

Ajustes orbitales y dinámica de sistemas planetarios extrasolares

Cristian Andrés Giuppone

Presentado ante la Facultad de Matemática, Astronomía y Física
como parte de los requerimientos para obtener el título de
Doctor en Astronomía de la

Universidad Nacional de Córdoba

Febrero, 2011

© Famaf - UNC, 2011

Director: Dr. Cristián Beaugé

Agradecimientos

Sin dudas la finalización de esta etapa de mi vida es consecuencia de un sin número de factores. La astronomía es también sin dudas una pasión para mí, algo que hago con mucho cariño y perseverancia. He disfrutado mucho de esta etapa de mi carrera, en perspectiva, tanto de los triunfos como de las frustraciones. Es una etapa durante la cual además de formarme como investigador, se ha moldeado parte de mi carácter. Como persona estoy seguro que no hubiera llegado a esta meta sin tantas personas que me rodearon y contuvieron en diversas ocasiones, a ellos quisiera especialmente agradecerles.

A mis padres por el apoyo brindado, incondicionales mamá y papá que desde aquellos primeros años en Córdoba entendieron que mi estudio es mi vocación; también a mis hermanos Ale y Leo porque están en esos momentos que se necesitan.

Sin dudas mi director Cristián Beaugé por recibirme para trabajar con él y porque supo darme ese apoyo que necesitaba, cuando más lo necesitaba. El fue quién me brindó un entorno cómodo de trabajo donde me sentí y siento libre de estar e investigar.

En mi entorno más íntimo, las personas que han estado conmigo estos años en todo momento Franco y Yamis (G y Y), por que ellos me demostraron lo importante de tener buenos amigos.

A mi grupo de trabajo, principalmente aquellos con los que más tiempo he compartido y deliberado: Jorge y Martín y a todos los integrantes del Chalet!

Al grupo de *Dinâmica de Sistemas Planetários e Mecânica Celeste* del IAG en São Paulo, especialmente los profesores Sylvio y Tatiana porque también participaron activamente en mi proceso de formación y a Adrián y Marcos con los cuales compartimos muchos momentos amenos y de discusión en nuestro trabajo. Esta interacción fue realizada en gran parte gracias a la financiación con un convenio SECYT-CAPES.

Al Observatorio/Famaf que me brindó un espacio y un lugar donde desarrollar mi carrera y finalmente a CONICET por que financió mi investigación.

..a todos....MUCHAS GRACIAS!!!!

RESUMEN

La determinación de parámetros orbitales de sistemas exoplanetarios a partir de datos de velocidad radial es un proceso complejo. Mientras que los primeros candidatos a exoplanetas eran rutinariamente observados durante varios períodos antes de ser publicados, en la actualidad los candidatos planetarios son publicados incluso cuando las observaciones cubren tan solo un período.

Un requisito fundamental para cualquier estudio dinámico es la disponibilidad de valores confiables para las masas y parámetros orbitales. Debido a que muchas veces los resultados de los ajustes orbitales son muy sensibles al método numérico y/o conjunto de datos, precisamos una estrategia que pueda identificar los extremos globales de la función residual y debemos estimar de alguna forma los niveles de confianza en los parámetros determinados. Las características altamente no-lineales de las ecuaciones involucradas impiden asumir errores Gaussianos en sus parámetros. Es por ello que las desviaciones estándar que se encuentran en publicaciones, generalmente muestran valores subestimados. La ausencia de una solución orbital confiable ha dado origen a estudios dinámicos más generales, que no se restringen sólo al análisis de una única condición inicial.

El presente trabajo explica y desarrolla estrategias de búsqueda de mínimos globales para funciones no-lineales como así también las estrategias de evaluación de errores. Particularmente centramos nuestra atención en el sistema HD82943 por dos razones: primero porque está demostrado que los mejores ajustes orbitales de este sistema muestran un comportamiento dinámicamente inestable, implicando que no es consistente con la configuración real de los planetas; segundo, porque ya fue estudiado por varios autores proveyendo una buena base con la cual comparar nuestros resultados.

Finalmente en la última parte del presente trabajo, abordamos la generación de datos sintéticos para planetas que se encuentran en resonancias de movimientos medios. Nuestros resultados permiten explicar algunas de las posibles características en la distribución de elementos orbitales, que indican tendencias a sobreestimar parámetros como excentricidades y amplitudes de oscilación de los ángulos resonantes.

La metodología desarrollada fue aplicada a otros sistemas planetarios y a otras técnicas de detección. El presente trabajo abrió muchas posibilidades de investigación que se encuentran en curso actualmente, como la determinación de parámetros planetarios en sistemas estelares binarios, análisis de detección para planetas en resonancia de movimientos medios 1:1 y análisis de las variaciones entre los tiempos consecutivos de tránsitos de un planeta, debido a la presencia un compañero no identificado.

Palabras Clave: Planetas extrasolares. Detección. Dinámica. Algoritmos genéticos. Minimización multiparamétrica no-lineal.

Clasificaciones:

97.82.-j Extrasolar planetary systems.

97.82.Cp Photometric and spectroscopic detection

95.75.-z Observation and data reduction techniques; computer modeling and simulation

02.50.Ng. Monte Carlo methods in probability theory and statistics

02.60.-x. Numerical methods (mathematics)

ABSTRACT

The determination of mass and orbital parameters for exoplanetary systems from radial velocity data is a complex process. While early planet candidates were routinely observed for multiple periods before publication, recently the planet candidates have been published when observations span only a single orbital period.

A fundamental requirement for any dynamical study is the availability of reliable values for these parameters. Since the outcome of orbital fits are usually very sensitive to numerical method and/or the data set, we need a strategy that can identify the global extreme for residual function and estimate the levels of confidence in the orbital parameter space. The highly nonlinear character of the equations, prevent the usual assumption of a Gaussian distribution of errors. This is why the standard deviations found in publications generally show underestimated values. The absence of a reliable orbital solution has given rise to more general dynamical studies, which are not restricted to the analysis of a single initial condition.

This work explains and develops strategies for finding global minima for non-linear functions as well as assessment strategies for errors evaluations for these problems. In particular we focus on the system HD82943 for two reasons: first because it is shown that best fits in the system systematically correspond to dynamically unstable behavior, implying that is not consistent with the actual configuration of the planets. Second, because it was previously studied by several authors providing a good base from which to compare our results.

Finally in the last part of this work, we address the generation of synthetic data for planets that are in mean-motion resonance. Our results explain some of the possible features in the distribution of orbital elements that show trends to overestimate parameters such as eccentricities and amplitudes of oscillation of the resonant angles.

The tools developed throughout this work were applied to other planetary systems and other detection techniques. This work has opened many possibilities for research that are currently under way, such as the determination of planetary parameters in tight binary stellar systems, analysis of detection of planets in 1/1 mean motion resonance and analysis of changes between consecutive transit times of a planet, due to the presence of an unidentified companion.

Índice general

1. Introducción	6
1.1. Descubrimientos de exoplanetas	7
1.2. ¿Qué es un exoplaneta?	9
1.3. Técnicas de detección de exoplanetas	11
1.3.1. Astrometría	11
1.3.2. Imagen Directa	13
1.3.3. Pulsar Timing	14
1.3.4. Microlensing	17
1.3.5. Tránsitos	19
1.3.5.1. Ecuaciones para resolver la curva de luz	21
1.3.5.2. Parámetros orbitales a partir de observables	23
1.3.6. Velocidades Radiales	25
1.3.6.1. Técnica con celda de absorción	26
1.3.6.2. Técnica de referencia simultánea	29
1.3.6.3. Errores de las mediciones y Jitter estelar	29
1.3.6.4. Velocidad Radial en función de parámetros orbitales	31
1.3.6.5. Expansiones Elípticas	36
1.3.7. Efecto Rossiter-McLaughlin	36
1.3.7.1. Ecuaciones para modelar el efecto RM	39
1.4. Efectos de Selección y Límites de detección	42
2. Obtención de los parámetros planetarios	45
2.1. Métodos de Fourier	47
2.1.1. Principios de Fourier	47
2.1.2. Transformada discreta de Fourier	49
2.1.3. Determinación de Parámetros orbitales con análisis de Fourier	51
2.1.4. Análisis de Fourier en datos no equiespaciados	53
2.1.5. Diferencias entre DFT y DCDFE	56
2.1.6. Filtrado de armónicos	57
2.1.7. Probabilidades de detección y límites de confianza	58
2.2. Métodos Cuadrados Mínimos	61
2.2.1. El Principio de Cuadrados Mínimos	63
2.2.2. Método del Gradiente Conjugado	67

2.2.3.	Método de Levenberg-Marquardt	68
2.2.4.	Algoritmo Genético	71
2.2.4.1.	Principios del Algoritmo Genético	72
2.2.4.2.	Algoritmo genético esquemáticamente	74
2.2.4.3.	Parámetros de inicialización	76
2.2.5.	Método Simplex	80
2.2.6.	Comparación Simplex Iterativo y Algoritmo Genético	81
2.2.7.	Simulated Annealing	82
2.3.	Ajustes Dinámicos	84
2.4.	Obtención de parámetros planetarios para el sistema HD82943	85
2.4.1.	Reseña histórica	85
2.4.2.	Velocidades radiales y mejores ajustes	88
3.	Estimación de Incertezas	90
3.1.	Intervalo de Confianza	91
3.2.	Matriz de Fisher	95
3.3.	Cadenas Markov con MonteCarlo	96
3.4.	Generación de Datos Sintéticos: Resampling	100
3.4.1.	Análisis dinámicos sobre las regiones	102
3.5.	Ajustes en una Grilla	104
3.6.	Ajustes Orbitales con conjuntos de datos reducidos: Jackknife	106
3.6.1.	Análisis dinámicos en soluciones Jackknife y Grilla	109
4.	Búsqueda de Planetas Adicionales	113
4.1.	Búsqueda de un tercer planeta en HD82943	114
4.2.	Soluciones de Laplace en HD82943	117
4.3.	Otros ejemplos de Resonancia de Laplace	120
5.	Efectos de Selección con Velocidades Radiales	125
5.1.	Detectabilidad de exoplanetas	126
5.1.1.	Detectabilidad de 1 exoplaneta	129
5.1.2.	Resultados de detectabilidad de 1 exoplaneta	132
5.2.	Sistemas de exoplanetas	135
5.2.1.	Detectabilidad de 2 Planetas Resonantes	137
5.2.2.	Planeta Interior más masivo	140
5.2.3.	Planeta Exterior más masivo	142
5.2.4.	Estimación de errores en los parámetros ajustados	144
5.2.4.1.	Planeta exterior más masivo	145
5.2.5.	Planetas de masas iguales	148
5.2.5.1.	Planeta exterior menos masivo	149
5.2.6.	Aplicación a exosistemas reales	150
6.	Conclusiones	153

Capítulo 1

Introducción

La pregunta acerca de la existencia de otros mundos ha estado presente en la historia de la humanidad por milenios, pero sólo la evidencia científica reciente pudo confirmar lo que muchos otros han anticipado: los planetas existen y son comunes fuera del Sistema Solar. El primer descubrimiento de otro sistema planetario fue en 1992 con la detección de dos planetas tipo terrestre (siendo sus masas de $2.8 M_{\oplus}$ y $3.4 M_{\oplus}$) orbitando el pulsar PSR 1257+12. Sin embargo, este hallazgo no recibió toda la atención que se podría haber esperado, probablemente debido al hecho de que estos dos planetas orbitan una “estrella muerta” muy diferente al Sol y poseen menos probabilidad de albergar vida en su vecindad.

Recién cuando Michel Mayor y Didier Queloz (1995) descubrieron el primer planeta extrasolar orbitando una estrella de secuencia principal, el mundo abrió los ojos a este nuevo hito histórico. Desde entonces se han sucedido en ritmo creciente los descubrimientos de nuevos planetas y hasta abril de 2010 se han detectado 386 sistemas planetarios que contienen un total de 453 cuerpos planetarios, con cuarenta y cinco sistemas que contienen más de un planeta. El catálogo actual de planetas extrasolares es actualizado en forma continua e incluye una inmensa cantidad de sistemas que poseen las más variadas características orbitales. Más allá de los esfuerzos en la búsqueda de planetas del tamaño terrestre en la actualidad, los datos observacionales en sistemas múltiples tienen implicaciones importantes para modelos de formación planetaria, proveyendo indicios importantes acerca del rol que tienen diferentes mecanismos así como su evolución dinámica.

Sin dudas es necesario el conocimiento de las técnicas que posibilitan dar las características del movimiento de estos nuevos mundos. Como es de esperar, estas técnicas de observación astronómica tienen asociadas errores y limitaciones. Es objetivo del presente capítulo describir tales técnicas y repasar cuales son las características planetarias que podemos obtener con ellas. Más adelante nos dedicaremos al análisis de los datos y a extraer la máxima cantidad de información disponible en los mismos.

1.1. Descubrimientos de exoplanetas

Las especulaciones de planetas orbitando otras estrellas datan de, al menos, principios del siglo XVIII. En el artículo *General Scholium*, publicado en 1713, Isaac Newton razona: “y si las estrellas fijas son los centros de otros sistemas parecidos, estos se formaron y están sujetos al dominio de sus estrellas”¹.

