langle x \rangle \cap \Lambda\}.$$

Otra definición equivalente es la que sigue: Sea $\beta = \{v_0, \ldots, v_m\} \subseteq \mathbb{C}^{m+1}$ una base de $\mathbb{C}^{m+1}$ como espacio vectorial, tal que $\Lambda = \Lambda(\beta)$. Sea $x \in \Lambda$ un punto no nulo y sean $\lambda_0, \ldots, \lambda_m \in \mathbb{Z}[i]$ los (únicos) enteros de Gauss tales que:

$$x = \lambda_0 v_0 + \ldots + \lambda_m x_m.$$

Entonces, x es visible desde el origen si y solo si

$$gcd_{\mathbb{Z}[i]}(\lambda_0, \ldots, \lambda_m) \in S^1(1),$$

donde $gcd_{\mathbb{Z}[i]}(\lambda_0, \ldots, \lambda_m)$ es el máximo común divisor de $\lambda_0, \ldots, \lambda_m$ en $\mathbb{Z}[i]$ (un dominio de factorización).
Utilizaremos las funciones definidas a continuación. Como es usual, $r_2(n)$ denota el número de puntos en $S^1(\sqrt{n})$, donde $S^1(\sqrt{n}) = \{a + bi \in \mathbb{Z}[i] : |a + bi| = \sqrt{n}\}$. Esto es,

$$r_2(n) := \sharp\{a + bi \in \mathbb{Z}[i] : |a|^2 + |b|^2 = n\}.$$

Esta función es bien conocida desde tiempos de Gauss. Está fuertemente relacionada con la factorización de n en $\mathbb{Z}$.
Sea $\mathbb{P}_m(\mathbb{C})$ el espacio proyectivo complejo y sea $\pi : \mathbb{C}^{m+1} \setminus \{0\} \longrightarrow \mathbb{P}_m(\mathbb{C})$ la proyección canónica. Sea $\mathbb{P}_m(\mathbb{Q}[i])$ el espacio proyectivo m–dimensional definido por el cuerpo de los racionales de Gauss. También denotamos por π la proyección canónica $\pi : \mathbb{Q}[i]^{m+1} \setminus \{0\} \longrightarrow \mathbb{P}_m(\mathbb{Q}[i])$. Observamos que la restricción $\pi_{|\mathbb{Z}[i]^{m+1} \setminus \{0\}}$ es también sobreyectiva. Para todo punto $x \in \mathbb{P}_m(\mathbb{Q}[i])$, definimos su altura como el mínimo de las normas de los puntos en $\pi^{-1}(\{x\}) \cap (\mathbb{Z}[i])^{m+1}$. En otras palabras, para todo $x \in \mathbb{P}_m(\mathbb{Q}[i])$ hay exactamente cuatro puntos visibles $\{x_1, -x_1, ix_1, -ix_1\} \subseteq \mathbb{Z}[i]^{m+1}$ tales que $\pi(x_1) = x$. Entonces, la altura de x se define como

$$H(x) = \|x_1\|_2.$$

Finalmente, definimos la *talla bit* del punto proyectivo $x \in \mathbb{P}_m(\mathbb{Q}[i])$ como el logaritmo de su altura.

$$bl(x) := \log_2 H(x).$$

Observamos también que $bl(x)$ es esencialmente equivalente al número de dígitos requeridos para representar el punto proyectivo x en una máquina de Turing.
Sea $\widetilde{\mathcal{W}} \subseteq \mathbb{C}^{m+1}$ un subconjunto definible como semi–algebraico mediante la identificación $\mathbb{C} \equiv \mathbb{R}^2$. Diremos que $\widetilde{\mathcal{W}}$ es un conjunto semi–algebraico complejo. Para todo número positivo H, denotamos por $N_{\mathbb{Z}[i]}(\widetilde{\mathcal{W}}, H)$ la siguiente cantidad:

$$N_{\mathbb{Z}[i]}(\widetilde{\mathcal{W}}, H) := \sharp\{\widetilde{\mathcal{W}} \cap \mathbb{Z}[i]^{m+1} \cap B_{m+1}(0, H) \setminus \{0\}\} =$$

$$\sharp\{x \in \widetilde{\mathcal{W}} \cap \mathbb{Z}[i]^{m+1} : 0 < \|x\|_2 \leq H\}.$$

Sea $\mathcal{W} \subseteq \mathbb{P}_m(\mathbb{C})$ un subconjunto del espacio proyectivo complejo, tal que el cono $\widetilde{\mathcal{W}} := \pi^{-1}(\mathcal{W}) \cup \{0\} \subseteq \mathbb{C}^{m+1}$ es un conjunto semi–algebraico complejo. Denotamos por $\mathcal{N}_{\mathbb{Z}[i]}(\mathcal{W}, H)$ el número de puntos en $\mathcal{W} \cap \mathbb{P}_m(\mathbb{Q}[i])$ de altura a lo más H.Esto es,

$$\mathcal{N}_{\mathbb{Z}[i]}(\mathcal{W}, H) = \sharp\{x \in \mathcal{W} \cap \mathbb{P}_m(\mathbb{Q}[i]) : H(x) \leq H\}.$$

El siguiente resultado relaciona $N_{\mathbb{Z}[i]}(\widetilde{\mathcal{W}}, H)$ y $\mathcal{N}_{\mathbb{Z}[i]}(\mathcal{W}, H)$.

Proposición 6.3.1 *Con las notaciones anteriores, para todo $H > 1$ se tiene:*

$$N_{\mathbb{Z}[i]}(\widetilde{\mathcal{W}}, H) = \sum_{1 \leq n \leq H^2} \mathcal{N}_{\mathbb{Z}[i]}(\mathcal{W}, H/\sqrt{n}) r_2(n)$$

Demostración.– Denotamos por $\widetilde{\mathcal{W}}(H)$ el conjunto de los puntos no nulos en $\widetilde{\mathcal{W}} \cap \mathbb{Z}[i]^{m+1} \cap B_{m+1}(0, H)$ y por $\widetilde{\mathcal{W}}^v(H)$el conjunto de los puntos visibles en $\mathbb{Z}[i]^{m+1}$ que pertenecen a $\widetilde{\mathcal{W}} \cap B_{m+1}(0, H)$. Observamos que

$$4\mathcal{N}_{\mathbb{Z}[i]}(\mathcal{W}, s) = \sharp\widetilde{\mathcal{W}}^v(s) \tag{6.5}$$

para todo número real $s \geq 1$. Ahora consideramos la unión disjunta

$$A := \bigcup_{n \leq H^2} \widetilde{\mathcal{W}}^v(H/\sqrt{n}) \times S^1(\sqrt{n}),$$

y definimos la siguiente aplicación:

$$\begin{array}{rccc} \varphi : & A & \longrightarrow & \widetilde{\mathcal{W}}(H) \\ & (y, \lambda) & \mapsto & \lambda y \end{array}$$

Observamos que φ está bien definida, porque para todo $\lambda \in S^1(\sqrt{n})$ y para todo $y \in \widetilde{\mathcal{W}}^v(H/\sqrt{n})$ tenemos que $\lambda y \in \widetilde{\mathcal{W}} \cap \mathbb{Z}[i]^{m+1}$ y

$$|\lambda| \|y\|_2 \leq \sqrt{n} H/\sqrt{n} = H.$$

Ahora demostramos que φ es sobreyectiva. Dado $x = (x_0, \ldots, x_n) \in \widetilde{\mathcal{W}}(H)$, sea $\lambda \in \mathbb{Z}[i]$ el máximo común divisor de las coordenadas de x. Es decir,

$$\lambda = gcd_{\mathbb{Z}[i]}(x_0, \ldots, x_n),$$

y definamos $y_i = \lambda^{-1} x_i \in \mathbb{Z}[i]$, $y = (y_0, \ldots, y_n) \in \mathbb{Z}[i]^{m+1}$. Como $\widetilde{\mathcal{W}}$ es un cono, tenemos $y \in \widetilde{\mathcal{W}}$. El punto y es visible y además $\|y\|_2 = \frac{\|x\|_2}{|\lambda|} \leq \frac{H}{|\lambda|}$. Eligiendo $n = |\lambda|^2 \in \mathbb{N}$, queda demostrado que φ es sobreyectiva.
Sea $x \in \widetilde{\mathcal{W}}(H), (y, \lambda) \in \varphi^{-1}(\{x\})$ tales que $\varphi(\lambda, y) = x$. Entonces, tenemos que

$$\varphi^{-1}(\{x\}) = \{(u^{-1}y, u\lambda) : u \in S^1(1)\}.$$

En efecto, si $\varphi(y, \lambda) = \varphi(z, \theta) = x$ entonces $\lambda y = \theta z$ y se tiene que $y, z \in \mathbb{Z}[i]^{m+1}$ son visibles. Por tanto, como definen el mismo punto proyectivo, existe alguna unidad $u \in \mathbb{Z}[i]^*$ tal que $z = uy$. Concluimos que $\theta = u^{-1}\lambda$, como queríamos.
Deducimos que $\sharp\varphi^{-1}(\{x\}) = 4, \forall x \in \widetilde{\mathcal{W}}(H)$ y, por tanto,

$$\frac{1}{4}\left(\sum_{n \leq H^2} \sharp\left[\widetilde{\mathcal{W}}^v(H/\sqrt{n}) \times S^1(\sqrt{n})\right]\right) = \sharp\widetilde{\mathcal{W}}(H).$$

En otras palabras,

$$\sum_{n \leq H^2} \frac{1}{4}\sharp(\widetilde{\mathcal{W}}^v(H/\sqrt{n}))r_2(n) = N_{\mathbb{Z}[i]}(\widetilde{\mathcal{W}}, H).$$

Utilizando la igualdad (6.5), concluimos

$$N_{\mathbb{Z}[i]}(\widetilde{\mathcal{W}}, H) = \sum_{n \leq H^2} \mathcal{N}_{\mathbb{Z}[i]}(\mathcal{W}, H/\sqrt{n})r_2(n),$$

y la proposición queda demostrada. ∎

Resumimos a continuación algunos hechos bien conocidos sobre series de Dirichlet, todos ellos incluidos en la excelente monografía [63]. Consideremos el Carácter primitivo de Legendre–Jacobi–Kronecker módulo 4, χ_4, definido como sigue:

$$\chi_4(n) = \begin{cases} 0 & si\ gcd(n,4) > 1 \\ 1 & si\ n \equiv 1\ mod\ 4 \\ -1 & si\ n \equiv 3\ mod\ 4 \end{cases}$$

Definimos entonces la siguiente serie de Dirichlet para $s > 1$:

$$L_4(s) := \sum_{n \geq 1} \frac{\chi_4(n)}{n^s} = \prod_p (1 - \chi_4(p)p^{-s})^{-1}$$

Utilizando la fórmula para la multiplicación de series de Dirichlet, observamos que

$$L_4(s)\left(\sum_{n\geq 1}\frac{\mu(n)\chi_4(n)}{n^s}\right)=\sum_{n\geq 1}\left(\sum_{d|n}\mu(n/d)\chi_4(n/d)\chi_4(d)\right)n^{-s}$$

$$=\sum_{n\geq 1}\frac{\chi_4(n)\sum_{d|n}\mu(d)}{n^s}=1.$$

Obtenemos por tanto el desarrollo en forma de serie de Dirichlet:

$$\frac{1}{L_4(s)}=\sum_{n\geq 1}\frac{\mu(n)\chi_4(n)}{n^s} \tag{6.6}$$

Sea ζ la función Zeta de Riemann, definida como sigue para $s>1$:

$$\zeta(s):=\sum_{n\geq 1}\frac{1}{n^s}$$

También conocemos la expresión en serie de Dirichlet de $\frac{1}{\zeta}$ (cf. [63, Teor. 287]) para $s>1$,

$$\frac{1}{\zeta(s)}=\sum_{n\geq 1}\frac{\mu(n)}{n^s}. \tag{6.7}$$

El siguiente resultado puede encontrarse por ejemplo en [63, Teor. 306]:

$$\sum_{n\geq 1}\frac{r_2(n)}{n^s}=4\zeta(s)L_4(s). \tag{6.8}$$

Presentamos el siguiente resultado:

Proposición 6.3.2 *Sea $m\geq 2$ un número natural y sea $\mathcal{W}\subset\mathbb{P}_m(\mathbb{C})$ un subconjunto medible Lebesgue del espacio proyectivo complejo. Sea $H>1$ un número real. Entonces, tenemos:*

$$\left|\mathcal{N}_{\mathbb{Z}[i]}(\mathcal{W},H)-\frac{\pi\nu_m[\mathcal{W}]}{4\zeta(m+1)L_4(m+1)(m+1)}H^{2m+2}\right|\leq$$

$$\beta(\widetilde{\mathcal{W}})\left[3(m+1)^2H^4+\frac{1}{4}\sum_{i=3}^{2m+1}K_i\binom{2m+2}{i}\zeta(i/2)^2H^i\right],$$

donde $\widetilde{\mathcal{W}}:=\pi^{-1}(\mathcal{W})\cup\{0\}\subseteq\mathbb{C}^{m+1}$ es el cono afín de $\mathcal{W}$ (con el origen incluido).

