

OPTIMIZACIÓN NO LINEAL CON RESTRICCIONES: PUNTOS DE FEKETE EN UNA ESFERA

E. Bendito¹, A. Carmona¹, A.M. Encinas¹ y J.M. Gesto²

1: Departament de Matemàtica Aplicada III
ETSECCPB, UPC

c/ Jordi Girona Salgado 1-3, 08034 Barcelona

e-mail: enrique.bendito,angeles.carmona,andres.marcos.encinas@upc.edu,

web: <http://www-ma3.upc.es/users/bencar>

2: Departament de Enginyeria del Terreny
ETSECCPB, UPC

c/ Jordi Girona Salgado 1-3, 08034 Barcelona

e-mail: jose.manuel.gesto@upc.edu,

web: <http://www-ma3.upc.es/users/bencar>

Palabras clave: Puntos de Fekete, optimización no lineal, restricciones no lineales, equilibrio electrostático.

Resumen. *Uno de los problemas más clásicos en el ámbito de la optimización de funcionales no lineales con restricciones no lineales es el cálculo de los denominados puntos de Fekete en la esfera unidad, que consiste en determinar la posición de n cargas eléctricas iguales sobre la esfera de manera que su energía potencial electrostática sea mínima.*

Este problema es considerado como uno de los paradigmas de la complejidad computacional. Con los algoritmos disponibles hasta la fecha no parece posible obtener configuraciones cercanas a mínimos para más de unos pocos miles de cargas, y eso haciendo uso de una considerable infraestructura de cálculo.

En esta comunicación presentamos un algoritmo robusto y eficiente para el problema de los puntos de Fekete en una esfera, que se basa en la integración numérica de la trayectoria cuasiestática seguida por cada carga sobre una esfera friccionante y donde el medio ofrece una resistencia viscosa infinita.

1. INTRODUCCIÓN

El problema de la obtención de los puntos de Fekete de orden n en una variedad S ocupa un lugar preeminente en la investigación matemática de las últimas décadas. En su versión original, consiste en determinar sobre un compacto $S \subset \mathbb{R}^2$ las posiciones de n puntos de manera que sea máximo el producto de sus distancias euclídeas dos a dos. Las n -uplas $\omega_n = \{x_1, \dots, x_n\}$, que satisfacen esa propiedad se conocen como *los puntos de Fekete de*

orden n en S . No es difícil comprobar que esas n -uplas minimizan en S el funcional

$$\mathcal{I}(\omega_n) = \sum_{1 \leq i < j \leq n} K(x_i, x_j), \quad (1)$$

donde $K(x_i, x_j) = -\ln|x_i - x_j|$ es el denominado *núcleo logarítmico* y $|x_i - x_j|$ es la distancia euclídea entre x_i y x_j . El valor $\mathcal{I}(\omega_n)$ representa, salvo un factor, la *energía potencial correspondiente al núcleo logarítmico*, ver [8], de la distribución ω_n cuando a todos sus puntos se les asigna un peso unidad.

Pueden obtenerse múltiples variantes del problema sin más que considerar diferentes núcleos en variedades arbitrarias. Un núcleo de particular interés es el *núcleo newtoniano*, que en el espacio tridimensional se define como $K(x_i, x_j) = |x_i - x_j|^{-1}$ y cuya energía potencial asociada $I(\omega_n)$ se denomina *energía potencial electrostática*. La interpretación electrostática del núcleo newtoniano corresponde a la situación en que cada dos puntos se repelen con una fuerza soportada por la recta que los une y de módulo inversamente proporcional al cuadrado de la distancia que los separa, es decir, como cargas eléctricas clásicas del mismo signo. Existen muchos otros núcleos relevantes sobre los que el lector interesado podrá encontrar información en [8].