Aseveraciones acerca de la detección de exoplanetas datan del siglo XIX. Una de las primeras involucra a la estrella binaria *70 Ophiuchi*. En 1855 el Capitán Jacob, trabajando para el Observatorio de Madras, reportó que las anomalías orbitales hacían “altamente probable” que exista un cuerpo planetario en este sistema. Más tarde, See (1896) de la Universidad de Chicago, afirmó que las anomalías orbitales probaban la existencia de un cuerpo oscuro en el sistema de *70 Ophiuchi* con período de 36 años alrededor de una de las estrellas. Sin embargo Ray Moulton seguidamente publicó un trabajo probando que dicho sistema de 3 cuerpos con aquellos parámetros orbitales sería altamente inestable (Sherril, 1999). Durante 1950-1960 Van de Kamp hizo una prominente serie de anuncios de detección, esta vez para planetas orbitando la estrella *Barnard*. Más tarde se corroboró que estos reportes iniciales eran erróneos. Recientemente Bailes et al. (1991) anunciaron descubrir un planeta en órbita alrededor del pulsar PSR 1829-10, usando variaciones en el tiempo entre los pulsos. El anuncio recibió una atención inusitada, pero pronto el equipo de trabajo se retractó (Lyne y Bailes, 1992).

El primer descubrimiento confirmado fue realizado en 1988 por el grupo de astrónomos canadienses Campbell, Walker y Yang. Sus observaciones de velocidad radial sugerían que un planeta orbitaba la estrella *Gamma Cephei*. Los descubridores se mantuvieron cautelosos acerca del anunciar una detección **planetaria** real y un escepticismo generalizado persistió en la comunidad astronómica por varios años. Esto fue fundamentalmente debido a que las observaciones estaban en el límite de las capacidades instrumentales de la época. Otra fuente de confusión fue que algunos posibles planetas podrían ser de hecho enanas marrones, objetos intermedios entre los planetas y las estrellas. Subsecuentes trabajos de Lawton y Wright (1989) y Walker et al. (1992) trataron de esclarecer la situación del sistema. Sin embargo, se debió esperar hasta 2003, donde el grupo de Hatzes confirmó la existencia de un planeta, con las técnicas mejoradas de velocidad radial.

A principios de la década del 90 los radio-astrónomos Wolszczan y Frail (1992) anunciaron el descubrimiento de planetas alrededor del pulsar PSR 1257+12. El descubrimiento fue rápidamente confirmado y es generalmente considerado la primera confirmación definitiva de exoplanetas. Se cree que estos planetas alrededor del pulsar se formaron de remanentes inusuales de la supernova que produjo el pulsar, en una segunda etapa de formación planetaria, o incluso podrían ser los remanentes rocosos de planetas gigantes gaseosos que sobrevivieron a la explosión de la supernova y luego cayeron en espiral hacia sus órbitas actuales.

El 6 de octubre de 1995 los astrónomos del Observatorio de Ginebra, Didier Queloz y Michael Mayor, anunciaron la primera detección definitiva de un exoplaneta orbitando una estrella de secuencia principal, *51 Pegasi*, inaugurando la era moderna del descubrimiento de exoplanetas. Los avances tecnológicos, más notables en las técnicas de alta resolución

¹el texto original puede ser consultado en:
[http://en.wikisource.org/wiki/The_Mathematical_Principles_of_Natural_Philosophy_\(1729\)/General_Scholium](http://en.wikisource.org/wiki/The_Mathematical_Principles_of_Natural_Philosophy_(1729)/General_Scholium).

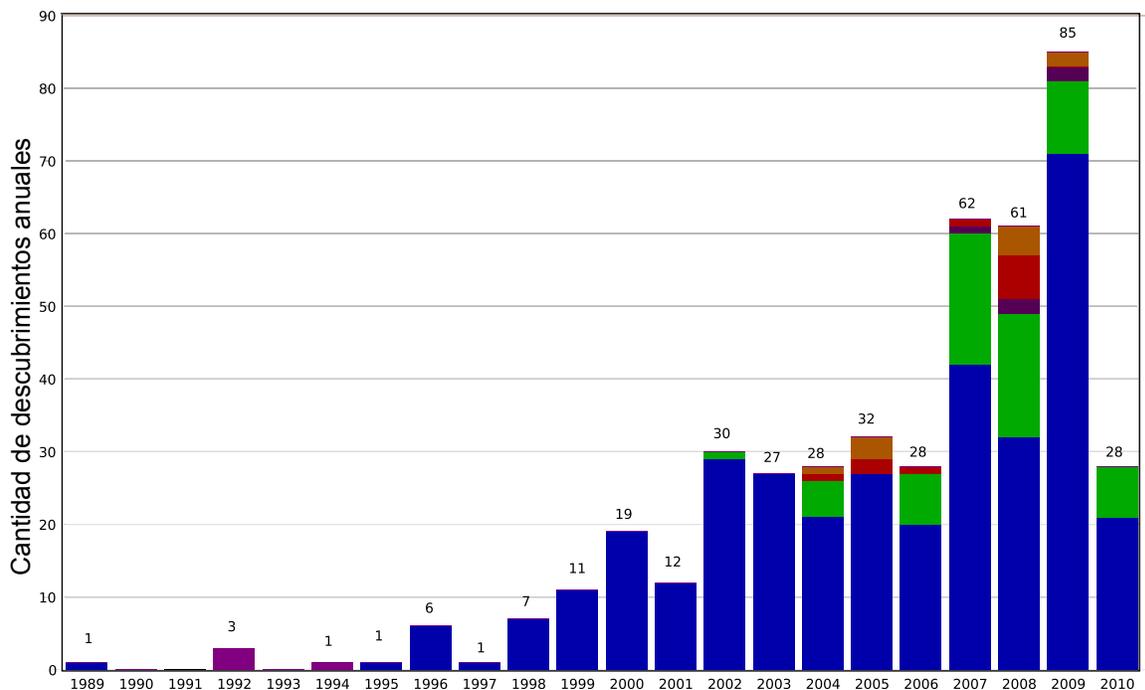


Figura 1.1: Histograma de detecciones anuales de exoplanetas hasta abril del 2010. En el año en curso se estima un total de 120 planetas a descubrir (ver texto para más detalle).

espectroscópica, abrieron el camino hacia una gran cantidad de nuevos descubrimientos.

Si bien las técnicas espectroscópicas dominaron los descubrimientos en los últimos 15 años, durante los últimos 4 años se han mejorado las técnicas de detección por tránsito y se esperan aún más descubrimientos una vez procesados los datos de las misiones espaciales Corot y Kepler.

Al 14 de abril de 2010 se conocen 379 sistemas planetarios totalizando 446 planetas y con 45 de esos sistemas planetarios con más de un cuerpo alrededor de la estrella central. Estos números se incrementan continuamente y la Figura (1.1) muestra la cantidad de descubrimientos por año desde las primeras detecciones. Los colores están asociados a las técnicas utilizadas, donde vemos que el azul (para velocidad radial) es ampliamente dominante con casi un 85 % del catálogo, seguido por el verde (para tránsitos) representando casi un 10 % de la muestra. Las otras técnicas como variaciones de tiempo de pulsar (5 planetas, en violeta), detección por imagen directa (11 planetas, en marrón) y microlente gravitacional (9 planetas, en naranja) tienen representación despreciable. En la actualidad dos páginas de internet mantienen actualizado el catálogo de exoplanetas, constituidas por recompilaciones de datos generales:

- **www.exoplanet.eu** a cargo de Jean Schneider del Observatorio de París.
- **www.exoplanets.org** actualizada por Jason Wright, Geoff Marcy y el consorcio del *California Planet Survey*.

Más allá de que el número total de exoplanetas no coincida entre ellas, cabe destacar que si

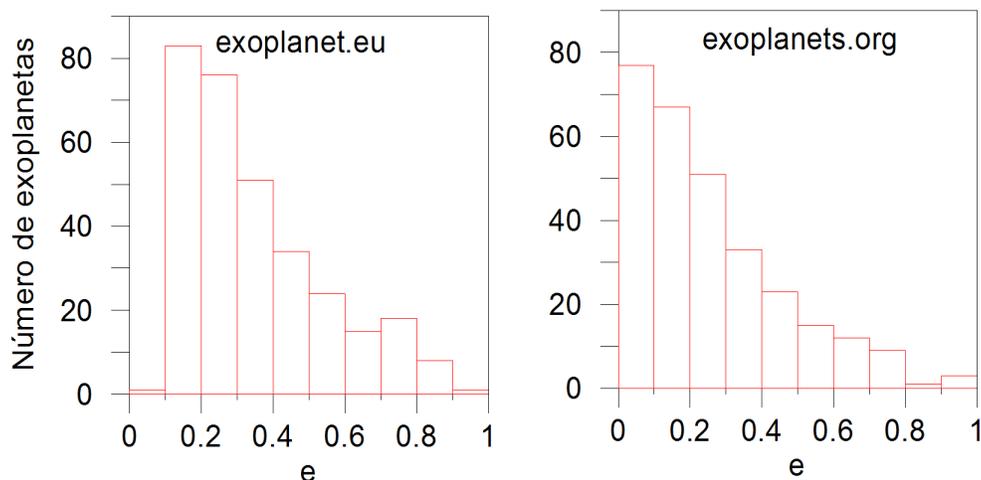


Figura 1.2: Histogramas de excentricidades de los exoplanetas según datos de los 2 catálogos mencionados. Sólo se tomaron exoplanetas detectados con Velocidades Radiales. El primer bin del gráfico de la izquierda muestra la mayor diferencia apreciable entre ambos catálogos (Realizado el 14/4/2010).

se comparan los parámetros orbitales entre ambas fuentes existen diferencias significativas; por lo tanto cuando se utilicen parámetros para realizar estadística se deben reconfirmar las fuentes de los datos. A modo de ejemplo mostramos en la Figura 1.2 los histogramas de todos los planetas detectados con técnica de velocidad radial al 14 de abril de 2010. Existe una diferencia apreciable en el primer bin del histograma: mientras que en el gráfico de la izquierda aparentemente no existen exoplanetas con excentricidades desde 0 hasta 0.1, en el histograma de la derecha se encuentran unos 75 exoplanetas dentro del mismo rango de excentricidades².

1.2. ¿Qué es un exoplaneta?

Desde los primeros descubrimientos de planetas alrededor de otras estrellas se planteó la necesidad de definir más precisamente que es un planeta. Se podría definir básicamente un exoplaneta como un planeta orbitando a una estrella distinta de nuestro Sol. Por lo tanto primero debemos definir que es una estrella.

Tradicionalmente se definía una estrella como: un objeto capaz de realizar fusión de hidrógeno en su interior. Sin embargo estrellas denominadas *enanas marrones* siempre desafiaron esta distinción: demasiado pequeñas para sostener la fusión del hidrógeno, aunque tienen la habilidad de fusionar el deuterio. No obstante debido a la escasez de tal isótopo, este proceso perdura sólo una ínfima parte de la vida de la estrella (unos 100 millones de años) y por lo tanto cuando son descubiertas han cesado dicha fusión mucho tiempo antes. La masa mínima de las enanas marrones es unas 13 masas de Júpiter (M_{jup}), para objetos de metalicidad solar. Spiegel et al. (2010) mostraron, mediante la utilización una variedad de modelos de objetos subestelares, que este límite puede variar desde unas $11M_{jup}$ hasta $16M_{jup}$ depen-

²Dicha discrepancia fue personalmente discutida con J. Wright quién no supo explicar su origen.

diendo de su metalicidad. Debido a que muchas *enanas marrones* orbitan otras estrellas y a que no producen reacciones nucleares al momento de detectarse deberían ser descritas como planetas según la definición del primer párrafo.

Recientemente, se encontraron objetos en cúmulos estelares jóvenes que tienen masas menores que las requeridas para sostener cualquier tipo de reacción. Ellos parecen *flotar* libres en el medio **intra-cúmulo**, debido a que no están ligados gravitacionalmente a otros miembros del mismo. Estos objetos han sido denominados “planetas flotantes libres”, porque según algunas hipótesis de su formación pueden haber sido eyectados del sistema estelar donde se formaron (Boss et al. 2003). Sin embargo, otra posibilidad es que su formación sea causada por un auto-colapso a partir de una nube primordial (Whitworth y Zinnecker, 2004).

Debido a estas ambigüedades, en 2003, la IAU (Unión Astronómica Internacional) definió en su grupo de planetas extrasolares (*WGESP*) lo siguiente:

En lugar de tratar de construir una definición detallada de qué es un planeta, designado a cubrir todas las posibilidades futuras, el *WGESP* acordó restringirse a si mismo en desarrollar una definición aplicable a los casos reclamados como una detección, es decir: los relevamientos de velocidad radial para compañeros de estrellas tipo solar (en su mayoría) y los relevamientos de imagen directas para *objetos flotantes libres* en cúmulos de estrellas jóvenes. A medida que se hagan nuevos descubrimientos en el futuro, el *WGESP* va a definir sus méritos individuales y circunstancias y tratará de ajustar los nuevos objetos en la definición de “planeta”, revisándola cuanto sea necesario. Este es un enfoque gradual de una definición que evoluciona, guiada por las observaciones que son las que definen todo finalmente. Enfatizando que es sólo una definición de trabajo, sujeta a cambios a medida que se aprenda más sobre el censo de compañeros de baja masa, se acordaron las siguientes afirmaciones:

- Los objetos con masas reales abajo del límite de masa para fusión termonuclear del deuterio (actualmente calculado en $13 M_{jup}$ para objetos de metalicidad solar), que orbiten estrellas o remanentes de estrellas son **planetas** sin importar cómo se hayan formado. La masa/tamaño mínima requerida para un objeto extrasolar para ser considerado un planeta debe ser la misma que la usada en nuestro Sistema Solar.
- Los objetos subestelares con masas superiores al límite de masas para la fusión termonuclear del deuterio son *enanas marrones*, sin importar cómo se formen o dónde se localicen.
- Los *objetos flotantes libres* en cúmulos estelares jóvenes, con masas inferiores al límite de fusión termonuclear del deuterio **no son planetas**, sino *sub-enanas marrones* o cualquier nombre que sea más apropiado (en algunos trabajos recientes se denominan *planemos*).