Demostración.– La demostración de este resultado es muy técnica. Definamos la serie de Dirichlet σ como sigue:

$$\sigma := \frac{1}{4\zeta(s)L_4(s)}$$

Expresamos σ en forma de serie de Dirichlet:

$$\sigma = \sum_{n\geq 1} \frac{a_n}{n^s}.$$

Observamos que tenemos dos opciones para definir la sucesión $\{a_n : n \in \mathbb{N}\}$. Por un lado, por la ecuación (6.8) y de nuevo la fórmula de multiplicación para series de Dirichlet tenemos que

$$1 = 4\zeta(s)L_4(s)\sigma = \sum_{n\geq 1} \frac{r_2(n)}{n^s} \sum_{n\geq 1} \frac{a_n}{n^s} = \sum_{n\geq 1} \left(\sum_{d|n} r_2(d) a_{\frac{n}{d}} \right) \frac{1}{n^s}.$$

Por tanto, podemos escribir:

$$\sum_{d|n} a_{\frac{n}{d}} r_2(d) = \begin{cases} 1 & \text{si } n = 1 \\ 0 & \text{en otro caso} \end{cases} \tag{6.9}$$

Por lo tanto, podemos escribir:

$$a_1 = \frac{1}{r_2(1)} = \frac{1}{4}$$

y para $n \geq 2$,

$$a_n = -\sum_{\substack{d|n \\ d\neq n}} a_d r_2\left(\frac{n}{d}\right)$$

La segunda manera de definir la sucesión $\{a_n : n \in \mathbb{N}\}$, equivalente a la primera, es utilizar la fórmula del producto de series de Dirichlet y las ecuaciones (6.6) y (6.7), obteniendo:

$$a_n = \frac{1}{4} \sum_{d|n} \mu(d)\mu\left(\frac{n}{d}\right) \chi_4\left(\frac{n}{d}\right) \tag{6.10}$$

Utilizaremos la sucesión $\{a_n : n \in \mathbb{N}\}$ que acabamos de definir para demostrar la proposición. Sea $\rho := \frac{1}{H^2}$ ese número real. Sean $f, g : (0, \infty) \longrightarrow \mathbb{Z}$ las funciones definidas como sigue. Para todo número real $s > 0$,

$$f(s) := N_{\mathbb{Z}[i]}(\widetilde{\mathcal{W}}, s^{-1/2}), \qquad g(s) := \mathcal{N}_{\mathbb{Z}[i]}(\mathcal{W}, s^{-1/2}).$$

Entonces, por la Proposición 6.3.1, tenemos:

$$f(\rho) = \sum_{n\leq H^2} g(n\rho) r_2(n).$$

Ahora, observamos que si $n > H^2$, entonces $\frac{H}{\sqrt{n}} < 1$ y

$$g(n\rho) = \mathcal{N}_{\mathbb{Z}[i]}\left(\mathcal{W}, \frac{H}{\sqrt{n}}\right) = 0.$$

Por lo tanto, no hacemos cambio alguno al escribir la suma como serie infinita

$$f(\rho) = \sum_{n\geq 1} g(n\rho) r_2(n). \tag{6.11}$$

Por otro lado, podemos invertir la igualdad (6.11), como sigue:

$$g(\rho) = \sum_{n\geq 1} f(n\rho) a_n,$$

donde $(a_n)_{n\geq 1}$ es la secuencia descrita en el comienzo de esta demostración. Para demostrar esta última igualdad observamos que (como en la demostración de [63, Teor. 268]):

$$\sum_{n\geq 1} f(n\rho)a_n = \sum_{n\geq 1} a_n \left(\sum_{k\geq 1} r_2(k) g(kn\rho)\right) =$$

$$= \sum_{m\geq 1} \left(\sum_{d|m} a_{\frac{m}{d}} r_2(d)\right) g(m\rho).$$

Por lo tanto, por la igualdad (6.9), tenemos

$$\sum_{n\geq 1} a_n f(n\rho) = g(\rho),$$

como queríamos. Ahora consideramos la siguiente expresión:

$$S := \left|\rho^{m+1} g(\rho) - \frac{\pi \nu_m[\mathcal{W}]}{4(m+1)\zeta(m+1)L_4(m+1)}\right|.$$

Por las expresiones obtenidas hasta ahora, concluimos que

$$S = \left|\sum_{n\geq 1} \frac{a_n}{n^{m+1}} \left[(n\rho)^{m+1} f(n\rho) - \frac{\pi\nu_m[\mathcal{W}]}{m+1}\right]\right|.$$

Distinguimos dos términos,

$$S_1 := \left|\sum_{n\leq H^2} \frac{a_n}{n^{m+1}} \left[(n\rho)^{m+1} f(n\rho) - \frac{\pi\nu_m[\mathcal{W}]}{m+1}\right]\right|$$

$$S_2 := \left| \sum_{n>H^2} \frac{a_n}{n^{m+1}} \frac{\pi \nu_m[\mathcal{W}]}{m+1} \right|.$$

y tenemos claramente que:

$$S \leq S_1 + S_2,$$

donde hemos utilizado que $f(n\rho) = N_{Z[i]}(\widetilde{\mathcal{W}}, \frac{H}{\sqrt{n}}) = 0$ si $n > H^2$. Acotaremos cada término por separado. Con respecto a S_1 tenemos:

$$S_1 \leq \sum_{n \leq H^2} \frac{|a_n|}{n^{m+1}} \left| (n\rho)^{m+1} f(n\rho) - \frac{\pi \nu_m[\mathcal{W}]}{m+1} \right|. \tag{6.12}$$

Ahora, observamos que $\frac{\pi \nu_m[\mathcal{W}]}{m+1} = \mathscr{L}^{m+1}[\widetilde{\mathcal{W}} \cap B_{m+1}(0,1)]$, donde $\mathscr{L}^m$ es la medida de Lebesgue $m+1$ dimensional. Por lo tanto,

$$\left| (n\rho)^{m+1} f(n\rho) - \frac{\pi \nu_m[\mathcal{W}]}{(m+1)} \right| =$$

$$\left(\frac{\sqrt{n}}{H}\right)^{2m+2} \left| N_{\mathbb{Z}[i]}(\widetilde{\mathcal{W}}, H/\sqrt{n}) - \mathscr{L}^{m+1}[\widetilde{\mathcal{W}} \cap B_{m+1}(0,1)] \left(\frac{H}{\sqrt{n}}\right)^{2m+2} \right|.$$

Como $\widetilde{\mathcal{W}} \subseteq \mathbb{C}^{m+1} \equiv \mathbb{R}^{2m+2}$ es un cono, esta última cantidad es igual a

$$\left(\frac{\sqrt{n}}{H}\right)^{2m+2} \left| N_{\mathbb{Z}[i]}(\widetilde{\mathcal{W}}, H/\sqrt{n}) - \mathscr{L}^{m+1}\left[\widetilde{\mathcal{W}} \cap B\left(0, \frac{H}{\sqrt{n}}\right)\right] \right|.$$

Por último, observamos que,

$$|N_{\mathbb{Z}[i]}(\widetilde{\mathcal{W}}, H/\sqrt{n}) - N(\widetilde{\mathcal{W}}, H/\sqrt{n})| \leq 1,$$

en las notaciones del Teorema 6.2.1 (pues la única diferencia entre ambas magnitudes es la aparición o no del origen). Por el Teorema 6.2.1, concluimos:

$$\left| (n\rho)^{m+1} f(n\rho) - \frac{\pi \nu_m[\mathcal{W}]}{(m+1)} \right| \leq$$

$$\left(\frac{\sqrt{n}}{H}\right)^{2m+2} \left(1 + \beta(\widetilde{\mathcal{W}}) \sum_{i=0}^{2m+1} K_i \binom{2m+2}{i} \left(\frac{H}{\sqrt{n}}\right)^i\right) =$$

$$\left(\frac{\sqrt{n}}{H}\right)^{2m+2} + \beta(\widetilde{\mathcal{W}}) \sum_{i=0}^{2m+1} K_i \binom{2m+2}{i} \left(\frac{\sqrt{n}}{H}\right)^{2m+2-i} =$$

$$\left(\frac{\sqrt{n}}{H}\right)^{2m+2} + \beta(\widetilde{\mathcal{W}}) \sum_{i=0}^{2m+1} K_i \binom{2m+2}{i} \left(\frac{\sqrt{n}}{H}\right)^{2m+2-i} =$$

$$(n\rho)^{m+1} + \beta(\widetilde{\mathcal{W}}) \sum_{i=0}^{2m+1} K_i \binom{2m+2}{i} (n\rho)^{m+1-\frac{i}{2}}. \tag{6.13}$$

Por otro lado, por la igualdad (6.10) tenemos:

$$|a_n| \leq \frac{1}{4} \sum_{d|n} \left| \mu(d)\mu\left(\frac{n}{d}\right) \chi_4\left(\frac{n}{d}\right) \right| \leq \frac{1}{4} d(n)$$

donde $d(n)$ es el número de divisores de n. Por el Teorema [63, 289] deducimos que para todo $i \geq 3$,

$$\sum_{n=1}^{H^2} \frac{|a_n|}{n^{\frac{i}{2}}} \leq \frac{1}{4} \sum_{n\geq 1} \frac{d(n)}{n^{i/2}} = \frac{1}{4} \zeta(i/2)^2.$$

Por otro lado, para $i = 0, 1, 2$, tenemos que

$$\sum_{n\leq H^2} d(n) \leq \sum_{n\leq H^2} n \leq \sum_{n\leq H^2} H^2 = \lfloor H^2 \rfloor H^2 \leq H^4 = \frac{1}{\rho^2};$$

$$\sum_{n\leq H^2} \frac{d(n)}{n^{1/2}} \leq \sum_{n\leq H^2} n^{1/2} \leq \lfloor H^2 \rfloor H \leq H^3 = \frac{1}{\rho^{3/2}};$$

$$\sum_{n\leq H^2} \frac{d(n)}{n} \leq \sum_{n\leq H^2} 1 = \lfloor H^2 \rfloor \leq H^2 = \frac{1}{\rho}.$$

Juntando estas estimaciones con las ecuaciones (6.12) y (6.13), concluimos:

$$S_1 \leq \frac{\beta(\widetilde{\mathcal{W}})}{4} \left(8(m+1)^2 \rho^{m-1} + \sum_{i=3}^{2m+1} K_i \binom{2m+2}{i} \rho^{m+1-i/2} \zeta(\frac{i}{2})^2 \right).$$

En cuanto al término S_2, se tiene:

$$S_2 \leq \frac{\pi \nu_m[\mathcal{W}]}{m+1} \sum_{n\geq \rho^{-1}} \frac{|a(n)|}{n^{m+1}}.$$

Podemos acotar esa cantidad (cuando $m \geq 2$) por:

$$S_2 \leq \frac{\pi \nu_m[\mathcal{W}]}{4(m+1)} \sum_{n\geq \rho^{-1}} \frac{n}{n^{m+1}} \leq \frac{\pi \nu_m[\mathcal{W}]}{4(m+1)} \left[\rho^m + \sum_{n\geq \rho^{-1}+1} \frac{1}{n^m} \right] \leq$$

$$\frac{\pi \nu_m[\mathcal{W}]}{4(m+1)} \left[\rho^m + \int_{t\geq \rho^{-1}+1} \frac{1}{t^m}\, dt \right] = \frac{\pi \nu_m[\mathcal{W}]}{4(m+1)} \left[\rho^m + \frac{\rho^{m-1}}{m-1} \right].$$

Hemos acotado por tanto S_1 y S_2. Concluimos que la siguiente es una cota superior para S.

$$\frac{\beta(\widetilde{\mathcal{W}})}{4}\left(8(m+1)^2\rho^{m-1}+\sum_{i=3}^{2m+1}K_i\binom{2m+2}{i}\rho^{m+1-i/2}\zeta(\frac{i}{2})^2\right)+$$

$$\frac{\pi\nu_m[\mathcal{W}]}{4(m+1)}\left[\rho^m+\frac{\rho^{m-1}}{m-1}\right].$$

Finalmente, observamos que S es exactamente la cantidad que queremos calcular en esta proposición, pero dividida por H^{2m+2}. Sustituyendo ρ por $\frac{1}{H^2}$ obtenemos la afirmación de la Proposición. Podemos acotar (groseramente) el término

$$2\beta(\widetilde{\mathcal{W}})(m+1)^2H^2+\frac{\pi\nu_m[\mathcal{W}]}{4(m+1)}\left(H^2+\frac{H^4}{m-1}\right)$$

por la cantidad

$$3\beta(\widetilde{\mathcal{W}})(m+1)^2H^4.$$

■

El siguiente corolario se sigue fácilmente de la Proposición 6.3.2, del mismo modo que deducíamos el Corolario 6.2.2 a partir del Teorema 6.2.1.

Corolario 6.3.3 *Sea $m \geq 2$ un número natural y sea $\mathcal{W} \subset \mathbb{P}_m(\mathbb{C})$ un subconjunto del espacio proyectivo. Sea $H \geq 4(m+1)^2$ un número real. Entonces, se tiene:*

$$\left|\mathcal{N}_{\mathbb{Z}[i]}(\mathcal{W},H)-\frac{\pi\nu_m[\mathcal{W}]}{4\zeta(m+1)L_4(m+1)(m+1)}H^{2m+2}\right|\leq$$

$$3\beta(\widetilde{\mathcal{W}})(m+1)K_{2m+1}H^{2m+1},$$

donde $\widetilde{\mathcal{W}} \subseteq \mathbb{C}^{m+1}$ es el cono afín de $\mathcal{W}$ (con el origen incluido).

Finalmente, utilizaremos el siguiente resultado técnico:

Lema 6.3.4 *Sean $A, B, C, D, \alpha_1, \alpha_2$ números reales positivos tales que:*

$$|A-B| \leq \alpha_1, \quad |C-D| \leq \alpha_2, \quad |A| \leq |C|.$$

Entonces, también se tiene que:

$$\left|\frac{A}{C}-\frac{B}{D}\right| \leq \frac{\alpha_1+\alpha_2}{|D|}.$$

Demostración.– Observamos simplemente que

$$\left|\frac{A}{C}-\frac{B}{D}\right| \leq \left|\frac{A}{C}-\frac{A}{D}\right| + \left|\frac{A}{D}-\frac{B}{D}\right| \leq A\left|\frac{D-C}{DC}\right| + \frac{|A-B|}{D} \leq$$

$$\frac{|D-C|}{D} + \frac{|A-B|}{D} \leq \frac{\alpha_2}{D} + \frac{\alpha_1}{D},$$

como queríamos. ∎

Teorema 6.3.5 *Sea $m \geq 2$ y sea $\mathcal{W} \subset \mathbb{P}_m(\mathbb{C})$ un subconjunto medible del espacio proyectivo. Sea $H > 4(m+1)^2$ un número real. Entonces, se tiene que:*

$$\left|\frac{\mathcal{N}_{\mathbb{Z}[i]}(\mathcal{W},H)}{\mathcal{N}_{\mathbb{Z}[i]}(\mathbb{P}_m(\mathbb{C}),H)} - \frac{\nu_m[\mathcal{W}]}{\vartheta_m}\right| \leq \frac{32\beta(\widetilde{\mathcal{W}})(m+1)^{3/2}}{H},$$

donde $\widetilde{\mathcal{W}}$ es el cono afín de $\mathcal{W}$ (con el origen incluido).