En relación a las diferentes versiones del problema de los puntos de Fekete se han obtenido una cantidad considerable de resultados tanto teóricos como numéricos, ver por ejemplo, [3, 5, 6, 7, 9, 10, 11] donde el problema ha sido completamente resuelto para ciertos núcleos y ciertos valores de n . Por otra parte, existen cientos de referencias a las aplicaciones prácticas del problema de los puntos de Fekete, que van desde la química y la biología a la electrostática, pasando por el cálculo numérico (esquemas de interpolación, construcción de mallas para elementos finitos, integración numérica) y el diseño asistido por ordenador, ver por ejemplo [1, 2, 9, 11, 13].

El problema de los puntos de Fekete en la esfera unidad está considerado como modelo de problema de optimización de un funcional altamente no lineal con restricciones no lineales. A pesar de la enorme capacidad de los ordenadores actuales y de la gran evolución que han experimentado en los últimos años las técnicas numéricas, este tipo de problemas constituye todavía uno de los paradigmas de la complejidad computacional. De hecho, el problema de la obtención de un algoritmo robusto y eficiente para el problema de los puntos de Fekete en una esfera figura en la lista de problemas abiertos elaborada por Smale, [12].

En general, sólo es razonable esperar un buen comportamiento de los algoritmos habituales de optimización con restricciones cuando éstas son lineales o pueden ser aproximadas de manera suficientemente exacta por restricciones lineales, e incluso en esos casos las velocidades de convergencia suelen resultar muy inferiores a las de los métodos de optimización libre o sin restricciones [9]. Así, muchos autores optan por transformar problemas como el de los puntos de Fekete en problemas sin restricciones mediante alguna parametrización de la variedad. Por ejemplo, en relación a la optimización de funcionales en la esfera, se ha recurrido las parametrizaciones estándar dadas por la *proyección*

estereográfica y por las *coordenadas esféricas*, [9, 13]. Esta manera de proceder permite utilizar las técnicas clásicas de optimización, como por ejemplo el *método del Gradiente*, el *método de los Gradientes Conjugados*, el *método de Newton* y toda la familia de *métodos quasi-Newton*. No obstante, también se han utilizado métodos más sofisticados como los denominados *métodos de Optimización Combinatoria*, entre los cuales figuran el *Simulated Annealing*, el *Tabu Search* y los *Algoritmos Genéticos*. Además, se han diseñado estrategias de optimización combinando los diferentes métodos disponibles con la intención de aprovechar lo mejor de cada uno de ellos en cada fase del cálculo.

En general, la dificultad numérica de este tipo de problemas va más allá de la aparición de múltiples extremos. En particular, para el problema de los puntos de Fekete en una esfera, existen publicaciones muy recientes en que se reconoce que la mera obtención de una posición cercana a un óptimo local para unos cientos de cargas requiere hacer uso de una importante infraestructura de cálculo y, con los medios actuales, se considera prácticamente imposible para unos pocos miles de cargas [5].

Presentamos aquí los fundamentos de un algoritmo para la obtención de los puntos de Fekete en una esfera. Mostraremos que dicho algoritmo proporciona posiciones cercanas a óptimos locales para configuraciones de varios miles de puntos en tiempos razonables y con un único procesador convencional tipo Pentium IV.

2. FUNDAMENTOS DEL ALGORITMO

Si M es una variedad e $\mathcal{I}: M \rightarrow \mathbb{R}$ es una función objetivo, las estrategias habituales para la obtención de extremos de \mathcal{I} parten de una parametrización de la variedad y, en el espacio de parámetros, tratan de localizar los óptimos entre los puntos que anulan el gradiente de \mathcal{I} . La localización de tales puntos suele efectuarse partiendo de una posición inicial dada, llevando a cabo un acercamiento progresivo a algún candidato a óptimo mediante la determinación de una *dirección de avance* y una magnitud de tal avance, denominada *paso*, en cada iteración. Este es el fundamento de métodos tales como el del Gradiente, Gradiente Conjugados y, en cierta forma, también de los métodos de Newton y quasi-Newton. Las dificultades y limitaciones de este tipo de métodos son bien conocidas, entre ellas, la indecibilidad de la convergencia en problemas generales y la excesiva dependencia de la parametrización escogida. Sin embargo, también existen métodos que evitan esta última dificultad prescindiendo de parametrizaciones, decidiendo en cada iteración tanto la dirección de avance como el paso, evaluando directamente tanto la función como su gradiente sobre la variedad, ver por ejemplo [9].