Al momento, 7 años después de la escueta definición de trabajo, no se ha avanzado sobre el tema y parece ser suficiente dicha definición. Por momentos la masa mínima está lejos del alcance de las capacidades actuales de detección, mientras que la masa máxima parece ser un buen punto de partida para definiciones futuras.

1.3. Técnicas de detección de exoplanetas

Si bien de forma teórica hay desarrolladas varias técnicas de detección, prácticamente el 95 % de los exoplanetas detectados son por técnicas de Velocidad Radial y Tránsitos (unos 440). Mientras tanto Imagen Directa (11), Microlente Gravitacional (*Microensing*, 9) y Variación en el tiempo de pulsaciones del pulsar (Pulsar Timing, 8) representan sólo el 5%. Existen otras técnicas como Astrometría, que aún no producen descubrimientos, pues las limitaciones tecnológicas no lo permiten. Cada una de las técnicas tiene sus puntos fuertes y débiles pero por el momento basta con decir que ninguna técnica por sí sola brinda la información total deseada y la combinación de varias es necesaria para estudiar distintos aspectos de los planetas extrasolares y sistemas exoplanetarios. Haremos un breve repaso por las técnicas haciendo especial hincapié en Velocidad Radial y Tránsitos debido a su importancia.

1.3.1. Astrometría

Este método consiste en medir precisamente la posición de la estrella en el cielo y medir como la misma cambia en el tiempo. Si la estrella posee un planeta, la interacción gravitatoria causará que la estrella se mueva alrededor del centro de masa del sistema. Debido a que la estrella es mucho más masiva que el planeta, frecuentemente el centro de masa se encuentra dentro del radio de la estrella. Si bien la técnica astrométrica es la más antigua de las desarrolladas para buscar exoplanetas y se hizo popular debido a su éxito en caracterizar estrellas binarias astrométricas, ninguno de los anuncios sobrevivió al escrutinio de otros astrónomos. Desafortunadamente los cambios en la posición estelar son tan pequeños y las distorsiones atmosféricas como los errores sistemáticos tan grandes que los telescopios en tierra no pueden producir mediciones lo suficientemente precisas. En 2002, el telescopio espacial Hubble, utilizó posiciones astrométricas para caracterizar un planeta previamente descubierto alrededor de la estrella *Gliese 876*.

Una de los principales potenciales de la técnica es su ventaja a ser más sensible en la detección de planetas con órbitas más alejadas, haciéndola complementaria a otros métodos. Sin embargo se requieren tiempos de observación largos (posiblemente décadas) para que puedan ser detectados y completar observaciones de un período. En contraposición debido a las limitaciones en la calidad de las observaciones sólo las estrellas más cercanas (20-25 pc) son factibles de ser observadas con esta técnica desde la Tierra.

La Figura 1.3 muestra el desplazamiento del Sol debido al efecto del movimiento planetario (de Júpiter y Saturno principalmente), visto a una distancia de 10 pc; esta figura se denomina “wobble” de la estrella. La amplitud de este desplazamiento es de 500 microsegundos de arco (*mas*). El desplazamiento debido a la Tierra presentaría una amplitud de 0.3mas a la misma distancia. En la misma figura, a la derecha se observa el wobble de *HD 33636*, siendo esta una de las mejores determinaciones astrométricas que se poseen, utilizando el Telescopio Espacial Hubble. Se observa que el movimiento detectado es aproximadamente un orden de magnitud mayor que el producido a nuestro propio Sol.

Cada observación de la estrella debe proporcionar coordenadas celestes (α, δ), pero en la imagen CCD poseen coordenadas estándares (ξ, η) que se encuentran en el plano tangente a la

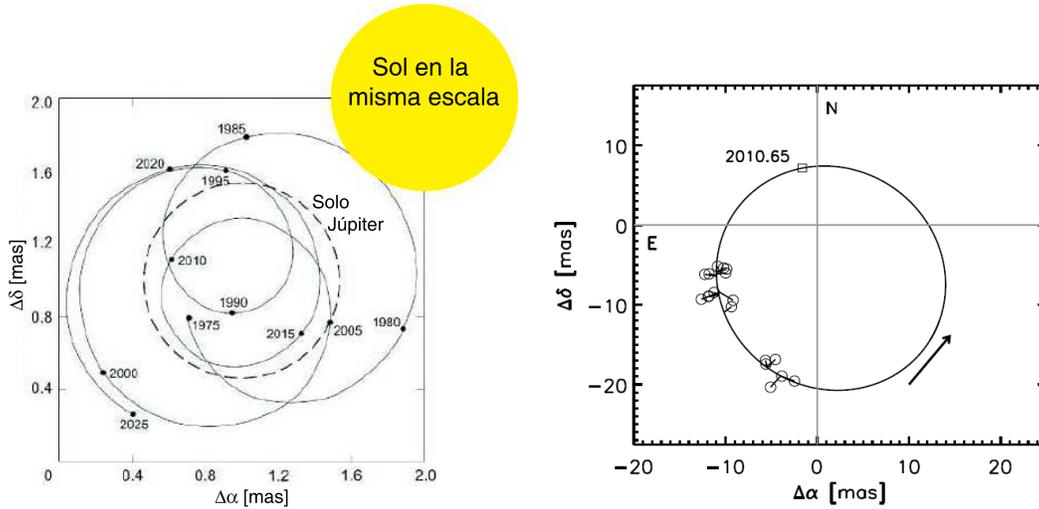


Figura 1.3: Izquierda: Movimiento astrométrico del centro solar (wobble) alrededor del baricentro del Sistema Solar debido a todos los planetas en un período de 50 años, desde una distancia de 30 años luz (Jones, 2008). El círculo a trazos muestra el movimiento del Sol sólo debido a Júpiter. El tamaño del disco del Sol es mostrado en la esquina superior derecha para comparación. Derecha: wobble de *HD 33636*, una estrella central de $1 M_{\odot}$ a 28.7 pc de distancia, se deduce la presencia de un cuerpo adicional subestelar de $142 M_{jup}$, que queda fuera de las definiciones de planeta. Reproducido de Bean et al (2007).

esfera celeste (ver, Kovalevsky, 1995). La relación entre el sistema de coordenadas estándares (ξ, η) y el sistema de coordenadas instrumentales medidas (x, y) puede ser establecida en forma polinómica (Bean et al. 2010) según:

$$\begin{aligned}
 x' &= x + LC_x(\text{MB-MV}) - XF_x, \\
 y' &= y + LC_y(\text{MB-MV}) - XF_y, \\
 \xi &= Ax' + By' + C - P_{\alpha}\pi - \mu_{\alpha}\Delta t - \text{ORBIT}_{\alpha}, \\
 \eta &= -Bx' + Ay' + F - P_{\delta}\pi - \mu_{\delta}\Delta t - \text{ORBIT}_{\delta}
 \end{aligned} \tag{1.1}$$

donde (MB-MV) es el color fotométrico, LC_x y LC_y son correcciones laterales de color; XF_x y XF_y son correcciones cruzadas de los filtros; A y B son las escalas de la placa y parámetros de rotación, mientras que C y F son corrimientos; μ_{α} y μ_{δ} son las componentes de los movimientos propios, Δt es la diferencia de tiempo a la época media; P_{α} y P_{δ} son los factores componentes del paralaje π ; y finalmente ORBIT_{α} y ORBIT_{δ} son las componentes astrométricas de la órbita perturbadora. La órbita astrométrica $\text{ORBIT}_{\alpha,\delta}$ es función de los parámetros orbitales período P , tiempo de pasaje por el pericentro τ_p , excentricidad e , semieje mayor a , posición del nodo ascendente Ω , inclinación i y argumento del pericentro ω .

La semiamplitud del wobble que describe una estrella de masa M_* y a una distancia del observador D , puede calcularse como:

$$\alpha = 3.00 \left(\frac{M_{\odot}}{M_*} \right) \left(\frac{m_p}{M_{\oplus}} \right) \left(\frac{a}{1\text{UA}} \right) \left(\frac{1\text{pc}}{D} \right) \mu\text{as}, \tag{1.2}$$

donde la perturbación es producida por un planeta de masa m_p , orbitando a una distancia a de su estrella. Esta fórmula posee el estándar de comparación: un planeta tipo terrestre que orbita a una estrella de masa solar, con su órbita a 1 UA, y ubicándose a una distancia de 10 pc del observador.

En 2009 se propuso la existencia del planeta VB 10b de masa $= 7M_{jup}$, orbitando alrededor de la estrella enana roja VB 10 (Pravdo y Shaklan, 2009). Sin embargo, las determinaciones actuales con *velocidad radial* han descartado la existencia de algún planeta alrededor de esta estrella (Bean et al. 2010).

1.3.2. Imagen Directa

Los planetas son una fuente de luz extremadamente débil comparada con las estrellas y tienden a perderse dentro del resplandor de la estrella que orbitan, siendo extremadamente difícil su detección directa.

Los telescopios pueden detectar imágenes de exoplanetas solo en circunstancias excepcionales, cuando el planeta es especialmente grande (mayor que Júpiter), ampliamente separado de su estrella y/o suficientemente caliente para emitir radiación infrarroja.

La evolución global de planetas aislados es caracterizada por su enfriamiento secular, que reduce su luminosidad en tres o cuatro órdenes de magnitud cuando tienen edades entre Myrs y pocos Gyrs (e.g., Burrows et al. 1997). La luminosidad estelar retrasa el enfriamiento, y añade una componente de luz reflejada a la emisión térmica. Esto crea una dicotomía en las estrategias de detección preferidas: los planetas jóvenes son más fácilmente detectables a través de su emisión térmica en longitudes de onda $\sim \mu m$; mientras, planetas viejos y fríos a varias UA (es decir, análogos a Júpiter) pueden probablemente ser observados reflejando la luz de su estrella en longitudes de onda visibles. El brillo de gigantes gaseosos fríos puede ser estimado relativamente fácil escalando con Júpiter, cuyo hemisferio iluminado es 3×10^{-9} menos brillante que el Sol en longitudes de onda visibles. Las separaciones angulares de los planetas extrasolares de sus estrellas son casi siempre $< 1''$. Esto significa que, en una imagen, se encuentran dentro del límite de seeing de observatorios desde tierra. Este seeing puede ser reducido, pero nunca completamente eliminado, con óptica adaptativa. La Figura 1.4 resume las separaciones angulares que se deben medir para detectar distintos tipos de exoplanetas (rocosos, gaseosos, etc) en función del **contraste** (es decir la razón de brillo entre la estrella y el planeta que la orbita).

Uno de los principales inconvenientes de este método es que los descubrimientos solo cuentan con un par de observaciones poco separadas temporalmente, por lo tanto es difícil confirmar la órbita y sólo se obtiene la distancia *osculadora*, que puede o no ser cercana al semieje de la órbita.

Sumado a ello, las distancias que nos separan de las estrellas observadas no son conocidas con absoluta precisión ³ por lo tanto las distancias del planeta a la estrella también tienen bastante incertidumbre. Conocida la distancia a la estrella d en parsecs y medida la separación angular de la estrella al planeta en segundos de arco, se obtiene la distancia instantánea del planeta a la estrella A en unidades astronómicas según:

$$A[\text{UA}] = d[\text{pc}] \cdot \gamma['] \quad (1.3)$$

³Los errores en las distancias a las estrellas más cercanas se conocen por paralajes y las determinaciones más precisas provienen de las mediciones de Hipparcos, cuyos errores están estimados en el 10 %.

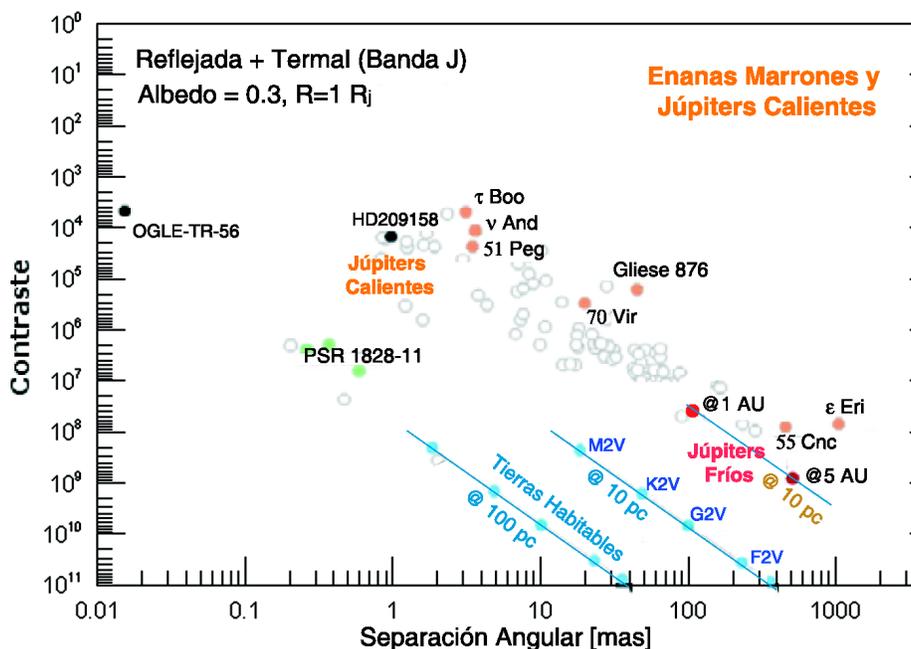


Figura 1.4: Posibles planetas detectables en el plano contraste vs separación angular. La dificultad de detección se incrementa desde arriba a la derecha (bajo contraste - gran separación) hacia la esquina inferior izquierda (alto contraste - pequeña separación). La posición esperada de exoplanetas conocidos se indica basados en modelos de sus atmósferas. Los análogos a la Tierra en zonas habitables son mostrados en 10 pc y 100 pc de distancias, para estrellas de distintos tipos espectrales. Adaptada de Quirrenbach (2006).