Demostración.– Por el Corolario 6.3.3, tenemos:

$$\left|\mathcal{N}_{\mathbb{Z}[i]}(\mathcal{W},H) - \frac{\pi\nu_m[\mathcal{W}]}{4\zeta(m+1)L_4(m+1)(m+1)}H^{2m+2}\right| \leq$$

$$3\beta(\widetilde{\mathcal{W}})(m+1)K_{2m+1}H^{2m+1}.$$

También por ese mismo corolario, tenemos:

$$\left|\mathcal{N}_{\mathbb{Z}[i]}(\mathbb{P}_m(\mathbb{C}),H) - \frac{\pi\vartheta_m}{4\zeta(m+1)L_4(m+1)(m+1)}H^{2m+2}\right| \leq$$

$$3\beta(\mathbb{C}^{m+1})(m+1)K_{2m+1}H^{2m+1} = 3(m+1)K_{2m+1}H^{2m+1}.$$

Por el Lema 6.3.4, concluimos que

$$\left|\frac{\mathcal{N}_{\mathbb{Z}[i]}(\mathcal{W},H)}{\mathcal{N}_{\mathbb{Z}[i]}(\mathbb{P}_m(\mathbb{C}),H)} - \frac{\nu_m[\mathcal{W}]}{\vartheta_m}\right| \leq$$

$$\frac{24\beta(\widetilde{\mathcal{W}})(m+1)^2K_{2m+1}H^{2m+1}L_4(m+1)\zeta(m+1)}{\pi\vartheta_m H^{2m+2}}.$$

Además,

$$\frac{K_{2m+1}}{\vartheta_m} = B\left(\frac{1}{2},m+1\right) \leq \frac{\sqrt{\pi}}{\sqrt{m+3/4}}.$$

La última desigualdad se sigue nuevamente de las desigualdades de Gautschi (véase [43]). Finalmente, observamos que

$$L_4(m+1)\zeta(m+1) \leq \zeta(m+1)^2 \leq \frac{3}{2}.$$

El teorema se sigue de estas estimaciones. ∎

También podemos escribir la versión en *talla bit* de este resultado, que es el Teorema 6.1.1 en la introducción de este capítulo (en efecto, basta sustituir en el Teorema 6.3.5 la altura absoluta H por 2^h, donde h es la talla bit).

6.4. Estimaciones discretas para los números de condicionamiento

A la luz del Teorema 6.1.1, podemos encontrar con facilidad las versiones discretas (más cercanas a la realidad computacional que las continuas) de los distintos resultados que hemos expuesto en secciones anteriores de esta memoria. En esta sección demostramos estos resultados discretos para los casos del condicionamiento lineal y no lineal, y también para el caso de la distancia a las variedades no lineales de corango dado.

6.4.1. El condicionamiento lineal

Recuperamos las notaciones la Sección 2.5 del Capítulo 2. Esto es, $n_2 \geq n_1 \geq r \geq 2$ son números naturales, el espacio $\mathbb{P}(\mathcal{M}_{n_1\times n_2}(\mathbb{C}))$ es el espacio proyectivo de las matrices de talla $n_1 \times n_2$, y $\Sigma^r_{\mathcal{M}}$ es la variedad proyectiva definida como sigue:

$$\Sigma^r_{\mathcal{M}} := \{A \in \mathbb{P}(\mathcal{M}_{n_1\times n_2}(\mathbb{C})) : rank(A) \leq r\}.$$

Llamamos talla bit de una matriz proyectiva racional $A \in \mathbb{P}(\mathcal{M}_{n_1\times n_2}(\mathbb{Q}))$ a la talla bit de A considerada como punto proyectivo en $\mathbb{P}_{n_1n_2-1}(\mathbb{C}) \equiv \mathbb{P}(\mathcal{M}_{n_1\times n_2}(\mathbb{C}))$. Además, para todo número real $h > 0$ denotaremos por

$$\mathbb{P}(\mathcal{M}_{n_1\times n_2}(\mathbb{C}))^{\mathbb{Q}[h]}$$

el conjunto de matrices de talla bit a lo sumo h.

Demostración del Corolario 6.1.2

Sea $\mathcal{W} \subset \mathbb{P}(\mathcal{M}_{n_1\times n_2}(\mathbb{C}))$ el subconjunto definido como:

$$\mathcal{W} := \left\{A \in \mathbb{P}(\mathcal{M}_{n_1\times n_2}(\mathbb{C})) : \kappa^{(r)}_D(A) > \frac{1}{\varepsilon}\right\}.$$

Por el Corolario 2.5.1, sabemos que

$$\frac{\nu_{n_1n_2-1}[\mathcal{W}]}{\vartheta_{n_1n_2-1}} \leq 2\left[\frac{e\ (n_1n_2-1)\sqrt{r}}{(n_1-r+1)(n_2-r+1)}\ \varepsilon\right]^{2(n_1-r+1)(n_2-r+1)}.$$

Por lo tanto, por el Teorema 6.1.1, basta comprobar que el cono afín $\widetilde{\mathcal{W}}$ de $\mathcal{W}$ satisface:

$$\beta(\widetilde{\mathcal{W}}) \leq (16r+1)^{2(2n_1n_2+2)}.$$

Por el Lema 6.2.4, es suficiente con demostrar que $\widetilde{\mathcal{W}}$ es la $2(n_1n_2)$–proyección de un conjunto semi–algebraico $(2, 2r)$–definible. Verificamos pues esta

condición. Por el Teorema 1.5.10, sabemos que una matriz afín $A \in \mathcal{M}_{n_1 \times n_2}$ está en $\mathcal{W}$ sí y sólo si existe una matriz $B \in \mathcal{M}_{n_1 \times n_2}$ tal que:

$$B \in \Sigma_{\mathcal{M}}^{r-1}, \quad d_{\mathbb{P}}(A, B) < \varepsilon. \tag{6.14}$$

La primera de las dos condiciones puede expresarse como

$$\det(M_r) = 0, \quad \text{para todo menor } M_r \text{ de tamaño } r \text{ de } B.$$

Como estamos tratando con matrices complejas, esto son $2\binom{n_1}{r}\binom{n_2}{r}$ ecuaciones reales de grado r en los coeficientes de B. Sumando los cuadrados de esas ecuaciones e igualando a cero, obtenemos que es equivalente a una sola ecuación de grado $2r$ en los coeficientes de B. En cuanto a la segunda de las condiciones de la expresión (6.14), podemos escribirla como

$$\sqrt{1 - \frac{|\langle A, B\rangle_F|^2}{\|A\|_F^2 \|B\|_F^2}} < \varepsilon.$$

Esta última condición puede escribirse por tanto como una única ecuación de grado 4 en los coeficientes de A y B.
Hemos expresado $\widetilde{\mathcal{W}}$ del modo

$$\widetilde{\mathcal{W}} = \{A \in \mathcal{M}_{n_1 \times n_2}(\mathbb{C}) : \exists B \in \mathcal{M}_{n_1 \times n_2}, EC_1, EC_2\},$$

donde EC_1 es una ecuación de grado $2r$ y EC_2 es una desigualdad de grado $4 \leq 2r$, ambas en los coeficientes de A y B. Por lo tanto, $\widetilde{\mathcal{W}}$ es la $2(n_1 n_2)$–proyección de un conjunto $(2, 2r)$ definible, lo que termina la demostración del corolario.

■

6.4.2. El condicionamiento no–lineal

Recuperamos ahora las notaciones de la Sección 1.4, en el caso cero–dimensional. Esto es, consideramos números naturales $n \geq m \geq r$ y una m–tupla de números naturales $(d) := (d_1, \ldots, d_m)$ tal que $d_i \geq 2$ para algún i. Entonces, denotamos por $\mathbb{P}(\mathcal{H}_{(d)}^m)$ el espacio proyectivo de los sistemas de m ecuaciones homogéneas de grados respectivos $d_1, \ldots, d_m$ en las incógnitas $X_0, \ldots, X_n$ con coeficientes complejos. Al igual que en secciones anteriores, consideraremos $\mathbb{P}(\mathcal{H}_{(d)}^m)$ con la estructura Riemanniana inducida por la métrica de Kostlan, definida en la Sección 1.4. El número de condicionamiento $\mu_{\text{norm}}^{(r)}(f, \zeta)$, definido para todo par (f, ζ) donde ζ es una solución de f, es como en la Sección 1.6. También denotaremos por $\Sigma_{(d)}^r$ el conjunto de sistemas de ecuaciones $f \in \mathbb{P}(\mathcal{H}_{(d)}^m)$ tales que el rango de la matriz diferencial $d_\zeta f$ en cualquier solución ζ de f es menor o igual que r. Consideramos también el condicionamiento del caso peor

$$\mu_{\text{worst}}^{(r)}(f) := \max_{\zeta \in V(f)} \mu_{\text{norm}}^{(r)}(f, \zeta).$$

Pasamos ahora a definir la talla bit representativa de un punto del espacio $\mathbb{P}(\mathcal{H}^m_{(d)})$. Para hacer esto utilizamos la aplicación Δ^{-1} (correspondiente a la matriz de Kostlan Δ definida en la Sección 1.4), de la que hemos hecho uso extensivo en otros capítulos de esta memoria. Recordemos que la aplicación Δ^{-1} viene dada por

$$\begin{array}{rccc} \Delta^{-1}: & (\mathbb{P}(\mathcal{H}^m_{(d)}), can) & \longrightarrow & \mathbb{P}(\mathcal{H}^m_{(d)}) \\ & f & \mapsto & \Delta^{-1} f, \end{array}$$

donde $(\mathbb{P}(\mathcal{H}^m_{(d)}), can)$ representa el espacio $\mathbb{P}(\mathcal{H}^m_{(d)})$ con la estructura Riemanniana canónica heredada de la métrica usual en el correspondiente espacio afín. Por el Lema 1.4.1, sabemos que Δ^{-1} es una isometría, con lo que conserva volúmenes. Sea $\Delta^{-1}\mathbb{P}_N(\mathbb{Q})$ la imagen mediante esta isometría del conjunto de los puntos racionales de $\mathbb{P}(\mathcal{H}^m_{(d)})$. Para todo sistema $f \in \Delta^{-1}\mathbb{P}_N(\mathbb{Q})$, definimos la *talla bit representativa* de f como la talla bit de $\Delta f \in \mathbb{P}(\mathcal{H}^m_{(d)})$, considerado como punto proyectivo (véase Sección 6.3). La talla bit representativa de un punto $f \in \mathbb{P}(\mathcal{H}^m_{(d)})$ así definida es esencialmente equivalente al número de cifras binarias necesarias para representar f a través de la isometría Δ^{-1}. Utilizaremos esta noción de talla por ser la que nos permite demostrar de un modo más directo los resultados que siguen. En lo que sigue, denotaremos por

$$\mathbb{P}(\mathcal{H}^m_{(d)})^{\mathbb{Q}_\Delta(h)}$$

el conjunto de los sistemas de ecuaciones en $\mathbb{P}(\mathcal{H}^m_{(d)})$ de talla bit representativa a lo más h.

Demostración del Corolario 6.1.3

Sea $\mathcal{W} \subset \mathbb{P}(\mathcal{H}_{(d)})$ el subconjunto definido como:

$$\mathcal{W} := \left\{ f \in \mathbb{P}(\mathcal{H}_{(d)}) : d_{\mathbb{P}}(f, \Sigma^r_{(d)}) < \varepsilon \right\}.$$

Como se indica en el Teorema 3.4.1, la probabilidad (continua) $\frac{\nu_\Delta[\mathcal{W}]}{\nu_{\mathbb{P}(\mathcal{H}_{(d)})}[\mathbb{P}(\mathcal{H}_{(d)})]}$ de que un sistema elegido al azar $f \in \mathbb{P}(\mathcal{H}_{(d)})$ verifique $d_{\mathbb{P}}(f, \Sigma^r_{(d)}) < \varepsilon$ es menor o igual que:

$$2\prod_{i=1}^{n}(d_i+1)\binom{n+1}{r}\binom{n}{r}\left(\frac{e\ N(r+1)\ d\ \varepsilon}{(n-r)^2}\right)^{2(n-r)^2}.$$

Por tanto, por el Teorema 6.1.1, basta comprobar que el cono afín $\widetilde{\mathcal{W}}$ de $\mathcal{W}$ satisface:

$$\beta(\widetilde{\mathcal{W}}) \leq (16(r+1)d+1)^{4(N+n+3)}.$$

Por el Lema 6.2.4, es suficiente con demostrar que $\widetilde{\mathcal{W}}$ es la $(2N+2n+4)$–proyección de un conjunto semi–algebraico $(2, 2(r+1)d)$–definible. Verificamos pues esta condición. En efecto, un punto $f \in \mathcal{H}_{(d)}$ está en $\widetilde{\mathcal{W}}$ si y solo si existe un sistema $g \in \mathcal{H}_{(d)}$ y un punto afín $\zeta \in \mathbb{C}^{n+1}$ tales que:

- $\|\zeta\|_2^2 = 1$,
- $g(\zeta) = 0$,
- $rank(d_\zeta g) \leq r$,
- $d_{\mathbb{P}}(f, g) < \varepsilon$.

