

40
2ej.



UNIVERSIDAD NACIONAL
AUTONOMA DE MEXICO

Facultad de Ciencias

PROPIEDADES ESTADISTICAS
DEL PROBLEMA DEL AGENTE
VIAJERO

T E S I S
Que para obtener el Titulo de
F I S I C O
p r e s e n t a
ALEXANDER VALLADARES Mc NELIS



FACULTAD DE CIENCIAS
DIVISION ESCOLAR

México, D. F.

1994

TESIS CON
FALLA DE ORIGEN



UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso

DERECHOS RESERVADOS © PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

CIUDAD UNIVERSITARIA



UNIVERSIDAD NACIONAL
AVENIDA DE
MEXICO

FACULTAD DE CIENCIAS
División de Estudios
Profesionales
Exp. Núm. 55

M. EN C. VIRGINIA ABRIN BATULE
Jefe de la División de Estudios Profesionales
Universidad Nacional Autónoma de México.
P r e s e n t e .

Por medio de la presente, nos permitimos informar a Usted, que habiendo
revisado el trabajo de tesis que realizo el pasante ALEXANDER
VALLADARES MC NELIS
con número de cuenta 8634176-7 con el título: "PROPIEDADES
ESTADISTICAS DEL PROBLEMA DEL AGENTE VIAJERO"

Consideramos que reúne los méritos necesarios para que pueda conti-
nuar el trámite de su Examen Profesional para obtener el título de
FISICO.

GRADO NOMBRE Y APELLIDOS COMPLETOS

DR. JORGE ANDRES FLORES VALDES
Director de Tesis
DR. OCTAVIO AUGUSTO NOVARO PERALOZA
DR. VICTOR MANUEL ROMERO ROCHIN
DR. THOMAS HENRY SELIGMAN SCHUCH
Suplente
M. EN C. VIRGINIA ABRIN BATULE
Suplente

FIRMA

Virginia Abrin Batule

Ciudad Universitaria, D.F., a 12 de mayo de 1994

PROPIEDADES ESTADISTICAS DEL PROBLEMA DEL AGENTE VIAJERO

... the nicest balancings between competitive advantages,
the smoothest compromises between the practical and the ideal.

Winston S. Churchill
My African Journey

A mis padres y a mi hermana por todo.

Agradezco al Dr. Jorge Flores Valdés por la dirección de este trabajo y al Dr. Thomas Seligman Schuch por su interés y participación.

A los Drs. Octavio Novaro Peñalosa y Víctor Romero Rochín y a la M. en C. Virginia Abrín Batule, sinodales de este trabajo, por sus útiles comentarios.

INDICE

INTRODUCCION	1
CAPITULO 1 EL PROBLEMA DEL AGENTE VIAJERO Y VIDRIOS DE ESPIN	3
1.1 EL PROBLEMA DEL AGENTE VIAJERO	3
1.2 VIDRIOS DE ESPIN	4
1.3 ANALOGIAS ENTRE EL PROBLEMA DEL AGENTE VIAJERO Y VIDRIOS DE ESPIN	13
CAPITULO 2 ALGUNOS METODOS PARA RESOLVER APROXIMADAMENTE EL PROBLEMA DEL AGENTE VIAJERO	25
2.1 METODO DE LIN	25
2.2 METODO DEL TEMPLADO SIMULADO	29
2.3 METODO JERARQUICO	32
CAPITULO 3 PROPIEDADES ESTADISTICAS DEL PROBLEMA DEL AGENTE VIAJERO	35
3.1 PROGRAMA DEL TEMPLADO SIMULADO	35
3.2 PROPIEDADES ESTADISTICAS DEL PROBLEMA DEL AGENTE VIAJERO	36
3.2.1 350 Muestras de 500 Ciudades Cada Una	40
3.2.2 8000 Muestras de 20 Ciudades Cada Una	49
3.2.3 350 Soluciones de Una Sola Muestra de 500 Ciudades	55
3.2.4 Dependencia con la Temperatura	60

CAPITULO 4 CONCLUSIONES	73
APENDICE 1 PROPIEDADES ESTADISTICAS DE ESPECTROS ALEATORIOS	77
A1.1 MATRICES ALEATORIAS	77
A1.2 PROPIEDADES ESTADISTICAS DE ESPECTROS ALEATORIOS	81
REFERENCIAS	97

INTRODUCCION

El problema del agente viajero es un ejemplo clásico del tipo de problemas que se estudian en el área de optimización combinatorial, la cual tiene como tema central de estudio el desarrollo de métodos eficientes para la obtención de máximos o mínimos de funciones que dependen de un gran número de variables. En su forma más simple, este problema consiste en encontrar el camino de mínima distancia que pase por N puntos distribuidos aleatoriamente dentro de un cuadrado de 1 por 1 de tal forma que se pase por cada punto una sola vez y se regrese al punto de origen. En este problema en particular, la función que se quiere minimizar es la distancia total recorrida y el espacio sobre el cual se efectúa la búsqueda del mínimo es el espacio que consta de todos los caminos diferentes, sin importar la dirección en la que éstos se recorran. Cuando el número N de puntos (o ciudades, si se piensa que un agente viajero tiene que realizar sus ventas en N ciudades distintas y quiere minimizar la distancia recorrida al visitar cada una de ellas) es grande, es muy difícil encontrar caminos que realmente sean de mínima distancia y debido a esto se recurre a diversos métodos que dan como resultado soluciones aproximadas.

El objetivo de este trabajo es estudiar algunas de las propiedades estadísticas que ciertas soluciones aproximadas del problema del agente viajero exhiben. Específicamente, se estudiarán propiedades estadísticas de las soluciones obtenidas usando el método del templado simulado, del cual se hablará más adelante. Para llevar a cabo este estudio se utilizará la teoría de las matrices aleatorias (la cual es muy útil tanto en física nuclear como en física atómica).

Aparte del interés que estas propiedades por sí mismas tienen, los resultados obtenidos pueden ser útiles en los

siguientes dos problemas. El primero es el estudio de los vidrios de espín. Existe una fuerte analogía entre el problema del agente viajero y los vidrios de espín, por lo que es de esperarse que información acerca de uno de estos problemas sea útil para el mejor entendimiento del otro. El segundo es la elaboración de un nuevo método, basado en estas propiedades estadísticas, que permita resolver aproximadamente el problema del agente viajero. Así, parte de la motivación del presente trabajo es su posible utilidad en estos dos problemas, aunque aquí sólo se hablará acerca de las propiedades estadísticas en sí y no de las implicaciones que dichas propiedades puedan tener en ellos.

El Capítulo 1 trata acerca del problema del agente viajero así como de algunas de las analogías que existen entre este problema y los vidrios de espín. En el Capítulo 2, se habla acerca de algunos de los métodos existentes que dan como resultado soluciones aproximadas. En particular, se habla acerca del método utilizado para la obtención de las soluciones del problema en cuestión. El Capítulo 3 trata acerca de la forma en que se utiliza la teoría de las matrices aleatorias para analizar este problema, así como de los resultados obtenidos. En el 4 se presentan las conclusiones obtenidas con base en los resultados obtenidos. Por último, en el Apéndice 1 se encuentra información acerca de la teoría de las matrices aleatorias y de las propiedades estadísticas que se acostumbra estudiar al utilizar dicha teoría.

CAPITULO 1

EL PROBLEMA DEL AGENTE VIAJERO Y VIDRIOS DE ESPIN

1.1 EL PROBLEMA DEL AGENTE VIAJERO

Como se mencionó en la Introducción, el problema del agente viajero⁽¹⁾, en su forma más simple, consiste en encontrar el camino de mínima distancia que pase por N puntos distribuidos aleatoriamente dentro de un cuadrado de lado 1 de tal forma que se pase por cada punto una sólo vez y se regrese al punto de origen. En general, las coordenadas x_i y y_i de cada punto son variables aleatorias con densidad de probabilidad uniforme. El número de caminos posibles es $(N-1)!/2$, ya que no importa ni el punto del cual se parte ni la dirección en la que el camino se recorre. Si N es grande, el número de caminos distintos es enorme y es sumamente impráctico calcular la distancia de todas las rutas para de ahí escoger la de mínima distancia (el tiempo de cómputo necesario para esto va como $e^{\text{const.} \cdot N}$). Debido a esto, se han diseñado distintos métodos⁽¹⁻⁶⁾ para encontrar soluciones que aunque no son de mínima distancia, son suficientemente buenas y rápidas de encontrar para todo propósito práctico.

El problema del agente viajero es un ejemplo del tipo de problemas que se estudian en el campo de optimización combinatorial⁽¹⁻²⁾. La investigación en esta área se enfoca a desarrollar técnicas eficientes para encontrar máximos o mínimos de funciones que dependen de muchas variables. La función que se quiere minimizar o maximizar se conoce como el costo o como la función objetivo y depende de la configuración de las distintas partes del sistema que se esté estudiando. En el caso del problema del agente viajero, el costo es la distancia total y ésta depende

de la distribución de las ciudades.

Resulta que existe una relación estrecha entre el campo de optimización combinatorial y la mecánica estadística^(2-4,7). En particular, se tiene que existen semejanzas notables entre algunos sistemas termodinámicos y ciertos problemas de optimización combinatorial, como es el caso con los vidrios de espín y el problema del agente viajero. Para poder estudiar esta relación, a continuación se mencionarán algunos aspectos acerca de los vidrios de espín.

1.2 VIDRIOS DE ESPIN

Un vidrio de espín⁽⁷⁻¹⁰⁾ es un conjunto de espines o momentos magnéticos cuyo estado a baja temperatura es uno de desorden congelado, a diferencia de los estados uniformes u ordenados de otros sistemas, como es el caso de materiales magnéticos convencionales. Para que existan tales estados, es necesario lo siguiente⁽⁸⁾: debe de haber competencia entre las interacciones entre espines en el sentido de que no haya ninguna configuración favorecida por todas ellas, lo cual se conoce como frustración y debe de haber cierta aleatoriedad en dichas interacciones.

Como ejemplos de vidrios de espín se tienen los cristales de metales nobles con impurezas de metales de transición magnéticos, como es el caso de Cu con bajas concentraciones de Mn (0.9%, por ejemplo). Lo que sucede en estos materiales es lo siguiente. Cada impureza polariza a los electrones de conducción del metal noble que se encuentran a su alrededor. Esta polarización es positiva a ciertas distancias de la impureza y negativa a otras. Debido a esto, las otras impurezas, al sentir el campo magnético local producido por los electrones de conducción polarizados, tratarán de alinearse a lo largo de la dirección de dicho campo. Así, debido a que las impurezas están distribuidas aleatoriamente dentro del cristal, algunas de las interacciones entre ellas serán positivas mientras que otras serán negativas. Esto, aunado al hecho de que dos impurezas interaccionarán más débilmente entre sí conforme más lejos se encuentren, trae como consecuencia que tanto la magnitud como el signo de las

interacciones de intercambio entre las impurezas sean aleatorias. Es evidente que estos sistemas presentan tanto frustración como aleatoriedad.

Uno de los modelos de vidrios de espín, el cuál está basado en lo que sucede en este tipo de materiales, consiste en espines S_i situados en sitios aleatorios R_i cuya interacción es del tipo RKKY (Ruderman, Kittel, Kasuya, Yosida)⁽¹¹⁾; es decir, $J(r) \propto \cos(ar)/r^3$. En la Figura 1.1 se muestra la interacción de intercambio $J(r)$ entre dos espines cualesquiera como función de la separación r entre ellos.

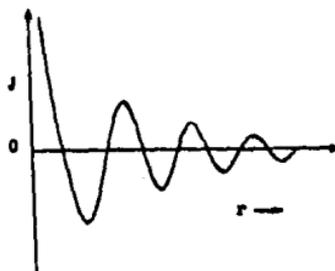


Fig. 1.1. Interacción de intercambio $J(r)$ del tipo RKKY como función de la separación r entre espines.

Ya que el comportamiento tipo vidrio de espín lo exhiben muchos sistemas además de los arriba mencionados, se han introducido modelos más simples para la descripción de todos ellos. Uno de estos modelos es el modelo SK (Sherrington y Kirkpatrick)⁽⁷⁻⁹⁾, el cuál incorpora muchas de las propiedades que dichos sistemas presentan. En este modelo, los espines S_i son espines de Ising que ocupan los sitios i de una red regular. Las interacciones J_{ij} entre ellos son variables aleatorias con densidad de probabilidad normal. Específicamente, el hamiltoniano está dado por

$$H = -(1/2) \sum_{i,j} J_{ij} S_i S_j - h \sum_i S_i \quad (1.1)$$

y la densidad de probabilidad asociada con las J_{ij} , por

$$P(J_{ij}) = 1/(2\pi\Delta_{ij})^{1/2} \exp(-J_{ij}^2/2\Delta_{ij}). \quad (1.2)$$

Aquí h es un campo externo uniforme y la varianza Δ_{ij} es independiente de los espines S_i y S_j ; es decir,

$$\Delta_{ij} = J^2/N, \quad (1.3)$$

siendo N el número total de espines. El que la varianza no dependa de S_i y S_j , trae como consecuencia que la interacción entre pares de espines es independiente de la distancia que los separa. Se hace notar que para una cierta muestra de material, las J_{ij} están fijas, pero variarán al ir de muestra en muestra.

Los vidrios de espín son un ejemplo de sistemas desordenados. Por sistema desordenado se entiende aquel sistema que no exhibe regularidades en lo que se refiere a ciertas características "particulares, como por ejemplo, estructura topológica o celular^(8,12).

Debido a la aleatoriedad presente en los sistemas desordenados, se puede pensar que la información que se obtenga al usar sólo una muestra va a ser información acerca de esa muestra en particular y no información general acerca del tipo de sistemas que se estén estudiando (en este caso, los vidrios de espín). Sin embargo, resulta que hay ciertas cantidades que van a fluctuar muy poco al ir de muestra en muestra si el tamaño de éstas es suficientemente grande; es decir, si el número de partículas N es suficientemente grande. A este tipo de cantidades se les conoce como cantidades autopromediantes. De igual forma que las fluctuaciones de la energía de un sistema en equilibrio termodinámico son del orden $O(N^{-1/2})$, las fluctuaciones de las cantidades autopromediantes, al pasar de una muestra a otra, van a ser de este mismo orden y por lo tanto serán despreciables si las muestras son suficientemente grandes. Debido a esto, al encontrar la media de cualquiera de estas cantidades sobre un ensemble de muestras, el valor obtenido será muy cercano al valor que se obtendría al considerar cualquier muestra en particular y es así como se puede obtener información general acerca del tipo de

sistemas que se estén considerando.

Como ejemplo de cantidades autopromediantes se tienen las cantidades extensivas, como lo son la energía interna y la energía libre. También, existen cantidades que no son autopromediantes como, por ejemplo, el campo interno local en un cierto punto en el caso de los vidrios de espín. Estas cantidades variarán de muestra en muestra sin importar el tamaño de éstas.

Regresando al modelo SK de los vidrios de espín, se tiene que la aleatoriedad que estos sistemas exhiben está contenida en el hecho de que las interacciones J_{ij} son variables aleatorias. Debido a esto, para obtener información que no dependa sólo de la muestra en cuestión, hay que hacer dos tipos de promedios de las diferentes cantidades en las que se tenga interés. Primero, hay que encontrar, para valores fijos de las J_{ij} , la media termodinámica de la cantidad en cuestión y segundo, hay que encontrar el valor esperado de esta media promediando sobre las diferentes configuraciones que se tienen al ir de muestra en muestra, esto es, al ir variando los valores de las J_{ij} . Para encontrar este valor esperado configuracional o sobre el desorden, como suele llamarse, hay que utilizar la densidad de probabilidad conjunta

$$P[J] = \prod_{\langle i,j \rangle} P(J_{ij}). \quad (1.4)$$

Cabe mencionar que esto no sólo se hace al estudiar vidrios de espín, sino que se hace al estudiar cualquier tipo de sistema desordenado. El valor esperado configuracional o sobre el desorden se denota como $[]_{av}$.

En el caso de sistemas no desordenados, se pueden calcular muchas cantidades de interés a partir de la energía libre. En el caso desordenado y por lo tanto en el caso de los vidrios de espín, primero se debe de encontrar la energía libre $F[J]$ correspondiente a una muestra en particular (en el modelo SK esto equivale a encontrar la energía libre $F[J]$ para valores fijos de las J_{ij}) y después se debe hacer el promedio configuracional para encontrar una energía libre F a partir de la cual se pueda obtener información general que sea independiente de las características

particulares de las muestras. Si k_B , la constante de Boltzmann, se toma igual a la unidad, para el modelo SK se tiene que

$$F[J] = -T \ln(Z[J]) \quad (1.5)$$

y

$$F = [F[J]]_{av} = -T \int \ln(Z[J]) dP[J], \quad (1.6)$$

siendo $Z[J]$ la función de partición correspondiente a valores fijos de las J_{ij} .

A partir de esta energía libre se pueden calcular diferentes cantidades extensivas de la misma forma que en el caso no desordenado; es decir, derivando F . Las cantidades que así se obtengan van a ser los valores esperados configuracionales de cantidades extensivas de una muestra dada en equilibrio termodinámico. Cabe recalcar que es el logaritmo de la función de partición y no la función de partición misma la que se promedia configuracionalmente, ya que el logaritmo es una cantidad extensiva mientras que la función de partición no lo es.

Al igual que en el caso de los ferromagnetos, los vidrios de espín sufren un cambio de fase^(8,10) a cierta temperatura T_c . Con $h = 0$ se tiene que arriba de T_c , $m_i = \langle S_i \rangle = 0$ para toda i , pero cuando $T < T_c$, $m_i \neq 0$. A diferencia de los ferromagnetos de Ising, en los cuales la magnetización M es igual a cero si $T > T_c$ y es diferente de cero si $T < T_c$, en los vidrios de espín $M = 0$ a cualquier temperatura debido a que los estados de baja temperatura son estados de desorden congelado. En el caso de los ferromagnetos de Ising, hay dos estados estables por debajo de T_c : uno en el cual el valor de la magnetización es M y otro en el cual es $-M$. En el caso de los vidrios de espín, hay muchos estados estables para $T < T_c$, que corresponden a ciertas configuraciones del sistema en las cuales los diferentes espines apuntan en distintas direcciones y no todos en la misma dirección, como en el caso ferromagnético. Este comportamiento es debido a la frustración y aleatoriedad presentes.

Estrictamente hablando, estos estados son estables si el número de partículas $N \rightarrow \infty$. Si N es finito, tanto en el caso de

los ferromagnetos como en el de los vidrios de espín, el sistema pasará de uno de estos estados a otro al transcurrir un tiempo suficientemente largo; es decir, estos estados son metaestables. En cambio, si N es infinito, el sistema requeriría una energía infinita para pasar de uno de estos estados a otro, de forma que estando en uno de ellos no podrá pasar a los demás. Debido a esto se dice que abajo de T_c y cuando $N \rightarrow \infty$ hay un rompimiento de la ergodicidad, ya que el sistema sólo podrá estar en una fracción de todos los estados a los que tendría acceso si la temperatura fuera mayor que T_c .

Si se grafica la energía libre F de un vidrio de espín a temperatura constante $T < T_c$ como función de m_1, \dots, m_N , ésta tendrá muchos valles o mínimos locales separados por barreras finitas. Estos valles corresponden a los estados metaestables del sistema. En la Figura 1.2 se muestra F como función de una de las m_i . Si $N \rightarrow \infty$, las barreras entre algunos de los valles se harán infinitas de tal forma que el sistema no podrá pasar de un valle a otro.



Fig. 1.2. Energía libre F de un vidrio de espín como función de una de las m_i .

A diferencia de los ferromagnetos, la superficie de energía libre es mucho más complicada en este caso. Según la teoría de Landau⁽¹³⁾, para un ferromagneto tipo Ising cerca del punto crítico y con $h = 0$, la energía libre, en la aproximación de campo medio, se puede expresar como

$$F = F_0(T) + \int ((1/2)a(T)M^2 + (1/4)u(T)M^4) d^3r, \quad (1.7)$$

en donde M es la magnetización en el punto r (independiente de r en esta aproximación) y $a(T) = a_0(T-T_c)/T_c$, con $a_0 > 0$. Si $a(T) > 0$, la expresión $V = (1/2)aM^2 + (1/4)uM^4$ tiene un mínimo en $M = 0$, mientras que si $a(T) < 0$, dicha expresión tiene dos mínimos, uno en M y otro en $-M$, tal y como se mencionó anteriormente. Esto se ilustra en la Figura 1.3.

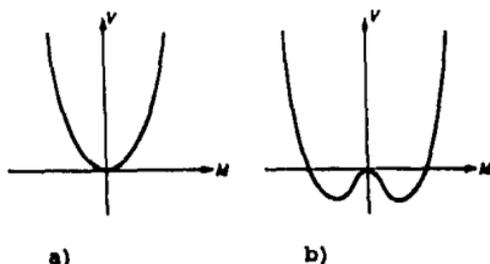


Fig. 1.3. a) V como función de M para $a(T) > 0$. b) V como función de M para $a(T) < 0$.

Si $T > T_c$, el valor de la magnetización cuando el ferromagneto está en equilibrio es 0 (ya que F tiene un mínimo en este punto), mientras que si $T < T_c$, este valor puede ser M ó $-M$.

Es interesante comparar la Figura 1.3b con la Figura 1.2. En el caso de los vidrios de espín, la superficie de energía libre a bajas temperaturas consta de muchos valles o mínimos locales, a diferencia de la superficie de los ferromagnetos, en la cual sólo hay dos mínimos globales. Arriba de T_c , la superficie de los vidrios de espín tendrá sólo un valle centrado en $m_1 = 0$, en analogía con la Figura 1.3a.

En el caso de los vidrios de espín y a temperaturas por debajo de T_c se tiene lo siguiente. Si se numeran los valles usando el índice a , la probabilidad de que al escoger cierta muestra ésta esté en el estado a es

$$P_a = \exp(-\beta F_a) / (\sum \exp(-\beta F_a)), \quad (1.8)$$

con F_a la energía libre del sistema cuando éste se encuentra en dicho estado. Una cantidad que mide el traslape entre dos estados diferentes de una misma muestra es

$$q_{ab} = (1/N) \sum_i m_i^a m_i^b, \quad (1.9)$$

donde m_i^z es el valor esperado termodinámico del espín S_i cuando la muestra está en el estado z . Como las diferentes muestras pueden estar en un gran número de estados distintos, q_{ab} puede tomar muchos valores diferentes, todos ellos entre 1 y -1. Debido a esto, es interesante saber cuál es la densidad de probabilidad asociada con estos traslapes. Esta está relacionada con la densidad de probabilidad P_a y está dada por

$$P_j(q) = \langle \delta(q - q_{ab}) \rangle = \sum_{ab} P_a P_b \delta(q - q_{ab}). \quad (1.10)$$

Esta densidad da la probabilidad de que para una muestra dada, el traslape q entre dos estados cualesquiera valga q_{ab} . De la ecuación 1.10 se ve que la probabilidad de que este traslape valga q_{ab} es $\sum_a P_a P_b$, donde esta suma significa suma sobre todos los pares ordenados de estados tales que q es igual a q_{ab} . Si resulta que para ningún par de estados q es igual a q_{ab} , esta probabilidad será igual a cero.

Es posible calcular el promedio configuracional $P(q)$ de la densidad de probabilidad $P_j(q)$. $P(q)$ da información acerca del número de estados o valles presentes en las superficies de energía libre asociadas con el tipo de sistemas en estudio cuando estos se encuentran a baja temperatura. Por ejemplo, si el tipo de sistemas en estudio sólo tienen dos estados posibles, la densidad de probabilidad será la suma de dos funciones delta, mientras que si el tipo de sistemas es tal que las superficies de energía libre presentan muchos valles, como en el caso de los vidrios de espín, $P(q)$ tendrá partes continuas, indicando un continuo de traslapes asociados con ciertos conjuntos de estados de estos sistemas.

Cabe mencionar que es posible construir una teoría de

campo medio para los vidrios de espín con base en el modelo SK. Sin embargo, aquí no se hablará en detalle acerca de esta teoría ya que esto implicaría desviarse mucho del tema central de este capítulo, el cual, como ya se mencionó, es el estudio de algunas de las analogías que existen entre el problema del agente viajero y los vidrios de espín. Lo único que se mencionará acerca de esta teoría es uno de sus resultados más importantes. Este consiste en que el espacio de configuración de cualquier muestra, esto es, el espacio N -dimensional cuyos ejes están etiquetados por las m_i , es ultramétrico⁽⁷⁻⁸⁾. Diversos experimentos parecen indicar que ésta es una de las propiedades características de los vidrios de espín, de forma que el modelo SK es bastante exitoso en el sentido de que reproduce, entre otras cosas, este hecho.

Un espacio es ultramétrico⁽⁷⁾ si en él se cumple la desigualdad

$$d(A,C) \leq \max\{d(A,B), d(B,C)\}, \quad (1.11)$$

siendo $d(X,Y)$ la distancia entre los puntos X y Y especificada con base en cierta métrica. Esta condición es parecida a la desigualdad del triángulo que se cumple en los espacios euclídeos, pero es más fuerte que ella. Todos los triángulos en un espacio ultramétrico son equiláteros, o isósceles siendo el lado diferente el de menor distancia.

Si se define la distancia entre dos estados en el espacio de configuración de una muestra dada como la fracción de espines que no coinciden, es decir, $d_{ab} = (1/2)(1 - q_{ab})$, la propiedad ultramétrica dice que si se escogen tres estados al azar, la distancia entre los tres es igual o bien la distancia de dos de ellos a un tercero es igual y mayor que la distancia entre ellos mismos. Esto indica que el espacio de configuración tiene cierta estructura de "árbol genealógico" en el que el grado de traslape entre dos estados está asociado con el grado de parentesco, como se muestra en la Figura 1.4.