En julio de 2004, un grupo de astrónomos anunciaba el primer descubrimiento de este tipo y que se denominó *2M1207b*, cuya estrella central es una enana marrón, luego fue confirmado y se estimó su distancia a la estrella en 30 UA (Chauvin et al. 2004).

Por otro lado, la existencia de discos de “residuos” (denominados *debris disks*) alrededor de estrellas, proveen un argumento contundente de la existencia de objetos sólidos en estos sistemas. Esto, en sí mismo, no provee información acerca de la existencia de planetas, debido a que en las últimas etapas de formación los objetos sólidos pueden no formarlos. Sin embargo es posible detectar exoplanetas indirectamente por su influencia en la morfología de los discos, como en el caso de *Fomalhaut* (Kalas et al. 2005). Los planetas suficientemente masivos abren un vacío en los discos, que pueden ser observados directamente con imagen o indirectamente analizando la distribución espectral de energía (SED). Si existe ausencia de material en alguna región del disco, el flujo con el que se contribuye es menor y debería haber una disminución en la SED. Un buen ejemplo de esto es HR 4796 donde existe un disco prominente observado en longitudes de onda infrarroja media e infrarroja cercana, que sugieren la interacción con un planeta (Telesco et al. 2000). Sin embargo, en otros casos, las observaciones del Telescopio Espacial Hubble probaron que las estructuras observadas como anillos son en realidad espirales, que podrían deberse a interacciones con una binaria cercana (particularmente el caso de *HD141569* estudiado en Clampin et al. 2003).

1.3.3. Pulsar Timing

Un pulsar es una estrella de neutrones: un pequeño y ultradenso remanente de una estrella que explotó como una supernova. El pulsar emite ondas de radio en un *cono de luz* muy

estrecho y posee un período de rotación extremadamente regular. Cuando el alineamiento con el observador es favorable, se puede observar un efecto de pulso, similar al que se observa con un faro. Estos pulsos son muy precisos y estables en el tiempo, por lo que es posible detectar pequeñas variaciones en su periodicidad. Si el pulsar posee un planeta, se moverá alrededor de su baricentro y la diferencia de tiempo entre pulsos puede revelar los parámetros de la órbita (Wolszczan y Frail, 1992).

Este método fue originalmente utilizado en el estudio de binarias que contenían un pulsar, pero es tan sensible que es capaz de detectar planetas mucho más chicos que con otros métodos, hasta centésimas de masas terrestres. También es capaz de detectar interacciones mutuas y perturbaciones entre varios miembros de un sistema planetario, revelando así, más información acerca de los parámetros orbitales. En 1992 Aleksander Wolszczan y Dale Frail descubrieron dos planetas alrededor del pulsar PSR 1257+12 con masas de 3.4 y 2.8 M_{\oplus} (masas terrestres). Su descubrimiento fue rápidamente confirmado, convirtiéndolo en el primer planeta detectado fuera de nuestro Sistema Solar.

Las observaciones del pulsar, son mediciones de los tiempos de arribo de los pulsos (TOAs, del inglés *Times of Arrival*). En el caso de un pulsar que forme parte de un sistema binario, los TOAs exhiben variaciones periódicas, resultantes del movimiento del pulsar alrededor del centro de masa. Tales variaciones pueden ser modeladas (Konacki et al. 2000) con la fórmula:

$$\Delta t = x \left[(\cos E - e) \sin \omega + \sqrt{1 - e^2} \sin E \cos \omega \right], \quad (1.4)$$

donde

$$\left[x = a \sin i / c, \quad E - e \sin E = n (t - \tau_p), \quad n = \frac{2\pi}{P}, \right] \quad (1.5)$$

y Δt es el adelanto/retraso adicional en el TOA, E es la anomalía excéntrica, a , e , ω , $\sin i$, P , τ_p son los elementos keplerianos de la órbita y c es la velocidad de la luz. Cuando un pulsar tiene N planetas que no interactúan, las variaciones TOA se convierten en (Konacki et al. 2000):

$$\Delta t = \sum_{j=1}^N x_j \left[(\cos E_j - e_j) \sin \omega_j + \sqrt{1 - e_j^2} \sin E_j \cos \omega_j \right]. \quad (1.6)$$

En la práctica los parámetros a modelar, por medio del ajuste de cuadrados mínimos a los datos, son: x_j , e_j , ω_j , P_j , τ_{pj} , es decir, la proyección del semieje mayor de la órbita del pulsar, excentricidad, argumento del pericentro, período orbital y tiempo del pasaje por el pericentro respectivamente.

La deducción de la dependencia con estos parámetros se puede realizar describiendo el movimiento del pulsar en un sistema de referencia baricéntrico, con el eje z del sistema dirigido hacia el baricentro del Sistema Solar y el plano xy del marco de referencia en el plano del cielo (en la Figura 1.5 se muestra la geometría de un sistema con dos planetas no coplanares).

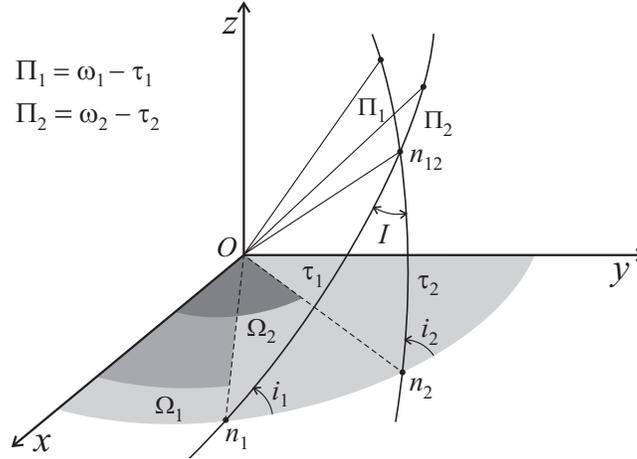


Figura 1.5: Geometría de un sistema con dos planetas. Los ángulos ω_1, ω_2 son los argumentos de pericentro, Ω_1, Ω_2 longitudes de nodo ascendente, i_1, i_2 inclinaciones de las órbitas de los planetas B y C ; τ_1 y τ_2 son los ángulos n_1On_{12} y n_2On_{12} respectivamente. Los ángulos I, τ_1, τ_2 pueden hallarse resolviendo el triángulo esférico $n_1n_{12}n_2$. Reproducido de Konacki et al. (2000).

Con estos ejes de coordenadas tenemos,

$$\Delta t = -\frac{1}{c}\mathbf{R}_0 \cdot \hat{\mathbf{z}}, \quad \text{con} \quad \mathbf{R}_0 = -\frac{1}{M_*} \sum_{i=1}^N m_i \mathbf{R}_i, \quad (1.7)$$

donde $\hat{\mathbf{z}}$ es el versor a lo largo del eje z del marco baricéntrico del sistema del pulsar y \mathbf{R}_j son las posiciones baricéntricas de los planetas. Asumiendo que:

$$\frac{m_j}{M_*} a_j \sin i_j = x_j c, \quad \frac{GM_*}{(1 + m_j/M_*)^2} = n_j^2 a_j^3, \quad (1.8)$$

con parámetros planetarios orbitales $a_j, n_j, e_j, \omega_j, \tau_{pj}$. Esta representación permite interpretar naturalmente el movimiento del pulsar como una superposición de movimientos elípticos de sus planetas alrededor del baricentro del sistema. Una teoría de perturbaciones analíticas se puede ver en el trabajo de Konacki et al. (2000).

Si se considera un solo planeta de masa m_p en una órbita circular con período P e inclinación i orbitando a una estrella de neutrones “canónica” de masa $1.4 M_\odot$, la amplitud de los residuos del timing τ puede ser escrita como (Cassen et al. 2006):

$$\tau = 1.2 m_s \left(\frac{m_p}{M_\oplus} \right) \left(\frac{P}{1 \text{ año}} \right)^{2/3} \sin i, \quad (1.9)$$

Las mediciones TOA son utilizables para detectar planetas de hasta $\sim 0.01 M_\oplus$.

Los residuos keplerianos dependen de $x_j = m_j \sin i_j$; esto significa que la masa del planeta y su inclinación orbital no se pueden determinar independientemente. En contraste la fuerza de la interacción mutua entre los planetas depende de sus masas. Integrando un sistema de N -cuerpos y modelando largas series de timing se pueden obtener los valores de masas e inclinaciones por separado (Konacki y Wolszczan, 2003).

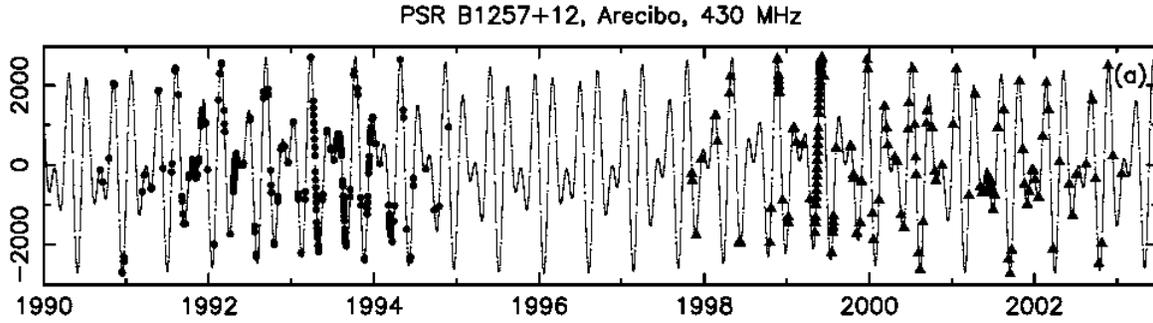


Figura 1.6: Mediciones y modelo ajustado a las variaciones TOA para el pulsar PSR B1257+12. El modelo en línea continua es superpuesto a las observaciones (puntos negros) y fue construido utilizando las variaciones keplerianas del planeta B y C . El eje vertical representa los TOA en μs . Reproducido de Konacki et al, 2000.

La Figura 1.6 muestra los tiempos de pulsación (TOAs) obtenidos con el observatorio de Arecibo desde 1990. El pulsar de 6.2 ms, PSR B1257+12, ha sido regularmente medido. Asumiendo una masa estándar de $1.4M_{\odot}$, las masas modeladas para los planetas B y C son 4.3 ± 0.2 y $3.9 \pm 0.2M_{\oplus}$ respectivamente. Las correspondientes inclinaciones orbitales obtenidas (53° y 47°) implican que son casi coplanares (Konacki y Wolszczan, 2003).

El principal inconveniente de la técnica de pulsar-timing es que los pulsares son relativamente raros, por lo tanto poco probable que un gran número de planetas se encuentren con este método. Además carecen de interés de estudio inmediato, pues la vida como la conocemos no podría sobrevivir en planetas que orbiten pulsares, debido a las radiaciones de altas energías que ellos reciben.

1.3.4. Microlensing

La teoría de relatividad general, permite explicar porque el camino recorrido por un haz de luz es curvado por la presencia de un campo gravitacional. En principio cualquier objeto masivo puede actuar como una lente, curvando la luz de un objeto de fondo y causando la magnificación temporaria de su brillo. Este efecto puede ser observado en la escala galáctica, donde por ejemplo un cúmulo masivo “cercano” actúa como lente de una galaxia lejana. En el caso de perfecto alineamiento la lente causará que el objeto de fondo aparezca como un anillo denominado *Anillo de Einstein*, de radio angular:

$$\theta_E = \sqrt{\frac{4GM_*(D_s - D_L)}{c^2 D_s D_L}} \quad (1.10)$$

donde D_L es la distancia a la lente, D_S es la distancia a la fuente (fondo), G es la constante de la gravedad, c la velocidad de la luz y M_* la masa de la lente. En el caso de un alineamiento imperfecto aparecen dos imágenes simples del objeto de fondo alrededor de la lente (ver Figura 1.7-a).