El algoritmo que presentamos se sitúa en el ámbito de estos últimos en el sentido de que prescinde de parametrizar la variedad, aunque precisa del conocimiento de su campo normal.

Las ideas de las que deriva el algoritmo provienen de la Física. En lugar de preguntarnos cuáles son las configuraciones de mínima energía potencial consideraremos, dada una configuración de partida no óptima, de qué manera pueden moverse las partículas, en el sentido mecánico del término, para alcanzar una configuración de mínima energía

potencial. Tras ello subyace la equivalencia que en sistemas mecánicos existe entre configuraciones de mínima energía potencial y configuraciones en equilibrio estático estable. Un sistema mecánico compuesto por partículas está en equilibrio si la fuerza resultante que actúa sobre cada una de esas partículas es nula. El equilibrio será estable si las fuerzas generadas por pequeñas perturbaciones en la posición tienden a devolver a cada partícula a su posición original, y entonces dicha configuración será un mínimo de la energía potencial del sistema. Así pues, dada una configuración de partida no óptima, deberán existir fuerzas no equilibradas actuando sobre las partículas, y esas fuerzas inducirán inevitablemente el movimiento de las mismas. Nos interesará conocer el carácter de esas fuerzas y de ese movimiento, y cuándo conducirá a las partículas a una posición de equilibrio.

Si $x(t) = (x_1(t), \dots, x_n(t))$ denota la posición del sistema de las n partículas, que supondremos de carga unidad, donde $x_j(t)$ representa la posición de la partícula j -ésima, entonces las ecuaciones para el movimiento del sistema sobre la superficie de la esfera están dadas por la identidad

$$x''(t) = -\nabla \mathcal{I}(x(t)) - cx'(t) + \Phi(x(t), x'(t)), \quad c \geq 0. \quad (2)$$

El término $-cx'(t)$ representa a una fuerza disipativa de origen viscoso y $\Phi = (\Phi_1, \dots, \Phi_n)$ es una fuerza adicional que tiene en cuenta la interacción de cada partícula con la esfera y que además satisface que

$$\Phi(x(t), x'(t)) = (\Phi_1(x(t), x'_1(t)), \dots, \Phi_n(x(t), x'_n(t))). \quad (3)$$

Por otra parte, para cada $i = 1, \dots, n$, la fuerza Φ_i puede descomponerse en sus componentes normal y tangencial a la esfera, esto es

$$\Phi_i(x(t), x'_i(t)) = \Phi_i^n(x(t), x'_i(t)) + \Phi_i^t(x(t), x'_i(t)). \quad (4)$$

En la identidad anterior, si $-\nabla \mathcal{I} = (F_1, \dots, F_n)$, para cada $i = 1, \dots, n$ la componente normal Φ_i^n se encarga de anular la componente de F_i normal a la esfera y de proporcionar la fuerza centrípeta, por lo que su expresión debe estar determinada por la identidad

$$\Phi_i^n(x(t), x'_i(t)) = -[|x'_i(t)|^2 + \langle F_i(x(t)), x_i(t) \rangle] x_i(t), \quad i = 1, \dots, n. \quad (5)$$