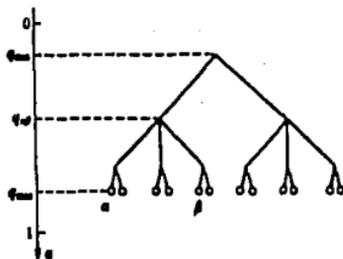


Fig. 1.4. Diagrama de la estructura de "árbol genealógico" que presentan los traslapes asociados con los estados de los vidrios de espín.

1.3 ANALOGIAS ENTRE EL PROBLEMA DEL AGENTE VIAJERO Y VIDRIOS DE ESPIN

Existen variaciones de la versión del problema del agente viajero hasta ahora mencionada⁽¹⁴⁻¹⁵⁾. Una de las más comunes consiste en considerar N ciudades en un espacio de D dimensiones y en lugar de especificar las coordenadas de todas las ciudades, se especifican las distancias entre todas ellas usando para esto cierta densidad de probabilidad, como puede ser una densidad de probabilidad uniforme o una densidad de probabilidad gamma. En este último caso, la densidad de probabilidad gamma utilizada está dada por

$$\rho_r(L) = (L^r e^{-L/r}) I_{(0, \infty)}(L), \quad (1.12)$$

siendo L la distancia entre ciudades, $r = D-1$ e $I_{(0, \infty)}(L)$ la función que vale 1 si $0 < L < \infty$ y cero si no.

Como en la primera versión las coordenadas de las ciudades están fijas, las distancias entre todas ellas están totalmente determinadas. En la segunda versión, sin embargo, lo que se hace es ignorar las correlaciones triangulares introducidas por las distancias euclídeas y se toma la distancia entre cada par de ciudades como una variable aleatoria independiente de todas las demás.

Como se verá a continuación, tanto la versión de

coordenadas fijas como la de distancias fijas tienen muchas cosas en común con los vidrios de espín.

Antes de proseguir, cabe mencionar que aunque ambas versiones son muy parecidas a los vidrios de espín en varios aspectos, existen ciertas propiedades de los vidrios de espín que sólo la versión de distancias fijas o la de coordenadas fijas presentan. Esta es la razón por la que en este capítulo se ha introducido la versión de distancias fijas, aunque en los siguientes capítulos sólo se trabajará con la versión de coordenadas fijas. Con esto, se tiene lo siguiente.

Como primer punto, se tiene el hecho de que ambas versiones del problema del agente viajero pueden ser tratadas como sistemas termodinámicos desordenados⁽⁷⁾.

En la primera versión, una muestra resulta ser un conjunto de N ciudades con coordenadas específicas. Al ir pasando de un conjunto de ciudades a otro se pasa de una muestra a otra. De igual forma que dentro del modelo SK cada muestra está caracterizada por ciertos valores específicos de las J_{ij} , en este caso cada muestra está caracterizada por las coordenadas de las N ciudades en cuestión. En la segunda versión, una muestra resulta ser un conjunto de N ciudades tales que las distancias L_{ij} entre todas ellas están definidas. En este caso, son precisamente estas distancias las que caracterizan a cada muestra.

Para cualquiera de las dos versiones, dado un camino inicial es posible construir otros caminos empleando para esto ciertos movimientos de los segmentos rectilíneos que unen, a lo largo del camino, a las distintas ciudades. Aunque se puede utilizar cualquier tipo de movimiento para ir construyendo nuevos caminos, existen dos tipos de movimientos que son utilizados con mucha frecuencia. La razón por la cual estos movimientos son tan comunes quedará clara en el siguiente capítulo.

El primer tipo⁽⁵⁾ consiste en escoger dos segmentos d_{ij} y d_{ki} y reemplazarlos por los segmentos d_{ik} y d_{ji} , como se muestra en la Figura 1.5a. El primer subíndice indica la ciudad de la que se parte y el segundo la ciudad a la que se llega, una vez que se ha escogido arbitrariamente la dirección en la que se está recorriendo el camino. El resultado de este movimiento es que la porción del camino que se encontraba entre las ciudades j y k

ahora se recorre en la dirección contraria. Es por esto que a este tipo de movimientos se les conoce como inversiones.

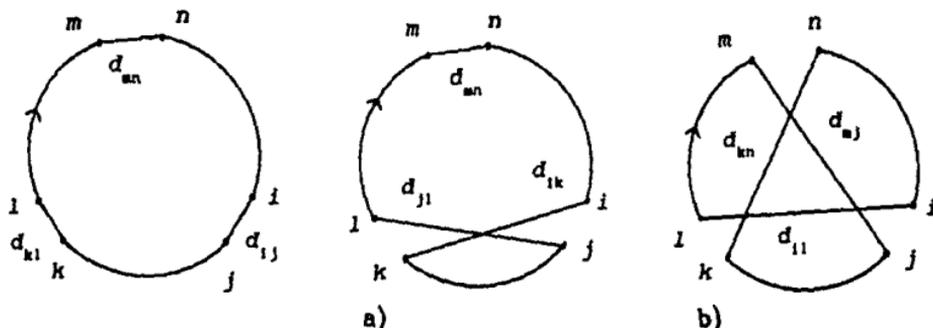


Fig. 1.5. a) Ilustración de una inversión. b) Ilustración de una inserción.

El segundo tipo de movimiento⁽⁵⁾ consiste en escoger tres segmentos d_{ij} , d_{kl} y d_{mn} al azar y una vez hecho esto insertar la porción del camino que se encuentra entre las ciudades j y k entre las ciudades m y n , procurando que esta porción que se insertó se siga recorriendo en el mismo sentido que en el que se recorría anteriormente. Esto se ilustra en la Figura 1.5b. Los tres segmentos escogidos al azar se reemplazan por los segmentos d_{nj} , d_{kn} y d_{il} , sin alterar la dirección en la que se recorren las otras partes del camino. Para poder realizar este movimiento, es necesario que el segmento d_{mn} no se encuentre en la porción del camino que está entre las ciudades j y k . A este tipo de movimientos se les conoce como inserciones.

Ahora bien, para tratar a cada muestra como un sistema termodinámico, se hace lo siguiente⁽²⁻⁴⁾. Se identifican los diferentes caminos posibles, para una muestra dada, como los diferentes estados del sistema. Se introduce un parámetro T que se identifica como la temperatura a la cuál está el sistema y se identifica la distancia de los diferentes caminos como la energía del sistema. Por último, se utiliza lo que se conoce como el algoritmo de Metrópolis⁽¹⁶⁾. Originalmente diseñado para simular numéricamente sistemas termodinámicos, este algoritmo consiste en lo siguiente: Se van efectuando los movimientos arriba mencionados o cualquier otro movimiento cuyo resultado sea un cambio de los

caminos en cuestión. Cada vez que se obtenga un nuevo camino a partir del camino anterior (comenzando todo el proceso a partir de cualquier camino escogido arbitrariamente), este nuevo camino se acepta con probabilidad 1 si su distancia es menor que la del anterior, mientras que sólo se acepta con probabilidad $\exp(-\Delta L/T)$ si su distancia es mayor, siendo ΔL la diferencia de distancia entre ambos caminos. Si se efectúan suficientes movimientos a una temperatura T y se utiliza este algoritmo, la densidad de probabilidad de que el sistema esté en cierto estado estará dada por una densidad de probabilidad canónica, en analogía con los sistemas termodinámicos en equilibrio. Los movimientos que se van efectuando son análogos a las interacciones internas y externas de cualquier sistema termodinámico, pues ocasionan que la muestra vaya adoptando diferentes estados.

La identificación de la distancia con la energía está de acuerdo con la noción de que para cualquier sistema termodinámico en equilibrio, la energía es mayor entre más alta sea la temperatura. Esto es, para cualquiera de las dos versiones se tiene que si la temperatura es muy alta se aceptarán caminos de mayor distancia con muy alta probabilidad y esto traerá como consecuencia que la distancia promedio de los diferentes caminos sea grande. De igual forma, si la temperatura es muy baja, la probabilidad de aceptar caminos de distancia mayor será casi cero y por esto la distancia promedio de los caminos, cuando el sistema está en equilibrio, será pequeña. Es precisamente por esto que se hace la identificación de la distancia con la energía.

Tanto para la versión de coordenadas fijas como para la de distancias fijas, es posible encontrar una energía libre F a partir de la cuál se pueden obtener diferentes cantidades termodinámicas, de la misma forma que con cualquier tipo de sistema termodinámico desordenado. Ahora bien, no es claro qué significan esta energía libre o las diferentes cantidades que de ella se derivan, pero lo que se quiere hacer notar es que se puede trabajar con las dos versiones como si fueran sistemas termodinámicos; es decir, se pueden encontrar diferentes cantidades de interés por los métodos habituales e inclusive se puede extraer cierta información de utilidad al analizar dichas cantidades, como se verá más adelante.

En la versión de coordenadas fijas⁽¹⁷⁾, la función de partición para una muestra I en particular está dada por

$$Z[I] = \sum_{\pi} \exp(-\beta L_{\pi}), \quad (1.13)$$

siendo β el inverso de la temperatura T , π cualquiera de las permutaciones posibles y L_{π} la distancia del camino asociado con la permutación π . La suma se realiza sobre el conjunto de todas las $(N-1)!/2$ permutaciones de las N ciudades en cuestión.

Ahora, para encontrar la energía libre, es necesario realizar el promedio sobre el desorden del logaritmo de la función de partición. Sin embargo, esto resulta ser bastante complicado no sólo para esta versión sino para cualquier sistema desordenado en general. Una de las cosas que se puede hacer es utilizar la aproximación $[\ln(Z[I])]_{av} \approx \ln([Z[I]_{av}])$. Esta aproximación es muy común⁽⁷⁻⁸⁾ en la teoría de sistemas desordenados y es válida si la temperatura es suficientemente alta. Esto es debido a que como para cualquier muestra la probabilidad de encontrarla en cualquiera de sus posibles estados es casi la misma si la temperatura es muy alta, el promedio de la función de partición va a ser casi igual a la función de partición de cualquiera de las muestras en particular. Esto trae como consecuencia que el promedio del logaritmo sea muy parecido al logaritmo asociado con una sola muestra y que dicho logaritmo, a su vez, sea muy parecido al logaritmo del promedio.

Utilizando esta aproximación se tiene que la energía libre está dada por⁽¹⁷⁾

$$F = -T \ln([Z[I]_{av}]) = -TN \ln(\langle \exp(-\beta d) \rangle), \quad (1.14)$$

siendo d la variable aleatoria correspondiente a la distancia entre pares de ciudades y $\langle f(d) \rangle$ el promedio de $f(d)$ sobre todos los pares de ciudades de una muestra en particular y sobre el ensemble de muestras.

A partir de esta energía se pueden calcular diferentes cantidades de interés. Por ejemplo⁽¹⁷⁾, el promedio sobre el desorden del promedio termodinámico de la distancia de los caminos

está dado por

$$L(\beta) = N \langle d \exp(-\beta d) \rangle / \langle \exp(-\beta d) \rangle \quad (1.15)$$

y el calor específico por

$$c(\beta) = \beta^2 N \langle d^2 \exp(-\beta d) \rangle / \langle \exp(-\beta d) \rangle - \beta^2 N (\langle d \exp(-\beta d) \rangle / \langle \exp(-\beta d) \rangle)^2. \quad (1.16)$$

Como ya se mencionó, estos resultados son válidos si la temperatura es suficientemente alta. A bajas temperaturas, los resultados numéricos encontrados para esta versión ya no concuerdan con estas ecuaciones⁽¹⁷⁾.

Pasando a la versión de distancias fijas con densidad de probabilidad dada por la ecuación (1.12), se tiene lo siguiente⁽¹⁴⁾.

Para temperaturas altas

$$\begin{aligned} [Z]_{av} &= \int \prod_{i < j} \rho(L_{ij}) dL_{ij} \sum_P \exp(-\beta \sum_i L_{P(i), P(i+1)}) = \\ &= N! (g(\beta))^N, \end{aligned} \quad (1.17)$$

siendo $g(\beta) = \int_0^\infty \rho(L) \exp(-\beta L) dL$, L_{ij} la distancia entre las ciudades i y j , $\rho(L_{ij})$ la densidad de probabilidad (1.12), P el conjunto de todas las permutaciones distintas asociadas con valores fijos de las L_{ij} y $L_{P(i), P(i+1)}$ la distancia entre las ciudades $P(i)$ y $P(i+1)$ al recorrer todas las ciudades en el orden especificado por P .

En este régimen de temperatura, la energía libre está dada por⁽¹⁴⁾

$$F = -TN \ln(N/e) - TN \ln(g(\beta)) + O(\ln(N)) \quad (1.18)$$

y la distancia promedio por⁽¹⁴⁾

$$L = -N \, d \ln(g(\beta)) / d\beta. \quad (1.19)$$

En el régimen de baja temperatura y mediante un análisis

apropiado, se puede obtener, entre otras cosas, el siguiente resultado relativo al promedio configuracional de los caminos de mínima distancia⁽¹⁴⁾:

$$(r/(r+1))a_r \leq \lim_{N \rightarrow \infty} [L_{\min}[I]]_{av} / ((r+1)N^{1-1/(r+1)}) \leq a_r, \quad (1.20)$$

con $a_r = \Gamma(1/(r+1))(r+1)!^{1/(r+1)} / (r(r+1))$. Si $r \rightarrow \infty$, se obtiene que

$$[L_{\min}[I]]_{av} / ((r+1)N^{1-1/(r+1)}) \approx 1/e; \quad (1.21)$$

es decir, en esta versión en particular es posible obtener el comportamiento asintótico de $[L_{\min}[I]]_{av}$ cuando N tiende a infinito.

Con esto queda claro que tanto la versión de coordenadas fijas como la de distancias fijas pueden ser consideradas como sistemas termodinámicos desordenados y pueden ser estudiadas usando los métodos habituales. Es más, este análisis puede llegar a dar información valiosa acerca de los caminos de mínima distancia, como ocurrió en el último caso.

Otra analogía importante entre el problema del agente viajero en cualquiera de las dos versiones y los vidrios de espín es que ambos^(7,15) exhiben frustración y aleatoriedad.

En el caso del problema del agente viajero hay aleatoriedad debido a que las coordenadas de las ciudades o las distancias entre todas ellas son aleatorias. La frustración, en ambas versiones, es debido a la competencia que existe entre el requerimiento de corto alcance de moverse a la ciudad más cercana a cada paso (para tratar de minimizar la distancia total) y el requerimiento de largo alcance de que el camino debe ser cerrado y debe de pasar por cada ciudad una sola vez.

Esta es una semejanza importante ya que muchas de las propiedades de los vidrios de espín son debidas precisamente a la existencia de estos dos ingredientes⁽⁸⁾.

Cabe mencionar que estrictamente hablando, el problema del agente viajero presenta aleatoriedad sólo si se consideran a las dos versiones en estudio como sistemas termodinámicos desordenados. En este contexto sí tiene sentido hablar acerca de

la aleatoriedad ya que las coordenadas o las distancias entre las ciudades van a ser aleatorias y sus valores cambiarán al ir de muestra en muestra. Sin embargo, si se tiene el problema de minimizar la distancia o costo total al recorrer las ciudades que conforman un mapa real, las coordenadas estarán fijas y no tiene sentido hablar acerca de aleatoriedad.

En la versión de coordenadas fijas $L(\beta)$ es una función continua de T para $T \geq 0$ ⁽¹⁷⁾. Esto indica que no hay ninguna transición de fase de primer orden. Sin embargo, el calor específico presenta un comportamiento peculiar para temperaturas entre $0.5/N^{1/2}$ y $1/N^{1/2}$, lo cual puede indicar que el sistema sufre un cambio de fase de segundo orden al bajar la temperatura. La ecuación (1.16) no presenta ningún comportamiento de este tipo, pero hay que recordar que esta ecuación sólo es válida para temperaturas altas. La forma en que se ha detectado dicho comportamiento para el calor específico es por medio de estudios numéricos.

En la versión de distancias fijas con densidad de probabilidad $(L^r e^{-L})/r! I_{(0, \infty)}(L)$, $r \rightarrow \infty$, el sistema sufre un cambio de fase⁽¹⁴⁾ muy parecido al que exhibe el modelo de energía aleatoria de los vidrios de espín⁽¹⁸⁾. Esta transición ocurre en el régimen de baja temperatura y es tal que abajo de la temperatura crítica $T_c = 1/e$ la entropía del sistema es cero; es decir, a bajas temperaturas el sistema está totalmente congelado. Como sería necesario desviarse mucho en el estudio del modelo de energía aleatoria de los vidrios de espín para poder entender en qué consisten las semejanzas entre estas transiciones de fase, no se mencionará más al respecto. El único punto que se recalca es que ciertas versiones del problema del agente viajero presentan cambios de fase a temperatura distinta de cero, en analogía con los vidrios de espín y dichas transiciones pueden ser muy similares a las de los vidrios de espín, como ocurre con esta versión.

Para el problema del agente viajero, el espacio sobre el cuál se quiere minimizar la distancia, esto es, el espacio de configuración asociado a cada muestra, es discreto y consiste de todas las permutaciones cíclicas de las ciudades, sin importar la dirección en la que se recorren los caminos asociados a ellas.

Numéricamente se encuentra que, para ambas versiones, la distancia como función de la configuración de la muestra tiene muchos mínimos locales. Algo similar ocurre con los vidrios de espín. Como se mencionó anteriormente, en el modelo SK la energía libre como función de las m_i , a baja temperatura, tiene muchos mínimos locales separados por barreras finitas (ver Figura 1.2). Esto indica que otra analogía entre el problema del agente viajero en cualquiera de sus dos versiones y los vidrios de espín es que ambos presentan una gran cantidad de estados metaestables o mínimos locales⁽⁷⁾.

Pasando a la versión de distancias fijas con densidad de probabilidad uniforme en el intervalo $(0,1)$, se ha encontrado numéricamente que el espacio de configuración a baja temperatura para cualquier muestra presenta un alto grado de ultrametricidad⁽¹⁵⁾, de manera análoga a lo que ocurre en el modelo SK de los vidrios de espín.

De igual forma que en el caso de los vidrios de espín se introduce la cantidad q_{ab} para medir el traslape entre dos estados distintos de cualquier muestra, en el caso del problema del agente viajero en cualquiera de sus versiones, q_{ab} se define como la fracción de segmentos que coinciden en dos caminos a y b correspondientes a una misma muestra.

La desigualdad ultramétrica en términos de traslapes es⁽⁷⁾

$$q_{ac} \geq \min\{q_{ab}, q_{bc}\}. \quad (1.22)$$

En un espacio ultramétrico, todos los triángulos cuyos lados están dados por q_{ab} , q_{ac} y q_{bc} , son equiláteros o isósceles con el lado de diferente longitud siendo el más grande. Así, una forma fácil de ver si el espacio de configuración, para el caso en cuestión, es ultramétrico o no, es analizar la frecuencia con que se cumple esta condición al generar un gran número de triángulos.

Si se grafica en un primer eje la longitud de los lados más pequeños, en un segundo eje la longitud de los lados intermedios y en un tercer eje la frecuencia con que se encuentran triángulos tales que la longitud de sus lados más pequeños estén dados por estos pares de valores, la superficie que se obtiene, si el espacio es ultramétrico, valdrá cero en todas partes excepto

para valores iguales de las longitudes de estos dos lados. Los resultados numéricos que se han obtenido para la versión en consideración se muestran en la Figura 1.6. De esta figura se deduce que el espacio de configuración presenta un alto grado de ultrametricidad.

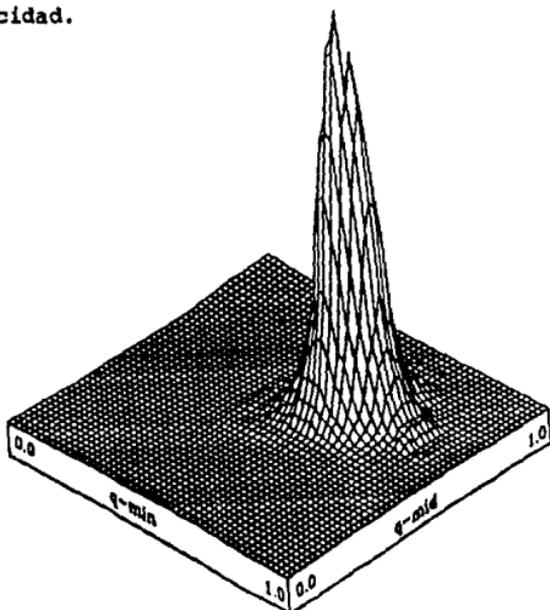


Fig. 1.6. Distribución de triángulos para la versión de distancias fijas con densidad de probabilidad uniforme.

Si N es muy grande, el número de diferentes valores de q_{ab} es enorme. Por esto, resulta interesante considerar la densidad de probabilidad $P_1(q)$ asociada con estos traslapes. De igual forma que en el caso de los vidrios de espín, estudios numéricos muestran⁽¹⁵⁾ que en esta versión $P_1(q)$ y por lo tanto $\{P_1(q)\}_{av}$ son prácticamente continuas aún con N relativamente pequeño. Esto indica, en concordancia con lo expuesto anteriormente, que para la versión en cuestión cualquier muestra tiene un gran número de estados metaestables.

Por último, cabe mencionar que es posible construir una teoría de campo medio para el problema del agente viajero, en analogía con las teorías de campo medio existentes para el caso de

los vidrios de espín. En la Ref. [19] se encuentra un estudio detallado acerca de dicha teoría.

Así, se ve que existe una relación muy fuerte entre los vidrios de espín y el problema del agente viajero en cualquiera de sus dos versiones. Es por esto que nueva información acerca de uno de estos campos puede ser muy útil en el otro, de igual forma que las técnicas utilizadas en un caso pueden ser aplicables en el otro.

CAPITULO 2

ALGUNOS METODOS PARA RESOLVER APROXIMADAMENTE EL PROBLEMA DEL AGENTE VIAJERO

Este capítulo tratará exclusivamente con el problema del agente viajero en la versión de coordenadas fijas. Como se mencionó en el Capítulo 1, cuando el número N de ciudades es muy grande resulta sumamente impráctico encontrar un camino que sea de mínima distancia. Esto es debido a que el tiempo de cómputo que se requiere para resolver exactamente el problema crece exponencialmente con N .⁽⁴⁾ El problema del agente viajero pertenece a lo que se conoce como la clase de problemas NP-completos⁽¹⁾, cuyo tiempo de cómputo tiene esta dependencia en N . Debido a esto, existen varios métodos que dan soluciones aproximadas en menor tiempo. Aquí se hablará acerca de tres, que son el método de Lin⁽⁵⁾, el método del templado simulado⁽²⁻⁴⁾ y el método jerárquico⁽⁶⁾.

2.1 METODO DE LIN

Antes de describir este método se definirá lo que son los caminos λ -óptimos⁽⁵⁾. Un camino es λ -óptimo si la distancia total es menor que la de cualquier otro camino que se pueda construir a partir de éste haciendo movimientos que involucren λ segmentos. A continuación se presentan algunas propiedades de los caminos λ -óptimos.

1. Cualquier camino es 1-óptimo.
2. Las siguientes propiedades son equivalentes:
 - a) Un camino es 2-óptimo.

- b) El camino es óptimo respecto a inversiones (de las cuales ya se habló en el capítulo anterior, junto con las inserciones).
- c) El camino no se intersecta en ningún punto.
3. Un camino es óptimo o de mínima distancia si y sólo si es N -óptimo, siendo N el número de ciudades.
 4. Si C_λ es el conjunto de todos los caminos λ -óptimos, $C_1 \supset C_2 \supset \dots \supset C_N$.
 5. Un camino es óptimo respecto a inversiones y a inserciones si y sólo si es 3-óptimo. Por camino óptimo respecto a inversiones e inserciones se entiende aquel camino tal que para cada k , el quitar cualquier porción del camino en el cual haya k ciudades y el invertirlo o insertarlo entre otras dos ciudades, da como resultado un camino cuya distancia total es mayor que la de éste.

Las propiedades 1, 3 y 4 son evidentes.

Respecto a la 2 la demostración es la siguiente. Como los únicos movimientos posibles que involucran dos segmentos son las inversiones, si un camino no es 2-óptimo, tampoco es óptimo respecto a inversiones. Ahora, si el camino no es óptimo respecto a inversiones, esto quiere decir que al realizar este tipo de movimientos se llegará, a fin de cuentas, a un camino 2-óptimo. Pero esto, a su vez, quiere decir que los pares de segmentos que intervinieron en dichos movimientos se intersectaban entre sí, ya que si esto no fuera cierto, lo único que se obtendría al hacer una inversión sería un camino de mayor distancia, lo cual indicaría que el camino original sí era óptimo respecto a inversiones. Esto se puede apreciar en la Figura 2.1.



Fig. 2.1. Diagrama de una inversión.

Así, si el camino no es óptimo respecto a inversiones esto indica que se intersecta en por lo menos un punto. Con base en la Figura 2.1, se ve que si el camino se intersecta en por lo menos un punto, es posible realizar una inversión para disminuir la distancia total y esto quiere decir que dicho camino no era 2-óptimo. Con esto se tiene que a), b) y c) son equivalentes.

La demostración de la 5 es la siguiente. Un camino no es 3-óptimo si y sólo si existen 3 segmentos d_{ij} , d_{kl} y d_{mn} tales que al cambiarlos por los segmentos d_{im} , d_{jk} y d_{nl} , por ejemplo, se obtiene un camino de menor distancia (las otras posibilidades se tratan de manera análoga). Esto se ilustra en la Figura 2.2.

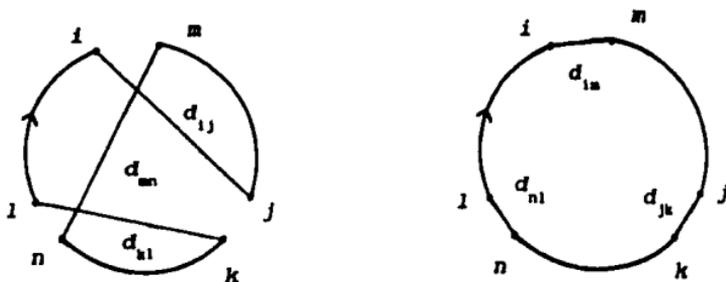


Fig. 2.2. Reemplazo de los segmentos d_{ij} , d_{kl} y d_{mn} por los segmentos d_{im} , d_{jk} y d_{nl} .