La primera condición es una igualdad real de grado 2. La segunda de las condiciones puede escribirse como n ecuaciones complejas de grado d. La tercera condición es equivalente a la anulación de $\binom{n+1}{r+1}\binom{n}{r+1}$ ecuaciones complejas de grado $(r+1)d$. Por tanto, las tres primeras condiciones juntas pueden expresarse como una única ecuación real de grado $2(r+1)d$. En cuanto a la última condición, observamos que es equivalente a

$$\sqrt{1 - \frac{|\langle f, g\rangle|_\Delta^2}{\|f\|_\Delta^2 \|g\|_\Delta^2}} < \varepsilon,$$

esto es una desigualdad de grado 4. De este modo, hemos expresado $\widetilde{\mathcal{W}}$ del modo

$$\widetilde{\mathcal{W}} = \{f \in \mathcal{H}_{(d)} : \exists g \in \mathcal{H}_{(d)}, \zeta \in \mathbb{C}^{n+1}, EC_1, EC_2\},$$

donde EC_1 es una ecuación de grado $2(r+1)d$ y EC_2 es una desigualdad de grado $4 \leq 2rd$, todas ellas en los coeficientes de f, g y ζ. Deducimos que $\widetilde{\mathcal{W}}$ es la $(2N+2n+4)$–proyección de un conjunto $(2, 2(r+1)d)$ definible, lo que termina la demostración del corolario. ∎

Demostración del Corolario 6.1.4

Sea $\mathcal{W} \subset \mathbb{P}(\mathcal{H}_{(d)}^m)$ el subconjunto definido como:

$$\mathcal{W} := \left\{ f \in \mathbb{P}(\mathcal{H}_{(d)}^m) : \mu_{\text{worst}}^{(m)}(f) > \frac{1}{\varepsilon} \right\}.$$

Como se indica en la demostración del Teorema 3.7.6, la probabilidad de que un sistema elegido al azar $f \in \mathbb{P}(\mathcal{H}_{(d)}^m)$ verifique $\mu_{\text{worst}}^{(m)}(f) > \frac{1}{\varepsilon}$ es menor o igual que:

$$2\mathcal{D}\left[10N^{1/2}mn^{1/2}d^{3/2}\right]^{2(n-m)}[6N^{1/2}mn^{1/2}\ \varepsilon]^4.$$

Por tanto, por el Teorema 6.1.1, basta comprobar que el cono afín $\widetilde{\mathcal{W}}$ de $\mathcal{W}$ satisface:

$$\beta(\widetilde{\mathcal{W}}) \leq (24rd+1)^{4(N+n+3)}.$$

Por el Lema 6.2.4, es suficiente con demostrar que $\widetilde{\mathcal{W}}$ es la $(2N+2n+4)$–proyección de un conjunto semi–algebraico $(2, 2md)$–definible. Verificamos pues esta condición. Por el Teorema 3.1.1, un sistema afín $f \in \mathcal{H}^m_{(d)}$ está en $\widetilde{\mathcal{W}}$ si y solo si existen un punto $\zeta \in \mathbb{C}^{n+1}$ y un sistema $g \in \mathcal{H}^m_{(d)}$ tales que:

- $\|\zeta\|_2^2 = 1$,
- $f(\zeta) = g(\zeta) = 0$,
- $rank(d_\zeta g) \leq m-1$,
- $d_{\mathbf{P}}(f,g) < \varepsilon$.

La primera condición es una igualdad real de grado 2. La segunda de las condiciones puede escribirse como $2n$ ecuaciones complejas de grado d. La tercera condición es equivalente a la anulación de $\binom{n+1}{m}$ ecuaciones complejas de grado md. Por tanto, las tres condiciones juntas pueden expresarse como una única ecuación real de grado $2md$. En cuanto a la última condición, es equivalente a

$$\sqrt{1 - \frac{|\langle f,g\rangle|^2_\Delta}{\|f\|^2_\Delta \|g\|^2_\Delta}} < \varepsilon,$$

esto es una desigualdad de grado 4. De este modo, hemos expresado $\widetilde{\mathcal{W}}$ del modo

$$\widetilde{\mathcal{W}} = \{f \in \mathcal{H}^m_{(d)} : \exists g \in \mathcal{H}^m_{(d)}, \zeta \in \mathbb{C}^{n+1}, EC_1, EC_2\},$$

donde EC_1 es una ecuación de grado $2md$ y EC_2 es una desigualdad de grado $4 \leq 2md$, todas ellas en los coeficientes de f, g y ζ. Deducimos que $\widetilde{\mathcal{W}}$ es la $(2N+2n+4)$–proyección de un conjunto $(2, 2md)$ definible, lo que termina la demostración del corolario. ∎

6.5. El problema 17 de Smale

Estructuramos la prueba del Teorema 6.1.5 en las dos subsecciones que siguen.

6.5.1. Homotopía y conjuntos semi–algebraicos

Sea $M := (M^0, \ldots, M^n) \in \mathcal{H}_{(1)} = \mathcal{M}_{n\times(n+1)}(\mathbb{C})$ una matriz no–singular, con columnas $M^0, \ldots, M^n \in \mathbb{C}^n$. Definimos el vector complejo

$$v(M) := (v(M)_0, \ldots, v(M)_n) \in \mathbb{C}^{n+1}$$

como sigue:

$$v(M)_i := \begin{cases} \det(M^1, \ldots, M^n) & \text{si i=0,} \\ (-1)^i \det(M^0, \ldots, M^{i-1}, M^{i+1}, \ldots, M^n) & 1 \leq i \leq n-1, \\ (-1)^n \det(M^0, \ldots, M^{n-1}) & \text{si i=n.} \end{cases}$$

Sea $\Omega(M)$ la matriz definida como sigue. Aplicamos el procedimiento de Gram–Schmidt a los $n+1$ vectores complejos $\{v(M), M_1, \ldots, M_n\}$, donde $M_1, \ldots, M_n$ son las filas de M. Sean $v_0^M, v_1^M, \ldots, v_n^M$ los vectores obtenidos mediante este procedimiento. Entonces, definimos:

$$\Omega(M) := \begin{pmatrix} v_0^M \\ \vdots \\ v_n^M \end{pmatrix}^t \in \mathcal{U}_{n+1}.$$

Obsérvese que para toda matriz no–singular $M \in \mathcal{H}_{(1)}$, se tiene que e_0 pertenece al núcleo de $M\Omega(M)$. En efecto,

$$M\Omega(M)e_0 = M(v_0^M)^t = \|v_0^M\|_2 M v(M)^t.$$

Ahora, la i–ésima coordenada del vector $Mv(M)^t$ es igual a

$$\sum_{j=0}^{n} m_{ij} v(M)_j = \sum_{j=0}^{n} (-1)^j m_{ij} \det(M^0, \ldots, M^{j-1}, M^{j+1}, \ldots, M^n) =$$

$$\det \begin{pmatrix} m_{i0} & \cdots & m_{in} \\ m_{00} & \cdots & m_{0n} \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ m_{i0} & \cdots & m_{in} \\ \cdots & \cdots & \cdots \\ m_{n0} & \cdots & m_{nn} \end{pmatrix} = 0,$$

por el desarrollo en la primera fila del determinante de una matriz. Sea $\mathcal{W} \subseteq \mathbb{R} \times \mathbb{C}^{N+1}$ un conjunto semi–algebraico complejo, tal que $\mathcal{W} \subseteq B_\infty(0,1) := \{z \in \mathbb{R} \times \mathbb{C}^{N+1} : \|z\|_\infty \leq 1\}$. Para todo número natural H, denotamos por $N_{\frac{1}{H}}(\mathcal{W})$ el siguiente número:

$$N_{\frac{1}{H}}(\mathcal{W}) := \sharp \left(\mathcal{W} \cap \frac{1}{H} \mathbb{Z}^{2N+3} \right),$$

donde $\frac{1}{H}\mathbb{Z}^{2N+3} := \left\{ \frac{1}{H} z : z \in \mathbb{Z}^{2N+3} \right\}$ es ese retículo en $\mathbb{R} \times \mathbb{C}^{N+1} \equiv \mathbb{R}^{2N+3}$. Sea $m \geq 1$ un natural positivo cualquiera, y sea $\{y_i : 1 \leq i \leq m\} \subseteq Y$ una colección finita de puntos de Y. Consideramos la discrepancia

$$\mathscr{D}_{\{y_1, \ldots, y_m\}} := \left| \frac{1}{m} \sum_{i=1}^{m} \widetilde{A_\varepsilon}(y_i) - E_Y[\widetilde{A_\varepsilon}] \right|,$$

donde $\widetilde{A_\varepsilon}$ es la función definida en la Sección 4.4. Para todo número real positivo $0 < \varepsilon$ consideramos el siguiente subconjunto de $Y \times \mathbb{S}_\Delta$:

$$R_\varepsilon := \{((t,h,M),f) \in Y \times \mathbb{S}_\Delta : \mathfrak{L}_\Delta(f, G_{(d)}(t,h,M), e_0) \cap (\Sigma'_{(d)})_\varepsilon \neq \emptyset\}$$

donde estamos utilizando las notaciones de la Proposición 4.2.3. Para ser más precisos, exigimos también que f y $G_{(d)}(t,h,M), e_0)$ no sean iguales ni antipodales. Entonces, para cada sistema $f \in \mathbb{S}_\Delta$, consideramos el conjunto $R_{\varepsilon,f}$ definido como sigue:

$$R_{\varepsilon,f} := \{(t,h,M) \in Y : ((t,h,M),f) \in R_\varepsilon\} \subseteq Y.$$

El siguiente lemma describe $R_{\varepsilon,f}$ como conjunto semi–algebraico complejo.

Lema 6.5.1 *Para todo número real positivo $\varepsilon > 0$ y para todo sistema $f \in \mathbb{S}_\Delta$, el conjunto $R_{\varepsilon,f}$ es la k–proyección de un conjunto semi-algebraico (s,d')–definible, donde*

$$\begin{aligned} k &= 3n+10, \\ s &\leq d^{O(nN^3)}, \\ d' &\leq d^{O(nN^3)}, \end{aligned}$$

y d es el máximo de los grados de los polinomios en f.

Demostración.– Observamos que un punto $((t,h,M),f) \in Y \times \mathbb{S}_\Delta$ está en R_ε si y sólo si se tiene que:

$$G_{(d)}(t,h,M) \neq \pm f, \tag{6.15}$$

y

$$\exists\ (s_1,s_2) \in S^1(\mathbb{R}), \zeta \in S^1(\mathbb{C}^{n+1}) \text{ tales que}$$

$$(s_1 f + s_2 G_{(d)}(t,h,M), \zeta) \in \mathfrak{L}_\Delta(f, G_{(d)}(t,h,M), e_0) \text{ y} \tag{6.16}$$

$$(s_1 f + s_2 G_{(d)}(t,h,M), \zeta) \in (\Sigma'_{(d)})_\varepsilon, \tag{6.17}$$

donde denotamos por el mismo símbolo el punto afín $\zeta \in S^1(\mathbb{C}^{n+1})$ y el punto proyectivo asociado $\mathbb{P}_n(\mathbb{C})$. Usando la descomposición en valores singulares y escribiendo $g(s_1,s_2,t,h,M) := s_1 f + s_2 G_{(d)}(t,h,M)$, obtenemos otra formulación equivalente de la condición (6.17):

$$\exists \mu \in \mathbb{R} : |\mu| < \varepsilon^2,$$

$$\det(\mu Id_n - \Delta(d)^{-1/2} T_\zeta g(s_1,s_2,t,h,M) T_\zeta g(s_1,s_2,t,h,M)^* \Delta(d)^{-1/2}) = 0,$$

donde $T_\zeta g(s_1,s_2,t,h,M)$ es la matriz diferencial de $g(s_1,s_2,t,h,M)$ restringida al espacio $\zeta^\perp$ en alguna base ortonormal. Equivalentemente, la condición (6.17) puede escribirse como:

$$\exists \mu \in \mathbb{R} : |\mu| < \varepsilon^2,$$

$$\det(\mu I d_n - \Delta(d)^{-1/2} d_\zeta g(s_1, s_2, t, h, M) d_\zeta g(s_1, s_2, t, h, M)^* \Delta(d)^{-1/2}) = 0,$$

de modo que hemos eliminado la dependencia en el ortogonal de ζ. Esto se debe a que los valores singulares de $d_\zeta g(s_1, s_2, t, h, M)$ y $T_\zeta g(s_1, s_2, t, h, M)$ coinciden. Observamos que para cualquier secuencia de números reales positivos $\lambda_1, \ldots, \lambda_n$, tenemos la siguiente igualdad:

$$\Omega \begin{pmatrix} \lambda_1 M_1 \\ \vdots \\ \lambda_n M_n \end{pmatrix} = \Omega(M). \tag{6.18}$$

Consideramos la siguiente secuencia de números positivos:

$$\begin{array}{rcl} \lambda_0 & := & \frac{1}{\|v(M)\|_2}, \\ \lambda_1 & := & \frac{1}{\|M_1 - \langle M_1, v_0^M \rangle_2 v_0^M\|_2}, \\ & \vdots & \\ \lambda_n & := & \frac{1}{\|M_n - \sum_{i=1}^{n-1} \langle M_n, v_i^M \rangle_2 v_i^M\|_2}. \end{array}$$

Entonces, por la ecuación (6.18), los vectores $v_0^M, \ldots, v_n^M$ que definen $\Omega(M)$ satisfacen las siguientes igualdades:

$$\begin{array}{rcl} v_0^M & = & \lambda_0 v(M), \\ v_1^M & = & \lambda_1 (M_1 - \langle M_1, v_0^M \rangle_2 v_0^M), \\ & \vdots & \\ v_n^M & = & \lambda_n (M_n - \sum_{i=1}^{n-1} \langle M_n, v_i \rangle_2 v_i). \end{array}$$