Por otra parte, la componente tangencial Φ_i^t representa las fuerzas disipativas de fricción que actúan sobre cada partícula y que dependen por tanto de las características de su contacto con la esfera. En este trabajo, el modelo de fricción que proponemos es una modificación del modelo clásico de *Mohr-Coulomb* y está descrito por las identidades

$$\Phi_i^t(x(t), x'_i(t)) = - \left[|F_i(x(t)) + |x'_i(t)|^2 x_i(t)| - \frac{1}{4} \right] \sin(\alpha_i(x(t))) \frac{x'_i(t)}{|x'_i(t)|}, \quad i = 1, \dots, n, \quad (6)$$

donde $\alpha_i(x(t))$ es el ángulo que forman los vectores $F_i(x(t)) + |x'_i(t)|^2 x_i(t)$ y $x_i(t)$. Para la formulación de este modelo se ha tenido en cuenta que en la esfera unidad

$$\left| F_i(x(t)) + |x'_i(t)|^2 x_i(t) \right| \geq \frac{1}{4}, \quad i = 1, \dots, n. \quad (7)$$

Si consideramos

$$\mathbf{u}_i(x(t)) = \frac{F_i(x(t)) - \langle F_i(x(t)), x_i(t) \rangle x_i(t)}{|F_i(x(t)) - \langle F_i(x(t)), x_i(t) \rangle x_i(t)|}, \quad i = 1, \dots, n, \quad (8)$$

puede demostrarse que las trayectorias

$$x_i^c(t) = \frac{1}{4c} \int_0^t \sin(\alpha_i(x(s))) \mathbf{u}_i(x(s)) ds, \quad i = 1, \dots, n, \quad (9)$$

convergen cuando $c \rightarrow \infty$ a una función z_i y que $z(t) = (z_1(t), \dots, z_n(t))$ es solución de la ecuación (2). Denominaremos *trayectoria cuasi-estática* a cada función obtenida por el anterior proceso de paso al límite cuando $c \rightarrow \infty$. Puede demostrarse que una trayectoria de este tipo es siempre de descenso de la energía potencial \mathcal{I} .

3. INTEGRACIÓN DE LAS TRAYECTORIAS CUASI-ESTÁTICAS

Una forma de entender las trayectorias cuasi-estáticas anteriores consiste en interpretarlas, en tanto que envolventes de los vectores $\sin(\alpha_i(x(s))) \mathbf{u}_i(x(s))$, como soluciones del sistema de ecuaciones diferenciales ordinarias

$$z'_i(s) = \sin(\alpha_i(z(s))) \mathbf{u}_i(z(s)), \quad i = 1, \dots, n. \quad (10)$$

Para la resolución numérica de (10), utilizaremos el método de Euler hacia delante, con lo que cada iteración se reduce a

$$z_i(s_k + \Delta s_k) = z_i(s_k) + \Delta s_k \sin(\alpha_i(z(s_k))) \mathbf{u}_i(z(s_k)), \quad i = 1, \dots, n. \quad (11)$$

Como nuestro principal objetivo es que el algoritmo anterior permita localizar un mínimo de la energía potencial, cuando a partir de una posición inicial dada el algoritmo determine un mínimo, diremos que ha convergido. Sin embargo, cabe también preguntarse si el algoritmo consigue integrar suficientemente bien las trayectorias cuasiestáticas, es decir, si dada una posición inicial, consigue localizar el único punto estacionario que corresponde a dicha posición inicial. Cuando a partir de una posición inicial dada el algoritmo consiga encontrar el punto estacionario que le corresponde, diremos que ha sido estable.

Parece obvio que cuando en la posición inicial haya partículas muy cercanas a otras se debería tener un especial cuidado durante la integración de las trayectorias: sería necesario evitar que las partículas se atropellasen rompiendo la suavidad del movimiento, y eso conduciría a unos incrementos Δs_k muy pequeños que, si se mantuvieran constantes, ralentizarían excesivamente el proceso de avance de las partículas sobre la esfera cuando ya se hubieran distribuido razonablemente bien. En definitiva, para conseguir un grado de eficiencia razonable habría que redefinir Δs_k en cada iteración y además, Δs_k debería depender del número de partículas. Una forma sencilla y cómoda de tener todo esto en cuenta es reparametrizar la ecuación (10) para obtener el siguiente sistema