Si la porción del camino que está entre las ciudades l e i se inserta entre las ciudades n y m , se encuentra un camino de menor distancia. Esto indica que si el camino original no es 3-óptimo, tampoco es óptimo respecto a inversiones y a inserciones. Ahora, si el camino no es óptimo respecto a estos dos movimientos, es posible realizar dichos movimientos hasta que se llegue a un camino que no se pueda mejorar. Pero para llegar a este camino se están usando movimientos que involucran 2 y 3 segmentos. Esto quiere decir que el camino original no es 3-óptimo. Así, se tiene que un camino es 3-óptimo si y sólo si es óptimo respecto a inversiones y a inserciones.

Cabe mencionar que las propiedades 2, 3 y 4 traen como consecuencia que el camino de mínima distancia no se intersecta.

Con esta información, a continuación se describirá el método de Lin. Este método es un ejemplo de un grupo de métodos que se conocen como iterativos⁽²⁾. Empezando con un camino escogido arbitrariamente, se van realizando inversiones e inserciones aleatoriamente y sólo se acepta un nuevo camino si su distancia es menor que la distancia del camino anterior. Este procedimiento se lleva a cabo hasta que ya no se pueda mejorar la solución; es decir, hasta que se haya encontrado un camino 3-óptimo. Cabe mencionar que para propósitos prácticos, lo que se hace es que se repite el procedimiento hasta que ya no se pueda mejorar el camino después de haber realizado un número razonable de iteraciones.

La razón por la cual se buscan caminos 3-óptimos y no, por ejemplo, caminos 2-óptimos, es la siguiente: Numéricamente se ha encontrado⁽⁵⁾ que, en promedio, la distancia total de un camino 3-óptimo es bastante menor que la de uno 2-óptimo y que la probabilidad de que un camino 3-óptimo sea óptimo es bastante más alta que la probabilidad de que uno 2-óptimo lo sea. También, cualquier camino 3-óptimo es óptimo respecto a inversiones. Por otra parte, también se ha encontrado numéricamente que el tiempo de cómputo necesario para encontrar caminos 4-óptimos es mucho más grande que el tiempo necesario para encontrar caminos 3-óptimos y que la probabilidad de que uno 4-óptimo sea de mínima distancia no es mucho mayor que en el caso de caminos 3-óptimos. Por lo tanto, los movimientos que se realizan son las inversiones y las inserciones ya que éstos permiten encontrar caminos 3-óptimos. El tiempo de cómputo necesario para obtener un camino 3-óptimo crece con N como $30N^3$.⁽⁵⁾

Para que la probabilidad P_o de encontrar un camino óptimo al usar este procedimiento sea más alta, éste se repite varias veces empezando desde caminos distintos. Si se hace esto un número k de veces suficientemente grande, la probabilidad $1-(1-P_o)^k$ de que entre los caminos 3-óptimos obtenidos esté el de mínima distancia se puede hacer tan cercana a uno como se desee. Numéricamente se ha encontrado que $P_o = 2^{-M/10}$ si el número de ciudades está entre 30 y 60, aproximadamente. Para un mayor número de ciudades se espera que esta probabilidad sea menor. Esta relación permite calcular el número de veces que se tiene que

repetir el procedimiento si se quiere que la probabilidad de encontrar el camino de mínima distancia sea muy alta.

Conforme el número de ciudades crece, lo que se puede hacer al ir repitiendo el procedimiento es usar información de caminos 3-óptimos ya obtenidos. Por ejemplo, se puede utilizar la información de qué segmentos aparecen en todos los caminos 3-óptimos obtenidos, ya que estos segmentos tienen una alta probabilidad de pertenecer al camino de mínima distancia. De esta forma se puede ir incorporando información ya adquirida de tal forma que se puede reducir el tiempo necesario para llegar a una solución. Con esta variación, el tiempo de cómputo se reduce apreciablemente y el método de Lin da buenos resultados para problemas en los que el número de ciudades no es muy grande (abajo de 150 aproximadamente)⁽⁵⁾.

2.2 METODO DEL TEMPLADO SIMULADO

Antes de describir este método, se hablará acerca de la forma en que un metal se temple. Si el sistema se calienta de tal forma que llegue a estar a una temperatura suficientemente alta y si ésta se va bajando lentamente de manera que al llegar cerca del punto de solidificación ésta se mantenga mucho tiempo alrededor de dicho punto, el sistema, al solidificarse, estará en estados cuya energía muy baja. Esto es lo que se conoce como templado y el resultado es un sólido cristalino y no uno amorfo, que es lo que se obtiene si la temperatura se baja rápidamente. La energía del sistema es mayor en el caso amorfo que en el caso cristalino. Es importante bajar la temperatura lentamente de tal forma que el sistema siempre esté en equilibrio; es decir, es importante dar suficiente tiempo para que los átomos puedan reorganizarse de tal forma que eventualmente se pueda llegar a estados cuya energía sea muy baja.

Cuando el sistema está en equilibrio a una temperatura T , la probabilidad de encontrarlo en un cierto estado con energía E está dada por $P = A \exp(-E/k_B T)$, siendo k_B la constante de Boltzmann y A una constante de normalización. Debido a esto, el sistema podrá adoptar estados con energía relativamente grande, aunque la probabilidad de que esto ocurra es muy pequeña. Esta

probabilidad se va haciendo más chica conforme baja la temperatura, pero será diferente de cero siempre y cuando $T > 0$. Debido a esto, el sistema, al estar en un estado de cierta energía, podrá salirse de él (al tener más energía en un instante dado) y llegar eventualmente a un estado de energía menor. Es así como al bajar la temperatura lentamente de manera que el sistema siempre esté en equilibrio es posible llegar a estados de energía muy baja.

Como se mencionó en el capítulo anterior, la versión de coordenadas fijas se puede tratar como un sistema termodinámico desordenado. Con base en esto, lo que se hace en el método del templado simulado⁽²⁻⁴⁾ es ir bajando lentamente la temperatura de la muestra en cuestión para tratar de llegar a un camino (estado) de mínima distancia (energía).

Específicamente el método consiste en lo siguiente⁽²⁻⁴⁾. A partir de una configuración arbitraria del sistema y escogiendo un valor de T suficientemente alto, se van realizando diferentes tipos de movimientos de manera que la configuración de la muestra vaya cambiando; es decir, de tal forma que el camino vaya cambiando. Estos movimientos pueden ser, por ejemplo, las inversiones e inserciones mencionadas anteriormente. Se aceptan las configuraciones de distancia menor todo el tiempo mientras que sólo se aceptan las de distancia mayor con probabilidad $\exp(-\Delta L/T)$; esto es, se utiliza el algoritmo de Metrópolis. Esto se hace muchas veces de tal forma que el sistema llegue a estar en equilibrio a la temperatura T . Después, se baja un poco la temperatura a un nuevo valor y se repite todo el procedimiento hasta que el sistema vuelva a estar en equilibrio a esta nueva temperatura. Se sigue bajando la temperatura lentamente y se detiene todo el proceso cuando el sistema esté templado. Los valores que T puede tomar, así como la forma en que se debe ir bajando la temperatura, forman el itinerario del templado y muchas veces éste se tiene que ir modificando hasta que se obtenga uno que dé buenos resultados. El sistema se considera como templado cuando después de realizar un número grande de movimientos a cierta temperatura, el camino ya no se puede mejorar.

Como ilustración, en la Figura 2.3 se muestran los caminos obtenidos para una muestra dada correspondientes a diferentes

temperaturas de un proceso de templado simulado⁽⁷⁾.

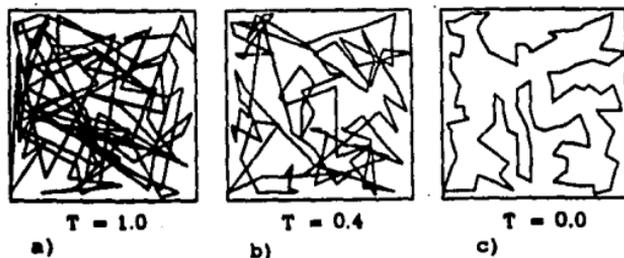


Fig. 2.3. a) $T = 1.0$, temperatura inicial. b) $T = 0.4$. c) $T = 0.0$, temperatura final.

Los métodos iterativos difieren del método del templado simulado en que^(2,4) en aquellos sólo se adoptan configuraciones de distancia menor mientras que en el del templado simulado también se adoptan configuraciones de distancia mayor. Los primeros se asemejan al caso en que no se temple el sistema sino que se enfría rápidamente. Esta diferencia tiene como consecuencia que en el método del templado simulado el sistema no se quede atrapado fácilmente en mínimos locales de altura menor que T aproximadamente, ya que siempre hay cierta probabilidad de que adopte una configuración de distancia mayor y al adoptarla, podrá salirse de dicho mínimo y llegar eventualmente a una mejor solución (ya sea el mínimo global o un mejor mínimo local).

Cabe mencionar que este método puede ser utilizado en una gran variedad de problemas. En particular, para instrumentarlo al tratar con problemas del área de optimización combinatorial⁽²⁻³⁾ se requiere lo siguiente:

1. Una descripción de las posibles configuraciones que el sistema en cuestión puede adoptar.
2. Una forma de producir y presentarle al sistema distintas opciones.
3. Una función que sea la que se quiere minimizar.
4. Un parámetro T y un itinerario del templado, el cual se formula experimentando y/o con base en otras consideraciones.

Esto, junto con el algoritmo de Metrópolis, permite utilizar este método en una gran cantidad de problemas.

En el caso del problema del agente viajero, el método del templado simulado da mejores resultados que los métodos iterativos ya que, como se mencionó arriba, el sistema puede adoptar configuraciones de mayor distancia lo cuál permite una búsqueda de mejores mínimos. El tiempo de cómputo necesario para obtener una solución crece con N como una pequeña potencia de N .⁽⁴⁾

2.3 METODO JERARQUICO

Este método se basa⁽⁶⁾ en la siguiente estrategia jerárquica. Se divide el cuadrado dentro del cuál están las ciudades en cuatro distritos, como se muestra en la Figura 2.4.

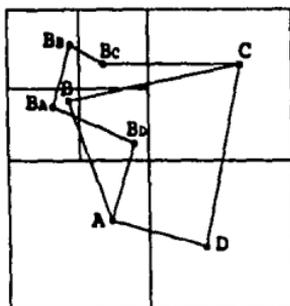


Fig. 2.4. Ilustración del método jerárquico.

Para cada distrito, se promedian las coordenadas de todas las ciudades que están dentro de él y la posición que resulte se toma como la posición del distrito en cuestión. Una vez hecho esto para los cuatro distritos, se encuentra el camino de mínima distancia para este nuevo problema, tal y como se muestra en la Figura 2.4. Después, se divide cada uno de estos distritos en 4 subdistritos y para cada subdistrito, se promedian las coordenadas de las ciudades que caigan dentro de él. Los puntos que resulten se toman como las posiciones de estos subdistritos. A continuación, se modifica en cada distrito la ruta existente de manera que se obtenga la ruta más corta que pase por las

posiciones de los cuatro subdistritos en los que está dividido dicho distrito (ver Figura 2.4). Una vez hecho ésto en cada uno de los distritos se vuelve a dividir en cuatro a cada subdistrito. Se sigue repitiendo todo el procedimiento hasta que el máximo número de ciudades que haya en cualquiera de las subdivisiones más pequeñas sea 1. De esta forma se llega a una solución del problema original.

Este método tiene mucho en común⁽⁶⁾ con las ideas del grupo de renormalización que se usa para analizar fenómenos críticos de sistemas termodinámicos.

Si el número de ciudades es muy grande, este método es mucho más rápido que el del templado simulado y da mejores resultados que él. Cabe mencionar que no se sabe de ninguna forma sistemática en que se puedan mejorar las soluciones obtenidas usando el método en cuestión, lo cual quiere decir que una vez obtenida cierta solución, no hay mucho más que se pueda hacer con ella.

También es posible dividir al cuadrado original no en 4 distritos sino en 9, ó 16, etc. Si se hace esto las soluciones mejoran un poco pero el tiempo necesario para obtenerlas aumenta bastante. El que mejoren algo es debido a que entre más finas sean las divisiones, el camino se vuelve más exacto en cada subdivisión. Por otra parte, no hay una gran mejoría debido a que la estrategia básica no se ve modificada; es decir, la forma en que se obtienen las soluciones es la misma.

CAPITULO 3

PROPIEDADES ESTADISTICAS DEL PROBLEMA DEL AGENTE VIAJERO

En este capítulo se verán algunas propiedades estadísticas que presentan las soluciones del problema del agente viajero obtenidas usando el método del templado simulado. Específicamente, se tratará el caso del problema del agente viajero en la versión de coordenadas fijas, de la cual ya se habló en los Capítulos 1 y 2. Antes de empezar, se hablará un poco acerca del programa de templado simulado que se utilizó para obtener dichas soluciones.

3.1 PROGRAMA DE TEMPLADO SIMULADO

En este programa, los movimientos que se utilizaron fueron los de Lin, esto es, se utilizaron las inversiones y las inserciones. La forma en que se escogió entre hacer una inversión o hacer una inserción, cada vez que se le presentaba una nueva configuración al sistema, fue aleatoria y la probabilidad de hacer cualquiera de estos dos movimientos fue igual a 0.5. La función que se trató de minimizar fue la distancia total $L = \sum_{i=1}^N ((x_i - x_{i+1})^2 + (y_i - y_{i+1})^2)^{1/2}$, donde el punto 1 coincide con el $N+1$ y N es el número de ciudades.

El itinerario del templado fue el siguiente. Empezando a una temperatura T de 0.5, ésta se redujo un 10% en cada paso. Se mantuvo fijo cada valor de T hasta que se le presentaran al sistema 800N opciones distintas o hasta que éste adoptara 10N nuevas configuraciones, lo que ocurriera primero. Se paró todo el proceso cuando ya no hubo ninguna mejoría apreciable para la cantidad de esfuerzo realizado; es decir, cuando el sistema ya no adquirió ninguna nueva configuración después de haberle presentado

800N opciones distintas. El programa fue tomado de la Ref. [4].

3.2 PROPIEDADES ESTADÍSTICAS DEL PROBLEMA DEL AGENTE VIAJERO

Las propiedades estadísticas consideradas en este capítulo son las mismas que las que se estudian al analizar espectros de energía utilizando la teoría de las matrices aleatorias⁽²⁰⁻²¹⁾. Aunque en el presente capítulo sólo se describirá brevemente en qué consisten estas propiedades, en el Apéndice 1 se habla con más detalle de aquellos tópicos de la teoría de las matrices aleatorias que son relevantes para el presente trabajo y en lo que sigue se utilizará la información ahí contenida para estudiar estadísticamente el problema del agente viajero.

Para utilizar la teoría de las matrices aleatorias en el caso del problema del agente viajero, lo que se hace es construir, a partir de cada muestra, un espectro aleatorio. Para ésto, se empieza de cualquier ciudad y se van marcando sobre una recta numérica las distancias de los diferentes segmentos que conforman el camino en cuestión; es decir, se marca sobre la recta la distancia de la primera ciudad a la segunda, a continuación y a partir del punto marcado, la distancia de la segunda a la tercera, y así hasta que se marque la distancia de la N -ésima a la primera. De esta forma, lo que se obtiene es un "espectro" con N niveles en el cual las distancias entre los niveles corresponden a las distancias entre las ciudades y los niveles en sí corresponden a las posiciones que tendrían las ciudades si el camino se cortara en cierta ciudad y se enderezara sobre la recta. Esto se muestra en la Figura 3.1.

La posición del nivel más alto de energía, así como el origen, corresponden a la ciudad de la cual se partió y a la cuál finalmente se llega al cerrarse el camino.

Es de esta forma como a partir de cada muestra distinta se puede construir un espectro aleatorio. Al analizar estos espectros, se pueden obtener diferentes propiedades estadísticas del problema del agente viajero. Estas propiedades son independientes de las ciudades que se escojan como origen al construir los espectros, debido a la forma "circular" que tienen

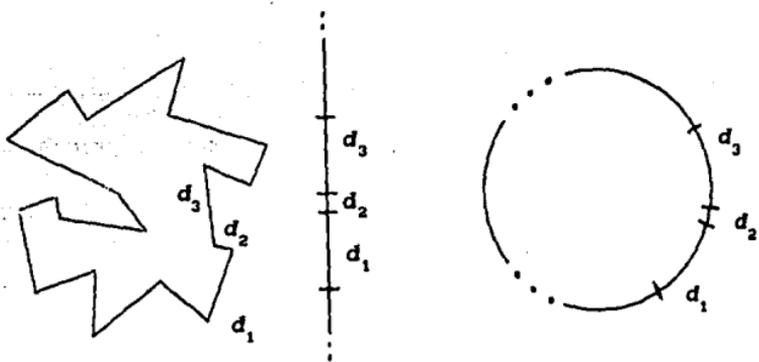


Fig. 3.1. Ejemplo de un espectro aleatorio asociado con el problema del agente viajero. Para cada muestra se construye un espectro a partir de la solución obtenida.

todos los caminos.

Cabe mencionar que todos los espectros utilizados en este trabajo se normalizaron dividiendo entre la media D de la distancia entre niveles vecinos, tal y como se menciona en el Apéndice 1 para los casos de espectros gaussianos y de tipo Poisson⁽²⁰⁻²¹⁾. Cada vez que se quiera comparar los espectros en cuestión con otros espectros aleatorios, será necesario normalizarlos de tal forma que la distancia promedio entre niveles vecinos sea igual a uno.

A partir de los espectros ya normalizados, se estudiaron las siguientes propiedades estadísticas⁽²⁰⁻²²⁾:

1. Densidades de probabilidad $p(k;s)$.

$p(k;s)$ es la densidad de probabilidad de que la distancia entre dos niveles cualesquiera con k niveles intermedios, $k = 0, 1, 2, \dots$, esté entre s y $s+ds$. A estas densidades también se les conoce como densidades de probabilidad de k -ésimos vecinos. Cuando $k \rightarrow \infty$, $p(k;s)$ tiende a una densidad de probabilidad normal con media $k+1$. Esto se debe al teorema central del límite, ya que la variable aleatoria s asociada con la densidad de probabilidad $p(k;s)$ es la suma de las $k+1$ variables aleatorias s_i correspondientes a los espaciamientos entre los $k+2$ niveles en cuestión (los dos de los extremos y

los k intermedios).

2. Desviación estándar $\sigma(k)$ de las densidades de probabilidad $p(k;s)$ como función de k .

Esta desviación está dada por $\sigma(k) = (\langle (s-\langle s \rangle)^2 \rangle)^{1/2}$, donde los promedios se realizan utilizando la densidad de probabilidad $p(k;s)$. Es una medida de la anchura de estas densidades de probabilidad, o dicho de otra forma, de la variabilidad de s alrededor de la media $\langle s \rangle = k+1$.

3. Varianza Σ^2 de la densidad de probabilidad $E(n;L)$ como función de L .

La función $E(n;L)$ da la probabilidad de encontrar n niveles dentro de un intervalo de longitud L escogido arbitrariamente. La varianza de esta densidad de probabilidad se denota como Σ^2 y está dada por $\Sigma^2(L) = \langle (n-\langle n \rangle)^2 \rangle$. Es una medida de qué tanto varían, en promedio, los valores que la variable aleatoria n toma alrededor de la media $\langle n \rangle$. Como $E(n;L)$ depende de la longitud L de los intervalos, todos sus momentos y en particular Σ^2 también dependerán de ella. A Σ^2 se le suele llamar varianza de número, ya que es la varianza del número de niveles que caen dentro de los diferentes intervalos de longitud L .

4. Coeficiente de sesgo γ_1 de la densidad de probabilidad $E(n;L)$ como función de L .

Este coeficiente está dado por $\gamma_1(L) = \mu_3/\mu_2^{3/2}$, siendo μ_r el r -ésimo momento central de $E(n;L)$ respecto a la media; es decir, $\mu_r(L) = \langle (n-\langle n \rangle)^r \rangle$. Al igual que Σ^2 , también depende de la longitud L de los intervalos ya que $E(n;L)$ depende de esta longitud. Aunque no muy confiable, es una medida de la asimetría de la densidad en consideración (ver Apéndice 1) y la forma en que depende de L va a ser característica del tipo de espectros que se estén analizando.

5. Coeficiente de exceso γ_2 de la densidad de probabilidad $E(n;L)$ como función de L .

Este coeficiente se define como $\gamma_2(L) = (\mu_4/\mu_2^2) - 3$. También

depende de la longitud L de los intervalos y es una medida de la picudez, alrededor de la media, de la densidad de probabilidad $E(n;L)$. Sin embargo, al igual que γ_1 y como se menciona en el Apéndice 1, la información que se obtiene de él no es muy confiable y su utilidad radica en que la forma en que depende de L va a ser característica del tipo de espectros del cual proviene.

6. Coeficiente de correlación ρ_{s_1, s_2} de dos espaciamentos consecutivos s_1 y s_2 .

Este coeficiente está dado por $\rho_{s_1, s_2} = (\sigma_{s_1, s_2}) / (\sigma_{s_1} \sigma_{s_2})$, siendo σ_{s_1, s_2} la covarianza de los espaciamentos s_1 y s_2 y σ_{s_1} y σ_{s_2} sus desviaciones estándar, respectivamente. Al igual que la covarianza, es una medida de la relación lineal que existe entre s_1 y s_2 , pero tiene la ventaja de que no depende de la variabilidad ni de s_1 ni de s_2 . Debido a esto, dicho coeficiente siempre cumple con la condición $-1 < \rho_{s_1, s_2} < 1$.

7. Media de la función $|F(y)|^2$.

Esta media es el promedio sobre el ensemble del módulo al cuadrado de las transformadas de Fourier $F(y)$ de los distintos espectros. Específicamente, si para cada espectro se escoge una región suficientemente grande y si se construye a partir de ésta la función $f(x) = \sum \delta_{x, x_1}$, la cual vale cero en todos lados excepto para aquellos valores de x que corresponden a los diferentes niveles de energía que caen dentro de la región en cuestión, la transformada de Fourier de $f(x)$ es precisamente la función $F(y)$. El promedio sobre el ensemble contendrá información acerca de las correlaciones entre niveles de energía muy separados entre sí, tal y como se explica en el Apéndice 1.

Estas cantidades se estudiaron para los siguientes tres casos: 350 muestras de 500 ciudades cada una, 8000 muestras de 20 ciudades cada una y 350 soluciones de una sola muestra de 500

ciudades. Los resultados obtenidos se compararon con los resultados correspondientes a espectros tipo Poisson y GOE. La elección de estos dos tipos de espectros se debió a lo siguiente. Ambos tipos son ejemplos clásicos o estándar de espectros aleatorios. Esto quiere decir, entre otras cosas, que sus diferentes propiedades estadísticas están muy bien estudiadas. También, ambos ocurren muy seguido al estudiar una gran variedad de sistemas. Los espectros Poisson, por un lado, aparecen con mucha frecuencia al contar el número de veces que ocurre cierto evento en un intervalo de tiempo (o espacio) determinado; los espectros GOE, por otro lado, aparecen al estudiar los espectros de energía de muchos átomos y núcleos pesados. Así, existe la posibilidad de que ambos espectros aparezcan al analizar otros tipos de sistemas diferentes. Aparte de esto, al comparar con estos dos tipos de espectros se está haciendo una comparación tanto con espectros correlacionados como con espectros no correlacionados. Los espectros Poisson son el ejemplo clásico de espectros en los cuales no existe ninguna correlación entre los diferentes niveles de energía mientras que los espectros GOE son un ejemplo muy común de espectros en los cuales sí existe correlación entre los distintos niveles.

Además de estudiar las propiedades estadísticas aquí presentadas para los tres casos arriba mencionados, también se analizó la forma en que estas propiedades cambiaban durante el proceso de templado simulado de las distintas muestras.

A continuación se hablará en detalle acerca de estos tres casos así como acerca de la dependencia con la temperatura de las diferentes propiedades estadísticas.

3.2.1 350 Muestras De 500 Ciudades Cada Una

A partir de 350 muestras de 500 ciudades cada una, se obtuvieron 350 espectros aleatorios. Para obtener las densidades de probabilidad $p(k;s)$ a partir de estos espectros se construyó, para cada valor de k (específicamente para $k = 0, 1, 2$ y 10), un histograma usando como datos las distancias s entre pares de niveles con k niveles intermedios. En las Figuras 3.2a, 3.2b, 3.2c y 3.2d se muestran dichos histogramas para $k = 0, 1, 2, \text{ y } 10$

respectivamente.