En el caso de detección planetaria con el método de lente gravitacional, una estrella con su planeta actúan como la lente y el objeto de fondo es una estrella distante. Este evento se

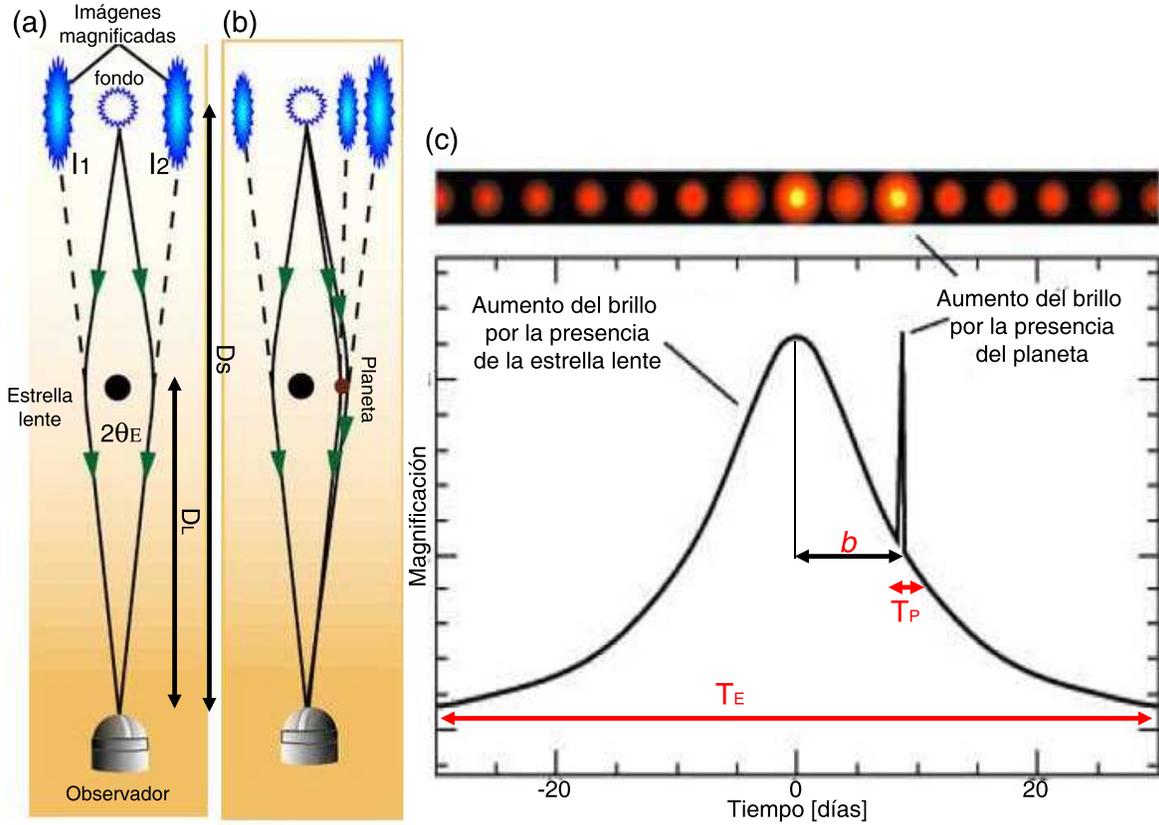


Figura 1.7: Geometría de evento de lente. (a). El alineamiento perfecto entre la estrella de fondo, la estrella lente y el observatorio. Si el alineamiento no es perfecto se forman dos imágenes I_1 e I_2 . (b). En el caso donde la estrella lente tiene un planeta. Adaptado de Bennet et al. 2006. (c). La observación de una de las imágenes I del evento de microlente. Cuando el planeta pasa por la región del anillo de Einstein, se producirían dos imágenes muy cercanas, debido a las limitaciones actuales en resolución espacial, se aprecia un aumento de brillo en la imagen.

denomina *microlente gravitatoria* o *microlensing* y se refieren al caso especial donde se crean múltiples imágenes pero están separadas por unos pocos milisegundos de arco y por lo tanto no son resueltas espacialmente con las capacidades actuales.

La deducción de las ecuaciones escapa a nuestro enfoque y puede ser consultada en Gaudi (2010), mientras que un análisis profundo de la técnica puede consultarse en Cassen et al. (2006). Simplemente mencionaremos que la curva de magnificación principal $Q(t)$, puede ser escrita en función de la distancia proyectada entre la estrella de fondo y la de primer plano, $u(t)$:

$$Q(t) = \frac{u^2(t) + 2}{u(t)\sqrt{u^2(t) + 4}} \quad (1.11)$$

Además, la relación de masa entre la estrella y el planeta μ , puede estimarse a partir de la duración total del evento principal t_E y de la duración de la anomalía planetaria t_p (ver la Figura 1.7-c), con la ecuación:

$$\mu = (t_p/t_E)^2 \quad (1.12)$$

Además, la diferencia en tiempo entre el pico principal y el de la anomalía del plane-

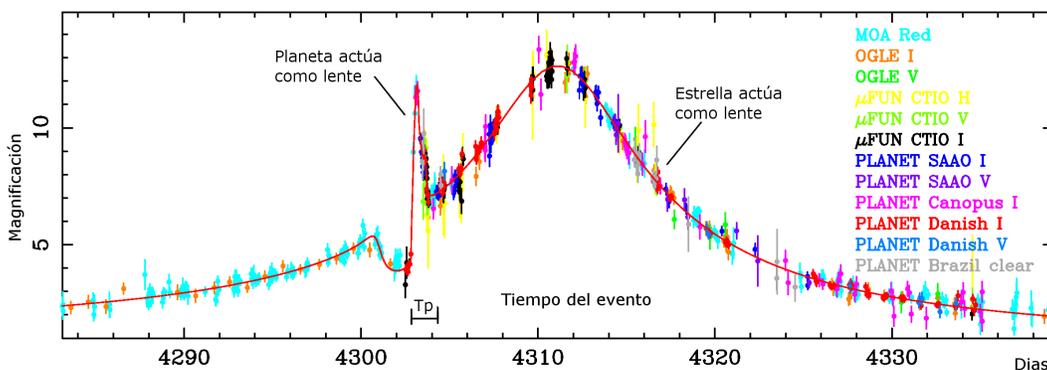


Figura 1.8: Observaciones de un evento de microlente OGLE-2007-BLG-368. Los distintos colores corresponden a distintos observatorios y la curva continua el modelo ajustado. Reproducción de Sumi et al. 2010.

ta, permite estimar la distancia proyectada entre el planeta y la estrella, denominada b . Sin embargo existe una degeneración en el problema y la solución a la curva de luz puede ser resuelta también por una distancia $1/b$ (Cassen et al. 2006). Otros parámetros que pueden ser determinados son el movimiento propio y paralaje π de la estrella que tiene el exoplaneta.

Los planetas masivos son más fáciles de detectar, porque es más fácil que su luz intercepte la fuente de la microlente. Sin embargo, la sensibilidad del método permite detectar planetas tipo terrestres. Una de las desventajas notables es que el método de *microlensing* no se puede repetir debido a que la probabilidad de alineamiento nunca ocurre nuevamente. También los planetas detectados tienden a estar varios miles de parsecs de distancia, por lo que seguimientos con otros métodos observacionales son imposibles. Sin embargo si se observa suficiente cantidad de estrellas de fondo (observando hacia el centro de la galaxia) con suficiente precisión, el método eventualmente revelará cuán comunes son los planetas tipo terrestres en la galaxia. Los eventos de microlensing son cortos, durando desde semanas o días, mientras la Tierra y las estrellas se mueven relativamente. Más de mil de estos eventos se han observado durante los últimos 10 años (Gaudi, 2010).

En 1991, los astrónomos Mao y Paczynski propusieron el método para buscar exoplanetas. Recién en 2002 el método fue exitoso, llevado a cabo por un grupo de astrónomos polacos (Udalski et al. 2002), aunque debido a las limitaciones del método no se obtuvo confirmación clara. La Figura 1.8 muestra la microlente de OGLE-2007-BLG-368 publicada en Sumi et al. (2010), que contó con la colaboración de 12 observatorios a lo largo del mundo. El evento fue producido por una estrella enana K, con masa $M_* = 0.64M_\odot$ ubicada a una distancia de $D_L=5.9$ Kpc. La anomalía planetaria es causada por un planeta de masa $m_p = 20M_\oplus$, a una distancia proyectada $d = 2.8$ UA de su estrella y cuyo semieje fue estimado en $a = 3.3$ UA.

1.3.5. Tránsitos

Cuando un sistema planetario está orientado en el espacio de tal forma que el plano de la órbita se encuentra cercano a la línea de visión del observador, los planetas periódicamente transitarán frente al disco estelar. Las observaciones fotométricas de estos eclipses pueden ser usadas para inferir parámetros físicos y orbitales de los planetas.

Para poner en práctica la técnica, la precisión fotométrica y el tiempo de muestreo de la curva de luz deben ser altos (mejor que ~ 0.005 mag y ~ 5 min en muestreo temporal). Las curvas de luz de alta precisión como las obtenidas por el Telescopio Espacial Hubble, permiten por ejemplo, estimar radios planetarios muy precisos (Brown et al. 2001). Además se pueden buscar satélites planetarios, anillos o planetas adicionales (Agol et al. 2005). Durante el tránsito se puede emplear espectroscopía de transmisión para estudiar la composición química y propiedades físicas de las atmósferas planetarias (Brown 2001). El estudio de los cambios relativos entre la profundidad de los tránsitos como función de la longitud de onda, es una manera de probar las propiedades de absorción de la atmósfera planetaria (Charbonneau, 2002a). También es posible detectar el achatamiento de los planetas, determinando su tasa de rotación (Seager y Hui, 2002).

En contraparte para cada candidato a planeta detectado por tránsito, se debe determinar su masa con la técnica de velocidad radial (V_r):

- Primero, porque varios objetos astronómicos de distinta masa, tienen los mismos tamaños (es decir sus radios). Estos objetos son: las *enanas M tardías* ($M \geq 80M_J$), las *enanas marrones* ($13M_J < M < 80M_J$), y los *planetas gigantes* ($M \leq 13M_J$), según se menciona en Seager y Mallén-Ornelas (2003).
- Segundo, porque es posible identificar erróneamente la presencia de un planeta, cuando en realidad se están observando las manchas estelares.

La curva de luz de un planeta que transita permite obtener una solución bajo ciertas suposiciones:

- i) que se tengan al menos dos tránsitos observados (para determinar el período)
- ii) que se asuma órbita circular, $e=0$
- iii) que la masa del planeta sea menor que la masa de la estrella, $m_p \ll M_*$
- iv) que la luz provenga de una estrella individual
- v) que los eclipses tengan fondos planos, es decir que el compañero transitó completamente frente al disco estelar

Respecto del punto (ii) recordemos que las búsquedas de tránsitos desde la Tierra esperan encontrar planetas de corto período, porque tienen la mayor probabilidad temporal de mostrar tránsitos. Se supone que los planetas de corto período tienen órbitas circulares ($e=0$), como resultado de la circularización tidal (Seager y Mallén-Ornelas, 2003). Respecto a la suposición (iv), se debe tener cuidado de no tener luz contaminante de una fuente adicional que se mezcle con la estrella. Finalmente para (v), los tránsitos planos aparecen en una banda donde el oscurecimiento del limbo es débil (dentro de los límites del muestreo temporal y el ruido observacional), como por ejemplo la banda I.

1.3.5.1. Ecuaciones para resolver la curva de luz

La geometría del tránsito, juntamente con la disminución de la curva de luz son mostradas en la Figura 1.9. En línea continua se marca como es la curva de luz que se produce durante el tránsito. También se marcan la profundidad del tránsito (ΔF), la duración total del tránsito (t_T) y la duración del tránsito entre el ingreso y egreso (t_F) (i.e., la parte “plana” de la curva de luz, cuando el planeta es totalmente superpuesto a su estrella). El primero, segundo, tercero y cuarto contactos son marcados para un planeta que se mueve de izquierda a derecha. También se muestran los radios de la estrella y planeta (R_* , R_p) y el parámetro de impacto b correspondiente a la inclinación orbital i . Parámetros de impacto b distintos (o inclinaciones distintas, i) producen distintas duraciones de los tránsitos como se muestra con la curva punteadas de la Figura 1.9.

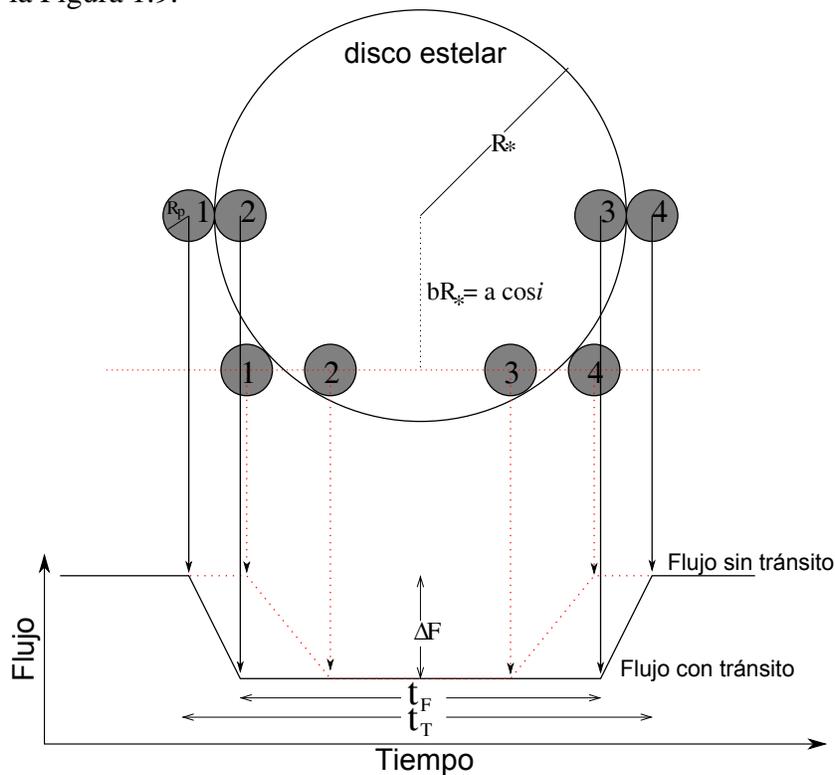


Figura 1.9: Esquema de los observables de la curva de tránsito. **Arriba.** Se dibujan los cuatro contactos de dos tránsitos distintos: uno que pasa por el centro de la estrella y el otro con parámetro de impacto b . **Abajo.** Curvas de luz que producen los tránsitos esquematizados en la parte superior. La línea continua describe el flujo observado para un planeta que transita por el centro de la estrella y la línea punteada corresponde al planeta que transita con parámetro de impacto b . Adaptado de Seager y Mallen-Ornelas, 2003.