$$z'_i(s) = \varphi(z(s)) \sin(\alpha_i(z(s))) \mathbf{u}_i(z(s)), \quad i = 1, \dots, n \quad (12)$$

cuyas trayectorias solución coinciden con las del sistema original (10). Ahora el esquema de integración queda

$$z_i(s_k + \Delta s_k) = z_i(s_k) + \Delta s_k \varphi(z(s_k)) \sin(\alpha_i(z(s_k))) \mathbf{u}_i(z(s_k)), \quad i = 1, \dots, n \quad (13)$$

Cuando $\varphi(z) = \min_{1 \leq p \leq q \leq n} \{|z_p - z_q|\}$, lo que significa que el producto $\Delta s_k \sin(\alpha_i(z(s_k)))$ representa el máximo porcentaje de la mínima distancia entre partículas que se permite avanzar a cada una de ellas en cada iteración, se ha observado que con independencia de n , tomar $\Delta s_k \leq 1$ garantiza la convergencia del método salvo para configuraciones de partida realmente extremas. No obstante, no es posible asegurar que el algoritmo sea estable, salvo para posiciones de partida muy buenas.

Es posible aumentar la robustez del algoritmo considerando funciones φ algo más complicadas. Por ejemplo, si se toma $\varphi(z) = \sqrt{\min_{1 \leq i \leq n} \{\cos \alpha_i(z)\} \min_{1 \leq p \leq q \leq n} \{|z_p - z_q|\}}$, la condición $\Delta s_k \leq 1$ garantiza la convergencia del método. Si se toma

$$\varphi(z) = \left[(0,5 - a) e^{-\frac{1}{2b^2} \max_{1 \leq i \leq n} \{\sin^2 \alpha_i(z)\}} + a \right] \min_{1 \leq p \leq q \leq n} \{|z_p(s) - z_q(s)|\}$$

con $a = 0,005$ y $b = 0,01$, se obtiene un esquema siempre convergente y estable excepto para configuraciones de partida extremadamente difíciles. Cada usuario puede por tanto adecuar el algoritmo a su orden de prioridades; si lo único que se persigue es la localización de un punto estacionario, debe tenerse en cuenta que asegurar la convergencia es sumamente simple.

Naturalmente, si se ejecuta al pie de la letra el algoritmo que acabamos de describir, las partículas abandonarán la esfera. Para asegurar que se verifican las restricciones del problema es necesario devolverlas a ella después de cada nuevo avance. Para conseguirlo haremos uso de un *algoritmo de retorno* que consiste simplemente en realizar la proyección ortogonal sobre la esfera. De esta forma, cada nuevo avance involucra los pasos

$$\begin{aligned} z_i(s_k + \Delta s_k) &= z_i(s_k) + \Delta s_k \varphi(z(s_k)) \sin(\alpha_i(z(s_k))) \mathbf{u}_i(z(s_k)), \quad i = 1, \dots, n, \\ z_i(s_k + \Delta s_k) &= \frac{z_i(s_k + \Delta s_k)}{|z_i(s_k + \Delta s_k)|}, \quad i = 1, \dots, n. \end{aligned} \quad (14)$$

Ya que las distancias recorridas por cada partícula con este algoritmo son pequeñas, proceder así no afecta ni la convergencia ni la estabilidad del método.

4. APLICACIÓN DEL MÉTODO

En la aplicación del método descrito anteriormente, hemos observado que después de una primera fase altamente no lineal de acomodación o acercamiento propiamente dicho, el algoritmo muestra un orden de convergencia lineal. Cuando se alcanza la tendencia lineal puede suponerse en general que ya se está suficientemente cerca de un mínimo local y que el método deja de ser, en principio, competitivo, en comparación con el de Newton.