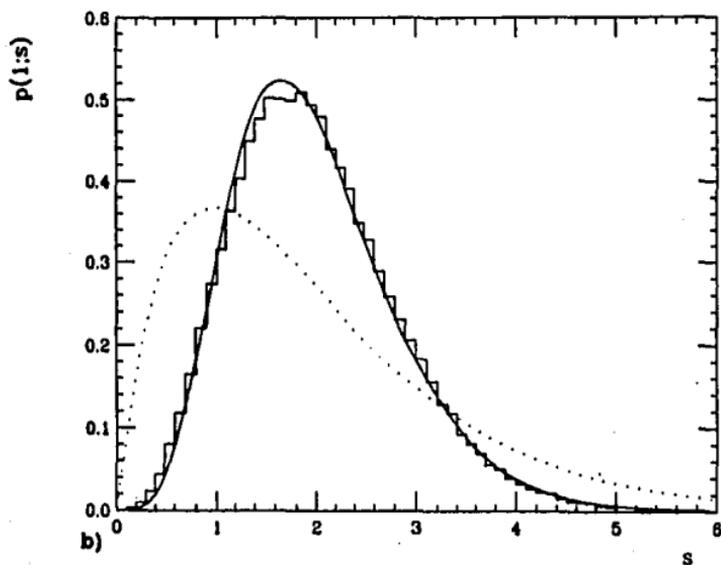
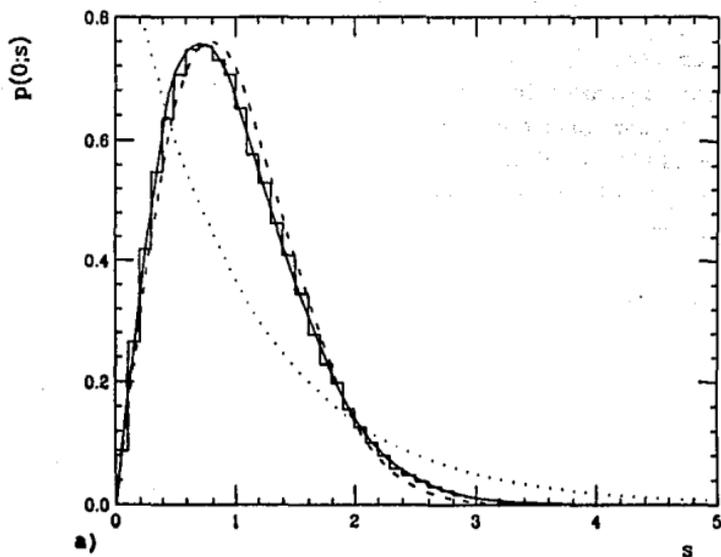
Por comparación, en la Figura 3.2a también se muestra, aparte de la densidad correspondiente al caso Poisson, la densidad de probabilidad de Wigner. Para el caso del problema del agente viajero la densidad de probabilidad de vecinos cercanos $p(0;s)$ está dada por

$$p(0;s) = as^b \exp(-cs^d) I_{(0,\infty)}(s), \quad (3.1)$$

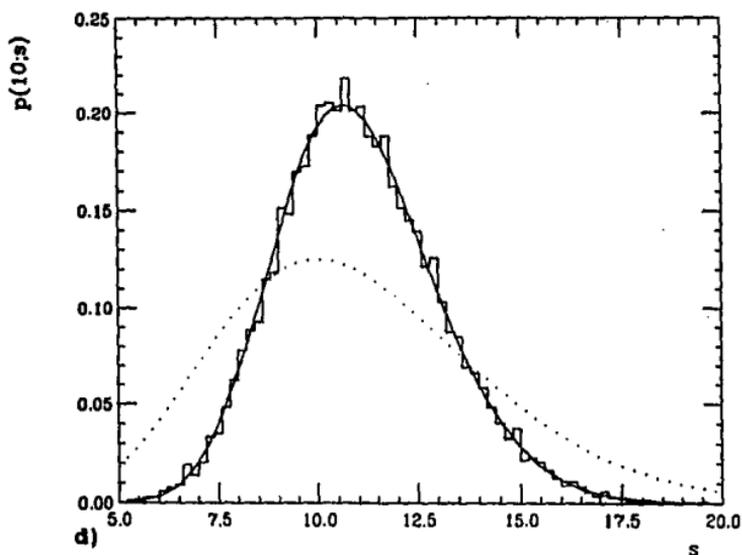
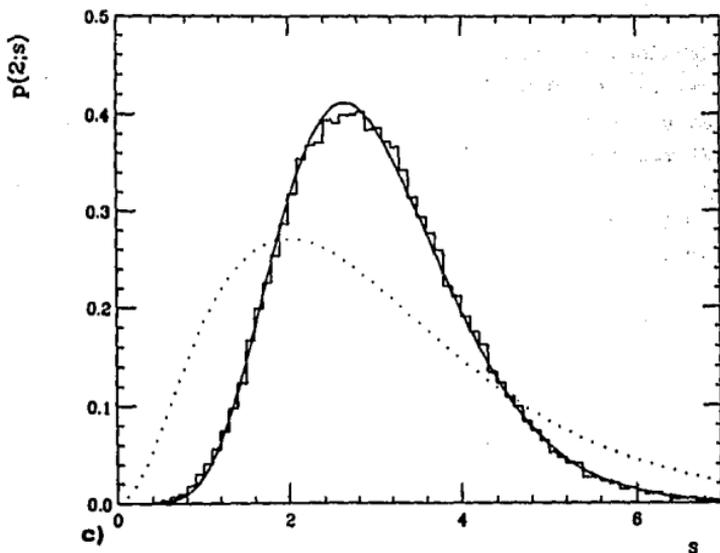
siendo $c = \Gamma((b+2)/d)^d / \Gamma((b+1)/d)^d$, $a = (dc^{(b+1)/d}) / \Gamma((b+1)/d)$, $b = 1.1657$ y $d = 1.4845$. Cabe mencionar que después de tratar con diversas funciones, ésta fué la forma funcional que mejor se ajustó al histograma y como se aprecia en la Figura 3.2a, el ajuste es muy bueno. El valor de a está determinado por la condición de normalización ($\int_0^\infty p(0;s) ds = 1$) y el valor de c por la condición de que la media $\langle s \rangle$ de la distancia entre niveles vecinos es igual a uno; esto es, la ecuación $\langle s \rangle = \int_0^\infty sp(0;s) ds = 1$ determina el valor de c . Los valores de b y d fueron encontrados usando el método de momentos⁽²³⁾. De la Figura 3.2a se ve que la probabilidad de que s sea muy chico o muy grande, respecto a $\langle s \rangle = 1$, es más grande para el problema del agente viajero que para espectros del ensemble ortogonal gaussiano. Para el caso Poisson, la probabilidad de que dos niveles estén separados por una distancia entre s y $s+ds$ disminuye conforme ésta aumenta.

De la Figura 3.2d se ve que cuando el número de niveles intermedios es grande, $p(k;s)$ se aproxima a una densidad de probabilidad normal con media $k+1$. Esto es debido al teorema central del límite (ver Apéndice 1).

Cabe mencionar que mientras que en el caso de $p(0;s)$ la curva sólida es la encontrada con base en el histograma (ecuación 3.1), en el caso de las densidades $p(k;s)$, $k = 1$, $k = 2$ y $k = 10$, las curvas sólidas son aproximaciones a las curvas reales; es decir, las curvas que se muestran son densidades de probabilidad gamma con media y varianza determinadas con base en los datos numéricos. Específicamente, lo que se hizo para cada uno de estos tres casos fué, primero, considerar que el histograma correspondía a una densidad de probabilidad gamma dada por $p(s) = (\lambda^{r+1} s^r e^{-\lambda s}) / \Gamma(r+1)$ y segundo, igualar la media y varianza



Figs. 3.2a, 3.2b. Densidad de probabilidad de k -ésimos vecinos $p(k;s)$ para espectros asociados con el problema del agente viajero (—). a) $k = 0$ y b) $k = 1$. Por comparación se muestra la densidad de probabilidad $p(0;s)$ para el caso del GOE (--) y las densidades $p(k;s)$, $k = 0, 1$, para el caso Poisson (—).



Figs. 3.2c, 3.2d. Densidad de probabilidad de k -ésimos vecinos $p(k;s)$ para espectros asociados con el problema del agente viajero (—). c) $k = 2$ y d) $k = 10$. Por comparación se muestran las densidades $p(k;s)$, $k = 2, 10$, para el caso Poisson (---).

de esta densidad de probabilidad con la media y varianza obtenida a partir de los datos numéricos. Esto dió como resultado un par de ecuaciones a partir de las cuales se obtuvieron valores para λ y r . Los valores encontrados fueron los siguientes: Para $k = 1$, $\lambda = 2.967$ y $r = 4.934$, para $k = 2$, $\lambda = 2.889$ y $r = 7.667$ y para $k = 10$, $\lambda = 2.798$ y $r = 29.780$. Es claro que estas densidades de probabilidad no son las que se obtienen a partir de la ecuación 3.1, pero son las que se usaron en lugar de las densidades de probabilidad reales ya que son menos complicadas que éstas. De las figuras se ve que estas aproximaciones son bastante buenas.

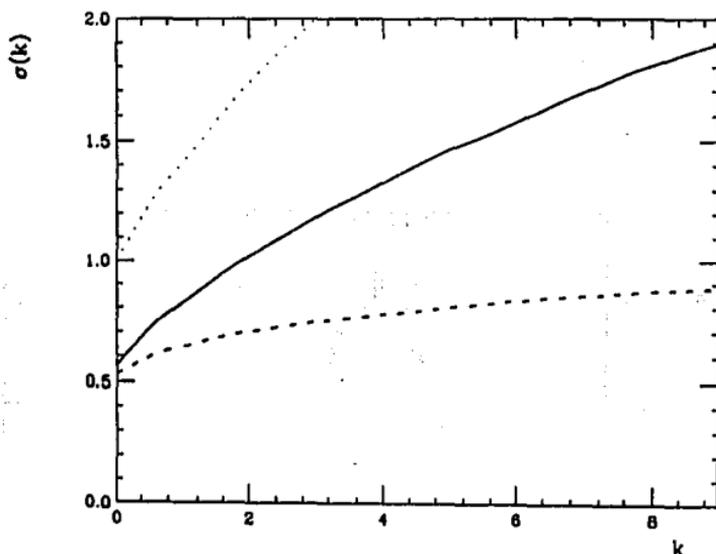


Fig. 3.3. Desviación estándar $\sigma(k)$ de las densidades de probabilidad de k -ésimos vecinos para el caso del problema del agente viajero (—). Por comparación, también se muestran las curvas correspondientes a espectros del tipo Poisson (—) y a espectros del GOE (- -).

Para calcular las desviaciones estándar de las densidades de probabilidad $p(k;s)$ a partir de los espectros en cuestión lo que se hizo fue, para cada valor de k , usar las distancias s entre pares de niveles con k niveles intermedios y usar la fórmula A1.6

del Apéndice 1. En la Figura 3.3 se muestra esta desviación estándar como función del número k de niveles intermedios. También aparecen las curvas correspondientes al caso Poisson y al caso del GOE.

De esta figura se puede ver que la desviación $\sigma(k)$ para el problema del agente viajero es mayor que la del GOE y menor que la del caso Poisson, independientemente del valor de k . Esto quiere decir que al escoger aleatoriamente un par de niveles entre los cuales hay k intermedios, la probabilidad de que la distancia s entre ellos esté alejada de $k+1$ disminuye al ir del caso Poisson al caso del problema del agente viajero, y al ir de éste al caso del GOE. Aunque en los tres casos la densidad de probabilidad $p(k;s)$ se aproxima a una normal con media $k+1$ cuando $k \rightarrow \infty$, las desviaciones estándar de estas densidades son diferentes en los tres casos, como se puede apreciar claramente en la Figura 3.3.

También se encontraron tanto la varianza como los coeficientes de sesgo y de exceso de la densidad de probabilidad $E(n;L)$ como función de L . Para esto, se dividió cada espectro en intervalos de longitud L y se contó el número n de niveles que había dentro de cada uno de estos intervalos. Después, con esta información, se encontraron los promedios adecuados con base en las fórmulas A1.7, A1.10, A1.11 y A1.12 del Apéndice 1. Esto se hizo para diferentes valores de L . Los resultados obtenidos, así como los resultados correspondientes a espectros del tipo Poisson y a espectros del GOE se muestran en las Figuras 3.4, 3.5 y 3.6.

En todas estas figuras se ve que la curva correspondiente al caso del problema del agente viajero cae entre las correspondientes a espectros Poisson y las correspondientes a espectros del GOE para valores de L mayores que o iguales a 1, aproximadamente.

Para la varianza $\Sigma^2(L)$, esto indica que, en promedio, el intervalo de valores que la variable aleatoria n toma alrededor de $\langle n \rangle$ es más grande en el caso del problema del agente viajero que en el caso del GOE, pero más chico que para un proceso de Poisson. Dicho de otra forma, al ir de intervalo en intervalo, las fluctuaciones alrededor de $\langle n \rangle$ van a ser más grandes en el caso de un proceso de Poisson y van a ir decreciendo al pasar a espectros del problema del agente viajero y finalmente a espectros del GOE.

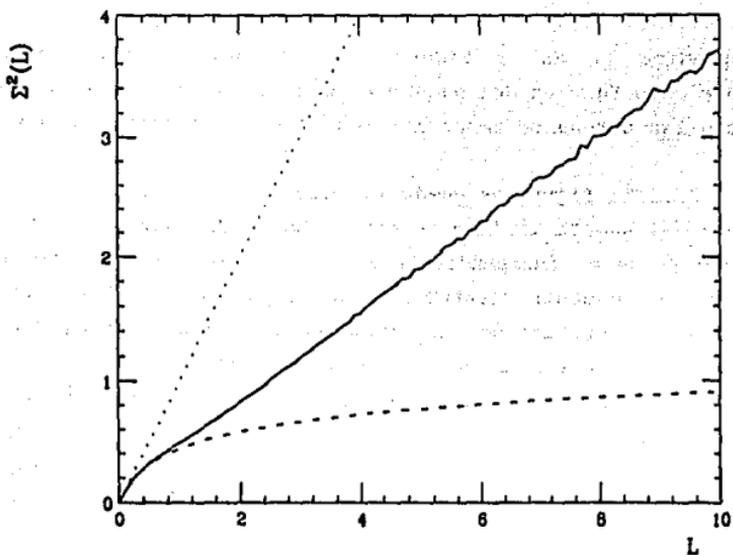


Fig. 3.4. Varianza de número Σ^2 como función de la longitud L de los intervalos para el problema del agente viajero (—) y para los casos GOE (- -) y Poisson (—·—).

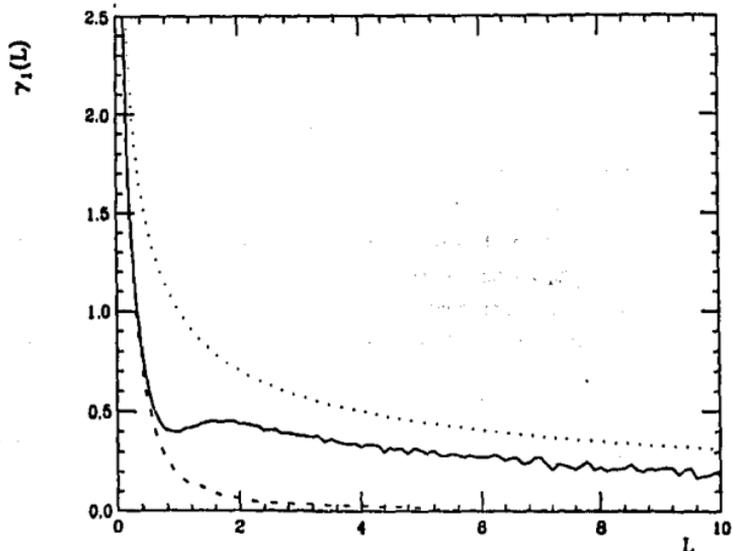


Fig. 3.5. Coeficiente de sesgo γ_1 como función de la longitud L de los intervalos para espectros del problema del agente viajero (—), del caso Poisson (—·—) y del GOE (- -).

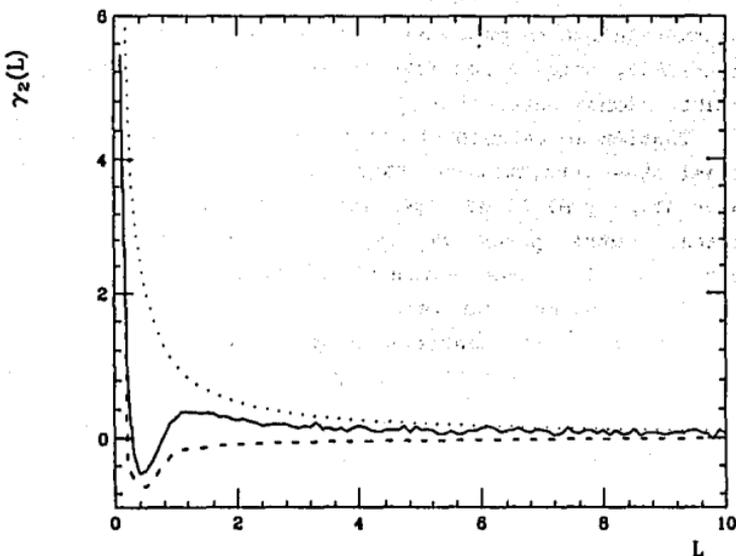


Fig. 3.6. Coeficiente de exceso γ_2 como función de la longitud L de los intervalos para espectros del problema del agente viajero (—), del GOE (- -) y del caso Poisson (—).

Para valores pequeños de L las tres curvas son muy parecidas. Esto indica que, localmente, la variabilidad del número de niveles es muy parecida en los tres casos, lo cual es de esperarse ya que localmente es muy raro que el número de niveles sea diferente de 0 ó de 1. Cabe notar que al igual que la curva de Poisson, la correspondiente al caso del problema del agente viajero crece linealmente con L para valores mayores que 1 (aproximadamente), aunque con diferente pendiente.

Cabe mencionar que entre más grande sea la variabilidad del número de niveles dentro de los diferentes intervalos, más grande serán las fluctuaciones alrededor de $k+1$ de la distancia entre pares de niveles con k niveles intermedios; es decir, estas dos cantidades están relacionadas entre sí⁽²¹⁾ y esto explica la semejanza entre las Figuras 3.3 y 3.4.

En el caso de $\gamma_1(L)$ y de $\gamma_2(L)$ se tiene que las tres curvas coinciden si $L \rightarrow 0$ ó si $L \rightarrow \infty$. Sin embargo, la forma de las tres curvas es diferente para valores intermedios de L . Respecto a la asimetría y picudez de las densidades de probabilidad $E(n;L)$,

la información que se puede obtener de las Figuras 3.5 y 3.6 no es muy confiable, como se menciona en el Apéndice 1 y por esto no se discutirán dichas características.

También se calculó el coeficiente de correlación entre dos espaciamientos consecutivos. Para hacer esto, se utilizaron las fórmulas A1.13 y A1.14 del Apéndice 1, formando el cociente de la covarianza entre pares de espaciamientos consecutivos y el producto de las desviaciones estándar asociadas con los espaciamientos pares e impares.

Para espectros asociados al problema del agente viajero se obtiene que $\rho_{s_1, s_2} = 0.0547$. Esto podría indicar que la longitud de un espaciamiento casi no influye en la longitud del espaciamiento que está arriba de él, esto es, se podría pensar que las variables aleatorias s_1 y s_2 son prácticamente independientes. Esto a pesar del hecho de que sus valores corresponden a segmentos del camino en cuestión, el cuál sí presenta cierto tipo de correlación debido a que debe de ser cerrado. Sin embargo, que el coeficiente de correlación de dos variables aleatorias sea igual a cero no necesariamente implica que éstas son independientes. Dos variables aleatorias pueden estar relacionadas entre sí y su coeficiente de correlación puede ser igual a cero⁽²¹⁾. Debido a esto, la evidencia en favor de la independencia de s_1 y s_2 no es conclusiva.

No hay contradicción entre el hecho de que las curvas de $\Sigma^2(L)$ para los casos del GOE y del problema del agente viajero son muy parecidas entre sí para valores de L relativamente pequeños y el hecho de que el coeficiente de correlación es igual -0.27 en el primer caso y casi 0 en el segundo. Esto es cierto independientemente de que para el caso en estudio s_1 y s_2 estén correlacionadas, ya que el comportamiento de la curva $\Sigma^2(L)$ no está totalmente determinado por las correlaciones de corto alcance.

En el caso de la función $\langle |F(y)|^2 \rangle$ lo que se hizo fue promediar el módulo al cuadrado de las transformadas de Fourier de los diferentes espectros. El resultado se muestra en la Figura 3.7.

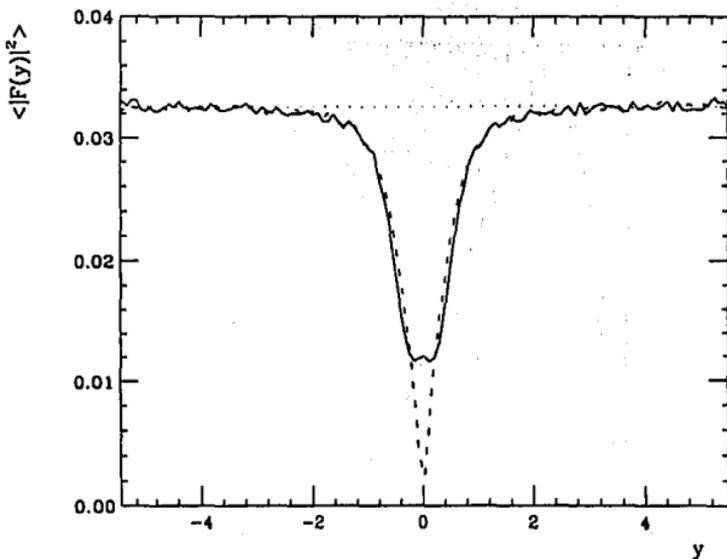


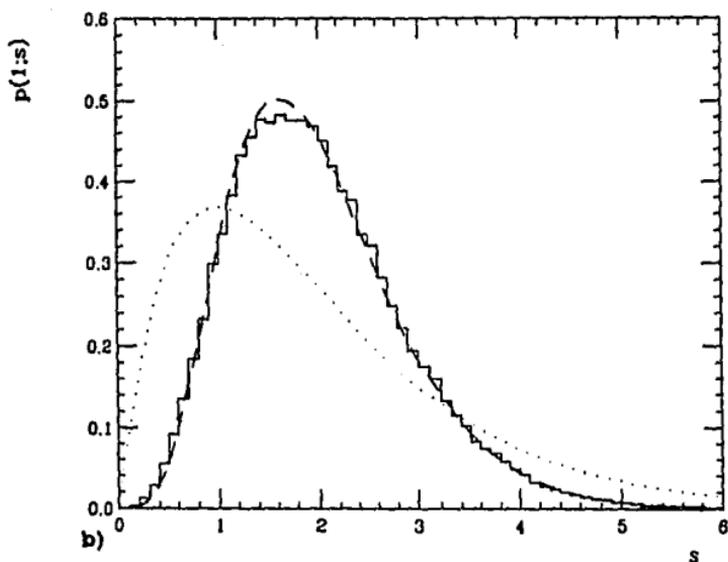
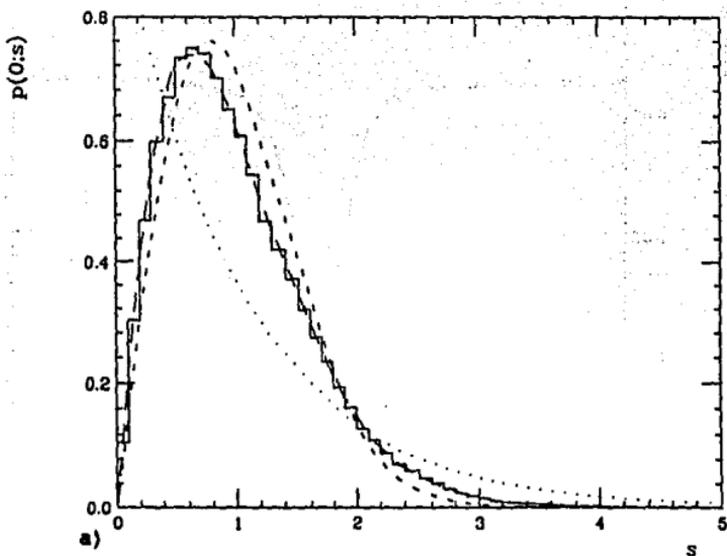
Fig. 3.7. Media del módulo al cuadrado de las transformadas de Fourier $F(y)$ de espectros correspondientes al problema del agente viajero (—). Por comparación también se muestran las curvas correspondientes al caso del GOE (- -) y al caso Poisson (- · -).

De esta figura se ve claramente que para el problema del agente viajero sí existen correlaciones de largo alcance, aunque son distintas a las existentes en el caso del GOE. Esta correlación es entendible debido a que la condición de que el camino debe de ser cerrado impone correlaciones entre regiones del espectro separadas entre sí. Como se mencionó en el Capítulo 1, esta condición, entre otras causas, es responsable del hecho de que el problema del agente viajero exhiba frustración.

3.2.2. 8000 Muestras De 20 Ciudades Cada Una

Ahora, para 8000 muestras de 20 ciudades cada una, se volvieron a calcular todas estas cantidades.

En la Figura 3.8 se muestran las densidades de probabilidad de k -ésimos vecinos.



Figs. 3.8a, 3.8b. Densidades de probabilidad de k -ésimos vecinos $p(k;s)$ para el problema del agente viajero. a) $k = 0$ y b) $k = 1$. También se muestra la densidad de probabilidad de primeros vecinos $p(0;s)$ para el GOE (- -) así como las densidades $p(0;s)$ y $p(1;s)$ para el caso Poisson (—).