Por lo tanto, podemos expresar cada coordenada de v_k^M como un polinomio en las variables $\{\lambda_j, m_{ij}\}_{\substack{i=1\ldots n \\ j=0\ldots n}}$, de grado acotado por $2^{2k}(n+1)$, y todo elemento de $\Omega(M)$ puede expresarse como un polinomio de grado a lo más $2^{2n}(n+1)$ en esas variables. Además, para todo $t \in [0,1]$, consideramos puntos $(t_1, t_2) \in S^1(\mathbb{R})$, y $l_1, l_2 \in \mathbb{R}$ tales que:

$$t_2 = \left(\frac{n^2+n}{N}\right)^{1/2} t^{\frac{1}{2n^2+2n}}, \quad l_1 := \frac{1}{\|h\|_2}, \quad l_2 := \frac{1}{\|M\|_F}.$$

Entonces, podemos escribir:

$$g(s_1, s_2, t, h, M) = s_1 f + s_2 \left[t_1 l_1 \Delta^{-1} h + t_2 \psi_{e_0}^{-1} \left(T_{e_0}(l_2 M \Omega(M)) \right) \right].$$

Deducimos que todo coeficiente de $g(s_1, s_2, t, h, M)$ puede expresarse como un polinomio de grado acotado por $2^{2n+1}(n+1)$ en las variables

$$t_1, t_2, l_1, l_2, s_1, s_2, \lambda_0, \ldots, \lambda_n, M, h.$$

Por lo tanto, la correspondiente expresión para los elementos de la matriz $d_\zeta g(s_1, s_2, t, h, M)$ es un polinomio de grado a lo más $2^{2n+1}(n+1)d$ en las variables de arriba más las variables de ζ. Deducimos que la igualdad

$$\det(\mu Id_n - \Delta(d)^{-1/2} d_\zeta g(s_1, s_2, t, h, M) d_\zeta g(s_1, s_2, t, h, M)^* \Delta(d)^{-1/2}) = 0,$$

puede expresarse como un polinomio de grado a lo más $2^{2n+2}(n+1)^2 d$ en las variables $\mu, s_1, s_2, t_1, t_2, l_1, l_2, h, M, \lambda_0, \ldots, \lambda_n, \zeta$. Concluimos que la condición (6.17) es equivalente a

$$\exists \mu \in \mathbb{R} : \mu^2 < \varepsilon^4, (P_1 = 0),$$

donde P_1 es un polinomio de grado acotado por $2^{2n+2}(n+1)^2 d$ en las variables

$$\mu, t_1, t_2, l_1, l_2, s_1, s_2, \zeta_0, \ldots, \zeta_n, \lambda_0, \ldots, \lambda_n, M, h.$$

Por lo tanto, la condición (6.17) añade un existencial, una desigualdad de grado 1 y una igualdad de grado acotado por $2^{2n+2}(n+1)^2 d$ en todas esas variables. Sin embargo, la finura con la que hemos hecho estos cálculos no resultará demasiado útil pues es la condición (6.16) la que marcará el grado y el número de polinomios necesarios para definir $R_{\varepsilon,f}$ como conjunto semi–algebraico. Lo que sí hemos obtenido es el número de variables cuantificadas que necesitamos: $3n + 10$.

Respecto a la condición (6.15), observamos que es equivalente a que el rango de la matriz formada por los coeficientes de f y $G_{(d)}(t, h, M)$ sea 2. Esto equivale a una desigualdad de grado acotado también por $2^{2n+2}(n+1)^2 d$.

Vamos ahora con la condición (6.16). Primero observamos que, levantando las soluciones a la esfera, el conjunto $\mathfrak{L}_\Delta(f, G_{(d)}(t, h, M), e_0)$ puede ser visto como una componente conexa del conjunto descrito a continuación:

$$\{(f', \zeta') \in \mathbb{S}_\Delta \times \mathbb{S}^1(\mathbb{C}^{n+1}) : f'(\zeta') = 0,$$

$$f' \in \mathfrak{L}_\Delta(f, G_{(d)}(t, h, M)), f' \neq -G_{(d)}(t, h, M)\},$$

donde $\mathfrak{L}_\Delta(f, g)$ es el círculo máximo que contiene a f y a g. Denotemos por B ese conjunto. Hemos visto que los coeficientes de $G_{(d)}(t, h, M)$ son polinomios de grado acotado por $2^{2n+1}(n+1)$ en las variables

$$t_1, t_2, l_1, l_2, \lambda_0, \ldots, \lambda_n, h, M.$$

La condición $f' \neq -G_{(d)}(t, h, M)$ puede escribirse como la no–anulación de la suma de cuadrados de las diferencias entre los coeficientes, equivale por tanto a una desigualdad de grado acotado por $2^{2n+2}(n+1)$. Dado que estamos suponiendo se satisface la condición (6.15), observamos que la condición $f' \in \mathfrak{L}_\Delta(f, G_{(d)}(t, h, M))$ equivale a

$$rank(f\ G_{(d)}(t, h, M)\ f') \leq 2,$$

donde $(f\ G_{(d)}(t,h,M)\ f') \in \mathcal{M}_{2N+2,3}(\mathbb{R})$ es visto como una matriz real. Ésta condición equivale a la anulación de $3\binom{2N+2}{2}$ ecuaciones de grado acotado por $2^{2n+1}(n+1)+1$ en las anteriores variables. Por tanto, el conjunto B se puede expresar como el conjunto de pares $(f',\zeta') \in \mathcal{H}_{(d)} \times \mathbb{C}^{n+1}$ que satisfacen EC_1, EC_2 donde EC_1 es una desigualdad de grado acotado por $2^{2n+2}(n+1)+1 \leq (2d)^{2n+2}(n+1)$ y EC_2 es una ecuación de grado menor que máx$\{2^{2n+2}(n+1)+1, 2(d+1)\} \leq (2d)^{2n+2}(n+1)$. Por [7] concluimos que toda componente conexa de B puede expresarse como un conjunto semi–algebraico libre de cuantificadores con un número de polinomios del orden de

$$2^{2N+n+9+1}[(2d)^{2n+2}(n+1)]^{O((2N+n+9)^3)} \leq d^{O(nN^3)},$$

con grados acotados por esa misma cantidad. Concluimos que la condición (6.16) puede expresarse como una condición semialgebraica en la que aparecen a lo sumo

$$d^{O(nN^3)}$$

polinomios de grado acotado por

$$2^{n+1}(n+1)d^{O(nN^3)} \leq d^{O(nN^3)}.$$

Esto termina la demostración del resultado. ∎

Lema 6.5.2 *Con las notaciones anteriores, para toda colección de puntos $\{y_i : 1 \leq i \leq m\} \subseteq Y$, se tiene la siguiente desigualdad:*

$$\mathscr{D}_{\{y_1,\ldots,y_m\}} \leq \frac{1}{\nu_\Delta[\mathbb{S}_\Delta]} \int_{f\in\mathbb{S}_\Delta} \left| \frac{1}{m}\sum_{i=1}^{m} \chi_{R_\varepsilon}(y_i,f) - \frac{\nu_Y[R_{\varepsilon,f}]}{\nu_Y[Y]} \right| \, d\mathbb{S}_\Delta.$$

Demostración.– Primero, observamos que para todo $y \in Y$, se tiene:

$$\widetilde{A_\varepsilon}(y) = \frac{1}{\nu_\Delta[\mathbb{S}_\Delta]} \int_{f\in\mathbb{S}_\Delta} \chi_\varepsilon(\mu_{\text{norm}}(f, G_{(d)}(y), e_0)) d\mathbb{S}_\Delta =$$

$$\frac{1}{\nu_\Delta[\mathbb{S}_\Delta]} \int_{f\in\mathbb{S}_\Delta} \chi_{R_\varepsilon}(G_{(d)}(y), f) d\mathbb{S}_\Delta.$$

Por lo tanto, $\mathscr{D}_{\{y_1,\ldots,y_m\}}$ es igual a

$$\frac{1}{\nu_\Delta[\mathbb{S}_\Delta]} \left| \frac{1}{m}\sum_{i=1}^{m} \int_{f\in\mathbb{S}_\Delta} \chi_{R_\varepsilon}(y_i,f) d\mathbb{S}_\Delta - \int_{y\in Y}\int_{f\in\mathbb{S}_\Delta} \chi_{R_\varepsilon}(y,f) d\mathbb{S}_\Delta \, dY \right|.$$

Por el Teorema de Fubini, esta última cantidad es igual a

$$\frac{1}{\nu_\Delta[\mathbb{S}_\Delta]} \left| \int_{f\in\mathbb{S}_\Delta} \left(\frac{1}{m}\sum_{i=1}^{m} \chi_{R_\varepsilon}(y_i,f) - \int_{y\in Y} \chi_{R_\varepsilon}(y,f)\ dY \right) d\mathbb{S}_\Delta \right|,$$

lo que concluye la prueba del lema. ∎

Lema 6.5.3 *Sea $H \geq (2N+3)^2$ un natural positivo. Sean k, s, d' los números del Lema 6.5.1. Con las notaciones anteriores, las siguientes desigualdades se satisfacen para todo sistema $f \in \mathbb{S}_\Delta$:*

$$\left| N_{\frac{1}{H}}(R_{\varepsilon,f}) - \nu_Y[R_{\varepsilon,f}]H^{2N+3}] \right| \leq d^{O(n^2N^3)}H^{2N+2},$$

$$\left| \sharp[Y^H] - \nu_Y[Y]H^{2N+3}] \right| \leq 2(2N+3)K_{2N+2}3^{N+1}H^{2N+2}.$$

Demostración.– Primero observamos que, dado que el cardinal de un conjunto discreto no cambia mediante homotecias,

$$\left| N_{\frac{1}{H}}(R_{\varepsilon,f}) - \nu_Y[R_{\varepsilon,f}]H^{2N+3}] \right| = \left| N(HR_{\varepsilon,f}) - \nu_Y[HR_{\varepsilon,f}] \right|,$$

donde $HR_{\varepsilon,f} := \{ty : Hy : y \in R_{\varepsilon,f}\}$ y $N(HR_{\varepsilon,f})$ es el número de puntos enteros en $HR_{\varepsilon,f}$. Además, dado que $R_{\varepsilon,f} \subseteq B_{2N+3}(0,\sqrt{3})$, esta última cantidad es igual a

$$\left| N(HR_{\varepsilon,f}, \sqrt{3}H) - \nu_Y[HR_{\varepsilon,f} \cap B_{2N+3}(0,\sqrt{3}H)] \right|,$$

donde $N(HR_{\varepsilon,f}, \sqrt{3}H) := \sharp[\mathbb{Z}^{2N+3} \cap HR_{\varepsilon,f} \cap B_{2N+3}(0,\sqrt{3}H)]$. Por el Corolario 6.2.2, sabemos que

$$\left| N(HR_{\varepsilon,f}, \sqrt{3}H) - \nu_Y[HR_{\varepsilon,f} \cap B_{2N+3}(0,\sqrt{3}H)] \right| \leq$$

$$\beta(HR_{\varepsilon,f})2(2N+3)K_{2N+2}3^{N+1}H^{2N+2}.$$

Además, $\beta(HR_{\varepsilon,f}) = \beta(R_{\varepsilon,f})$. Por los lemas 6.5.1 y 6.2.4, sabemos que

$$\beta(R_{\varepsilon,f}) \leq \left(d^{O(nN^3)} \right)^{6n+24} \leq d^{O(n^2N^3)}.$$

Finalmente, también podemos escribir

$$2(2N+3)K_{2N+2}3^{N+1}d^{O(n^2N^3)} \leq d^{O(n^2N^3)},$$

de donde se sigue el primer enunciado del Lema. En cuanto al segundo, procedemos de igual manera, pero esta vez utilizamos el hecho de que Y es un conjunto convexo, por lo que $\beta(Y) = 1$. ∎

El siguiente resultado se sigue de los lemas 6.5.3 y 6.3.4.

Lema 6.5.4 *Sea $H \geq (2N+3)^2$ un natural positivo, y sea $f \in \mathbb{S}_\Delta$ un sistema. Con las notaciones anteriores, tenemos:*

$$\left| \frac{N_{\frac{1}{H}}(R_{\varepsilon,f})}{\sharp[Y^H]} - \frac{\nu_Y[R_{\varepsilon,f}]}{\nu_Y[Y]} \right| \leq \frac{1}{H} \frac{d^{O(n^2N^3)} + 12^{2N+4}}{\nu_Y[Y]}.$$

En particular,

$$\left| \frac{N_{\frac{1}{H}}(R_{\varepsilon,f})}{\sharp[Y^H]} - \frac{\nu_Y[R_{\varepsilon,f}]}{\nu_Y[Y]} \right| \leq \frac{d^{O(n^2N^3)}}{H}.$$

6.5.2. Demostración del Teorema 6.1.5.

El Teorema 6.1.5 es consecuencia del Corolario 6.5.5 que exponemos a continuación.