La Figura 1 muestra la evolución de $\log(\max_{1 \leq i \leq n} \{\sin \alpha_i\})$, que se ha considerado un buen indicador del error, en función del número de iteraciones.

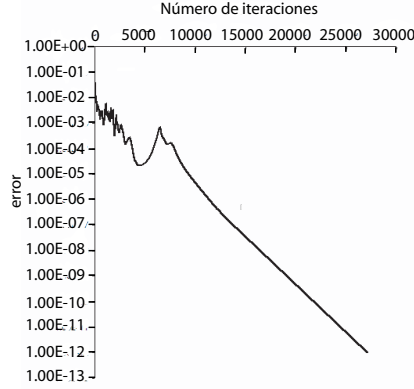


Figura 1: Comportamiento del algoritmo.

En concreto, la Figura 1 corresponde a la obtención de un mínimo local para la energía potencial electrostática de 1000 partículas sobre la esfera unidad, partiendo de una configuración generada aleatoriamente ajustándose a una densidad de probabilidad uniforme sobre su superficie y exigiendo únicamente convergencia. El esquema utilizado fue el correspondiente a la elección $\varphi(z) = \min_{1 \leq p \leq q \leq n} \{|z_p - z_q|\}$.

Utilizando un único computador convencional con procesador tipo Pentium IV de 2,8 GHz y 512 MB de memoria RAM, el tiempo necesario para conseguir un error inferior a 10^{-12} fue de algo menos de una hora. No obstante, el tiempo necesario para alcanzar la tendencia lineal fue de apenas 20 minutos. La Figura 2 muestra, a izquierda y derecha respectivamente, las posiciones inicial y final en este caso.

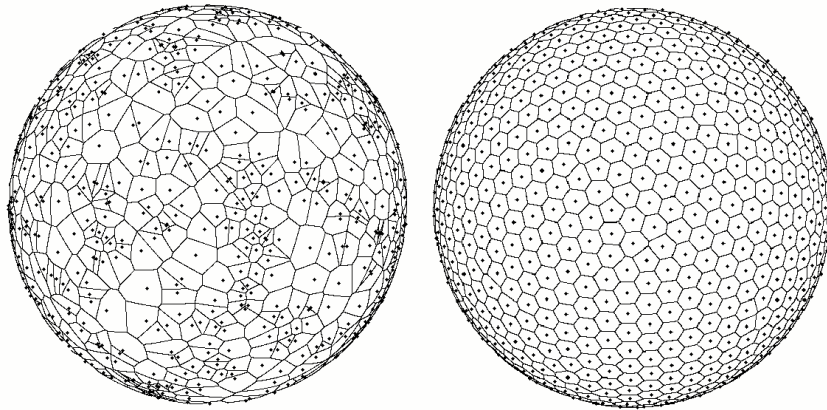


Figura 2: Configuraciones inicial y final y celdas de Dirichlet, para $n = 1000$.

La Figura 3 muestra una configuración de tamaño $n = 5000$ después de alcanzar la tendencia lineal, lo que requirió aproximadamente 20 horas con el mismo procesador.