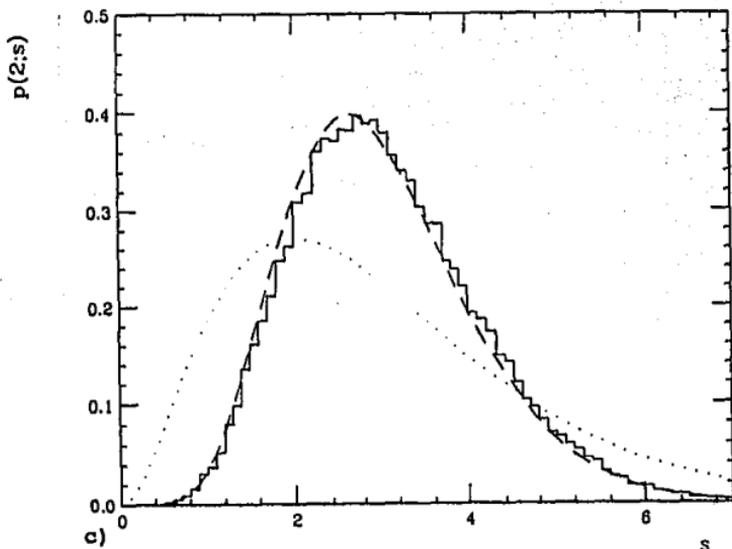
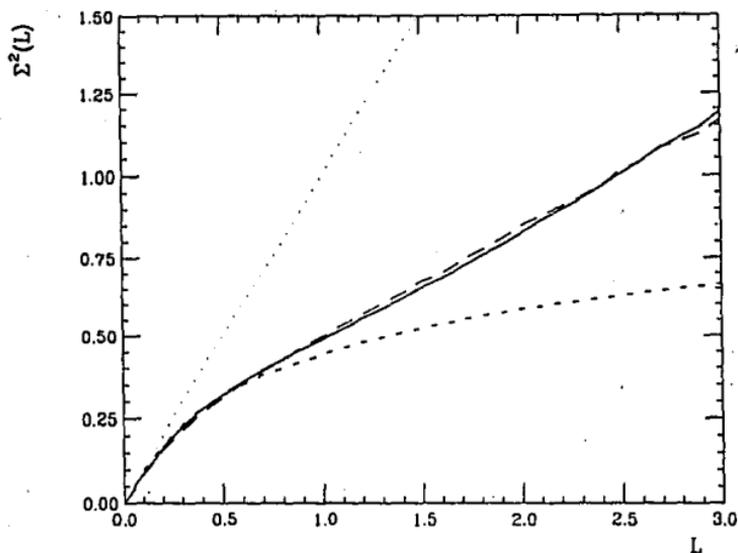
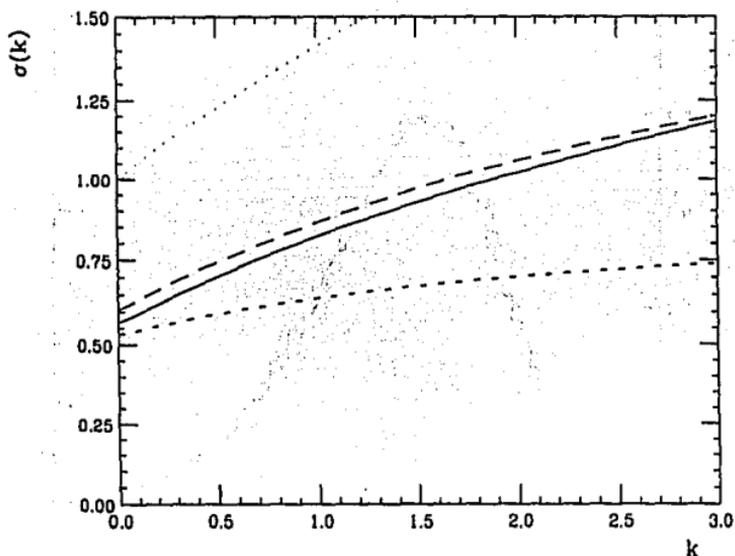


Fig. 3.8c. Densidad de probabilidad $p(2;s)$ para el problema del agente viajero (—). También se muestra esta densidad para el caso de espectros tipo Poisson (—).

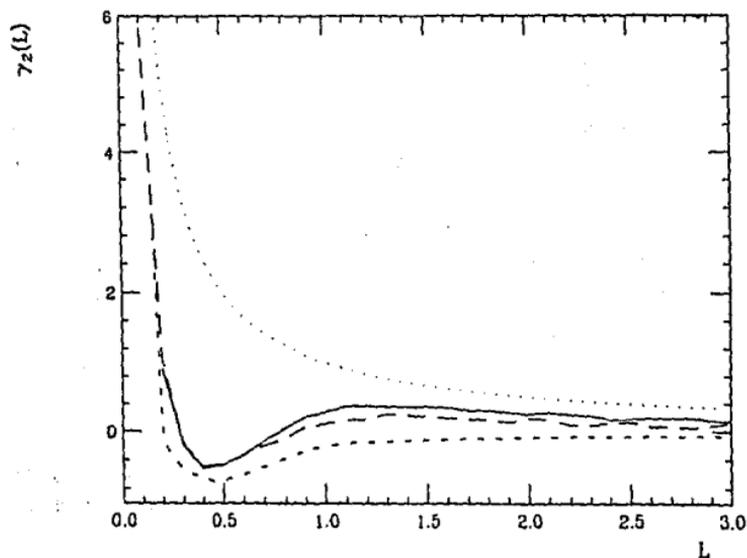
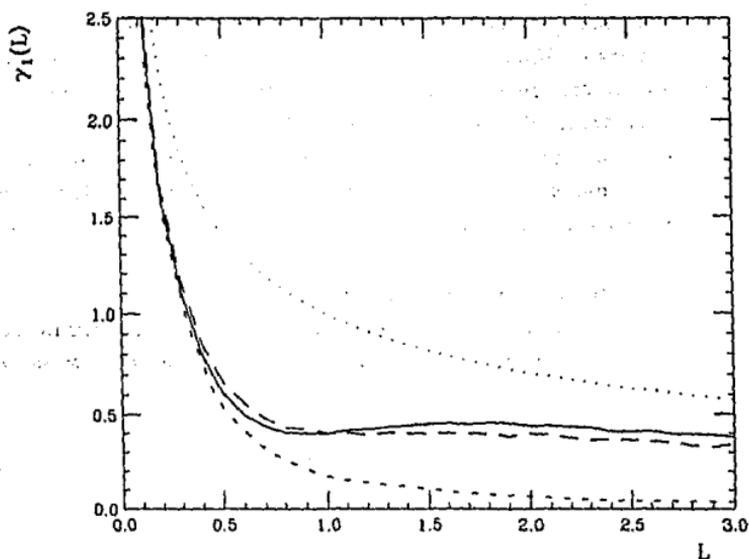
En este caso, la forma de la función $p(0;s)$ está dada, otra vez, por la ecuación 3.1, pero ahora los valores de b y d son 1.0671 y 1.3418 respectivamente. De las Figuras 3.2 y 3.8 se ve que las densidades de probabilidad $p(k;s)$, $k = 0$, $k = 1$ y $k = 2$, correspondientes a muestras de 500 ciudades y a muestras de 20 ciudades respectivamente coinciden bastante bien.

Al igual que para el caso de muestras con 500 ciudades, la curva correspondiente a $p(0;s)$ fue obtenida con base en el histograma, mientras que las curvas $p(1;s)$ y $p(2;s)$ son aproximaciones a las curvas reales. Otra vez, las curvas utilizadas son densidades de probabilidad gamma y la razón por la que se usaron es que son menos complicadas que las densidades de probabilidad reales. En este caso, $\lambda_1 = 2.682$, $r_1 = 4.364$, $\lambda_2 = 2.683$ y $r_2 = 7.050$.

Respecto a $\sigma(k)$, $\Sigma^2(L)$, $\gamma_1(L)$ y $\gamma_2(L)$, las curvas obtenidas se encuentran en las Figuras 3.9, 3.10, 3.11 y 3.12 respectivamente, junto con las correspondientes a las de los casos GOE y Poisson. Por comparación, también se muestran las curvas



Figs. 3.9 y 3.10. Desviación estándar $\sigma(k)$ de las densidades de probabilidad de k -ésimos vecinos y varianza de número Σ^2 como función de la longitud L de los intervalos, respectivamente, para los casos GOE (- -), Poisson (—) y problema del agente viajero: 8000 de 20 (— —) y 350 de 500 (—).



Figs. 3.11 y 3.12. Coeficiente de sesgo γ_1 y coeficiente de exceso γ_2 , respectivamente, como función de la longitud L de los intervalos. GOE (— · —), Poisson (····) y problema del agente viajero: 8000 de 20 (—) y 350 de 500 (— —).

obtenidas con base en las 350 muestras de 500 ciudades.

En todas ellas se ve que el comportamiento en el caso de 20 ciudades es bastante similar al comportamiento correspondiente a las 350 muestras de 500 ciudades. Sin embargo y aunque en las figuras no se muestra, para valores de k y L más grandes que 3 empiezan a haber fuertes desviaciones entre las curvas correspondientes a estos dos casos. Estas desviaciones se deben a lo siguiente.

Los espectros correspondientes a muestras de 20 ciudades son mucho más cortos que los correspondientes a muestras de 500 ciudades. Debido a esto, para valores grandes de L y k se están considerando, en el primer caso, regiones del orden de la mitad de los espectros, mientras que en el segundo caso, se están considerando regiones apenas del orden de $1/50$ de la longitud total. Debido a esto, las propiedades estadísticas en el caso de las muestras de 20 ciudades se van a ver alteradas para estos valores de k y L por diferentes factores (como por ejemplo, las correlaciones de largo alcance) que tiene que ver con el hecho de que ya no se están considerando regiones locales de los caminos.

En lo que se refiere al coeficiente de correlación, para el caso de las 8000 muestras éste resulta ser igual a 0.0219; esto es, dicho coeficiente resulta ser del mismo orden que el asociado con las muestras de 500 ciudades, pero como ya se mencionó con anterioridad, esto no significa que s_1 y s_2 sean independientes.

Por último, cabe mencionar que también en el caso de 20 ciudades existen correlaciones de largo alcance, como lo indican las transformadas de Fourier. Esto es debido, otra vez, a la condición de que el camino debe de ser cerrado.

Con esto se ve que existe un gran intervalo, en lo que al número de ciudades se refiere, tal que si las muestras caen dentro de él, entonces las propiedades estadísticas locales de las soluciones del templado simulado serán bastante parecidas independientemente del número de ciudades. Además, todas estas soluciones presentan correlaciones de largo alcance debido a la condición de que los caminos deben de ser cerrados.

3.2.3 350 Soluciones De Una Sola Muestra De 500 Ciudades

Los resultados que a continuación se presentan son los correspondientes a los espectros obtenidos al aplicar el método del muestreo simulado a una sola muestra, empezando cada vez de configuraciones diferentes.

En la Figura 3.13 se muestra el histograma correspondiente a la densidad de probabilidad $p(0;s)$. Por comparación, también se muestra la densidad de probabilidad 3.1.

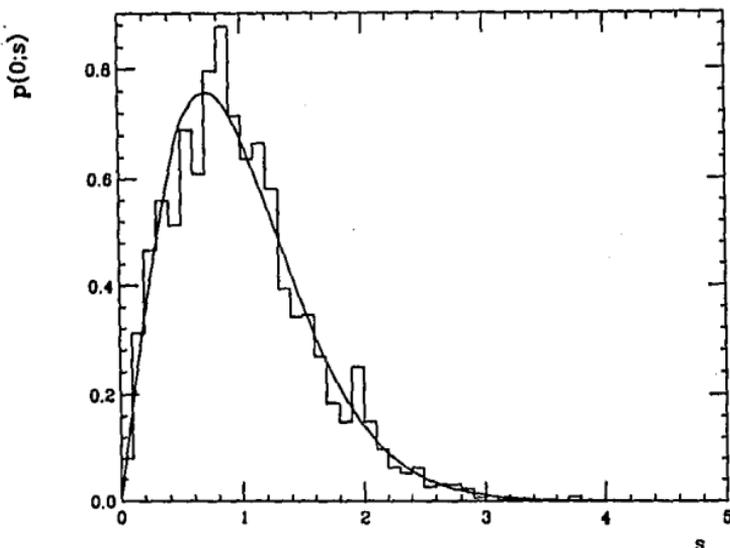
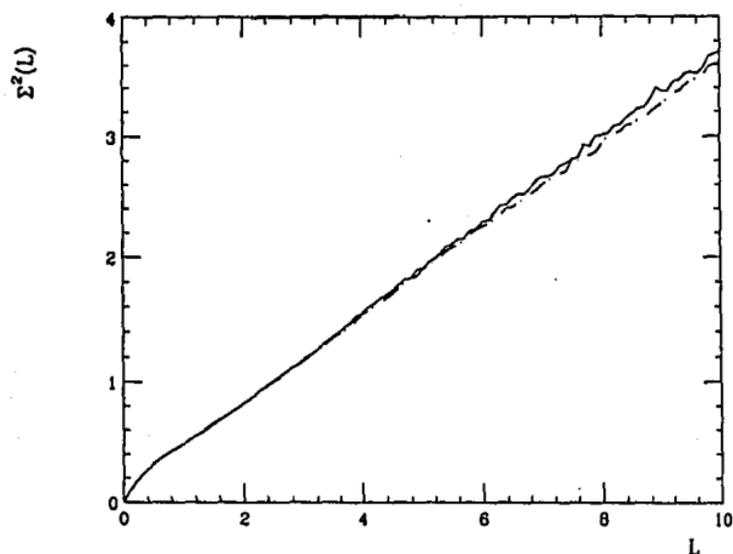
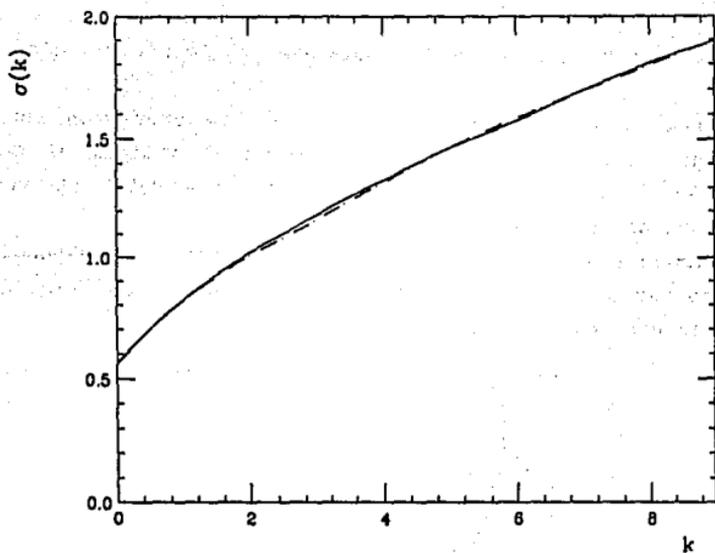
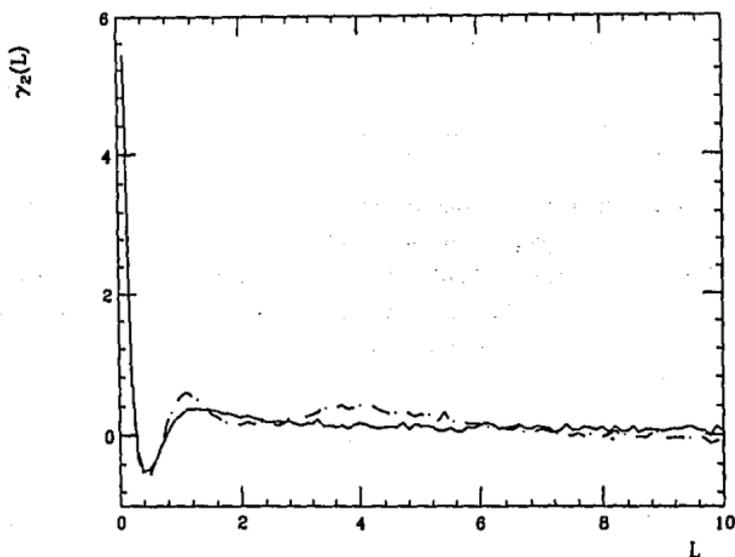
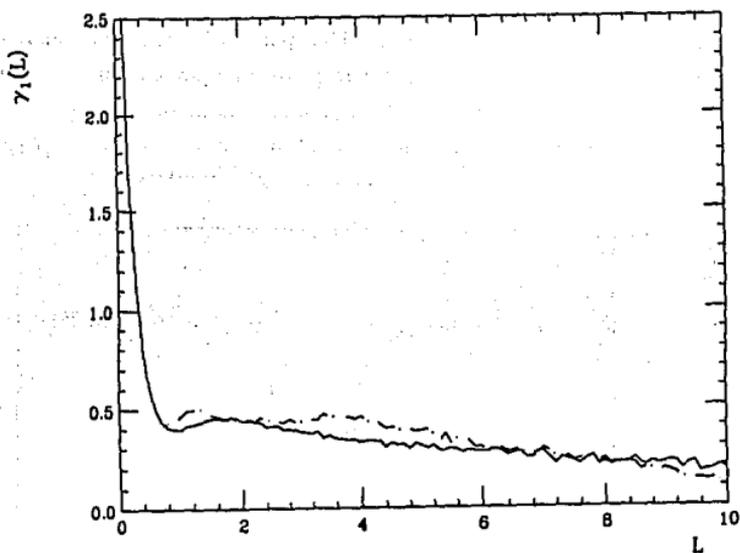


Fig. 3.13. Histograma de la densidad de probabilidad de primeros vecinos para 350 soluciones de una sola muestra de 500 ciudades. La curva que se muestra es la correspondiente a la densidad de probabilidad de primeros vecinos en el caso de 350 muestras distintas de 500 ciudades cada una.

En las Figuras 3.14, 3.15, 3.16 y 3.17 aparecen, respectivamente, la desviación estándar $\sigma(k)$, la varianza $\Sigma^2(L)$, el coeficiente de sesgo $\gamma_1(L)$ y el coeficiente de exceso $\gamma_2(L)$ tanto para el caso de 350 muestras distintas como para el caso de una sola muestra.



Figs. 3.14 y 3.15. Desviación estándar $\sigma(k)$ de las densidades de probabilidad de k -ésimos vecinos y varianza de número Σ^2 como función de la longitud L de los intervalos para el caso del problema del agente viajero: 350 muestras distintas de 500 ciudades cada una (—) y 350 soluciones de una sola muestra (— · —).



Figs. 3.16 y 3.17. Coeficiente de sesgo γ_1 y coeficiente de exceso γ_2 , respectivamente, como función de la longitud L de los intervalos para 350 muestras distintas de 500 ciudades cada una (—) y para 350 soluciones de una sola muestra también de 500 ciudades (— · —).

El coeficiente de correlación para el caso de una sola muestra es igual a 0.0678. Es del mismo orden que el que se obtuvo para 350 muestras distintas de 500 ciudades cada una.

En la Figura 3.18 se muestra la función $\langle |F(y)|^2 \rangle$ correspondiente a las soluciones de una sola muestra.

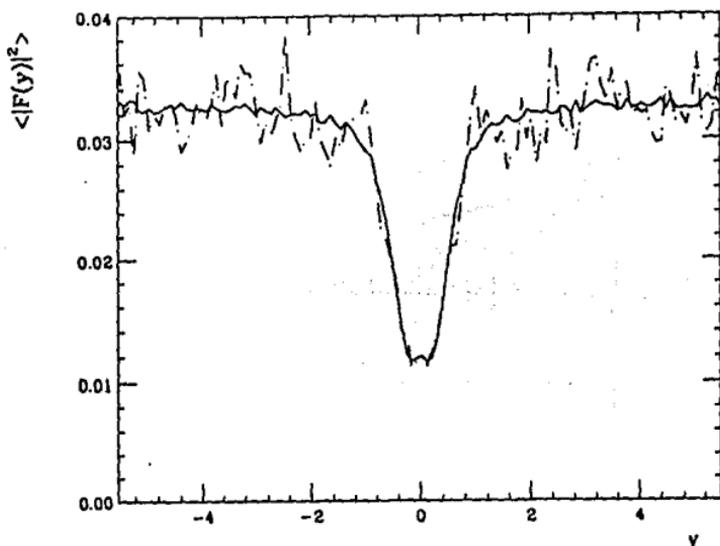


Fig. 3.18. Media del módulo al cuadrado de la transformada de Fourier $F(y)$ de espectros asociados con el problema del agente viajero. La curva sólida (—) corresponde a 350 muestras distintas de 500 ciudades mientras que la otra (— · —) corresponde a 350 soluciones de una sola muestra también de 500 ciudades.

Respecto a la Figura 3.13 se tiene lo siguiente. Al obtener 350 soluciones de una sola muestra usando el método del templado simulado, se están obteniendo un gran número de caminos distintos. Estos caminos serán muy parecidos en lo que se refiere a su longitud, pero el hecho de que se empiece cada vez desde configuraciones diferentes trae como consecuencia la existencia de un gran número de caminos finales (aunque esto no quiere decir que nunca sea posible obtener dos caminos iguales empezando desde diferentes condiciones iniciales). Con base en esto se podría

pensar que el histograma debería ser muy parecido a la curva correspondiente a 350 muestras distintas. Sin embargo, esto no es así. La diferencia existente es debida a lo siguiente. Para una sola muestra, existe una serie de segmentos que van a aparecer en muchas de las 350 soluciones. Estos segmentos tienen una muy alta probabilidad de pertenecer al camino óptimo o de mínima distancia. En otras palabras, existen ciertos valores de la variable aleatoria s que se repiten con mucha frecuencia. Debido a esto, las frecuencias relativas de esta variable y por lo tanto el histograma correspondiente a la densidad de primeros vecinos se van a ver alterados y esta es precisamente la causa de las discrepancias entre la curva y el histograma.

De las Figuras 3.14, 3.15, 3.16, y 13.7 se ve que las curvas correspondientes a una sola muestra son muy parecidas a las de 350 muestras distintas. Esto es debido a la gran diversidad de soluciones para el caso de una sola muestra. Aunque haya varios segmentos que se repitan con alta frecuencia, el número de segmentos que no o casi no se repiten es grande, de forma que al realizar los diferentes promedios los valores que se obtengan no se van a ver alterados por la alta repetición de algunos de los segmentos. Esta repetición sólo tiene como consecuencia que el número "efectivo" de datos a partir de los cuales se obtienen dichos promedios disminuya. Esto, sin embargo, no es el caso con el histograma, ya que lo que se está viendo ahí es precisamente la frecuencia relativa de s , de manera que ahí sí importa esta repetición.

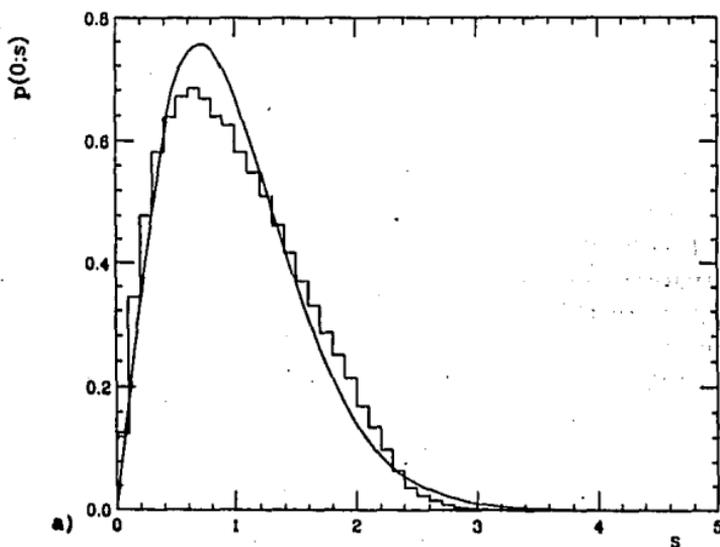
Esto mismo es lo que pasa con el coeficiente de correlación y por eso ambos son del mismo orden.

De la Figura 3.18 se ve que ambas curvas tienen la misma forma, aunque las fluctuaciones son mucho más grandes para el caso de una sola muestra. Esta similitud es debido, otra vez, a que para el caso de una sola muestra existe una gran diversidad de caminos finales. Las fluctuaciones grandes para valores de y alejados del origen (valores cercanos a cero en los espectros originales) tienen que ver con el hecho de que el número efectivo de datos a partir de los cuales se está encontrando el promedio de las transformadas para estos valores de y se ve disminuído por la repetición de ciertos segmentos.

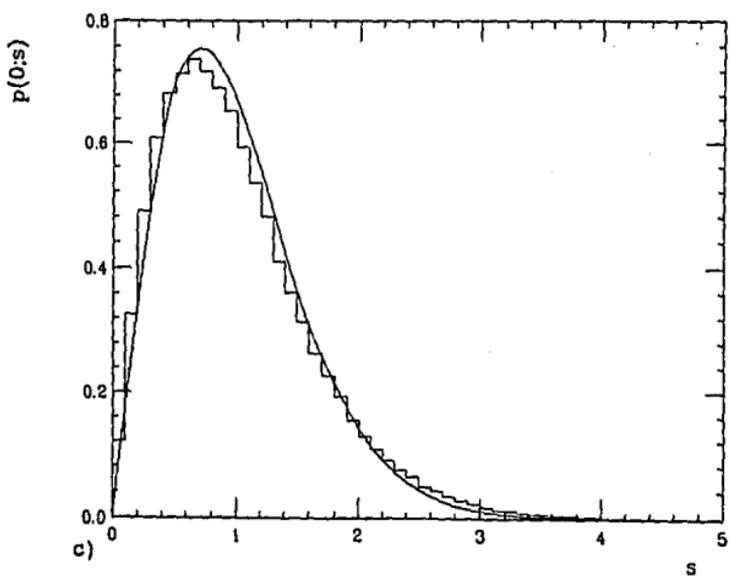
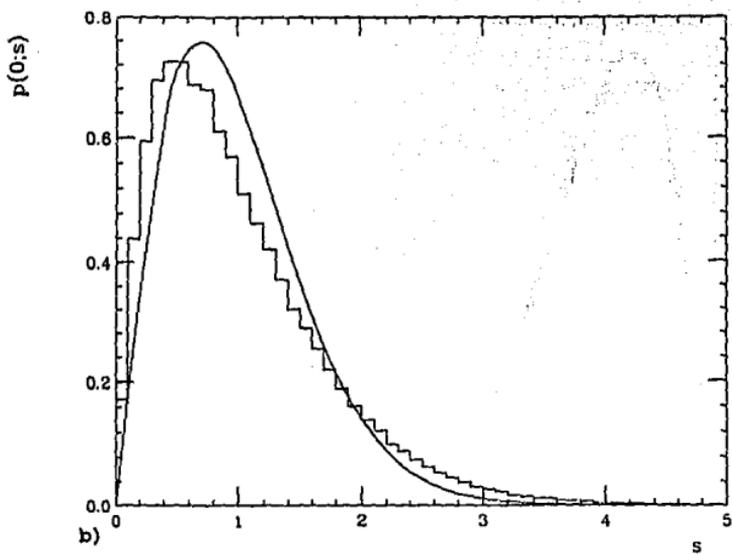
De esta forma, lo que se concluye con base en estos resultados es que el conjunto de soluciones obtenidas con el método del templado simulado para el caso de una sólo muestra es prácticamente equivalente, en lo que respecta a ciertas propiedades estadísticas, al ensemble de soluciones correspondientes a diferentes muestras. Las discrepancias se deben a la alta repetición de algunos de los segmentos. Esta similitud indica, como es sabido, que el espacio de configuración de cualquier muestra tiene un gran número de mínimos locales.