Corolario 6.5.5 *Sea $H \geq (2N+3)^2$ un natural positivo. Tenemos la siguiente desigualdad:*

$$\left| \frac{1}{\sharp[Y^H]} \sum_{y \in Y^H} \widetilde{A_\varepsilon}(y) - E_Y[\widetilde{A_\varepsilon}] \right| \leq \frac{d^{O(n^2N^3)}}{H}. \tag{6.19}$$

En particular, sea $\delta(\varepsilon) := (\frac{10000}{3} n^{7/2}(n+1)^{3/2}N^2 d^{3/2})^{1/2}\varepsilon$. Entonces, existe una constante universal $C > 0$ tal que si

$$H \geq d^{Cn^2N^3} H_1,$$

para algún número real positivo $H_1 \geq 1$, entonces se tienen las siguientes dos desigualdades:

$$\frac{1}{\sharp[Y^H]} \sum_{y \in Y^H} \widetilde{A_\varepsilon}(y) \leq \delta(\varepsilon)^2 + \frac{1}{H_1},$$

y

$$\frac{1}{\sharp[Y^H]} \sharp \left\{ y \in Y^H : \widetilde{A_\varepsilon}(y) \geq \delta(\varepsilon) \right\} \leq \delta(\varepsilon) + \frac{1}{\delta(\varepsilon) H_1}.$$

Demostración.– La desigualdad (6.19) es consecuencia directa de los lemas 6.5.2 y 6.5.4. Por la Proposición 4.4.4, sabemos que

$$E_Y[\widetilde{A_\varepsilon}] \leq \delta(\varepsilon)^2,$$

de donde se sigue la segunda de las desigualdades. La última desigualdad es consecuencia de la desigualdad de Markov. ∎

Para terminar la demostración del Teorema 6.1.5, cambiamos cada aparición de ε en el Corolario 6.5.5 por

$$\left(\frac{10000}{3} n^{7/2}(n+1)^{3/2}N^2 d^{3/2} \right)^{-1/2} \varepsilon.$$

Obtenemos pues que existe una constante universal $C > 0$ tal que para todo $H > 0$ tal que

$$H \geq d^{Cn^2N^3} 2^{h_1},$$

para algún $h_1 > 0$, se tiene que

$$\frac{1}{\sharp(Y^H)} \sharp\{y \in Y^H : \widetilde{A}_{(\frac{10000}{3} n^{7/2}(n+1)^{3/2}N^2 d^{3/2})^{-1/2}\varepsilon}(y) \geq \varepsilon\} \leq \varepsilon + \frac{1}{\varepsilon 2^{h_1}}.$$

Tomemos $h_1 = 2\log_2 \varepsilon^{-1}$. Entonces,

$$\frac{1}{\sharp(Y^H)}\sharp\{y \in Y^H : \widetilde{A}_{(\frac{10000}{3}n^{7/2}(n+1)^{3/2}N^2d^{3/2})^{-1/2}\varepsilon}(y) \geq \varepsilon\} \leq 2\varepsilon.$$

Finalmente, por el Teorema 4.2.6, si un punto $y \in Y$ satisface

$$\widetilde{A}_{(\frac{10000}{3}n^{7/2}(n+1)^{3/2}N^2d^{3/2})^{-1/2}\varepsilon}(y) < \varepsilon,$$

entonces con probabilidad mayor o igual que $1 - \varepsilon$, el número de pasos de homotopía de NHD con par inicial $(G_{(d)}(y), e_0)$ está acotado por

$$60000n^{7/2}(n+1)^{3/2}N^2d^3\varepsilon^{-2}.$$

Esto termina la demostración del Teorema 6.1.5.

Apéndice A

La Estructura Riemanniana del Espacio Proyectivo Complejo

El espacio proyectivo complejo es una variedad Riemanniana bien conocida. Sin embargo, resulta difícil encontrar un texto en el que se resuman sus propiedades más elementales sin recurrir a un lenguaje especializado. En las páginas que siguen describimos de modo elemental las propiedades de este espacio que se utilizan a lo largo de la memoria.
Dado que estamos interesados en distintas estructuras Riemannianas, escogemos una forma de definirlas que las reúna a todas. Sea pues $\mathbb{C}^{k+1}$ un espacio vectorial complejo de dimensión k y sea $\langle , \rangle$ un producto hermitiano en $\mathbb{C}^{k+1}$. Esto es, una forma bilineal en $\mathbb{C}^{k+1}$ que es definida positiva y tal que

$$\langle v, w\rangle = \overline{\langle w, v\rangle},$$

donde $\bar{\cdot}$ denota conjugación compleja. Queda así definida una norma en $\mathbb{C}^{k+1}$ del modo usual: Para todo vector $v \in \mathbb{C}^{k+1}$, se tiene que

$$\|v\|_2 := \langle v, v\rangle^{1/2}.$$

Sea $S(\mathbb{C}^{k+1})$ la esfera en $\mathbb{C}^{k+1}$ para la norma que acabamos de definir. Esto es,

$$S(\mathbb{C}^{k+1}) := \{\underline{x} \in \mathbb{C}^{k+1} : \|x\|_2 = 1\}.$$

La esfera es una variedad diferenciable de dimensión impar $2k+1$, y por tanto no es una variedad compleja. Sin embargo, posee una estructura de variedad Riemanniana real heredada de $\mathbb{C}^{k+1}$. Pasemos ahora a la estructura Riemanniana de $\mathbb{P}_k(\mathbb{C})$. Recordemos que el espacio proyectivo $\mathbb{P}_k(\mathbb{C})$ se define como el conjunto de clases de equivalencia en $\mathbb{C}^{k+1} \setminus \{0\}$ módulo la relación de equivalencia

$$\underline{x} \sim \underline{y} \Longleftrightarrow \exists \lambda \in \mathbb{C} : \underline{x} = \lambda \underline{y}.$$

Denotamos por π la aplicación canónica,

$$\begin{array}{rccc} \pi: & \mathbb{C}^{k+1}\setminus\{0\} & \longrightarrow & \mathbb{P}_k(\mathbb{C}) \\ & \underline{x} & \mapsto & \pi(\underline{x}). \end{array}$$

También consideramos la restricción a $S(\mathbb{C}^{k+1})$, que en el caso de que $\langle\cdot,\cdot\rangle$ es el producto hermitiano usual se denomina fibración de Hopf,

$$\begin{array}{rccc} p:=\pi\mid_{S(\mathbb{C}^{k+1})}: & S(\mathbb{C}^{k+1}) & \longrightarrow & \mathbb{P}_k(\mathbb{C}) \\ & \underline{x} & \mapsto & p(\underline{x}). \end{array}$$

Entonces, como se observa en [48, Prop. 2.28], [48, ex. 2.29], existe una única estructura Riemanniana en $\mathbb{P}_k(\mathbb{C})$ tal que p es una submersión Riemanniana, esto es:

- p es sobreyectiva,
- La diferencial $d_{\underline{x}}p$ es sobreyectiva en todo punto $\underline{x}\in S(\mathbb{C}^{k+1})$, y
- La restricción de $d_{\underline{x}}p$ al ortogonal de $Ker(d_{\underline{x}}p)$ es una isometría lineal.

En la tercera de estas propiedades, por ortogonal de $Ker(d_{\underline{x}}p)$ queremos decir el complemento ortogonal (para $\langle\cdot,\cdot\rangle_{\mathbb{R}}:=\mathbb{R}e(\langle\cdot,\cdot\rangle)$ del núcleo de $d_{\underline{x}}p$. Esto es, el espacio horizontal de p en x.
Esta estructura Riemanniana cuya existencia queda garantizada por la Proposición 2.28 de [48] es la que denominamos "estructura Riemanniana inducida por $\langle\cdot,\cdot\rangle$". La siguiente proposición permite identificar el espacio tangente $T_x\mathbb{P}_k(\mathbb{C})$ con el ortogonal a x para $\langle\cdot,\cdot\rangle$.

Proposición A.0.6 *Sea $x\in\mathbb{P}_k(\mathbb{C})$ un punto proyectivo. Sea $\underline{x}\in p^{-1}(x)$ un representante de x tal que $\|\underline{x}\|_2=1$. Consideremos el espacio*

$$x^{\perp}:=\{v\in\mathbb{C}^{k+1}:\langle x,v\rangle=0\},$$

y sea $\varphi_{\underline{x}}$ la siguiente aplicación:

$$\begin{array}{rccc} \varphi_{\underline{x}}: & x^{\perp} & \longrightarrow & \mathbb{P}_k(\mathbb{C}) \\ & \underline{y} & \mapsto & \pi(\underline{x}+\underline{y}). \end{array}$$

Entonces, $\varphi_{\underline{x}}$ es una isometría en 0. En otras palabras, la aplicación diferencial $d_0\varphi_{\underline{x}}$ es una isometría lineal.

Demostración.– Tenemos el siguiente diagrama

$$\begin{array}{ccc} & & S^{2k+1} \\ & & \Big\downarrow{\scriptstyle p} \\ x^{\perp} & \xrightarrow{\varphi_{\underline{x}}} & \mathbb{P}_k(\mathbb{C}) \end{array}$$

Sean $v, w \in x^{\perp}$ dos vectores, y veamos que

$$\langle d_0\varphi_{\underline{x}}(v), d_0\varphi_{\underline{x}}(w)\rangle_{T_x\mathbb{P}_k(\mathbb{C})} = \langle v, w\rangle$$

Denotemos por $D := (d_{\underline{x}}p) \mid_{Ker(d_{\underline{x}}p)^{\perp}}$ la restricción de $d_{\underline{x}}p$ al ortogonal de $Ker(d_{\underline{x}}p)$, que según hemos visto es una isometría lineal. Por lo tanto,

$$\langle d_0\varphi_{\underline{x}}(v), d_0\varphi_{\underline{x}}(w)\rangle_{T_x\mathbb{P}_k(\mathbb{C})} = \langle D^{-1}d_0\varphi_{\underline{x}}(v), D^{-1}d_0\varphi_{\underline{x}}(w)\rangle,$$

y basta con demostrar que

$$D^{-1}d_0\varphi_{\underline{x}}(v) = v, \quad D^{-1}d_0\varphi_{\underline{x}}(w) = w.$$

De hecho, ambas afirmaciones son equivalentes, pues v y w son vectores cualesquiera. Además, D^{-1} y $d_0\varphi_{\underline{x}}$ son biyectivas y v pertenece al dominio de D, luego basta comprobar que

$$v = (d_0\varphi_{\underline{x}})^{-1}D(v) = d_x(\varphi_{\underline{x}})^{-1}d_{\underline{x}}p(v) = d_{\underline{x}}(\varphi_{\underline{x}}^{-1} \circ p)(v).$$

Ahora, $\varphi_{\underline{x}}^{-1} \circ p$ es una aplicación entre la esfera y un subespacio lineal de $\mathbb{C}^{k+1}$, luego podemos utilizar análisis clásico para calcular su diferencial:

$$d_{\underline{x}}(\varphi_{\underline{x}}^{-1} \circ p)(v) = \lim_{t\mapsto 0} \frac{\varphi_{\underline{x}}^{-1} \circ p\left(\frac{\underline{x}+tv}{\|\underline{x}+tv\|_2}\right) - \varphi_{\underline{x}}^{-1} \circ p(\underline{x})}{t} =$$

$$\lim_{t\mapsto 0} \frac{\varphi_{\underline{x}}^{-1}(\pi(\underline{x}+tv)) - \varphi_{\underline{x}}^{-1}(\pi(\underline{x}))}{t} = \lim_{t\mapsto 0} \frac{tv-0}{t} = v,$$

lo que termina la demostración. ∎

El siguiente sencillo resultado es útil en numerosas ocasiones, pues permite hacer uso de las propiedades simétricas del espacio proyectivo. En particular, tiene como consecuencia el Lema 1.4.1, que ha sido utilizado en numerosas ocasiones a lo largo de esta memoria, y también demuestra el hecho de que las matrices unitarias actúan isométricamente sobre el espacio proyectivo complejo.

Lema A.0.7 *Sean $\langle\cdot,\cdot\rangle_1$ y $\langle\cdot,\cdot\rangle_2$ dos productos hermitianos en $\mathbb{C}^{k+1}$. Sea*

$$\psi : (\mathbb{C}^{k+1}, \langle\cdot,\cdot\rangle_1) \longrightarrow (\mathbb{C}^{k+1}, \langle\cdot,\cdot\rangle_2)$$

una isometría lineal, esto es,

$$\langle v, w\rangle_1 = \langle\psi(v), \psi(w)\rangle_2, \quad \forall v, w \in \mathbb{C}^{k+1}.$$

Sea $\mathbb{P}_1$ el espacio $\mathbb{P}_k(\mathbb{C})$ con la estructura hermitiana heredada de $\langle\cdot,\cdot\rangle_1$ y sea $\mathbb{P}_2$ el espacio $\mathbb{P}_k(\mathbb{C})$ con la estructura hermitiana heredada de $\langle\cdot,\cdot\rangle_2$. Entonces, la aplicación (que denotamos por el mismo símbolo)

$$\begin{array}{rccc} \psi: & \mathbb{P}_1 & \longrightarrow & \mathbb{P}_2 \\ & x & \mapsto & \pi(\psi(\pi^{-1}(x))) \end{array}$$

es también una isometría.