Existen 5 ecuaciones que describen completamente la curva de luz del planeta en tránsito (para su deducción consultar Cassen et al. (2006) y Seager y Mallen-Ornelas, (2003)), considerando una curva de luz producida por dos esferas que pasan una delante de la otra:

$$\Delta F \equiv \frac{\text{Flujo}_{sin\ \text{tránsito}} - \text{Flujo}_{con\ \text{tránsito}}}{\text{Flujo}_{sin\ \text{tránsito}}} = \left(\frac{R_p}{R_*} \right)^2; \quad (1.13)$$

$$t_F = \frac{P}{\pi} \sin^{-1} \left(\frac{R_*}{a} \left[\frac{\left(1 - \frac{R_p}{R_*}\right)^2 - \left(\frac{a}{R_*} \cos i\right)^2}{1 - \cos^2 i} \right]^{1/2} \right); \quad (1.14)$$

$$t_T = \frac{P}{\pi} \sin^{-1} \left(\frac{R_*}{a} \left[\frac{\left(1 + \frac{R_p}{R_*}\right)^2 - \left(\frac{a}{R_*} \cos i\right)^2}{1 - \cos^2 i} \right]^{1/2} \right); \quad (1.15)$$

$$P^2 = \frac{4\pi^2 a^3}{G(M_* + m_p)}; \quad (1.16)$$

$$R_* = kM_*^x, \quad (1.17)$$

Las primeras tres ecuaciones ((1.13)–(1.15)) describen la geometría del tránsito en términos de los observables. La Ecuación (1.16) es la Tercera Ley de Kepler. La Ecuación (1.17) expresa la relación masa-radio estelar en función de dos parámetros: k y x . k es un coeficiente para cada tipo de secuencia estelar (secuencia principal, gigantes, etc.) y x describe la ley de potencia de la secuencia ($x \simeq 0.8$ para estrellas de secuencia principal F–K (Cox 2000)).

Más allá que resulte simple del esquema determinar los observables (Figura 1.9), las observaciones reales lucen un poco más ruidosas (Figura 1.10) y debe realizarse una buena modelización de datos.

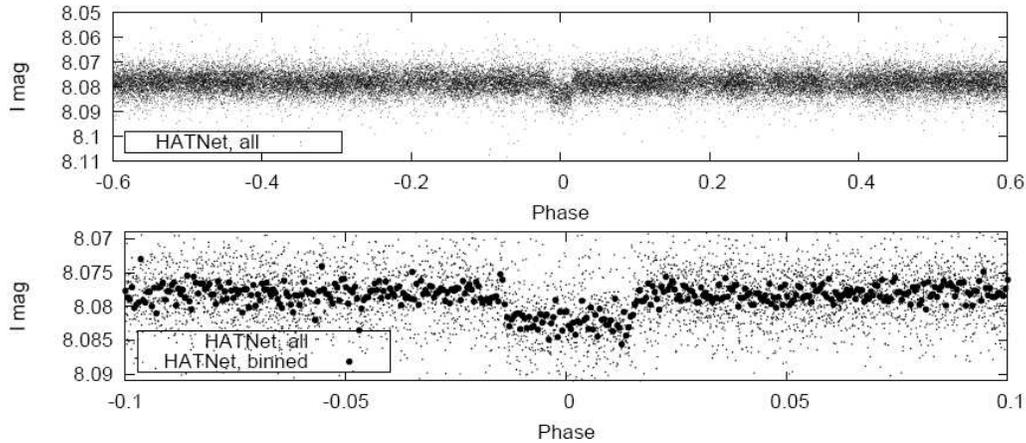


Figura 1.10: El panel superior muestra los datos crudos de los proyectos *HATNet* y *WHAT*, que constituyen una curva de luz con 26400 observaciones en banda *I*, mostradas en fase con un período $P = 5.63341$ días (Bakos et al. 2007, para HD 147506). La profundidad del tránsito es del orden de 5 mmag, con una relación señal-ruido de 26. El panel de abajo muestra los mismos datos bineados con un tamaño de bin de $\phi = 0.0005$.

Cada tránsito observado brinda 5 parámetros a determinar: la disminución en el brillo ΔF , la duración total del tránsito t_T , la duración entre el ingreso y egreso t_F y definir el tiempo del máximo de la ocultación respecto a algún origen arbitrario.

1.3.5.2. Parámetros orbitales a partir de observables

Si se despejan las ecuaciones de la sección anterior se obtiene:

i) La relación de los radios entre el planeta y la estrella, a partir de la Ecuación (1.13),

$$\frac{R_p}{R_*} = \sqrt{\Delta F}; \quad (1.18)$$

ii) El parámetro de impacto b , se define como la distancia proyectada entre el centro de la estrella durante el medio del tránsito en unidades de R_* (ver Figura 1.9), y que puede ser obtenido combinando (1.14) y (1.18),

$$b \equiv \frac{a}{R_*} \cos i = \left[\frac{(1 - \sqrt{\Delta F})^2 - \frac{\sin^2 \frac{t_F \pi}{P}}{\sin^2 \frac{t_T \pi}{P}} (1 + \sqrt{\Delta F})^2}{1 - \frac{\sin^2 \frac{t_F \pi}{P}}{\sin^2 \frac{t_T \pi}{P}}} \right]^{1/2}; \quad (1.19)$$

iii) La razón entre a/R_* , que puede ser obtenida directamente de la ecuación (1.15),

$$\frac{a}{R_*} = \left[\frac{(1 + \sqrt{\Delta F})^2 - b^2 (1 - \sin^2 \frac{t_T \pi}{P})}{\sin^2 \frac{t_T \pi}{P}} \right]^{1/2}; \quad (1.20)$$

iv) La densidad estelar ρ_* , puede ser derivada de la ecuación anterior y la Tercera Ley de Kepler con $m_p \ll M_*$ (Ecuación (1.16)),

$$\frac{\rho_*}{\rho_\odot} \equiv \frac{M_*/M_\odot}{(R_*/R_\odot)^3} = \left[\frac{4\pi^2}{P^2 G} \right] \left[\frac{(1 + \sqrt{\Delta F})^2 - b^2 (1 - \sin^2 \frac{t_T \pi}{P})}{\sin^2 \frac{t_T \pi}{P}} \right]^{3/2}. \quad (1.21)$$

Los 5 parámetros físicos R_* , M_* , i , a , y R_p pueden ser derivados de las soluciones de arriba usando una ecuación adicional: la relación masa-radio estelar (Ecuación (1.17)).

La masa estelar M_* , se calcula utilizando la Ecuación (1.21) y (1.17):

$$\frac{M_*}{M_\odot} = \left[k^3 \frac{\rho_*}{\rho_\odot} \right]^{\frac{1}{1-3x}}. \quad (1.22)$$

El radio estelar,

$$\frac{R_*}{R_\odot} = k \left(\frac{M_*}{M_\odot} \right)^x = \left[k^{1/x} \frac{\rho_*}{\rho_\odot} \right]^{\frac{x}{(1-3x)}}; \quad (1.23)$$

El radio orbital a se puede derivar usando M_* y de la tercera Ley de Kepler con $m_p \ll M_*$,

$$a = \left[\frac{P^2 G M_*}{4\pi^2} \right]^{1/3}; \quad (1.24)$$

Con la Ecuación (1.19), la inclinación orbital es:

$$i = \cos^{-1} \left(b \frac{R_*}{a} \right), \quad (1.25)$$

Finalmente, el radio planetario es

$$\frac{R_p}{R_\odot} = \frac{R_*}{R_\odot} \sqrt{\Delta F} = \left[k^{1/x} \frac{\rho_*}{\rho_\odot} \right]^{\frac{x}{(1-3x)}} \sqrt{\Delta F}. \quad (1.26)$$

Para estrellas de la secuencia principal $k = 1$ y $x \approx 0.8$, en cuyo caso $\frac{R_p}{R_\odot} = \left(\frac{\rho_*}{\rho_\odot} \right)^{-0.57} \sqrt{\Delta F}$.

Alternativamente se puede aproximar la curva de luz en función de la distancia proyectada entre la estrella y el planeta ($d \equiv d(t)$). Cuando se tiene una fuente uniforme, la razón entre el flujo oscurecido y no oscurecido está dada por $F^e(p, z) = 1 - \lambda^e(p, z)$, deducida en Mandel y Algol (2002), donde:

$$\lambda^e(p, z) = \begin{cases} 0 & 1 + p < z \\ \frac{1}{\pi} \left[p^2 \kappa_0 + \kappa_1 - \sqrt{\frac{4z^2 - (1+z^2-p^2)^2}{4}} \right] & |1-p| < z \leq 1+p \\ p^2 & z \leq 1-p \\ 1 & z \leq p-1, \end{cases} \quad (1.27)$$

donde d es la distancia centro del planeta-estrella, R_p el radio del planeta, R_* el radio estelar, $z = d/R_*$ la separación normalizada de los centros, $p = R_p/R_*$ la razón de tamaños, $\kappa_1 = \cos^{-1}[(1-p^2+z^2)/2z]$, $\kappa_0 = \cos^{-1}[(p^2+z^2-1)/2pz]$. Geométricamente $\lambda^e(p, z)$ es el área superpuesta entre dos círculos de radio R_* y R_p cuyos centros están separados z unidades.

El oscurecimiento límbico causa que una estrella tenga un brillo central más intenso que en los bordes, en lugar de comportarse como una fuente de luz uniforme. Esto produce que la parte plana de la curva de luz (marcada con t_F en la Fig. 1.9), tenga una pequeña “panza”. Claret, (2000) propuso una ley de oscurecimiento límbico para calcular curvas de luz más precisas, que ajusta una amplia gama de rangos de modelos estelares y bandas de observación con la forma:

$$I(r) = 1 - \sum_{n=1}^4 c_n (1 - \mu^{n/2}), \quad (1.28)$$

donde $\mu = \cos \theta = \sqrt{1-r^2}$, r es la coordenada radial normalizada medida sobre el disco de la estrella y $I(r)$ es la intensidad específica como función de r o μ con condición de contorno $I(0) = 1$ (ver Mandel y Algol (2002) para más detalles).

Recientemente se mostró que los planetas con órbitas excéntricas deberían ser más fáciles de detectar (Barnes, 2007). El mismo autor ha demostrado que si se ajusta el radio estelar con los otros parámetros, se produce un error sistemático en las determinaciones resultantes. Si se analizan las aceleraciones orbitales, se obtienen diferencias entre las duraciones del tiempo de ingreso y egreso, que permiten obtener la excentricidad e de la órbita usando sólo mediciones fotométricas. Por el momento este efecto no es medible con la tecnología disponible (Barnes, 2007).

1.3.6. Velocidades Radiales

La velocidad con la que una estrella se mueve en la dirección radial desde el baricentro del sistema solar, es decir su velocidad radial, puede deducirse del desplazamiento de la posición de las líneas espectrales a través del efecto Doppler.

Al comienzo de este trabajo en 2006, casi la totalidad de los exoplanetas eran indirectamente detectados midiendo el movimiento de la estrella a la que orbitaban, mediante las técnicas de velocidad radial (V_r). Las técnicas de V_r empleadas fueron mejorando. En 1988 donde se iniciaron los programas observacionales, los errores en medición eran de 20 m/s. En la actualidad gracias a mejoras, tanto tecnológicas como técnicas, han disminuido los errores nominales hasta 20 cm/s. Estos errores son calculados a partir del análisis del corrimiento Doppler de las líneas espectrales, sin embargo es un error formal y existe una fuente adicional de errores. Recordemos que la Tierra, provoca cambios en la velocidad del Sol su velocidad que tienen una amplitud de sólo 9 cm/s, mientras que Júpiter induce una perturbación de 12.5 m/s.

Nuestro trabajo, es realizado analizando las velocidades radiales. Sin embargo, mucha de la teoría subyacente de modelización de datos es aplicable a cualquiera de las otras técnicas de detección. Si bien en este trabajo no realizamos el análisis Doppler de las observaciones, es sumamente ilustrativo conocer las bases de la reducción de los datos observacionales ya que nuestro análisis tiene en cuenta las características de la obtención de los mismos (como ser, errores Gaussianos en las mediciones, indeterminación de la velocidad absoluta del baricentro de la estrella, etc).

La espectroscopía Doppler es un método de detección indirecto que utiliza la medición de la luz estelar para medir la influencia gravitacional de un planeta sobre su estrella. En los sistemas planetarios, la estrella y el planeta orbitan su baricentro común. De acuerdo al efecto Doppler-Fizeau (efecto Doppler de ahora en más), la luz emitida por una fuente que se aproxima (que se aleja) al observador sufre un corrimiento hacia longitudes de onda más cortas (largas). Según el trabajo de Anne Eggenberger y Stéphane Udry (2010), en su forma relativística más simple, el corrimiento puede expresarse como:

$$z = \frac{\lambda - \lambda_0}{\lambda_0} = \frac{1 + V_r/c}{\sqrt{1 - \mathbf{V}^2/c^2}} - 1, \quad (1.29)$$

donde z es llamado “redshift” (en sentido general), λ y λ_0 son las longitudes de onda observada y en reposo, V_r la velocidad radial del baricentro, c es la velocidad de la luz y \mathbf{V} la velocidad total del movimiento del baricentro del sistema.

Los métodos tradicionales de medición de velocidad radial consisten en observar una estrella y seguidamente un espectro de comparación. La luz de la lámpara de comparación brinda las líneas espectrales en reposo para el observador, mientras que por efecto Doppler, la estrella tiene sus líneas espectrales corridas. Este método es repetido en distintos momentos a lo largo del año y de esta forma se tienen las velocidades radiales observadas. La porción del espectro visible de estrellas tipo F, G, K y enanas M contiene una considerable porción de líneas de absorción metálicas, que se usan para medir este efecto. La figura 1.11 muestra una

parte del espectro de una estrella tipo solar que posee al menos un planeta, 51 Peg, ilustrando las múltiples líneas de absorción en un espectro.