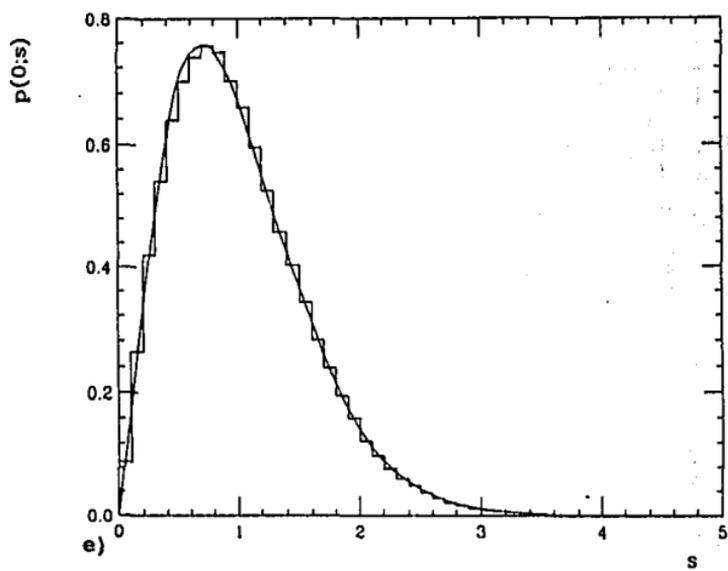
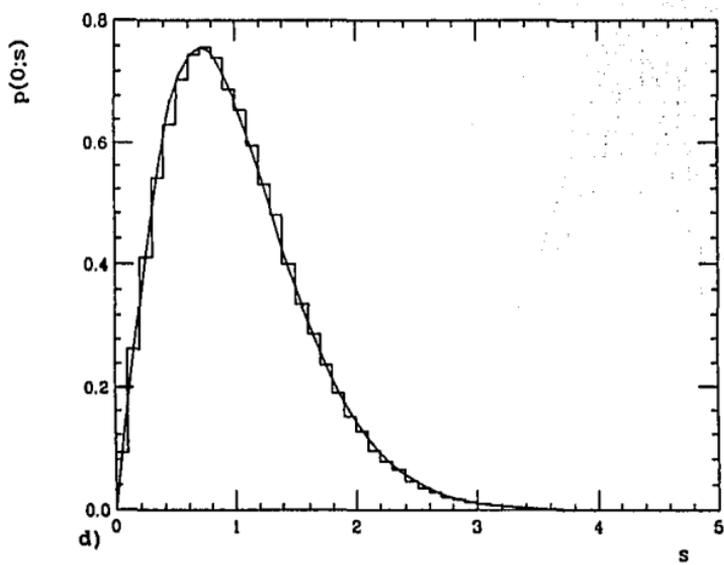
3.2.4 Dependencia Con La Temperatura

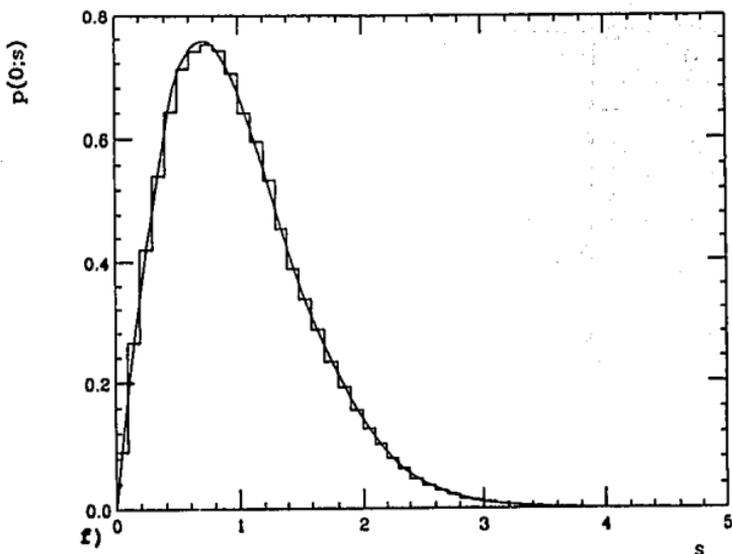
En esta sección se muestran las diferentes cantidades estadísticas hasta ahora consideradas para diferentes etapas del proceso de templado simulado. Específicamente, se muestran dichas cantidades para 350 muestras de 500 ciudades cada una.



Figs. 3.19. Histogramas de la densidad de probabilidad de primeros vecinos. Cada gráfica corresponde a cierta temperatura del proceso de templado simulado: a) $T_1 = T_1 = 0.5$, b) $T_2 = 0.125$, c) $T_3 = 0.0139$, d) $T_4 = 0.0080$, e) $T_5 = 0.0039$ y f) $T_6 = 0.0006$. La curva que aparece en todas las figuras es la densidad de primeros vecinos a la temperatura final $T_f = 0.00001$.







En las Figuras 3.19 se muestra el histograma correspondiente a la densidad de probabilidad de primeros vecinos para las siguientes temperaturas: a) $T_1 = T_1 = 0.5$, b) $T_2 = 0.125$, c) $T_3 = 0.0139$, d) $T_4 = 0.0080$, e) $T_5 = 0.0039$ y f) $T_6 = 0.0006$, siendo T_1 la temperatura inicial. La curva sólida corresponde a la densidad de primeros vecinos a la temperatura final $T_f = T_7 = 0.00001$.

De estas figuras se ve que la forma en que la densidad de primeros vecinos se aproxima a su forma templada presenta cierta oscilación; es decir, al ir bajando la temperatura, primero se carga hacia la izquierda y después se empieza a mover hacia la derecha hasta adquirir su forma final.

Estas figuras también indican que la densidad de primeros vecinos es sensible a la longitud de los diferentes caminos, ya que entre más alta sea la temperatura, más grande será la distancia promedio L de estos, tal y como se ilustra en la Figura 3.20.

En las Figuras 3.21, 3.22, 3.23 y 3.24 se muestran $\sigma(2)$, $\Sigma^2(2.0)$, $\gamma_1(2.0)$ y $\gamma_2(1.5)$ como función de la temperatura. La razón por la cual no se muestran las curvas completas para cada una de estas temperaturas es porque en muchos casos éstas

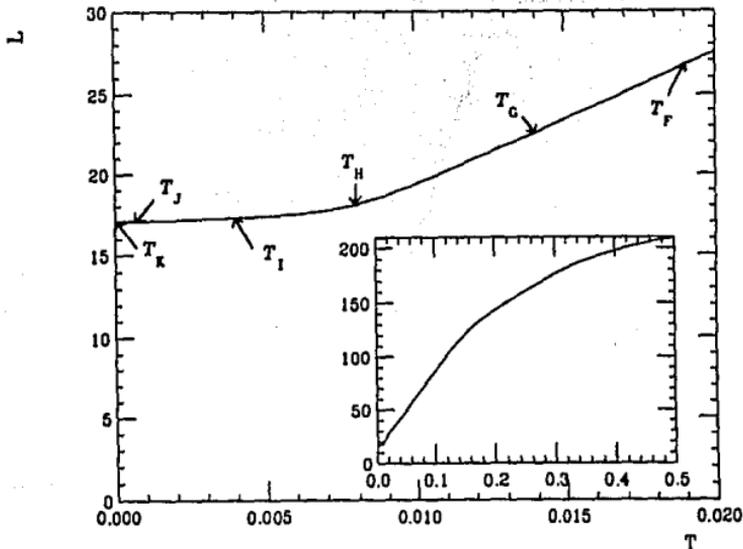
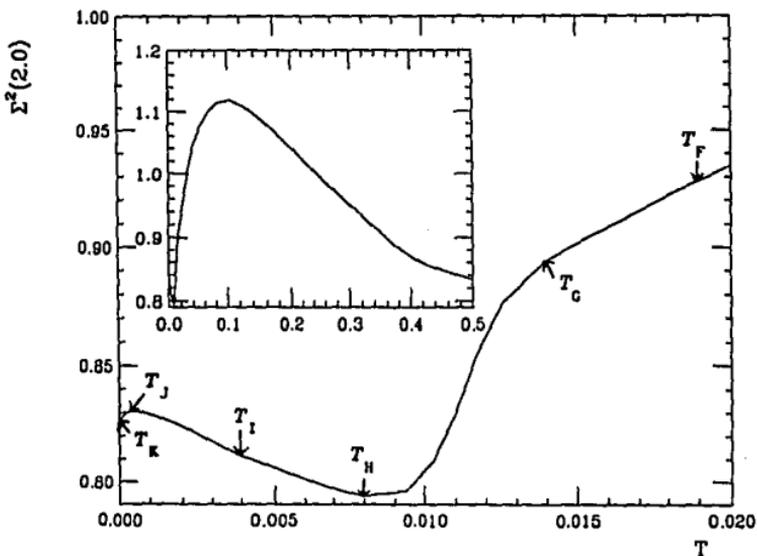
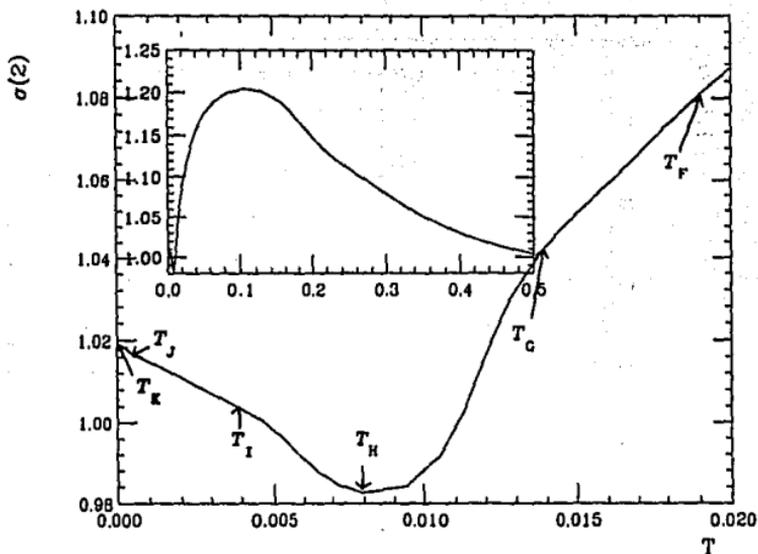


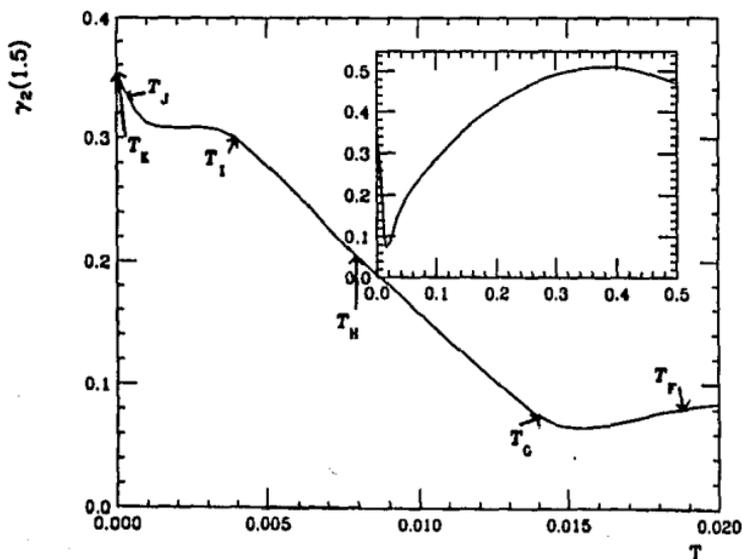
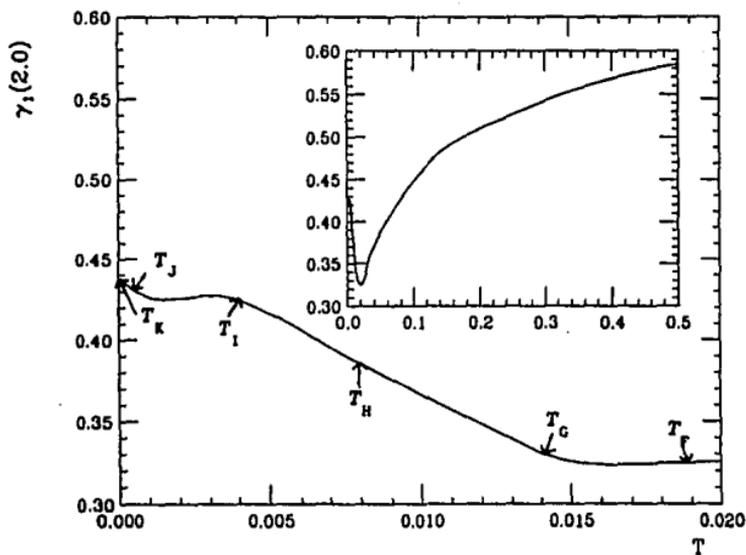
Fig. 3.20. Distancia promedio L de los caminos correspondientes a 350 muestras de 500 ciudades cada una como función de la temperatura T a la que se encuentran dichas muestras durante el proceso de templado simulado. $T_I = T_A = 0.5$, $T_B = 0.365$, $T_C = 0.240$, $T_D = 0.125$, $T_E = 0.0323$, $T_F = 0.0190$, $T_G = 0.0139$, $T_H = 0.0080$, $T_I = 0.0039$, $T_J = 0.0006$ y $T_r = T_K = 0.00001$.

están unas encima de las otras y no se alcanzan a distinguir claramente. La elección de los puntos $\sigma(2)$, $\Sigma^2(2.0)$, etc. es totalmente arbitraria y los otros puntos se comportan de forma similar.

A partir de estas figuras se ve que la forma en que estas cuatro cantidades se aproximan a sus valores templados no es monotónica, sino que hay cierta oscilación en todo el proceso. Lo mismo ocurre con los otros puntos, de manera que al ir disminuyendo la temperatura, las cuatro curvas en consideración van subiendo y bajando hasta que llegan a sus formas finales. Es más, para cualquier de estas cuatro funciones ($\sigma(k)$, $\Sigma^2(L)$, $\gamma_1(L)$ y $\gamma_2(L)$), puede llegar a haber dos temperaturas para las cuales las curvas correspondientes a éstas coincidan entre sí, de manera que la forma templada de la $\sigma(k)$, por ejemplo, no corresponde



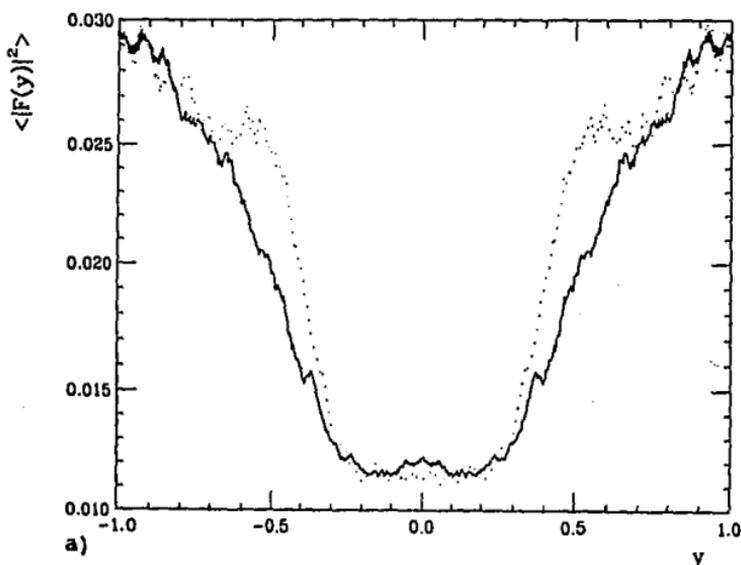
Figs. 3.21 y 3.22. Desviación estándar $\sigma(2)$ de la densidad de probabilidad $p(2;s)$ y varianza de número $\Sigma^2(2.0)$, respectivamente, como función de la temperatura. $T_1 = T_A = 0.5$, $T_B = 0.365$, $T_C = 0.240$, $T_D = 0.125$, $T_E = 0.0323$, $T_F = 0.0190$, $T_G = 0.0139$, $T_H = 0.0080$, $T_I = 0.0039$, $T_J = 0.0006$ y $T_r = T_K = 0.00001$.



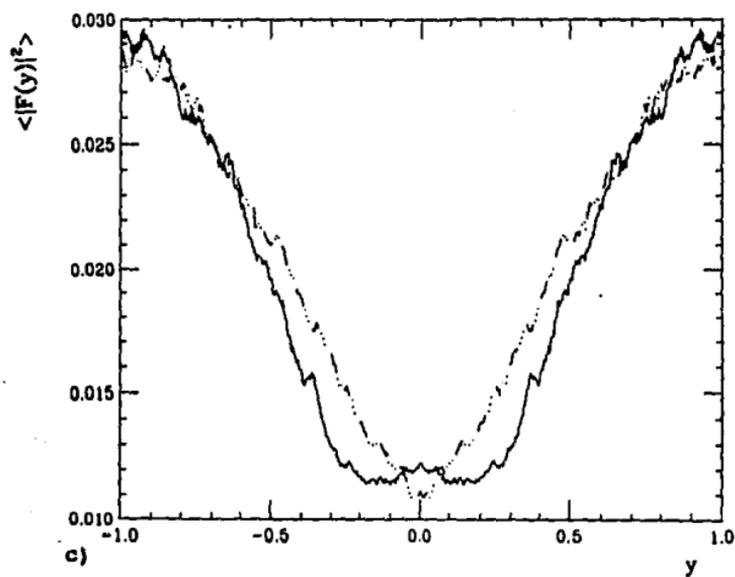
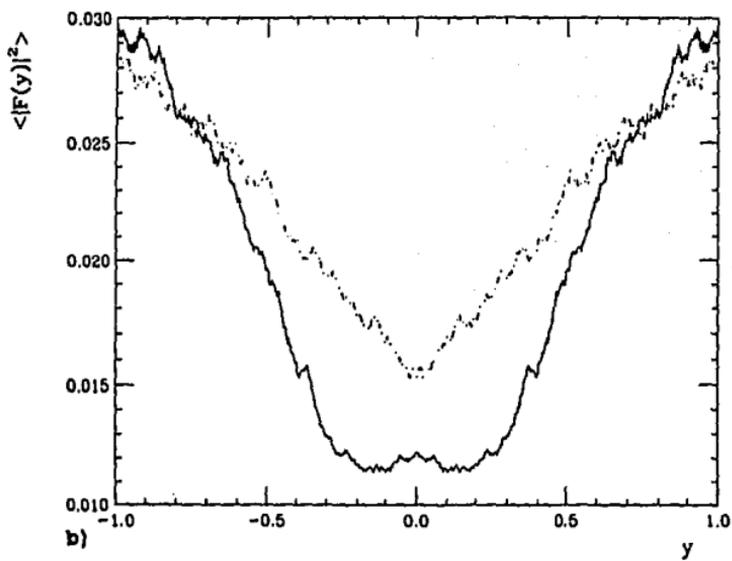
Figs. 3.23 y 3.24. Coeficiente de sesgo $\gamma_1(2.0)$ y coeficiente de exceso $\gamma_2(1.5)$, respectivamente, como función de la temperatura. $T_I = T_A = 0.5$, $T_B = 0.365$, $T_C = 0.240$, $T_D = 0.125$, $T_E = 0.0323$, $T_F = 0.0190$, $T_G = 0.0139$, $T_H = 0.0080$, $T_I = 0.0039$, $T_J = 0.0006$ y $T_r = T_k = 0.00001$.

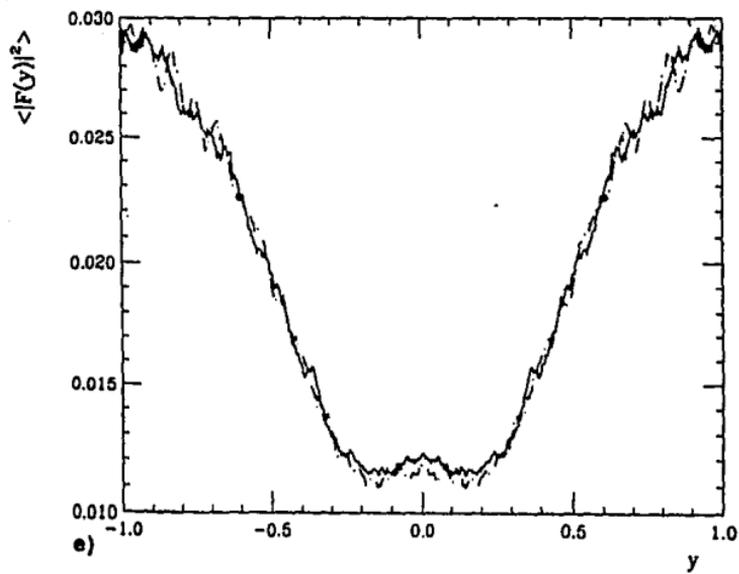
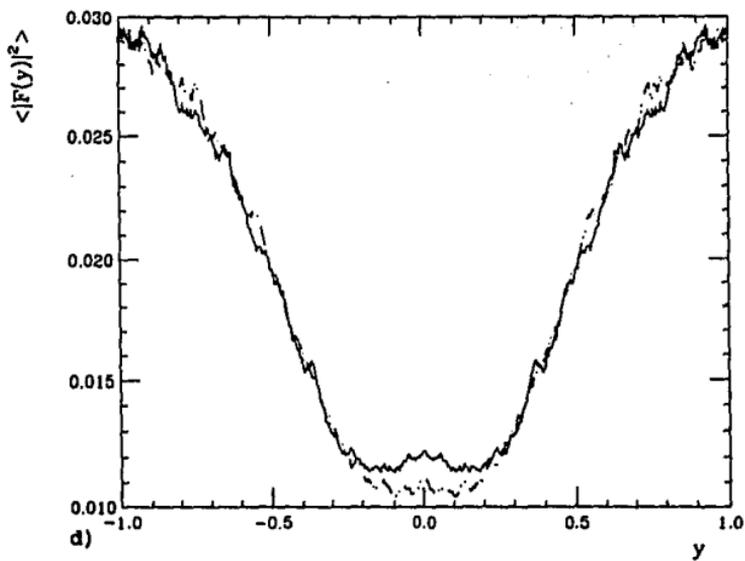
sólamente a la temperatura final, y lo mismo para las otras tres cantidades. Como a cada temperatura le corresponde una longitud promedio L , esto indica que la forma final de cualquiera de estas cantidades considerada por sí sola no caracteriza a las soluciones de mínima distancia encontradas con el templado simulado. Sin embargo, las temperaturas para las cuales $\sigma(k)$ adquiere la misma forma que la forma templada no coinciden con las temperaturas para las cuales $\gamma_1(L)$ lo hace, por ejemplo, de forma que estas cuatro cantidades en conjunto si caracterizan a las soluciones finales obtenidas con el método del templado simulado. Esto se ve claramente en las Figuras 3.21, 3.22, 3.23 y 3.24.

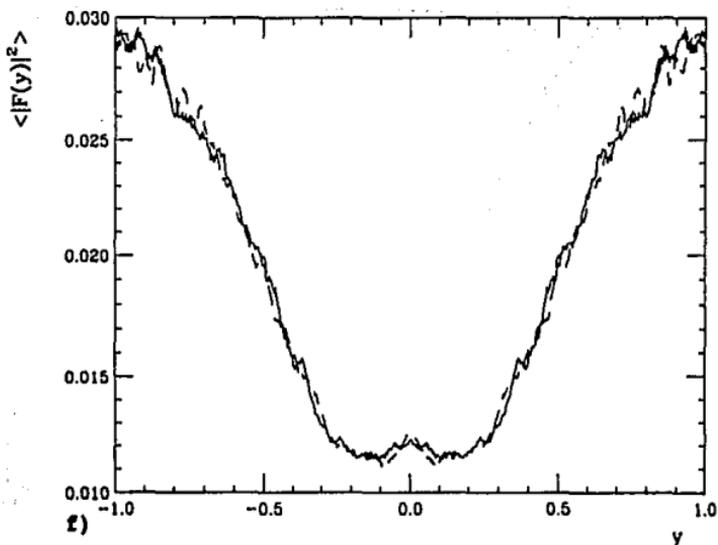
Pasando a la media de las transformadas de Fourier, ésta se muestra en las Figuras 3.25 como función de la temperatura.



Figs. 3.25. Media del módulo al cuadrado de las transformadas de Fourier $F(y)$ para 350 muestras de 500 ciudades. Cada curva corresponde a una temperatura del proceso de templado simulado. a) $T_1 = T_1 = 0.5$ (—), b) $T_2 = 0.125$ (- · · -), c) $T_3 = 0.0139$ (— · · · —), d) $T_4 = 0.0080$ (— · · · —), e) $T_5 = 0.0039$ (— · —) y f) $T_6 = 0.0006$ (— —). La curva sólida corresponde a la forma ya templada.







En estas figuras se aprecia que la forma de la función $\langle |F(y)|^2 \rangle$ va cambiando conforme baja la temperatura. A $T = 0.5$, la forma de esta función es parecida a la forma final. Conforme baja la temperatura, la curva va subiendo más y más hasta que a partir de cierto punto empieza a bajar otra vez hasta llegar a su forma final. De nuevo se ve que el proceso es oscilatorio y no monotónico.