Demostración.– Primero, comprobamos trivialmente que ψ está bien definida como aplicación proyectiva, pues $\pi(\psi(\pi^{-1}(x)))$ no depende del representante $\pi^{-1}(x)$ elegido. Sea S_1 la esfera en $\mathbb{C}^{k+1}$ para el producto hermitiano $\langle\cdot,\cdot\rangle_1$ y sea S_2 la esfera para el producto hermitiano $\langle\cdot,\cdot\rangle_2$. Entonces, se tiene claramente que ψ transforma S_1 en S_2. Además, como ψ es una isometría entre los espacios afines, también lo es entre subvariedades diferenciables de estos espacios, luego

$$\psi : S_1 \longrightarrow S_2$$

es también una isometría. Por tanto, la situación es

$$\begin{array}{ccc} S_1 & \xrightarrow{\psi} & S_2 \\ \Big\downarrow{\scriptstyle p_1} & & \Big\downarrow{\scriptstyle p_2} \\ \mathbb{P}_1 & \xrightarrow{\psi} & \mathbb{P}_2 \end{array}$$

donde estamos denotando por p_1 y p_2 las respectivas restricciones de π a S_1 y a S_2. Dado que p_i, $i = 1, 2$ es una submersión Riemanniana (pues así hemos definido las respectivas estructuras de $\mathbb{P}_1$ y $\mathbb{P}_2$), basta con comprobar que para todo punto $\underline{x} \in S_1$, se tiene

$$d_{\underline{x}}\psi(Ker(d_{\underline{x}}p_1)) = Ker(d_{\psi(\underline{x})}p_2).$$

Por otro lado, sabemos que

$$Ker(d_{\underline{x}}p_1) =< \sqrt{-1}\underline{x} >, \qquad Ker(d_{\psi(\underline{x})}p_2) =< \sqrt{-1}\psi(\underline{x}) >,$$

donde $< \cdot >$ denota el espacio vectorial real generado por el vector correspondiente. Además, considerando ψ como definida entre los espacios afines, tenemos que

$$d_{\underline{x}}\psi(Ker(d_{\underline{x}}p_1)) = \sqrt{-1}d_{\underline{x}}\psi(\underline{x}) = \sqrt{-1}\psi(\underline{x}) = Ker(d_{\psi(\underline{x})}p_2),$$

Queda demostrado pues que

$$d_{\underline{x}}\psi(Ker(d_{\underline{x}}p_1)) = Ker(d_{\psi(\underline{x})}p_2),$$

y con ello el lema. ∎

Finalmente, como en toda variedad Riemanniana, en el espacio $\mathbb{P}_k(\mathbb{C})$ con la estructura heredada de $\langle\cdot,\cdot\rangle$ podemos medir volúmenes, calcular integrales y medir longitudes de curvas. Como consecuencia, hay una noción de distancia Riemanniana, definida como el ínfimo de las longitudes de curvas uniendo dos puntos. Denotamos esa distancia por d_R. Cuando $\langle\cdot,\cdot\rangle$ es el producto usual, la distancia Riemanniana obtenida recibe el nombre de distancia de Fubini–Study. Nosotros utilizamos con frecuencia a lo largo de la memoria

otra distancia prácticamente equivalente a ésta en valores pequeños: La distancia proyectiva, definida como el seno de la distancia Riemanniana. Esto es, para dos puntos cualesquiera $x, y \in \mathbb{P}_k(\mathbb{C})$, definimos

$$d_{\mathbf{P}} := \operatorname{sin} d_R(x, y).$$

Obsérvese que, aunque no aparezca en la notación, las distancias d_R y $d_{\mathbf{P}}$ dependen del producto hermitiano $\langle \cdot, \cdot \rangle$ que hallamos escogido. Recordamos una elegante fórmula que relaciona más directamente este producto hermitiano con la distancia proyectiva. Su demostración puede encontrarse, por ejemplo, en [14]:

$$d_{\mathbf{P}}(x, y) = \min_{\pi(\underline{x})=x, \pi(\underline{y})=y} \sqrt{1 - \frac{\langle \underline{x}, \underline{y} \rangle_{\mathbb{R}}^2}{\|\underline{x}\|_2^2 \|\underline{y}\|_2^2}} =$$

$$\sqrt{1 - \frac{|\langle x, y \rangle|^2}{\|x\|_2^2 \|y\|_2^2}} = \min_{\lambda \in \mathbb{C}} \frac{\|x - \lambda y\|_2}{\|x\|_2}.$$

Bibliografía

[1] V. Arnold, A. Varchenko, and S. Goussein-Zadé. *Singularités des aplications différentiables*. Éditions Mir, Moscou, 1986.

[2] J.M. Azaïs and M. Wschebor. On the roots of a random system of equations. The theorem on Shub and Smale and some extensions. *Found. Comput. Math.*, 5(2):125–144, 2005.

[3] B. Bank, M. Giusti, J. Heintz, and G. M. Mbakop. Polar varieties, real equation solving, and data structures: the hypersurface case. *J. Complexity*, 13(1):5–27, 1997.

[4] B. Bank, M. Giusti, J. Heintz, and G. M. Mbakop. Polar varieties and efficient real elimination. *Math. Z.*, 238(1):115–144, 2001.

[5] B. Bank, M. Giusti, J. Heintz, and L. M. Pardo. Generalized polar varieties and an efficient real elimination procedure. *Kybernetika (Prague)*, 40(5):519–550, 2004.

[6] B. Bank, M. Giusti, J. Heintz, and L. M. Pardo. Generalized polar varieties: geometry and algorithms. *J. Complexity*, 21(4):377–412, 2005.

[7] S. Basu, R. Pollack, and M.F. Roy. Complexity of computing semialgebraic descriptions of the connected components of a semi-algebraic set. In *Proceedings of the 1998 International Symposium on Symbolic and Algebraic Computation (Rostock)*, pages 25–29 (electronic), New York, 1998. ACM.

[8] T. Becker and V. Weispfenning. *Gröbner bases*, volume 141 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1993. A computational approach to commutative algebra, In cooperation with Heinz Kredel.

[9] C. Beltrán and L.M. Pardo. Upper bounds on the distribution of the condition number of singular matrices. *C. R. Math. Acad. Sci. Paris*, 340(12):915–919, 2005.

[10] C. Beltrán and L.M. Pardo. Estimates on the distribution of the condition number of singular matrices. *Found. Comput. Math.*, To appear, 2006.

[11] C. Beltrán and L.M. Pardo. On Smale's 17th problem: A uniform algorithm in probabilistic polynomial time. *Found. Comput. Math.*, Submitted, 2006.

[12] C. Beltrán and L.M. Pardo. On the complexity of non–universal polynomial equation solving: old and new results. In *Foundations of Computational Mathematics: Santander 2005. L. Pardo, A. Pinkus, E. Süli, M. Todd editors.*, pages 1–35. Cambridge University Press, 2006.

[13] C. Beltrán and L.M. Pardo. On the probability distribution of singular varieties of given corank. *J. Symbolic Comput.*, To appear, 2006.

[14] L. Blum, F. Cucker, M. Shub, and S. Smale. *Complexity and real computation*. Springer-Verlag, New York, 1998.

[15] L. Blum, M. Shub, and S. Smale. On a theory of computation and complexity over the real numbers: NP-completeness, recursive functions and universal machines. *Bull. Amer. Math. Soc. (N.S.)*, 21(1):1–46, 1989.

[16] W. D. Brownawell. Bounds for the degrees in the Nullstellensatz. *Ann. of Math. (2)*, 126(3):577–591, 1987.

[17] W. Bruns and U. Vetter. *Ein Algoritmus zum Auffinden der Basiselemente des Restklassenringes nach einen nulldimensionallen Polynomideal*. Ph.D. Thesis. Leopold–Franzens Universität, Innsbruck, 1965.

[18] W. Bruns and U. Vetter. *Determinantal rings*, volume 1327 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1988.

[19] L. Caniglia, A. Galligo, and J. Heintz. Borne simple exponentielle pour les degrés dans le théorème des zéros sur un corps de caractéristique quelconque. *C. R. Acad. Sci. Paris Sér. I Math.*, 307(6):255–258, 1988.

[20] D. Castro, M. Giusti, J. Heintz, G. Matera, and L. M. Pardo. The hardness of polynomial equation solving. *Found. Comput. Math.*, 3(4):347–420, 2003.

[21] D. Castro, J. L. Montaña, L. M. Pardo, and J. San Martín. The distribution of condition numbers of rational data of bounded bit length. *Found. Comput. Math.*, 2(1):1–52, 2002.

[22] D. Castro, L. M. Pardo, K. Hägele, and J. E. Morais. Kronecker's and Newton's approaches to solving: a first comparison. *J. Complexity*, 17(1):212–303, 2001.

[23] D. Castro, L. M. Pardo, and J. San Martín. Systems of rational polynomial equations have polynomial size approximate zeros on the average. *J. Complexity*, 19(2):161–209, 2003.

[24] K.K. Choi. On the distribution of points in projective space of bounded height. *Trans. Amer. Math. Soc.*, 352(3):1071–1111, 2000.

[25] D. Cox, J. Little, and D. O'Shea. *Ideals, varieties, and algorithms*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, second edition, 1997.

[26] J. A. Cuesta-Albertos and M. Wschebor. Some remarks on the condition number of a real random square matrix. *J. Complexity*, 19(4):548–554, 2003.

[27] J. A. Cuesta-Albertos and M. Wschebor. Condition numbers and extrema of random fields. In *Seminar on Stochastic Analysis, Random Fields and Applications IV*, volume 58 of *Progr. Probab.*, pages 69–82. Birkhäuser, Basel, 2004.

[28] H. Davenport. On a principle of Lipschitz. *J. London Math. Soc.*, 26:179–183, 1951.

[29] J. P. Dedieu, X. Gourdon, and J. C. Yakoubsohn. Computing the distance from a point to an algebraic hypersurface. In *The mathematics of numerical analysis (Park City, UT, 1995)*, volume 32 of *Lectures in Appl. Math.*, pages 285–293. Amer. Math. Soc., Providence, RI, 1996.

[30] J.P. Dedieu. Approximate solutions of numerical problems, condition number analysis and condition number theorem. In *The mathematics of numerical analysis (Park City, UT, 1995)*, volume 32 of *Lectures in Appl. Math.*, pages 263–283. Amer. Math. Soc., Providence, RI, 1996.

[31] J.P. Dedieu. Estimations for the separation number of a polynomial system. *J. Symbolic Comput.*, 24(6):683–693, 1997.

[32] J.P. Dedieu. Approximate solutions of analytic inequality systems. *SIAM J. Optim.*, 11(2):411–425 (electronic), 2000.

[33] J.P. Dedieu. Newton's method and some complexity aspects of the zero-finding problem. In *Foundations of computational mathematics (Oxford, 1999)*, volume 284 of *London Math. Soc. Lecture Note Ser.*, pages 45–67. Cambridge Univ. Press, Cambridge, 2001.

[34] J.P. Dedieu. *Points Fixes, Zéros et la Méthode de Newton.* Collection Mathématiques et Applications. Springer, to appear 2006.

[35] J.P. Dedieu and M. Shub. Multihomogeneous Newton methods. *Math. Comp.*, 69(231):1071–1098 (electronic), 2000.

[36] J.P. Dedieu and M. Shub. On simple double zeros and badly conditioned zeros of analytic functions of n variables. *Math. Comp.*, 70(233):319–327, 2001.

[37] J. Dégot. A condition number theorem for underdetermined polynomial systems. *Math. Comp.*, 70(233):329–335, 2001.

[38] M. Demazure. *Catastrophes et bifurcations.* Paris:Ellipses, Paris, 1989.

[39] J. W. Demmel. The geometry of ill-conditioning. *J. Complexity*, 3(2):201–229, 1987.

[40] J. W. Demmel. The probability that a numerical analysis problem is difficult. *Math. Comp.*, 50(182):449–480, 1988.

[41] A. Edelman. Eigenvalues and condition numbers of random matrices. *SIAM J. Matrix Anal. Appl.*, 9(4):543–560, 1988.

[42] A. Edelman. On the distribution of a scaled condition number. *Math. Comp.*, 58(197):185–190, 1992.

[43] N. Elezović, C. Giordano, and J. Pečarić. The best bounds in Gautschi's inequality. *Math. Inequal. Appl.*, 3(2):239–252, 2000.

[44] H. Federer. Some theorems on integral currents. *Trans. Amer. Math. Soc.*, 117:43–67, 1965.

[45] H. Federer. *Geometric measure theory.* Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York Inc., New York, 1969.

[46] F. J. Flaherty. The volume of a tube in complex projective space. *Illinois J. Math.*, 16:627–638, 1972.

[47] W. Fulton. *Intersection theory*, volume 2 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3).* Springer-Verlag, Berlin, 1984.

[48] S. Gallot, D. Hulin, and J. Lafontaine. *Riemannian geometry.* Universitext. Springer-Verlag, Berlin, 1987.

[49] M.R. Garey and D.S. Johnson. *Computers and Intractability: A Guide to the Theory of NP–Completness.* W.H. Freeman, San Francisco, California., 1979.

[50] M. Giusti, J. Heintz, K. Hägele, J. L. Montaña, J. E. Morais, and L. M. Pardo. Lower bounds for Diophantine approximations. *J. Pure Appl. Algebra*, 117/118:277–317, 1997.

[51] M. Giusti, J. Heintz, J. E. Morais, J. Morgenstern, and L. M. Pardo. Straight-line programs in geometric elimination theory. *J. Pure Appl. Algebra*, 124(1-3):101–146, 1998.

[52] M. Giusti, J. Heintz, J. E. Morais, and L. M. Pardo. When polynomial equation systems can be "solved" fast? In *Applied algebra, algebraic algorithms and error-correcting codes (Paris, 1995)*, volume 948 of *Lecture Notes in Comput. Sci.*, pages 205–231. Springer, Berlin, 1995.

[53] M. Giusti, J. Heintz, J.E. Morais, and L.M. Pardo. Le rôle des structures de données dans les problèmes d'élimination. *C. R. Acad. Sci. Paris Sér. I Math.*, 325(11):1223–1228, 1997.

[54] M. Giusti, G. Lecerf, and B. Salvy. A Gröbner free alternative for polynomial system solving. *J. Complexity*, 17(1):154–211, 2001.

[55] M. Giusti, G. Lecerf, B. Salvy, and J.C. Yakoubsohn. On location and approximation of clusters of zeros of analytic functions. *Found. Comput. Math.*, 5(3):257–311, 2005.

[56] M. Giusti, G. Lecerf, B. Salvy, and J.P. Yakoubsohn. On location and approximation of clusters of zeros: case of embedding dimension one. *Found. Comp. Mathematics*, to appear, 2005.

[57] M. Giusti and E. Schost. Solving some overdetermined polynomial systems. In *Proceedings of the 1999 International Symposium on Symbolic and Algebraic Computation (Vancouver, BC)*, pages 1–8 (electronic), New York, 1999. ACM.

[58] Gene H. Golub and Charles F. Van Loan. *Matrix computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, third edition, 1996.

[59] M. Golubitsky and V. Guillemin. *Stable mappings and their singularities*. Springer-Verlag, New York, 1973. Graduate Texts in Mathematics, Vol. 14.

[60] A. Gray. Volumes of tubes about complex submanifolds of complex projective space. *Trans. Amer. Math. Soc.*, 291(2):437–449, 1985.