Los mejores resultados con las técnicas tradicionales no superan la precisión de 200 m/s, dejando fuera de esta forma la posible detección de exoplanetas. Sin embargo según Eggenberger y Udry (2010), existen errores sistemáticos cuyas causas son:

- Movimiento del fotocentro en la ranura del espectrógrafo: en unidades Doppler, el ancho de la ranura es de unos pocos km/s, significando que errores más grandes que 1 m/s ocurren cuando el fotocentro se mueve sólo pocos 10^{-3} anchos de ranura. En la práctica, los movimientos del fotocentro debido a errores del guiado, foco y fluctuaciones del seeing son del orden de 10^{-1} anchos de ranura. Una solución a este problema es usar fibra óptica, para uniformizar la luz de la estrella en la entrada del espectrógrafo.
- Variaciones del índice de refracción y flexiones termodinámicas: una solución parcial al problema es estabilizar y controlar el espectrógrafo en presión y temperatura, aunque algunos corrimientos no se pueden evitar; otra solución es utilizar una calibración simultánea en longitud de onda para monitorizarlos. En estos instrumentos un cambio de 1° K en la temperatura produce un cambio neto de ~ 90 m/s, mientras que un cambio de presión de 1 mbar produce un corrimiento de ~ 300 m/s.
- Tiempos de exposición y corrección baricéntrica: las velocidades radiales clásicas son corregidas al baricentro del Sistema Solar. Las principales contribuciones del movimiento baricéntrico de la Tierra son la rotación diurna de la Tierra (1-2 m/s, dependiendo de la latitud del observatorio) y la traslación terrestre (~ 30 km/s). Al nivel necesario del m/s, el movimiento del Sol alrededor del baricentro del Sistema Solar y de la Tierra alrededor del baricentro Tierra-Luna son significativas. Para obtener velocidades radiales es necesario conocer las efemérides precisas del Sistema Solar y el valor del tiempo en el instante medio de cada observación, con error menor que 30 s. Vale la pena mencionar el hecho que la luz toma unos 1000 segundos (0.01 días) para ir desde un punto al otro diametralmente opuesto en la superficie de la Tierra. Esto significa que el tiempo de un fenómeno no será el mismo en todos los puntos de la Tierra y por lo tanto estas correcciones al tiempo deben ser aplicadas.

En la actualidad, la alta precisión necesaria conllevó al desarrollo de dos tipos de técnicas de alta precisión que reducen estos errores sistemáticos: una técnica que utiliza una celda de *celda de absorción* con yodo (I_2) y la otra denominada de *referencia simultánea*. Ambas usan el mismo tipo de espectrógrafo, pero difieren en diseño y modos de operación.

1.3.6.1. Técnica con celda de absorción

El método fue introducido en Marcy y Butler (1992) y consiste básicamente en hacer atravesar la luz de la estrella por una celda de absorción de I_2 superimprimiendo en el CCD ambos espectros a la vez. El espectro compuesto resultante entra al espectrómetro donde es convolucionado con el perfil del instrumento y dispersado casi linealmente sobre el CCD. El principio de análisis tiene en cuenta que, el *corrimiento observado del espectro estelar*

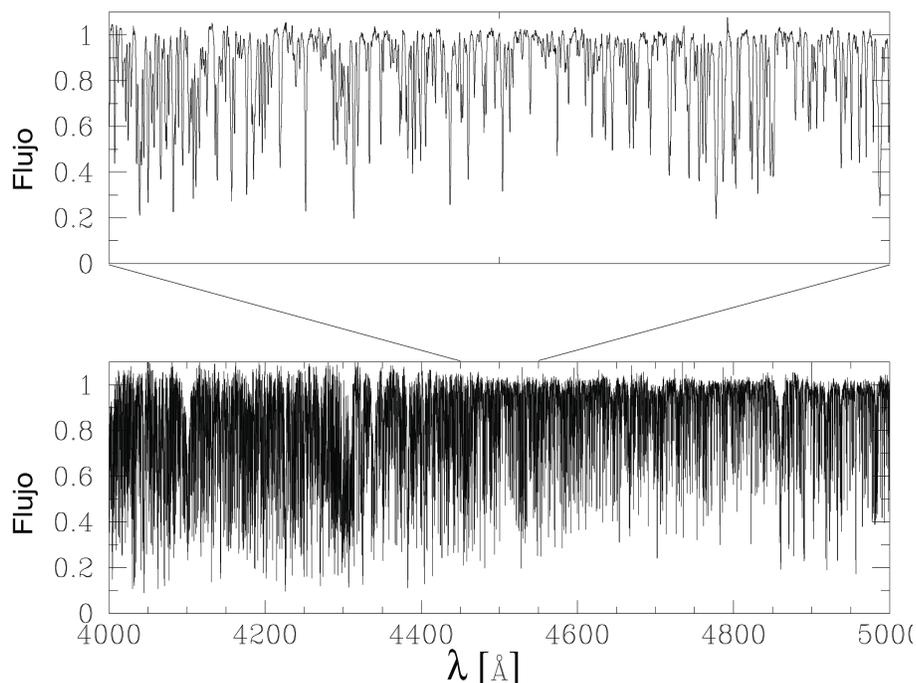


Figura 1.11: Espectro de la estrella con planetas 51 Peg, ilustrando las múltiples líneas de absorción en su espectro. Reproducido de Eggenberger y Udry (2010).

consiste en 2 partes: el corrimiento Doppler propio y un corrimiento pequeño espúreo. El corrimiento Doppler espúreo es representado enteramente por el corrimiento de las líneas de I_2 , que luego es aplicado como corrección al corrimiento observado del espectro estelar.

Este método corrige los errores causados por los diferentes caminos ópticos que atraviesan el haz de luz de la estrella y el de la lámpara de comparación.

La gran cantidad de información Doppler contenida en estos espectros (alrededor de 300 líneas de absorción estelar) requiere un análisis que es en general suficiente para tratar todas las correcciones necesarias⁴.

Para tener idea de los límites de precisión alcanzados, basta mencionar que para una estrella de magnitud visual $V=13$ observada con el observatorio Lick, el corrimiento en el CCD de las líneas espectrales de solo 0.01 pixeles en el detector corresponde a un error de medición de 25 m/s.

El análisis Doppler riguroso debe considerar entonces el modelo completo del espectro observado, incluyendo el corrimiento del espectro estelar ($\Delta\lambda_s$), el corrimiento de las líneas de I_2 sobreimpresas ($\Delta\lambda_{I_2}$), y la PSF del espectrógrafo. El espectro observado es compuesto por dos funciones, dependientes de la longitud de onda:

⁴debido a que las observaciones necesarias son realizadas en distintos momentos, existe una flexión del espectrógrafo, movimientos del detector, distintos focos en las ranuras e iluminación imperfecta del espectrógrafo que causa distorsiones en la dispersión del espectro al igual que en la respuesta del detector a la señal recibida denominada PSF (del inglés, Point Spread Function).

- el espectro estelar intrínseco I_s
- la función de transmisión T_{I_2} de la celda de absorción I_2 .

El producto de estas dos funciones es convolucionado dentro del espectrógrafo con la PSF e integrado sobre el pixel del CCD para reproducir el espectro observado. El espectro tomado a través de la celda de absorción, $I^{obs}(\lambda)$, se puede modelar como (Marcy y Butler, 1992):

$$I^{modelo} = k[I_s(\lambda + \Delta\lambda_s) * T_{I_2}(\lambda + \Delta\lambda_{I_2})] \otimes PSF, \quad (1.30)$$

donde la constante k es proporcional al nivel de exposición de la observación. Las incógnitas a determinar son: $\Delta\lambda_s$, $\Delta\lambda_{I_2}$ y k , determinados por el métodos de cuadrados mínimos ajustando la observación con la respuesta teórica de la ecuación (1.30). Este proceso se realiza en todo el espectro pero seleccionando cientos de pequeñas porciones del espectro de alrededor de 2 Å denominadas “chunks”.

El corrimiento Doppler corregido $\Delta\lambda$ es simplemente dado por:

$$\Delta\lambda = \Delta\lambda_s - \Delta\lambda_{I_2}, \quad (1.31)$$

Que es convertido en velocidad a través de una expresión más precisa para el corrimiento de las líneas espectrales (Marcy y Butler, 1992):

$$\lambda = \lambda_0 \frac{(1 + \beta \cdot \cos\theta)(1 + \rho_g)}{n(1 - \beta^2)^{1/2}}, \quad (1.32)$$

donde λ_0 es la longitud de onda en reposo, $\lambda = \lambda_0 + \Delta\lambda$, $\beta = v/c$ es la velocidad topocéntrica respecto de la velocidad de la luz, θ es el ángulo entre la línea de visión y el vector velocidad topocéntrico de la estrella, ρ_g es el enrojecimiento gravitacional del haz estelar y n es el índice de refracción del aire en el espectrógrafo. Ignorar ρ_g y n causa errores relativos en la velocidad menores que 1 m/s. Los términos de la expresión relativística deben ser tenidos en cuenta a fin de evitar errores de varios metros por segundo. Una de las desventajas de este método, es que se obtienen corrimientos Doppler relativos al modelo, **no velocidades absolutas**. Sin embargo los enormes errores sistemáticos son eliminados (Eggenberger y Udry, 2010).

Las 2 funciones necesarias en la Ecuación (1.30) son determinadas *a priori* en la puesta a punto del espectrógrafo:

- I_s se obtiene observando la estrella sin la celda de I_2 . Esto brinda $I_s \otimes PSF$, no la I_s deseada. Por lo tanto se realiza la deconvolución usando la transformada de Fourier para obtener el parámetro deseado.
- La función de transmisión de la celda de I_2 (i.e. T_{I_2}), se obtiene observando solamente la celda de I_2 . Esta observación se hace bajo estrictas normas y su espectro se encuentra disponible bajo pedido. La función de transmisión T_{I_2} provee dos elementos importantes en el proceso de modelado. Primero brinda una precisa escala de longitudes de onda. Segundo, brinda un espectro de referencia con el cual determinar la PSF del espectrómetro.

Los espectrógrafos que buscan planetas, basados en esta técnica, son: HAMILTON (Observatorio Lick, USA), HIRES (Telescopio Keck, Mauna Kea, USA), UCLES (AAT, Siding Spring, Australia), HRS (Telescopio HET, Davis Mountain, USA) y UVES (Telescopio VLT, Paranal, Chile). La precisión de estos instrumentos indica que es posible obtener errores internos de hasta 2.5 m/s.

1.3.6.2. Técnica de referencia simultánea

La técnica de referencia simultánea involucra el uso de 2 fibras ópticas que alimentan al espectrógrafo: la fibra óptica del objeto (que tiene la luz estelar) y la fibra óptica de referencia que recoge la información de longitud de onda de referencia. La longitud de onda de referencia utiliza una lámpara de Thorio-Argon (Th-Ar). El método a veces se conoce como técnica simultánea de Thorio, exige mucha precisión en el instrumento y el análisis Doppler es directo.

Una característica del método es que se implementa en un espectrógrafo, con muy alta estabilidad mecánica y térmica. El uso de fibra óptica reduce inestabilidades en la variación de la iluminación y variaciones del perfil instrumental (PSF), que no se pueden corregir en la calibración simultánea. La estabilización térmica, implica que ambos caminos ópticos de ambos haces de luz son similares en el espectrógrafo. La velocidad radial es obtenida con correlación cruzada entre el espectro observado y máscaras que consisten en plantillas binarias (*templates* con valores 0 y 1) formados por rectángulos. Esquemáticamente la función de correlación es construida moviendo la posición de la máscara, sobre una ventana centrada en la velocidad radial de la estrella (más detalles en Baranne et al. 1996, Pepe et al 2002). Las máscaras usadas para correlación son basadas en espectros de alta relación señal/ruido de estrellas brillantes o bien de espectros sintéticos.

Los espectrógrafos que buscan exoplanetas basados en esta técnica son: ELODIE (Telescopio de 1.93 m, Haute Provence, Francia, no funciona desde 2006), CORALIE (Telescopio Leonard Euler, La Silla, Chile), HARPS (Telescopio ESO 3.6 m, La Silla, Chile) y SOPHIE (Telescopio 1.93 m, Haute Provence, Francia). La precisión de estos instrumentos actualmente es superior a 1 m/s.

1.3.6.3. Errores de las mediciones y Jitter estelar

El análisis Doppler provee una medición formal de los errores, sin embargo puede haber un error adicional más difícil de caracterizar, proveniente de un ruido intrínseco en la velocidad del centro de masa de la estrella.

Entre las fuentes de **errores sistemáticos** en las mediciones se incluyen los problemas en el modelado de las líneas espectrales y la caracterización del espectrómetro.