Algo similar ocurre con el coeficiente de correlación. Los valores, como función de T , son los siguientes: $\rho_{T_A} = 0.0353$, $\rho_{T_B} = 0.0162$, $\rho_{T_C} = 0.0050$, $\rho_{T_D} = 0.0053$, $\rho_{T_E} = -0.0237$, $\rho_{T_F} = -0.0429$, $\rho_{T_G} = -0.0442$, $\rho_{T_H} = -0.0006$, $\rho_{T_I} = 0.0362$, $\rho_{T_J} = 0.0527$ y $\rho_{T_K} = 0.0547$. Otra vez, la forma en que está cantidad llega a su valor templado no es monotónica.

En resumen, la forma en que todas las cantidades estadísticas aquí presentadas se aproximan a sus formas templadas es oscilatoria y no monotónica.

De esta forma, con esto se concluye el estudio de las propiedades estadísticas de las soluciones del problema del agente viajero. Las conclusiones que se derivan de los resultados obtenidos se presentan en el siguiente capítulo.

CAPITULO 4

CONCLUSIONES

Con base en los resultados expuestos en el capítulo anterior las conclusiones son las siguientes:

1. Las soluciones del problema del agente viajero obtenidas por medio del método del templado simulado presentan regularidades estadísticas, como lo muestra el comportamiento de las diferentes propiedades estudiadas.
2. Estas propiedades son diferentes a las de los espectros aleatorios correspondientes al GOE o a procesos Poisson. De hecho, también son diferentes a las del GUE y GSE (ver Ref. [20]). Esto no es sorprendente, ya que hay muchos sistemas para los cuales los "espectros" asociados a ellos no son ni poissonianos ni gaussianos. Sin embargo, es interesante comparar con estos dos tipos de espectros ya que éstos aparecen en el estudio de una gran variedad de sistemas.
3. Localmente, las propiedades estadísticas encontradas son muy parecidas para muestras que caen dentro de un gran intervalo en lo que al número de ciudades se refiere. Esto indica que lo que determina la estructura local de este tipo de espectros no es tanto el número de ciudades sino el hecho de que se está trabajando con caminos "óptimos".
4. La probabilidad de que al obtener cierta solución mediante el método del templado, la distancia entre ciudades vecinas a lo largo del camino esté entre s y $s+ds$ está dada por

$$p(0;s) = as^b \exp(-cs^d) I_{(0,\infty)}(s), \quad (4.1)$$

siendo $a = dc^{(b+1)/d} / \Gamma((b+1)/d)$, $c = \Gamma((b+2)/d) / \Gamma((b+1)/d)^d$ y b y d las constantes encontradas en el capítulo anterior. Se hace notar que para el caso Poisson, $a = 1$, $b = 0$, $c = 1$ y $d = 1$ mientras que para la densidad de Wigner, $a = \pi/2$, $b = 1$, $c = \pi/4$ y $d = 2$.

5. Las curvas correspondientes a $\sigma(k)$, $\Sigma^2(L)$, $\gamma_1(L)$ y $\gamma_2(L)$ son las que aparecen en las Figuras 3.3, 3.4, 3.5 y 3.6 respectivamente. Los valores de dichas cantidades para L y k pequeños (menores que 3) son válidos para muestras con diferente número de ciudades. La información que estas cantidades dan en lo que se refiere a los caminos y no a los espectros es clara. Por ejemplo, si se obtiene por medio del templado una solución para cierta muestra y si el camino obtenido se normaliza, el valor $\sigma(0)$ de la Figura 3.3 indicará si es despreciable o no la probabilidad de que dos ciudades vecinas a lo largo del camino se encuentren separadas por una distancia entre s y $s+ds$.

Cabe mencionar que tanto $p(0;s)$ como estas cuatro cantidades tienen que ver con los espectros ya normalizados. Sin embargo, basta multiplicar s o L por la media D entre ciudades vecinas para así obtener la información en términos de las distancias reales.

6. El coeficiente de correlación de dos espaciamientos consecutivos a lo largo de los caminos obtenidos en el templado es muy cercano a cero. Sin embargo, la evidencia en favor de la independencia de estos segmentos en lo que se refiere a su longitud no es conclusiva, ya que el coeficiente puede ser igual a cero aunque la longitud de los espaciamientos esté correlacionada.
7. Existen correlaciones entre regiones de los caminos muy separadas entre sí, tal y como lo indican las figuras de las

transformadas de Fourier. Esto era de esperarse, y es debido a la condición de que todos los caminos deben de ser cerrados.

8. El conjunto de soluciones obtenidas por medio del método del templado simulado para una sola muestra es prácticamente equivalente, en lo que se refiere a la $\sigma(k)$, $\Sigma^2(L)$, $\gamma_1(L)$, $\gamma_2(L)$ y $\rho_{s_1 s_2}$, al ensemble de soluciones correspondientes a distintas muestras. En lo que se refiere a la densidad $p(0;s)$ y a la función $\langle |F(y)|^2 \rangle$, existen ciertas discrepancias entre ambos casos, aunque la forma burda de las curvas es la misma. La similitud entre los dos casos en consideración se debe a que para el caso de una sola muestra, el hecho de que se empiece cada vez desde configuraciones iniciales diferentes trae como consecuencia la existencia de un gran número de caminos finales distintos; esto es, se tiene algo muy parecido en ambos casos debido a esta diversidad de soluciones finales. Sin embargo, existe una diferencia importante. En el caso de una sola muestra, va a haber una serie de segmentos que aparecerán en casi todos los caminos finales obtenidos. Esta repetición tiene como resultado que el número "efectivo" de datos a partir de los cuales se encuentran las diferentes cantidades estadísticas disminuya. Esta disminución no va a afectar a $\sigma(k)$, $\Sigma^2(L)$, $\gamma_1(L)$, $\gamma_2(L)$ o $\rho_{s_1 s_2}$, ya que todavía hay un gran número de segmentos que casi no se van a repetir. Sin embargo, la alta repetición de algunos segmentos sí va a afectar tanto a $p(0;s)$ como a la función $\langle |F(y)|^2 \rangle$. En el caso de $p(0;s)$, dicha repetición afecta porque es precisamente la frecuencia relativa de s la que se está graficando. En el caso de la función $\langle |F(y)|^2 \rangle$, la alta repetición afectará lejos del origen, ya que ahí sí es importante la disminución en el número "efectivo" de datos. Así, ésta es la causa de las discrepancias tanto para $p(0;s)$ como para $\langle |F(y)|^2 \rangle$. Estos resultados indican, como es sabido, que el espacio de configuración de cualquier muestra presenta un gran número de mínimos locales.

9. La forma en que las diferentes cantidades estadísticas se aproximan a sus formas templadas es oscilatoria y no monotónica. Si se considera únicamente a la $\sigma(k)$, su forma templada no va a caracterizar a las soluciones obtenidas con el método del templado simulado, ya que existen otros caminos de distancia mayor que se comportan de igual forma en lo que a esta cantidad se refiere. Lo mismo va a suceder con la $\Sigma^2(L)$, $\gamma_1(L)$, $\gamma_2(L)$ o $\rho_{1,2}$. Sin embargo, estas cinco cantidades en conjunto sí van a caracterizar a las soluciones de mínima distancia encontradas con el método del templado. Tanto la densidad $p(0;s)$ como la función $\langle |F(\gamma)|^2 \rangle$ también caracterizarán por sí solas a las soluciones en cuestión. Cabe mencionar que la razón por la cual existen estas oscilaciones es desconocida y se requerirá de un análisis mucho más detallado para entender por qué ocurren.

Aparte del interés que estos resultados tiene por sí mismos, éstos pueden ser útiles en el estudio de otros campos, tal y como se mencionó en la Introducción.

Por último, es difícil saber que tanto varían estos resultados al tratar con caminos de mínima distancia. Esto es debido a las grandes complicaciones existentes para obtener soluciones que realmente sean óptimas.

APENDICE 1
PROPIEDADES ESTADISTICAS DE ESPECTROS ALEATORIOS

A1.1 MATRICES ALEATORIAS

En este apéndice se estudiarán, mediante la teoría de matrices aleatorias, las propiedades estadísticas de los espectros de energía asociados a diferentes sistemas. La teoría de las matrices aleatorias se ocupa⁽²⁰⁻²²⁾ de las propiedades que exhiben las matrices cuyos elementos son variables aleatorias. Esta teoría se utiliza tanto en física nuclear como en física atómica⁽²⁰⁻²¹⁾ y, en particular, es muy útil al estudiar los espectros de energía de núcleos o átomos pesados.

Al estudiar este tipo de sistemas, lo que se desea es poder resolver la ecuación de Schrödinger exactamente para así obtener tanto el espectro como las eigenfunciones del sistema en cuestión. Muchas veces sólo se tiene interés en la parte discreta del espectro o en una porción de ésta, de manera que en estos casos, lo que se hace es trabajar sólo con la parte de la matriz hamiltoniana asociada con la porción del espacio de estados del sistema en la cual se tiene interés. Aunque esto simplifica mucho el estudio de dichos sistemas, el resolver la ecuación resultante es muy difícil debido a la gran complejidad que estos sistemas presentan. Tanto el número de niveles de energía como la densidad de estados correspondientes a energías de excitación altas son tan grandes que resulta difícil analizar individualmente los diferentes estados⁽²⁰⁾.

Debido a esto, lo que se hace es reemplazar la parte de la matriz hamiltoniana con la cual se está trabajando por una matriz aleatoria. Tanto las densidades de probabilidad de los elementos

de matriz como la matriz misma deben de cumplir ciertos requisitos impuestos por las simetrías que el sistema presente. Ahora, una vez hecha esta substitución y por medio de la teoría de las matrices aleatorias, se puede obtener cierta información acerca de los eigenvalores en los cuales se tiene interés. Cabe mencionar que al hacer esto, no es posible obtener información detallada acerca de los conjuntos de niveles asociados con las submatrices originales para ningún núcleo o átomo en particular. En su lugar, lo que se obtiene es información acerca de la apariencia general de estos conjuntos para cualquier núcleo o átomo que sea lo suficientemente complejo como para no poder ser descrito detalladamente⁽²⁵⁾.

De igual forma que en la mecánica estadística se trabaja con diferentes ensembles dependiendo del tipo de sistema, en este caso se trabaja con diferentes ensembles de matrices dependiendo de ciertas propiedades que el sistema en consideración presente. En la mecánica estadística se obtienen los valores de diferentes cantidades de interés al promediar sobre el ensemble adecuado. En este caso también se obtiene información sobre los sistemas de interés al realizar diferentes promedios sobre los ensembles. Si se tiene interés en los niveles de energía asociados con una parte del hamiltoniano original y si el número de estos niveles es suficientemente grande, se postula⁽²⁵⁾ que la información que se obtenga al realizar promedios sobre el ensemble correspondiente describirá adecuadamente a dicha porción del espectro del sistema particular que se esté estudiando.

Específicamente, las submatrices con las cuales se trabaja son las siguientes. En general, existen ciertas cantidades que se conservan independientemente del grado de excitación del sistema, como por ejemplo, la magnitud J del momento angular total o la paridad π . Debido a esto, es posible encontrar eigenfunciones simultáneas de estas cantidades y si se representa la matriz hamiltoniana en esta base, ésta tendrá la forma mostrada en la Figura A1.1.

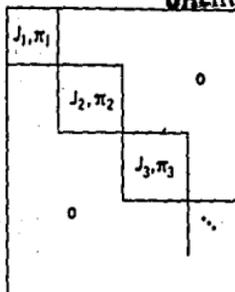


Fig. A1.1. Estructura de la matriz hamiltoniana representada en una base adecuada.

Cada submatriz o bloque de la diagonal corresponde a valores fijos de los números cuánticos asociados con las cantidades conservadas. Todos los elementos de matriz que están fuera de estos bloques son iguales a 0. Son precisamente estas submatrices con las que se acostumbra trabajar. Dependiendo de la porción del espectro en la que se tenga interés, se trabajará o con una de éstas o con varias. Ahora, se sustituye la submatriz en la cual se tiene interés por una matriz aleatoria adecuada (dependiendo de las propiedades y simetrías del sistema) y a partir de esta última se obtiene información acerca de la porción del espectro en cuestión.

Si el sistema que se está estudiando es invariante ante reflexiones en el tiempo y si el valor J de la magnitud del momento angular total es entero, las submatrices correspondientes serán reales y simétricas (siempre y cuando la matriz hamiltoniana original se represente en una base adecuada). Estas submatrices seguirán siendo reales y simétricas en otras bases siempre y cuando la transformación $\psi \rightarrow R\psi$ de una base a otra sea tal que la matriz R sea ortogonal. Así, en estos casos el ensemble de matrices que se usa para obtener información acerca del sistema que se está estudiando es el ensemble ortogonal gaussiano (GOE). Este ensemble está definido sobre el espacio de matrices reales y simétricas T_{10} y cumple con las siguientes dos condiciones:

1. Es invariante bajo cada transformación del espacio T_{10} sobre sí mismo, del tipo $H \rightarrow R^T H R$, siendo R cualquier matriz ortogonal y H la submatriz en cuestión.
2. Los elementos de matriz H_{ij} , $i \neq j$, son estadísticamente independientes.

Cabe mencionar que si el sistema es invariante ante rotaciones, la forma de las submatrices en una base adecuada también será real y simétrica, independientemente del valor J de la magnitud del momento angular total, de tal forma que el ensemble que se usa en estos casos es, otra vez, el GOE.

Si el sistema es invariante ante reflexiones en el tiempo y es tal que el valor de J es semi-entero pero no es invariante ante rotaciones, las submatrices correspondientes, en una base adecuada, serán matrices hermitianas autoduales⁽²⁰⁾. Estas submatrices seguirán siendo de este tipo siempre y cuando la transformación $\psi \rightarrow W\psi$ de una base a otra sea tal que W sea una matriz simpléctica⁽²⁰⁾. En este caso, el ensemble que se utiliza es el ensemble simpléctico gaussiano (GSE). Este ensemble está definido sobre el espacio T_{4c} de matrices hermitianas autoduales y cumple con las siguientes propiedades:

1. El ensemble es invariante ante transformaciones de T_{4c} sobre sí mismo, del tipo $H \rightarrow W^R H W$, siendo W cualquier matriz simpléctica y H la submatriz en consideración.
2. Varias de las componentes linealmente independientes de H también son estadísticamente independientes⁽²⁰⁾.

Si el sistema no es invariante ante reflexiones en el tiempo, las diferentes submatrices serán matrices hermitianas sin la necesidad de que éstas sean reales o autoduales. Así, en este caso el ensemble que se usa para describir al sistema es el ensemble unitario gaussiano (GUE), ya que una matriz hermitiana en cierta base lo seguirá siendo en cualquier otra base que esté relacionada con la primera mediante una transformación unitaria. Este ensemble está definido sobre el espacio T_{2c} de matrices hermitianas y cumple, de manera semejante a los dos casos

anteriores, con las siguientes condiciones:

1. Es invariante bajo transformaciones del tipo $H \rightarrow U^{-1}HU$ del espacio T_{2c} sobre sí mismo, siendo U cualquier matriz unitaria y H la submatriz en cuestión.
2. Ciertas componentes linealmente independientes de H también son estadísticamente independientes⁽²⁰⁾.

Las condiciones que cada ensemble cumple son tales que, en los tres casos, la densidad de probabilidad conjunta $P(H)$ de los elementos de matriz independientes está dada por⁽²⁶⁾

$$P(H) = \exp(-a\text{Tr}(H^2) + b\text{Tr}(H) + c), \quad (\text{A1.1})$$

siendo a real y positiva y b y c reales. Es por esto que a estos tres ensembles se les llama gaussianos.

De esta forma, al estudiar cierto sistema en particular y dependiendo de las características que presente, se puede obtener información acerca de sus niveles de energía al trabajar con el ensemble de matrices apropiado, a pesar de que dicho sistema sea muy complejo y por ello resulte muy difícil obtener los eigenvalores y eigenfunciones de manera exacta.

Ya que los niveles de energía de las matrices que conforman los diferentes ensembles son aleatorios, es posible estudiar algunas de sus propiedades estadísticas. El comportamiento estadístico de dichos niveles va a ser característico del ensemble al cual pertenezcan. En la siguiente sección se hablará acerca de las propiedades estadísticas que se suelen estudiar al trabajar con los espectros de estos ensembles.

A1.2 PROPIEDADES ESTADÍSTICAS DE ESPECTROS ALEATORIOS

A continuación se verá en que consisten las diferentes propiedades estadísticas que se suelen estudiar al trabajar con distintos espectros aleatorios. Cabe mencionar que es posible estudiar estas propiedades al trabajar con cualquier tipo de espectro aleatorio, sin la necesidad de que éste corresponda a alguno de los ensembles arriba mencionados. Por supuesto, el

comportamiento estadístico de estos espectros será diferente al de los espectros gaussianos. Como ejemplo de espectros aleatorios no gaussianos se tienen los espectros asociados a procesos de Poisson. Antes de discutir las propiedades estadísticas comúnmente estudiadas, se hablará un poco acerca de ciertos aspectos tanto de los espectros gaussianos así como de los de tipo Poisson.

Un espectro tipo Poisson es un espectro en el cual la probabilidad $E(n;L)$ de encontrar n niveles dentro de un intervalo de longitud L está dada por la densidad de probabilidad de Poisson $(e^{-\nu L}(\nu L)^n)/n! I_{(0,1,\dots)}(n)$, siendo ν el número de niveles promedio por unidad de intervalo. Si la distancia promedio entre los niveles es igual a uno, esto es, si $\nu = 1$, la densidad de probabilidad resultante estará dada por

$$E(n;L) = (e^{-L})/n! I_{(0,1,\dots)}(n). \quad (A1.2)$$

Debido a la relación que existe entre las densidades de probabilidad gamma y Poisson⁽²⁴⁾, se tiene que si $\nu = 1$, la densidad de probabilidad $p(k;s)$ de que la distancia entre dos niveles con k niveles intermedios esté entre s y $s+ds$, está dada por la densidad de probabilidad gamma

$$p(k;s) = (s^k e^{-s})/k! I_{(0,\infty)}(s). \quad (A1.3)$$

Los espectros tipo Poisson o poissonianos contienen, en principio, un número infinito de niveles. La densidad de niveles es constante en cualquier región del espectro y en promedio es igual a ν . Para estos espectros no existe ninguna correlación entre los diferentes niveles de energía; es decir, la posición de cualquier nivel es independiente de las posiciones de todos los demás⁽²⁴⁾.

Respecto a los espectros asociados con las matrices aleatorias de los ensembles gaussianos, se tiene lo siguiente. El número de niveles correspondiente a cada submatriz es del orden de N , siendo N la dimensión de ésta. En los tres casos, este número es muy grande ya que existe un gran número de niveles correspondientes a valores particulares de los números cuánticos conservados (tanto en el caso nuclear como en el caso atómico). Si

$N \rightarrow \infty$ la densidad de niveles en los tres casos está dada por⁽²⁰⁾

$$\sigma(x) = \begin{cases} (1/\pi) (2N-x^2)^{1/2}, & |x| < (2N)^{1/2}, \\ 0, & |x| > (2N)^{1/2}. \end{cases} \quad (\text{A1.4})$$

A esta ecuación se le conoce como la ley semicircular de Wigner. Cerca del origen la densidad es casi constante y alcanza su máximo valor. En lo que sigue, sólo se considerará esta región de densidad constante al hablar sobre los ensembles gaussianos. En el caso Poisson, se considerará cualquier región del espectro ya que la densidad en todo éste es constante. En general, es habitual reescalar los espectros de tal forma que la media de la distancia entre niveles vecinos sea igual a uno; esto es, se divide la escala de energía entre la media D de los espaciamientos entre niveles vecinos obteniéndose de esta forma espectros reescalados adimensionales. Con esta información, se procederá a continuación con la explicación de las propiedades estadísticas de espectros aleatorios ya normalizados.

Las densidades de probabilidad que generalmente se estudian al analizar espectros aleatorios son las siguientes⁽²⁰⁾: densidad de probabilidad $p(k;s)$ de que la distancia entre dos niveles escogidos al azar con k niveles intermedios esté entre s y $s+ds$ y la densidad de probabilidad $E(n;L)$ de que dentro de un intervalo de longitud L escogido al azar haya n niveles. Para el caso Poisson estas densidades están dadas por las ecuaciones A1.3 y A1.2 respectivamente.

En lo que se refiere a la densidad $E(n;s)$, habitualmente sólo se analizan sus primeros momentos centrales respecto a la media⁽²⁰⁾, de manera que debido a esto, aquí sólo se hablará acerca de estos momentos⁽²⁰⁾ y no acerca de la densidad de probabilidad como tal.

A la densidad de probabilidad $p(k;s)$ se le llama densidad de probabilidad de k -ésimos vecinos. La información que se obtiene de ella es evidente.

En el caso Poisson, $p(k;s)$ está dada por la densidad de probabilidad $((e^{-s}s^k)/k!)I_{(0,\infty)}(s)$, como ya se mencionó. Si $k = 0$, esta densidad se reduce a $e^{-s}I_{(0,\infty)}(s)$.

En el caso del GOE se tiene lo siguiente. La densidad de probabilidad $p(0;s)$ se puede aproximar bastante bien por la densidad

$$p(s) = (\pi/2)s \exp(-(\pi/4)s^2) I_{(0,\infty)}(s). \quad (A1.5)$$

A esta densidad de probabilidad se le conoce como⁽²⁰⁾ densidad de probabilidad de Wigner. Es un caso particular de la densidad de probabilidad $as^b \exp(-cs^d) I_{(0,\infty)}(s)$, $a = (dc^{(b+1)/d})/\Gamma((b+1)/d)$. Si $b = 1$, $c = \pi/4$ y $d = 2$, lo que se obtiene es precisamente la función de Wigner. La forma analítica de las densidades de probabilidad $p(k;s)$, $k \neq 0$, es bastante complicada y por ello no se presentarán en este apéndice. Un estudio detallado de ellas aparece en la Ref. [20].

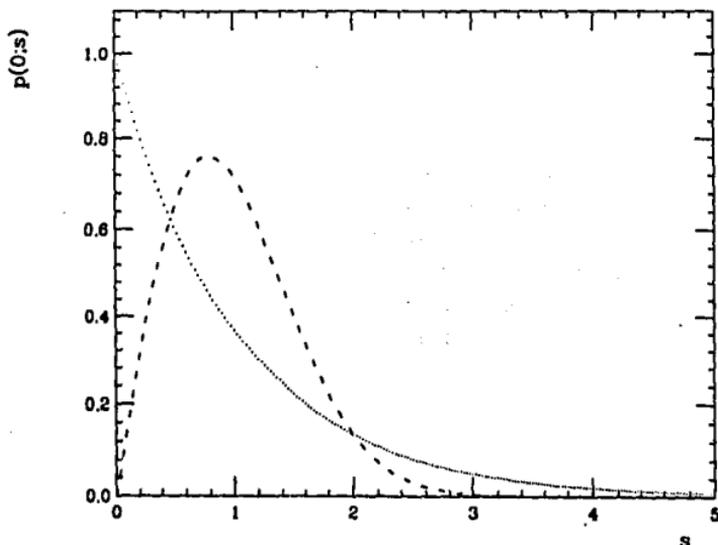


Fig. A1.2. Densidad de probabilidad de primeros vecinos $p(0;s)$ para los casos Poisson (—) y GOE (- -).

Tanto para el caso Poisson como para el del GOE, cuando $k \rightarrow \infty$, $p(k;s)$ tiende a una densidad de probabilidad normal. Esto es debido al teorema central del límite⁽²⁴⁾, ya que la variable aleatoria s asociada con $p(k;s)$ es la suma de las $k+1$ variables

aleatorias s_1 correspondientes a espaciamentos entre $k+2$ niveles; esto es, el par de niveles escogidos al azar y los k niveles intermedios. En la Figura A1.2 se muestran las densidades de probabilidad $p(0;s)$ para estos dos casos.

La desviación estándar $\sigma(k)$ de las densidades de probabilidad $p(k;s)$ está dada por

$$\sigma(k) = (\langle (s - \langle s \rangle)^2 \rangle)^{1/2}. \quad (\text{A1.6})$$

Esta desviación es una medida de qué tan grande es el intervalo alrededor de la media dentro del cuál caen con alta probabilidad los valores de s . Si el espaciamento promedio entre niveles es igual a uno, la media de s resulta ser igual a $k+1$. Si $\sigma(k)$ es grande, esto indica que al ir escogiendo aleatoriamente pares de niveles con k niveles intermedios, la variabilidad de la distancia que los separa será muy grande. En cambio, si $\sigma(k)$ es pequeña, esto indica que dicha variabilidad será muy pequeña; es decir, las desviaciones de s respecto a $k+1$ serán muy chicas.

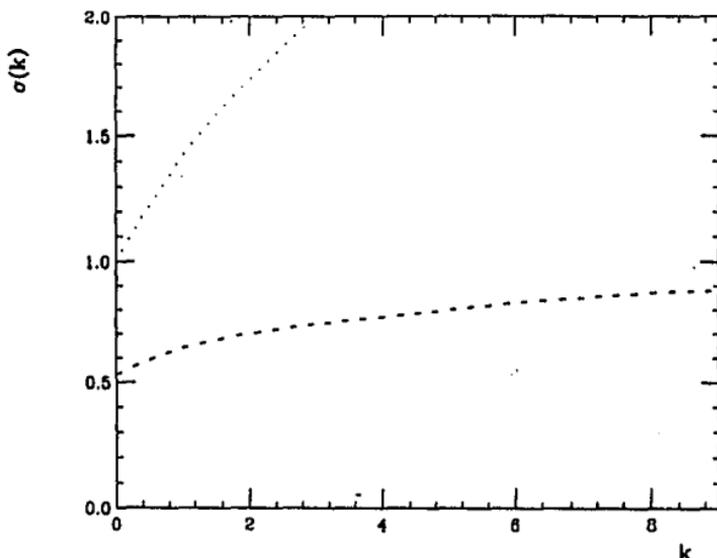


Fig. A1.3. Desviación estándar $\sigma(k)$ de las densidades de probabilidad de k -ésimos vecinos $p(k;s)$ como función de k . GOE (- -) y Poisson (—).