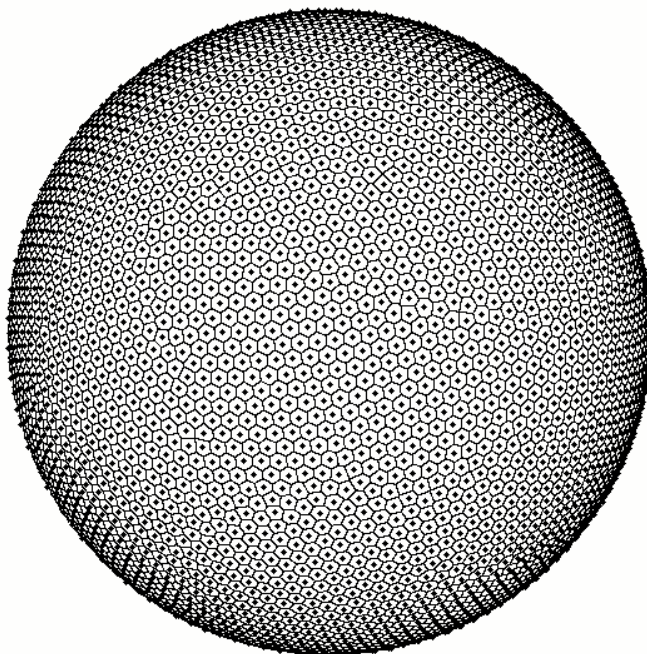


Figura 3: Configuración final y celdas de Dirichlet, para $n = 5000$.

En base a la experimentación numérica llevada a cabo con diferentes valores de n y utilizando diferentes posiciones de partida en cada caso es posible concluir que el coste computacional del algoritmo es $\mathcal{O}(n^2 \ln n)$; en cuanto a la demanda de memoria, es obvio que se reduce a $\mathcal{O}(n)$. Teniendo en cuenta los tiempos de cálculo necesarios para alcanzar la tendencia lineal correspondientes a los ejemplos anteriores, queda claro que con una buena infraestructura de cálculo pueden abordarse sin dificultad problemas de gran tamaño. Por otra parte, la robustez del algoritmo ha quedado corroborada en absolutamente todas las pruebas realizadas, habiéndose estudiado casos con valores de n superiores a 50000, ver [4].

Agradecimientos

Este trabajo ha sido parcialmente financiado por la CICYT, proyecto BFM2003-06014, y por la ETSECCPB.

REFERENCIAS

- [1] M. Atiyah, P. Sutcliffe, The geometry of points particles, *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.*, **458** (2002), 1089-1115.

- [2] M. Atiyah, P. Sutcliffe, Polyedra in Physics, Chemistry and Geometry, *Milan J. Math.*, **71** (2003), 33-58.
- [3] B. Bergersen, D. Boal, M. Pálffy-Muhoray, Equilibrium configurations of particles on a sphere: the case of logarithmic interactions, *J. Phys. A: Math. Gen.*, **27** (1986), 2579-2586.
- [4] J.M. Gesto, *Estimación de los puntos de Fekete en la esfera unidad*. Tesina Fin de Carrera, UPC, (2003).
- [5] D.P. Hardin, E.B. Saff, Discretizing Manifolds via Minimum Energy Points, *Notices Amer. Math. Soc.*, **51** (2004), 1186-1194.
- [6] J. Korevaar, M.A. Monterie, Approximation of the equilibrium distribution by distributions of equal point charges with minimal energy, *Trans. Amer. Math. Soc.*, **350** (1998), 2329-2348.
- [7] A.B.J. Kuijlaars, E.B. Saff, Asymptotics for minimal discrete energy on the sphere, *Trans. Amer. Math. Soc.*, **350** (1998), 523-538.
- [8] N.S. Landkof, *Foundations of Modern Potential Theory*, Springer-Verlag, (1972).
- [9] K.J. Nurmela, Constructing Spherical Codes by Global Optimization Methods, *Helsinki University of Technology, Series A: Research Reports*, **32**, (1995).
- [10] E.A. Rakhmanov, E.B. Saff, Y. Zhou, Minimal discrete energy on the sphere, *Math. Res. Lett.*, **1** (1994), 647-662.
- [11] E.B. Saff, A.B.J. Kuijlaars, Distributing many points on a sphere, *Math. Intelligencer*, **19** (1997), 5-11.
- [12] S. Smale, Mathematical Problems for the Next Century, *Math. Intelligencer*, **20** (1998), 7-15. Versión española en *Gaceta de la RSME*, **3** (2000), 413-434.
- [13] Y. Zhou, *Arrangements of points on the sphere*, Ph.D. Department of Mathematics, University of South Florida, (1995).