[61] A. Gray. *Tubes*, volume 221 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, second edition, 2004.

[62] K. Hägele, J. E. Morais, L. M. Pardo, and M. Sombra. On the intrinsic complexity of the arithmetic Nullstellensatz. *J. Pure Appl. Algebra*, 146(2):103–183, 2000.

[63] G. H. Hardy and E. M. Wright. *An introduction to the theory of numbers*. The Clarendon Press Oxford University Press, New York, fourth edition, 1979.

[64] J. Harris. *Algebraic geometry*, volume 133 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1992. A first course.

[65] R. Hartshorne. *Algebraic geometry*. Springer-Verlag, New York, 1977. Graduate Texts in Mathematics, No. 52.

[66] J. Heintz. Definability and fast quantifier elimination in algebraically closed fields. *Theoret. Comput. Sci.*, 24(3):239–277, 1983.

[67] J. Heintz, G. Matera, L. M. Pardo, and R. Wachenchauzer. The intrinsic complexity of parametric elimination methods. *Electron. J. SADIO*, 1(1):37–51 (electronic), 1998.

[68] J. Heintz, G. Matera, and A. Waissbein. On the time-space complexity of geometric elimination procedures. *Appl. Algebra Engrg. Comm. Comput.*, 11(4):239–296, 2001.

[69] J. Heintz and J. Morgenstern. On the intrinsic complexity of elimination theory. *J. Complexity*, 9(4):471–498, 1993.

[70] G. Hermann. Die Frage der endlich vielen Schritte in der Theorie der Polynomideale. *Math. Ann.*, 95(1):736–788, 1926.

[71] N.J. Higham. *Accuracy and stability of numerical algorithms*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 2002.

[72] D. Hilbert. Über theorie der Algebraischen Formen. *Math. Ann.*, 36:473–534, 1890.

[73] H. Hironaka. Resolution of singularities of an algebraic variety over a field of characteristic zero. I, II. *Ann. of Math. (2) 79 (1964), 109–203; ibid. (2)*, 79:205–326, 1964.

[74] M.W. Hirsch. *Differential topology*. Springer-Verlag, New York, 1976. Graduate Texts in Mathematics, No. 33.

[75] R. Howard. The kinematic formula in Riemannian homogeneous spaces. *Mem. Amer. Math. Soc.*, 106(509):vi+69, 1993.

[76] W. Kahan. Huge generalized inverses of rank-deficient matrices. Unpublished Manuscript, 2000.

[77] L. Kantorovich. Sur la méthode de Newton. XXVIII:104–144, 1949.

[78] M.H. Kim. Topological complexity of a root finding algorithm. *J. Complexity*, 5(3):331–344, 1989.

[79] P. Koiran. Approximating the volume of definable sets. In *36th Annual Symposium on Foundations of Computer Science (Milwaukee, WI, 1995)*, pages 134–141. IEEE Comput. Soc. Press, Los Alamitos, CA, 1995.

[80] P. Koiran. Hilbert's Nullstellensatz is in the polynomial hierarchy. *J. Complexity*, 12(4):273–286, 1996. Special issue for the Foundations of Computational Mathematics Conference (Rio de Janeiro, 1997).

[81] J. Kollár. Sharp effective Nullstellensatz. *J. Amer. Math. Soc.*, 1(4):963–975, 1988.

[82] T. Krick and L. M. Pardo. A computational method for Diophantine approximation. In *Algorithms in algebraic geometry and applications (Santander, 1994)*, volume 143 of *Progr. Math.*, pages 193–253. Birkhäuser, Basel, 1996.

[83] T. Krick, L.M. Pardo, and M. Sombra. Sharp estimates for the arithmetic Nullstellensatz. *Duke Math. J.*, 109(3):521–598, 2001.

[84] L. Kronecker. Grundzüge einer arithmetischen theorie de algebraischen grössen. *J. reine angew. Math.*, 92:1–122, 1882.

[85] E. Kunz. *Introduction to commutative algebra and algebraic geometry.* Birkhäuser Boston Inc., Boston, MA, 1985.

[86] P. Lelong. Propriétés métriques des variétés analytiques complexes définies par une équation. *Ann. Sci. École Norm. Sup. (3)*, 67:393–419, 1950.

[87] G. Malajovich. *On the complexity of path-following Newton algorithms for solving systems of polynomial equations with integer coefficients.* 1993. PhD Thesis. Univ. Rio de Janeiro.

[88] G. Malajovich. On generalized Newton algorithms: quadratic convergence, path-following and error analysis. *Theoret. Comput. Sci.*, 133(1):65–84, 1994. Selected papers of the Workshop on Continuous Algorithms and Complexity (Barcelona, 1993).

[89] G. Malajovich and J.M. Rojas. Polynomial systems and the momentum map. In *Foundations of computational mathematics (Hong Kong, 2000)*, pages 251–266. World Sci. Publishing, River Edge, NJ, 2002.

[90] G. Malajovich and M. Rojas. Random sparse polynomial systems. 2000. Selected papers of the Workshop on Continuous Algorithms and Complexity (Barcelona, 1993).

[91] Gregorio Malajovich and J. Maurice Rojas. High probability analysis of the condition number of sparse polynomial systems. *Theoret. Comput. Sci.*, 315(2-3):524–555, 2004.

[92] E. W. Mayr and A. R. Meyer. The complexity of the word problems for commutative semigroups and polynomial ideals. *Adv. in Math.*, 46(3):305–329, 1982.

[93] J. L. Montaña, J. E. Morais, and L. M. Pardo. Lower bounds for arithmetic networks. II. Sum of Betti numbers. *Appl. Algebra Engrg. Comm. Comput.*, 7(1):41–51, 1996.

[94] T. Mora. *Solving polynomial equation systems. I*, volume 88 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2003.

[95] F. Morgan. *Geometric measure theory: A beginner's guide*. Academic Press Inc., Boston, MA, 1988.

[96] D. Mumford. *Algebraic geometry. I.* Springer-Verlag, Berlin, 1976. Complex projective varieties, Grundlehren der Mathematischen Wissenschaften, No. 221.

[97] J.v. Neumann and H. H. Goldstine. Numerical inverting of matrices of high order. *Bull. Amer. Math. Soc.*, 53:1021–1099, 1947.

[98] I. Newton. *Analysis per Quantitatum Series, Fluxiones, ac Diferentias: cum Enumeratione Linearum Tertii Ordinis*. Ex Officina Pearsoniana, Londres, 1711. Capítulo De Analysi per Aecuationes Numero Terminorum Infinitas.

[99] L. M. Pardo and J. San Martin. Deformation techniques to solve generalised Pham systems. *Theoret. Comput. Sci.*, 315(2-3):593–625, 2004.

[100] L.M. Pardo. How lower and upper complexity bounds meet in elimination theory. In *Applied algebra, algebraic algorithms and error-correcting codes (Paris, 1995)*, volume 948 of *Lecture Notes in Comput. Sci.*, pages 33–69. Springer, Berlin, 1995.

[101] L.M. Pardo. Geometric elimination requires exponential running time, 2000. Plenary talk. Effective Methods in Algebraic Geometry, MEGA'2000.

[102] J. Renegar. On the cost of approximating all roots of a complex polynomial. *Math. Programming*, 32(3):319–336, 1985.

[103] J. Renegar. On the efficiency of Newton's method in approximating all zeros of a system of complex polynomials. *Math. Oper. Res.*, 12(1):121–148, 1987.

[104] J. Renegar. Is it possible to know a problem instance is ill-posed? Some foundations for a general theory of condition numbers. *J. Complexity*, 10(1):1–56, 1994.

[105] J. M. Rojas. Dedekind Zeta functions and the complexity of Hilbert's Nullstellensatz. *Math ArXiv preprint math.NT/0301111*. Paper corresponding to semi–plenary talk at FoCM'02, Minneapolis.

[106] J. M. Rojas. Computational arithmetic geometry I. Sentences nearly in the polynomial hierarchy. *J. Comput. System Sci.*, 62(2):216–235, 2001. Special issue on the Fourteenth Annual IEEE Conference on Computational Complexity (Atlanta, GA, 1999).

[107] W. Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, second edition, 1974. McGraw-Hill Series in Higher Mathematics.

[108] L.A. Santaló. *Integral geometry and geometric probability*. Addison-Wesley Publishing Co., Reading, Mass.-London-Amsterdam, 1976. Encyclopedia of Mathematics and its Applications, Vol. 1.

[109] I. R. Shafarevich. *Basic algebraic geometry.* Springer-Verlag, Berlin, second edition, 1994.

[110] M. Shub. Some remarks on Bezout's theorem and complexity theory. In *From Topology to Computation: Proceedings of the Smalefest (Berkeley, CA, 1990)*, pages 443–455, New York, 1993. Springer.

[111] M. Shub and S. Smale. Computational complexity: on the geometry of polynomials and a theory of cost. II. *SIAM J. Comput.*, 15(1):145–161, 1986.

[112] M. Shub and S. Smale. Complexity of Bézout's theorem. I. Geometric aspects. *J. Amer. Math. Soc.*, 6(2):459–501, 1993.

[113] M. Shub and S. Smale. Complexity of Bezout's theorem. II. Volumes and probabilities. In *Computational algebraic geometry (Nice,*

1992), volume 109 of *Progr. Math.*, pages 267–285. Birkhäuser Boston, Boston, MA, 1993.

[114] M. Shub and S. Smale. Complexity of Bezout's theorem. III. Condition number and packing. *J. Complexity*, 9(1):4–14, 1993. Festschrift for Joseph F. Traub, Part I.

[115] M. Shub and S. Smale. Complexity of Bezout's theorem. V. Polynomial time. *Theoret. Comput. Sci.*, 133(1):141–164, 1994. Selected papers of the Workshop on Continuous Algorithms and Complexity (Barcelona, 1993).

[116] M. Shub and S. Smale. Complexity of Bezout's theorem. IV. Probability of success; extensions. *SIAM J. Numer. Anal.*, 33(1):128–148, 1996.

[117] S. Smale. On the efficiency of algorithms of analysis. *Bull. Amer. Math. Soc. (N.S.)*, 13(2):87–121, 1985.

[118] S. Smale. Newton's method estimates from data at one point. In *The merging of disciplines: new directions in pure, applied, and computational mathematics (Laramie, Wyo., 1985)*, pages 185–196. Springer, New York, 1986.

[119] S. Smale. Mathematical problems for the next century. In *Mathematics: frontiers and perspectives*, pages 271–294. Amer. Math. Soc., Providence, RI, 2000.

[120] P. Stănică. Good lower and upper bounds on binomial coefficients. *JIPAM. J. Inequal. Pure Appl. Math.*, 2(3):Article 30, 5 pp. (electronic), 2001.

[121] G. W. Stewart. On the early history of the singular value decomposition. *SIAM Rev.*, 35(4):551–566, 1993.

[122] G. W. Stewart and J. G. Sun. *Matrix perturbation theory*. Computer Science and Scientific Computing. Academic Press Inc., Boston, MA, 1990.

[123] G. Stolzenberg. *Volumes, limits, and extensions of analytic varieties*. Lecture Notes in Mathematics, No. 19. Springer-Verlag, Berlin, 1966.

[124] E. Study. Kurzeste wege in complexen gebiete. *Math. annalen*, 60:321–377, 1905.

[125] B. Sturmfels. *Gröbner bases and convex polytopes*, volume 8 of *University Lecture Series*. American Mathematical Society, Providence, RI, 1996.

[126] L.N. Trefethen and D. Bau, III. *Numerical linear algebra*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.

[127] A. M. Turing. Rounding-off errors in matrix processes. *Quart. J. Mech. Appl. Math.*, 1:287–308, 1948.

[128] W. V. Vasconcelos. *Computational methods in commutative algebra and algebraic geometry*, volume 2 of *Algorithms and Computation in Mathematics*. Springer-Verlag, Berlin, 1998. With chapters by David Eisenbud, Daniel R. Grayson, Jürgen Herzog and Michael Stillman.

[129] J. Verschelde. Toric Newton method for polynomial homotopies. *J. Symbolic Comput.*, 29(4-5):777–793, 2000. Symbolic computation in algebra, analysis, and geometry (Berkeley, CA, 1998).

[130] W. Vogel. *Lectures on results on Bezout's theorem*, volume 74 of *Tata Institute of Fundamental Research Lectures on Mathematics and Physics*. Published for the Tata Institute of Fundamental Research, Bombay, 1984.

[131] A. Weinstein. Almost invariant submanifolds for compact group actions. *J. Eur. Math. Soc. (JEMS)*, 2(1):53–86, 2000.

[132] J. H. Wilkinson. *The algebraic eigenvalue problem*. Clarendon Press, Oxford, 1965.

[133] R. A. Wolf. The volume of tubes in complex projective space. *Trans. Amer. Math. Soc.*, 157:347–371, 1971.

[134] J.C. Yakoubsohn. A universal constant for the convergence of Newton's method and an application to the classical homotopy method. *Numer. Algorithms*, 9(3-4):223–244, 1995.

[135] J.C. Yakoubsohn. Finding zeros of analytic functions: α theory for secant type methods. *J. Complexity*, 15(2):239–281, 1999.

[136] J.C. Yakoubsohn. Contraction, robustness, and numerical path-following using secant maps. *J. Complexity*, 16(1):286–310, 2000. Real computation and complexity (Schloss Dagstuhl, 1998).

[137] J.C. Yakoubsohn. Finding a cluster of zeros of univariate polynomials. *J. Complexity*, 16(3):603–638, 2000. Complexity theory, real machines, and homotopy (Oxford, 1999).

[138] J.C. Yakoubsohn. Simultaneous computation of all the zero-clusters of a univariate polynomial. In *Foundations of computational mathematics (Hong Kong, 2000)*, pages 433–455. World Sci. Publishing, River Edge, NJ, 2002.