Tanto en el caso Poisson como en el del GOE $\sigma(k)$ aumenta con k . Esto se muestra en la Figura A1.3. En esta figura los únicos valores de $\sigma(k)$ que son significativos son los correspondientes a k entero, ya que k es el número de niveles intermedios.

Como ya se mencionó, si $k \rightarrow \infty$ la densidad de probabilidad $p(k;s)$ tiende, en ambos casos, a una normal estandarizada. Aunque en ambos casos la media de estas densidades va a ser igual a $k+1$, las desviaciones estándar van a diferir entre sí, como se aprecia en la Figura A1.3.

Pasando a la densidad de probabilidad $E(n;L)$, se tiene lo siguiente⁽²⁰⁾. Como $E(n;L)$ depende de la longitud L de los intervalos, sus momentos obviamente también dependerán de ella. Debido a esto, lo que se acostumbra estudiar es la forma en que dichos momentos dependen de L . Específicamente, aquí se hablará acerca de la varianza Σ^2 , coeficiente de sesgo γ_1 y coeficiente de exceso γ_2 de esta densidad como funciones de L .

La varianza Σ^2 está dada por

$$\Sigma^2(L) = \langle (n - \langle n \rangle)^2 \rangle. \quad (\text{A1.7})$$

Es una medida de qué tanto varían, en promedio, los valores que la variable aleatoria n toma alrededor de la media $\langle n \rangle$. Si Σ^2 es grande, esto quiere decir que al escoger aleatoriamente un intervalo de longitud L es muy probable encontrar un número de niveles n alejado del promedio $\langle n \rangle$, mientras que si Σ^2 es chico, es muy poco probable encontrar dentro de este intervalo un número de niveles alejado de dicho promedio. A $\Sigma^2(L)$ se le conoce como varianza de número, ya que es la varianza del número n de niveles que caen dentro de los intervalos de longitud L .

En el caso Poisson, Σ^2 es la varianza de la densidad de probabilidad A1.2; es decir,

$$\Sigma^2 = L. \quad (\text{A2.8})$$

En el caso del GOE, la forma de esta varianza está dada por

$$\begin{aligned} \Sigma^2 &= (2/\pi^2)(\ln(2\pi L) + \gamma - (\pi^2/8) - \text{Ci}(2\pi L)) + \\ &+ (4L/\pi) \int_{\pi L}^{\infty} (\text{sen}\xi/\xi)^2 d\xi + (1/\pi^2) (\int_{\pi L}^{\infty} (\text{sen}\xi/\xi) d\xi)^2 = \\ &= (2/\pi^2)(\ln(2\pi L) + 1 + \gamma - (\pi^2/8)) + O(L^{-1}), \end{aligned} \quad (\text{A2.9})$$

siendo γ la constante de Euler y $\text{Ci}(x)$ la función coseno-integral; esto es, $\text{Ci}(x) = -\int_x^{\infty} (\cos\xi/\xi) d\xi$. En el primer caso, la varianza aumenta linealmente con L mientras que en el segundo, aumenta logarítmicamente. En la Figura A1.4 se muestra Σ^2 como función de L para ambos casos.

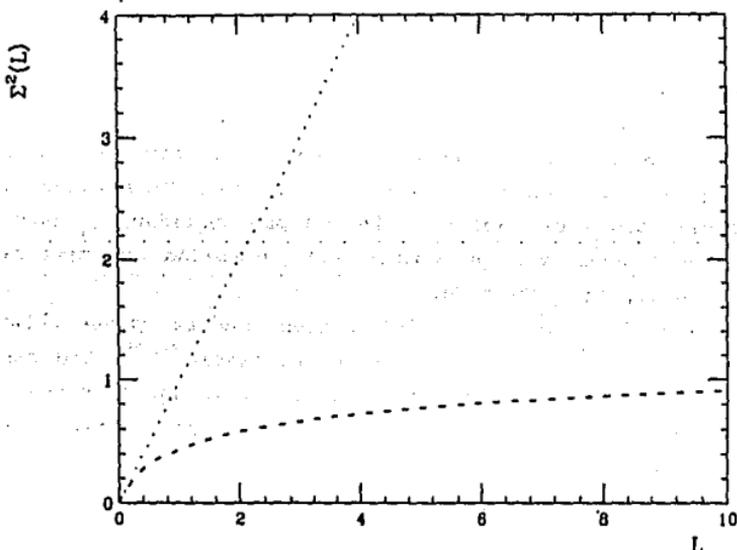


Fig. A1.4 Varianza de número Σ^2 como función de la longitud L de los intervalos para los casos Poisson (—) y GOE (- -).

Para valores chicos de L , la varianza es muy parecida en los dos casos, mientras que para valores más grandes, la asociada al caso Poisson es mayor. Esto quiere decir que dentro de un cierto intervalo de longitud $L > 1$, es más probable encontrar un número de niveles n alejado de $\langle n \rangle$ en el caso Poisson que en el del GOE.

El coeficiente de sesgo γ_1 está dado por

$$\gamma_1(L) = \mu_3/\mu_2^{3/2}, \quad (\text{A1.10})$$

siendo μ_r el r -ésimo momento central respecto a la media; esto es,

$$\mu_r(L) = \langle (n - \langle n \rangle)^r \rangle. \quad (\text{A1.11})$$

Si una densidad de probabilidad es simétrica respecto a su media, μ_3 , y por lo tanto γ_1 , serán iguales a cero. Sin embargo, una densidad no simétrica puede tener un coeficiente de sesgo igual a cero, de manera que el que γ_1 sea igual a cero no es indicio de que la densidad de probabilidad sea simétrica.

Ahora, en general, si la densidad de probabilidad está abultada hacia la izquierda de su media, es decir, si tiene una cola larga que se extiende hacia la derecha, γ_1 será mayor que cero, mientras que si esta densidad está abultada hacia la derecha, γ_1 será menor que cero. Sin embargo, dependiendo de las particularidades de cada densidad de probabilidad, γ_1 podrá ser positivo o negativo aun cuando esta densidad no presente la asimetría arriba mencionada.

Debido a esto, la información que se puede obtener a partir de este coeficiente no es muy confiable^(24,28). Sin embargo, γ_1 es útil al analizar espectros aleatorios ya que la forma en que depende de la longitud L del intervalo es característica del tipo de espectros con los que se esté trabajando.

En la Figura A1.5 se muestra dicho coeficiente como función de L tanto para el GOE como para espectros Poisson.

Para espectros del tipo Poisson $\gamma_1(L)$ es igual a $L^{-1/2}$, como se aprecia en la figura.

El coeficiente de exceso γ_2 es igual a

$$\gamma_2(L) = \langle \mu_4 / \mu_2^2 \rangle - 3. \quad (\text{A1.12})$$

Es una medida de la picudez alrededor de la media de las densidades de probabilidad⁽²⁴⁾, siendo positivo si la densidad es más picuda que una normal estandarizada, para la cual $\gamma_2 = 0$, y negativo si es menos picuda que ésta. Sin embargo, presenta los mismos problemas que el coeficiente de sesgo γ_1 en el sentido de que no mide adecuadamente esta picudez. A pesar de esto y al igual que γ_1 , su utilidad radica en el hecho que la forma en que depende

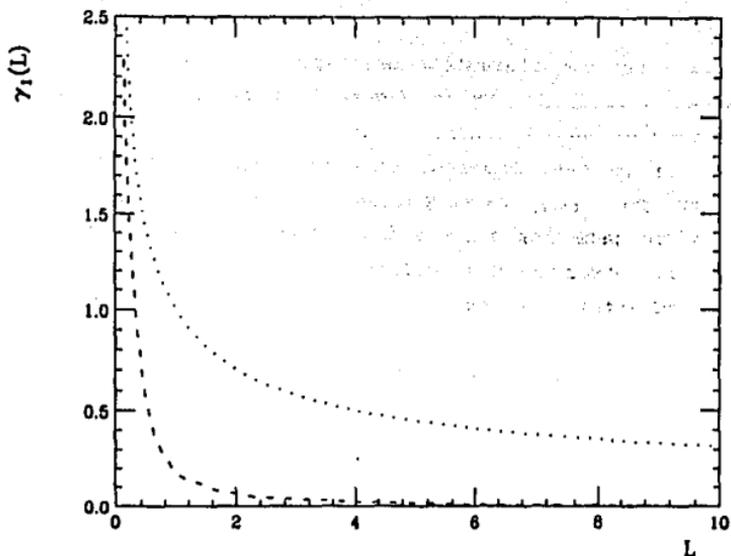


Fig. A1.5. Coeficiente de sesgo γ_1 como función de la longitud L de los intervalos para espectros del GOE (- -) y del caso Poisson (—).

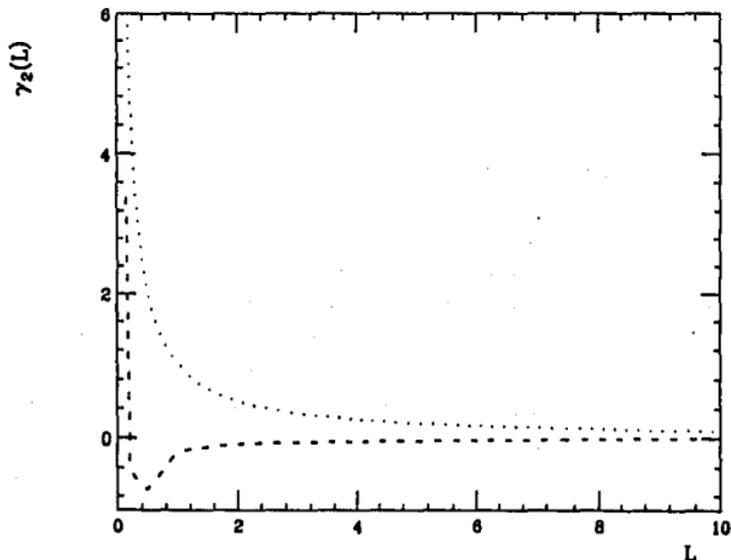


Fig. A1.6. Coeficiente de exceso γ_2 como función de la longitud L de los intervalos para los casos Poisson (—) y GOE (- -).

de L va a ser característica del tipo particular de espectros que se estén analizando. En la Figura A1.6 se muestra $\gamma_2(L)$ para el GOE y para el caso Poisson.

En el caso Poisson, este coeficiente es igual a L^{-1} . Al igual que con $\gamma_1(L)$, el comportamiento de $\gamma_2(L)$ cuando $L \rightarrow \infty$ ó a 0 es el mismo para los casos Poisson y GOE. Sin embargo, en el caso de γ_2 se observa una diferencia radical entre las curvas correspondientes a estos dos casos para valores de L alrededor de 0.5.

Aparte de todas estas cantidades, otra propiedad estadística que se acostumbra estudiar es el coeficiente de correlación $\rho_{s_1, s_2}^{(20)}$ de dos espaciamientos consecutivos s_1 y s_2 . Específicamente,

$$\rho_{s_1, s_2} = (\sigma_{s_1, s_2}) / \sigma_{s_1} \sigma_{s_2}, \quad (\text{A1.13})$$

siendo σ_{s_1, s_2} la covarianza de las variables aleatorias s_1 y s_2 y σ_{s_1} y σ_{s_2} sus desviaciones estándar, respectivamente. La covarianza de dos variables aleatorias X y Y está dada por⁽²⁴⁾

$$\sigma_{X, Y} = \langle (X - \mu_X)(Y - \mu_Y) \rangle \quad (\text{A1.14})$$

y es una medida de la relación lineal que existe entre ellas en el siguiente sentido. Si $X - \mu_X$ y $Y - \mu_Y$ tienden a tener el mismo signo con alta probabilidad, entonces la covarianza será positiva, mientras que si tienden a tener el signo contrario con alta probabilidad, $\sigma_{X, Y}$ será negativa. Es en este sentido en el que la covarianza mide la relación lineal entre X y Y , aunque su magnitud no tiene mucho sentido pues depende de la variabilidad tanto de X como de Y . Para resolver este último problema, lo que se hace es dividir entre el producto de las desviaciones estándar individuales, obteniéndose de esta forma el coeficiente de correlación. Si X y Y son variables independientes, entonces su coeficiente de correlación será igual a cero⁽²⁴⁾, como es fácil ver de la ecuación A1.14. Cabe mencionar que dicho coeficiente siempre cumple con la condición $-1 < \rho_{X, Y} < 1$.

De esta forma, si ρ_{s_1, s_2} es mayor que cero, esto indica que si el espaciamiento s_1 es mayor (menor) que la media $\langle s \rangle = 1$, s_2 será mayor (menor) que 1 con alta probabilidad. De igual forma, se tiene que si ρ_{s_1, s_2} es menor que cero y si s_1 es mayor (menor) que 1, s_2 será menor (mayor) que 1 con alta probabilidad.

En el caso del GOE el coeficiente de correlación es igual a -0.27 , mientras que en el caso Poisson es igual a cero ya que no hay correlación entre ninguno de los espaciamientos. En el caso del GOE, esto indica que si un espaciamiento es mayor (menor) que 1, el de arriba será menor (mayor) que 1 con alta probabilidad.

Para estudiar las correlaciones entre niveles de energía que están muy separados entre sí lo que se hace es analizar las transformadas de Fourier de los espectros aleatorios, ya que éstas contienen información sobre dichas correlaciones⁽²³⁾.

Las funciones $f(x)$ a partir de las cuales se obtienen las transformadas de Fourier son las siguientes: En cada espectro, se escoge arbitrariamente una región suficientemente grande y a partir de ésta se construye una función (la función $f(x)$) que vale cero para cualquier valor de x excepto para aquellos valores que corresponden a los diferentes niveles de energía que caen dentro de dicha región, para los cuales vale uno; es decir, la función $f(x)$ está dada por

$$f(x) = \sum_i \delta_{x, x_i}, \quad (\text{A1.15})$$

siendo las x_i los valores de las energías de los diferentes niveles que caen dentro del intervalo de interés.

Específicamente, lo que se estudia para obtener información acerca de las correlaciones de largo alcance es el promedio sobre el ensemble del módulo al cuadrado de las transformadas de Fourier $F(y)$.

La función $f(x)$ consta⁽²³⁾ de una envolvente rectangular multiplicada por una función aleatoria asociada con los niveles de energía del espectro. Debido a esto, la función $\langle |F(y)|^2 \rangle$ tiene dos componentes. La primera es la asociada con la transformada de Fourier de la envolvente y por esto es proporcional a $(\text{sen}(ny)/y)^2$, siendo n el número de niveles dentro de la región de

interés. Esta componente sólo contribuye cerca del origen y su altura es proporcional a n^2 . La otra componente es la asociada con la transformada de la función aleatoria y es proporcional a n . En lo que sigue, se ignorará la primera componente ya que es la que proviene de la envolvente y no del espectro en sí.

Si existen correlaciones entre los diferentes niveles de energía, la forma de la función $\langle |F(y)|^2 \rangle$ se verá afectada cerca del origen. Esto es debido a lo siguiente. Para estudiar las posiciones y correlaciones entre los diferentes niveles de energía se puede trabajar con las funciones de correlación $R_1(x_1)$ y $R_2(x_1, x_2)$, siendo $R_n(x_1, \dots, x_n)$ la densidad de probabilidad de que haya un nivel alrededor de x_1 , otro alrededor de x_2 , etc⁽¹⁹⁾. $R_2(x_1, x_2)$ se puede expresar como

$$R_2(x_1, x_2) = R_1(x_1)R_1(x_2)(1 - Y_2(x_1, x_2)), \quad (\text{A1.16})$$

siendo $Y_2(x_1, x_2)$ la llamada función cluster de dos puntos⁽²⁰⁾. Si Y_2 es cero, el que haya un nivel alrededor de x_1 es independiente del que haya o no otro nivel alrededor de x_2 . En cambio, si Y_2 es diferente de cero, el que haya un nivel alrededor de x_1 sí tiene que ver con el que haya o no otro nivel alrededor de x_2 ; esto es, el que Y_2 sea distinta de cero indica que los niveles están correlacionados.

Ahora, las funciones de correlación junto con la ecuación A1.16 contienen información acerca de los espectros en el sentido de que indican cual es el comportamiento promedio de los diferentes pares de niveles en lo que se refiere a sus posiciones y a la correlación entre ellos. Pero la función $\langle f(x) \rangle$ también contiene esta información ya que es el promedio sobre el ensemble de grandes porciones de los espectros y por lo tanto indica cual es el comportamiento promedio de las posiciones y correlaciones entre los diferentes niveles de energía. Así, es de esperarse que las transformadas de Fourier de las funciones de correlación contengan información similar a la contenida en la función $\langle |F(y)|^2 \rangle$.

Debido a esto, si existe correlación entre los diferentes niveles de energía entonces la componente de $\langle |F(y)|^2 \rangle$ asociada con el espectro contendrá un factor $(1-b(y))$, siendo $b(y)$ la

transformada de Fourier de Y_2 y esto traerá como consecuencia que la forma de $\langle |F(y)|^2 \rangle$ se vea modificada cerca del origen en relación con la forma que tendría si Y_2 , y por lo tanto $b(y)$, fueran iguales a cero.

También, resulta que la forma de $\langle |F(y)|^2 \rangle$ alrededor del origen está muy relacionada con la forma de la varianza $\Sigma^2(L)$ para valores grandes de L ⁽²³⁾. De hecho, existen relaciones entre la ordenada y la pendiente de $\langle |F(y)|^2 \rangle$ y la forma de $\Sigma^2(L)$ para valores de L grandes. Debido a esto, se deduce que la función $\langle |F(y)|^2 \rangle$ cerca del origen contiene información acerca de lo que ocurre en grandes porciones de los espectros; esto es, contiene información acerca de las correlaciones de largo alcance entre los diferentes niveles de energía.

De esta forma se concluye que la forma de $\langle |F(y)|^2 \rangle$ cerca del origen indica si existen correlaciones de largo alcance entre los niveles de energía que están dentro de la región de interés y es por esto que se estudian las transformadas de Fourier de los espectros aleatorios.

En las Figuras A1.7 y A1.8 se muestra la función $\langle |F(y)|^2 \rangle$ para el caso Poisson y para el caso del GOE, respectivamente.

Como en el caso Poisson no hay correlaciones entre los diferentes niveles de energía, $Y_2(x_1, x_2)$ y $b(y)$ son iguales a cero. En el caso del GOE, $Y_2(x_1, x_2)$ es diferente de cero, al igual que $b(y)$ y es por esto que la forma de $\langle |F(y)|^2 \rangle$ es diferente. El agujero que aparece indica que existen correlaciones de largo alcance entre los diferentes niveles de energía. En ambos casos se ha excluido la parte asociada con la transformada de la envolvente. Esto no afecta la forma de las curvas, ya que esta componente sólo contribuye para valores de y muy cercanos al origen.

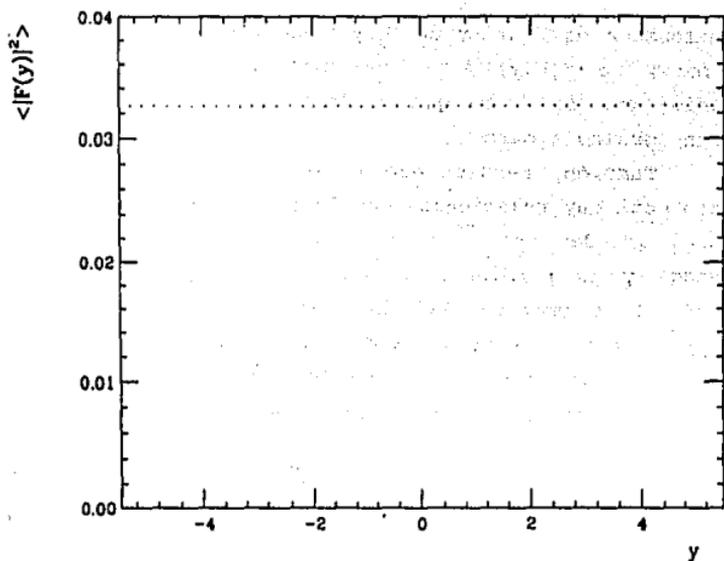


Fig. A1.7. Media del módulo al cuadrado de las transformadas de Fourier $F(y)$ de espectros de tipo Poisson (—).

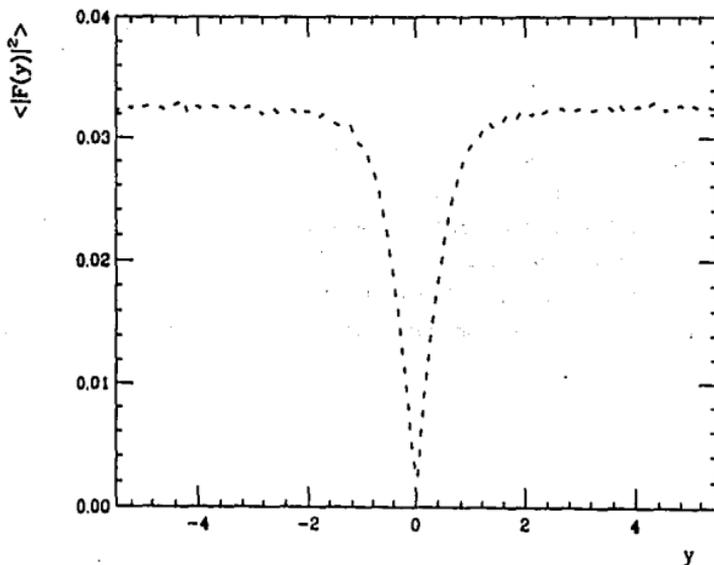


Fig. A1.8. Media $\langle |F(y)|^2 \rangle$ del módulo al cuadrado de las transformadas de Fourier de espectros del GOE (- -).

Con esto se ve que al estudiar las cantidades estadísticas aquí presentadas se puede obtener información acerca de varias de las propiedades de espectros aleatorios en general, como lo son la variabilidad de la distancia entre niveles, la variabilidad del número de niveles dentro de cierto intervalo o las correlaciones de corto y largo alcance, por mencionar algunas de ellas. De esta forma se puede conocer cuál es la apariencia general de dichos espectros y ésto en turno ayuda a conocer ciertas propiedades de los sistemas a los que están asociados.

Por último cabe mencionar que existen otras cantidades de interés aparte de las arriba mencionadas⁽²¹⁻²²⁾ las cuales también dan información acerca de espectros aleatorios. Sin embargo, para los propósitos de este trabajo las más importantes son las que aquí se han presentado.

REFERENCIAS

- [1]. *The Traveling Salesman Problem, A Guided Tour of Combinatorial Optimization*, Eds. E. L. Lawler, J. K. Lenstra, A. H. G. Rinoooy Kan and D. B. Shmoys, John Wiley & Sons, Chichester, 1985.
- [2]. S. Kirkpatrick, C. D. Gelatt, Jr., M. P. Vecchi, *Science*, **220** (1983) 4598.
- [3]. S. Kirkpatrick, *J. of Statistical Physics*, **34** (1984) 975.
- [4]. W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, *Numerical Recipes in Fortran, The Art of Scientific Computing* (Cambridge University Press, New York, 1992).
- [5]. S. Lin, *The Bell System Technical Journal*, **44** (1965) 2245.
- [6]. T. Sun, P. Meakin and T. Jøssang, *Physica A*, **199** (1993) 232.
- [7]. G. Venkataraman and G. Athithan, *Pramāna-J. Phys.*, **36** (1991) 1.
- [8]. K. H. Fischer & J. A. Hertz, *Cambridge Studies in Magnetism, Spin Glasses 1* (Cambridge University Press, Cambridge, 1993).
- [9]. D. Chowdhury, *Spin Glasses and Other Frustrated Systems* (World Scientific, Singapore, 1986).
- [10]. R. B. Stinchcombe, *Cuadernos del Instituto de Investigaciones en Materiales 4, Fenómenos Críticos en Sistemas Puros y Desordenados* (UNAM, México D.F., 1984).
- [11]. N. W. Ashcroft, N. D. Mermin, *Solid State Physics* (Holt, Rinehart and Winston, New York, 1976).
- [12]. I. M. Lifshits, S. A. Gredeskul and L. A. Pastur, *Introduction to the Theory of Disordered Systems* (John Wiley & Sons, New York, 1988).

- [13]. K. Huang, *Statistical Mechanics* (John Wiley & Sons, New York, 1987).
- [14]. J. Vannimenus and M. Mézard, *J. Physique Lett.*, **45** (1984) L-1445.
- [15]. S. Kirkpatrick and G. Toulouse, *J. Physique*, **46** (1985) 1277.
- [16]. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, *The Journal of Chemical Physics*, **21** (1953) 1087.
- [17]. E. Bonomi and J. Lutton, *SIAM Review*, **26** (1984) 551.
- [18]. B. Derrida, *Phys. Rev. B.*, **24** (1981) 2613.
- [19]. H. Orland, *J. Physique Lett.*, **46** (1985) L-763.
- [20]. M. L. Mehta, *Random Matrices* (Academic Press, Boston, 1991).
- [21]. T. A. Brody, J. Flores, J. B. French, P. A. Mello, A. Pandey, S. S. M. Wong, *Reviews of Modern Physics*, **53** (1981) 385.
- [22]. O. Bohigas, *Random Matrix Theories and Chaotic Dynamics* (Elsevier Science Publishers, 1991).
- [23]. *Lecture Notes in Physics 263, Quantum Chaos and Statistical Nuclear Physics*, Eds. T. H. Seligman and H. Nishioka, Springer-Verlag, Berlin, 1986.
- [24]. A. M. Mood, F. A. Graybill, D. C. Boes, *Introduction to the Theory of Statistics* (Mc Graw-Hill, Singapore, 1974).
- [25]. F. J. Dyson, *J. Math. Phys.*, **3** (1962) 140.
- [26]. C. E. Porter and N. Rosenzweig, *Ann. Acad. Sci. Fennicae, Serie A, VI Physica*, **44** (1960) 1.
- [27]. *Statistical Properties of Nuclei*, Ed. J. B. Garg, Plenum Publishing Corporation, New York.
- [28]. E. Kreyszig, *Introducción a la Estadística Matemática* (Limusa, México, 1989).