Entre las fuentes de **errores intrínsecos**, el *jitter estelar* es su causante y se piensa que su origen es la combinación de movimientos convectivos en la superficie, actividad magnética y rotación (Saar y Donahue, 1997). Esta cantidad depende de las propiedades estelares como tasa de rotación y tipo espectral pero es típica de 1-5 m/s para estrellas cromosféricamente

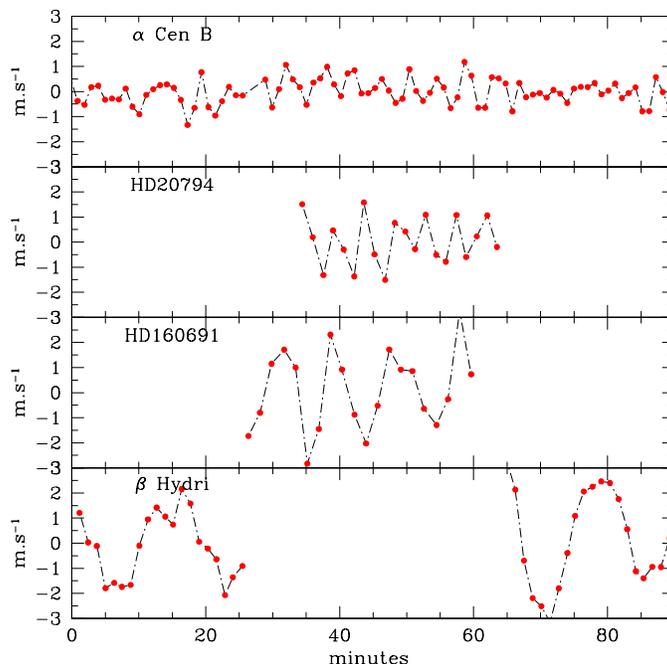


Figura 1.12: Mediciones de velocidad radial en HARPS sobre 4 estrellas de tipo solar. Son claramente visibles las oscilaciones en todas ellas y sus amplitudes son comparables. Reproducido de Eggenberger y Udry (2010).

estables (Cumming et al, 2008 y referencias en el artículo). Desafortunadamente a veces no es fácil identificar estas variaciones espúreas y en algunos casos como en HD 165435 tienen una amplitud de 80 m/s (Queloz et al. 2001) y están correlacionados con la variabilidad fotométrica. Los estudios sistemáticos de la actividad estelar y variación de las velocidades han demostrado que es difícil identificar correlaciones con los parámetros astrofísicos (Eggenberger y Udry, 2010).

El Sol, por ejemplo, oscila en varios modos de baja amplitud, cada uno caracterizado por una señal menor que 20 cm/s y período de tiempo entre 3 y 15 minutos. Estas fluctuaciones son debidas a movimientos estocásticos convectivos, cercanos a la superficie solar. En la Figura 1.12 aparecen observaciones recientes que muestran este comportamiento para varias estrellas de tipo solar. Si bien los modos individuales de oscilación son del orden de 10-50 cm/s para estrellas de secuencia principal, su interferencia constructiva puede producir variaciones hasta 10 veces mayores en tiempos de escala de minutos. A pesar que estas oscilaciones pueden ser interesantes desde el punto de vista estructural son un ruido *molesto* cuando se realiza la búsqueda de exoplanetas.

Por otro lado, los compañeros adicionales de baja masa indetectables, pueden ser una fuente adicional de variabilidad en la velocidad. Se encuentran enmascarados en el ruido propio de las velocidades. El límite de detección de planetas varía según el programa observacional. Es comúnmente aceptado como límite de detección una amplitud de la señal $K=10$ m/s, aunque programas específicos han llevado este límite a $K=2$ m/s (HD 156668b, con una masa equivalente a $4 M_{\oplus}$, Howard et al. 2010), e incluso $K=0.8$ m/s (HD 10180b $1.35 M_{\oplus}$, Lovis et al. 2010). Una primera estimación basada en el relevamiento con el instrumento HARPS sugiere que un 30 % de las estrellas enanas G y K tienen un planeta con masas menores que

$30 M_{\oplus}$ y períodos $P < 50$ días (Mayor et al. 2008). Estos planetas tienen una amplitud de señal muy baja y con otros instrumentos de menor precisión, pueden ser confundidos con el comportamiento del jitter estelar.

Teniendo en cuenta lo mencionado anteriormente, la incerteza σ_i en cada velocidad radial medida, puede ser caracterizada como la suma cuadrática de los errores sistemáticos e intrínsecos, definida por:

$$\sigma_i^2 = \varepsilon_{i_0}^2 + \sigma_J^2, \quad (1.33)$$

donde ε_{i_0} es dado por el proceso de la determinación del valor de la velocidad radial y σ_J^2 es el valor del jitter estelar. A lo largo del trabajo se adoptan los valores sugeridos de jitter estelar y son útiles porque de esa forma se “suavisa” el peso individual de cada dato. En nuestros trabajos se buscaron mejores ajustes probando distintos valores de jitter, y las soluciones encontradas no varían significativamente. La única excepción que encontramos es en el sistema HD 73526, donde este peso sobre las observaciones cambia drásticamente las soluciones.

1.3.6.4. Velocidad Radial en función de parámetros orbitales

Veamos ahora como obtener los parámetros planetarios en función de la velocidad radial de una estrella. La velocidad de una estrella respecto al observador puede ser descompuesta según dos direcciones, una de ellas en la dirección del observador y que es llamada **Velocidad Radial** y la otra proyectada sobre el plano tangente a la esfera celeste, indirectamente medida a través de técnicas astrométricas.

La orientación de la órbita de un sistema extrasolar en el espacio está definida según muestra la Figura 1.13. La orientación del eje x está dada en la dirección a γ , que es el nodo donde el movimiento del planeta es dirigido hacia el observador. El eje y es también tangente a la esfera celeste ortogonal al anterior y z es dirigido en la línea de la visual alejándose del observador. Por convención la órbita está orientada de tal forma que el argumento del pericentro está dado por $\omega + \pi$, i es la inclinación del plano de la órbita con respecto al cielo.

La expresión exacta para la velocidad radial que produce un planeta en su órbita alrededor de la estrella de masa M_* , puede ser deducida fácilmente en función de sus parámetros orbitales (ver Beaugé et al. 2007b) y está dada por la expresión:

$$V_r = K[\cos(f + \omega) + e \cos(\omega)] + \mathbf{V}_0 \quad (1.34)$$

donde K es la amplitud de la variación de la velocidad radial, la excentricidad es e , el argumento de pericentro ω . \mathbf{V}_0 es el desplazamiento arbitrario del origen. Se encuentra escrito en forma vectorial, pues es necesario una componente (V_0) por cada conjunto de datos que se disponga (sea un observatorio y/o instrumento diferente, ver Sección 1.3.6.1).

La amplitud K puede ser escrita como:

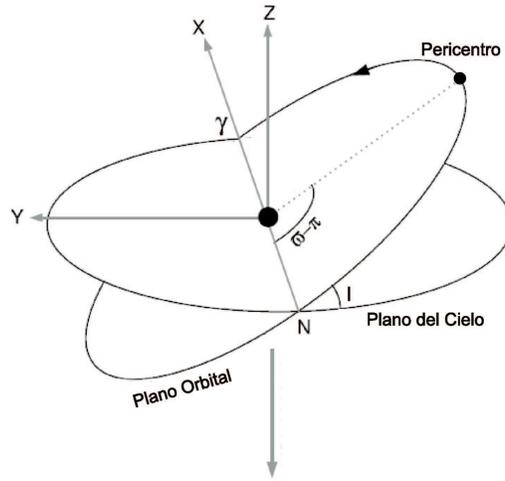


Figura 1.13: Marco de referencia astrocéntrico rotado, donde se muestra el plano orbital y el plano tangente a la esfera celeste. El origen de los ángulos es el punto γ y el planeta (círculo negro) se encuentra en el pericentro de la órbita. Reproducido de Beaugé et al. (2007b).

$$K = \frac{m}{m + M_*} \frac{2\pi a}{P} \frac{\sin i}{\sqrt{1 - e^2}} \quad (1.35)$$

donde P es el período, a el semieje, m la masa planetaria y M_* la masa estelar.

La anomalía verdadera f se relaciona con la anomalía excéntrica E de la siguiente manera:

$$\operatorname{tg} \left[\frac{f}{2} \right] = \operatorname{tg} \left[\frac{E}{2} \right] \sqrt{\frac{1 + e}{1 - e}}, \quad (1.36)$$

y luego con la ecuación de Kepler obtenemos la anomalía media:

$$E - e \sin E = \frac{2\pi}{T}(t - \tau) \quad (1.37)$$

donde τ es el tiempo del pasaje por el pericentro, por convención el origen de la órbita.

Resumiendo, se pueden identificar 6 parámetros planetarios que determinan la curva de velocidad radial y definen el conjunto de **parámetros primarios**:

- K , la amplitud de la velocidad radial
- P , período orbital
- e , excentricidad
- ω , argumento de pericentro

- τ , pasaje de tiempo por el pericentro
- V_0 , velocidad radial del baricentro respecto a un 0 arbitrario y asociado al conjunto de observaciones (ver explicación en 1.3.6.1)

A fin de disminuir errores numéricos, el tiempo de pasaje por el pericentro τ es dado en unidades de tiempo a partir de la primer observación disponible. El origen temporal de las observaciones es desplazado con el mismo objetivo *JD-2400000*.

Utilizando la tercera ley de Kepler:

$$n^2 a^3 = G(m + M_*) \quad (1.38)$$

donde $n = 2\pi/P$ es la velocidad angular media alrededor de la órbita y G es la constante de Gravitación Universal.

Los **parámetros secundarios** utilizados para estudiar la dinámica del sistema son: m , la masa del planeta y a , el semieje de la órbita. La masa del planeta debe ser obtenida combinando (1.35) y (1.38) para obtener la expresión:

$$\frac{\left(m \cdot \sin(i)\right)^3}{(m + M_*)^2} = \frac{P}{2\pi G} K^3 (1 - e^2)^{3/2}, \quad (1.39)$$

que en términos computacionales puede ser resuelta iterativamente. Debemos hacer una salvedad, ya que estrictamente hablando, se determina el parámetro $m \cdot \sin(i)$, y aunque la inclinación no se puede determinar directamente se asume una inclinación de 90° por lo que se tiene una **masa mínima** para cada planeta. Luego, con la ecuación (1.38) se puede calcular el valor del semieje, a . Los parámetros secundarios tienen, de por sí, asociado un error intrínseco porque dependen explícitamente de la masa de la estrella (determinada indirectamente, espectroscópica o fotométricamente, y cuyo error es de un 8 %).

Si aproximamos $m \ll M_*$, se obtiene una expresión para la masa planetaria mínima:

$$m \cdot \sin(i) \simeq \left(\frac{P}{2\pi G}\right)^{1/3} K M_*^{2/3} \sqrt{1 - e^2} \quad (1.40)$$

El desconocimiento de la inclinación de la órbita es un problema a la hora de estudiar la dinámica. En sistemas de múltiples planetas, se pueden calcular las razones de masas y semiejes. Si suponemos que los planetas son coplanares, podemos aproximar el cociente entre las masas:

$$\frac{m_j \sin(j)}{m_i \sin(i)} \simeq \frac{m_j}{m_i} \quad (1.41)$$

que no es afectado por la inclinación. **Esto es ventajoso desde el punto de vista dinámico, ya que se ha demostrado que el comportamiento de ambos planetas dentro de la resonancia depende principalmente de la razón de masas entre ellos y sólo débilmente de las masas individuales (ver Beaugé et al. 2003, 2006 para más detalles).**

El ajuste de velocidad radial con el modelo kepleriano, permite determinar la presencia de un compañero planetario, dando 4 de los 6 elementos planetarios para describir una órbita (los otros 2 son la longitud del nodo ascendente Ω y la inclinación orbital i). Eggenberger y Udry (2010) afirman que, la masa real del planeta y la masa mínima no difieren más de $m_p \leq 2m_p \sin(i)$, en casi el 90 % de los exoplanetas. Sin embargo no especifican cómo se estiman estos valores. Dado que el método por el cuál se puede determinar la inclinación i es mediante tránsito (1.3.5) y que la geometría del problema de tránsitos favorece detectarlos cuando $i \sim 90^\circ$, este valor puede contener un alto sesgo observacional.

Si consideramos un planeta en órbita circular, alrededor de una estrella de masa solar, la Ec. (1.35) se simplifica como:

$$K[m/s] = 28.57 \cdot m \cdot \sin(i) [M_{\text{jup}}] P^{-1/3} [\text{años}] \quad (1.42)$$

Debido a que la probabilidad de detectar un planeta, depende esencialmente del valor de la amplitud de la velocidad K , esta expresión indica que la medición Doppler favorece la detección de sistemas planetarios con planetas masivos y de corto período.

En el caso de tener más de un planeta, y cuando ellos no evidencien interacción, se pueden sumar las contribuciones keplerianas que produce cada uno por separado. Por lo tanto cada señal está caracterizada por M parámetros ($M = 5 \cdot n_{\text{pla}} + n_{\text{obs}}$), teniendo en cuenta el número de planetas (n_{pla}) y el número de conjuntos observacionales (n_{obs}).

La suma de las contribuciones individuales de los planetas ($V_r(i)$), está dada por la suma de ecuaciones de la forma de la Ecuación (1.34), es decir

$$\mathbf{V}_r = \sum_{i=1}^{n_{\text{pla}}} V_r(i). \quad (1.43)$$

Nombraremos los parámetros orbitales desde el planeta más interno al más externo por razones de simplicidad. Existe cierto criterio generalizado, de nombrar los planetas desde el más interno en orden alfabético, empezando por la letra b . Sin embargo cuando se descubren exoplanetas previamente no detectados, puede ocurrir que se encuentre más próximo a la estrella central y por lo tanto la regla alfabética pierde sentido (e.g. 55 Cnc e , Fischer et al. 2008)

En el caso de sistemas exoplanetarios donde la excentricidad de alguno de los planetas es cercana a cero ($e_i \sim 0$), a fin de evitar la indefinición del ángulo ϖ_i , es conveniente usar variables regularizadas:

$$k = e \cos [\varpi] \quad (1.44)$$

$$h = e \sin [\varpi] \quad (1.45)$$

Para fijar algunas ideas, la Figura 1.14 muestra cómo cada uno de los parámetros primarios modificados cambia la forma de la velocidad radial.

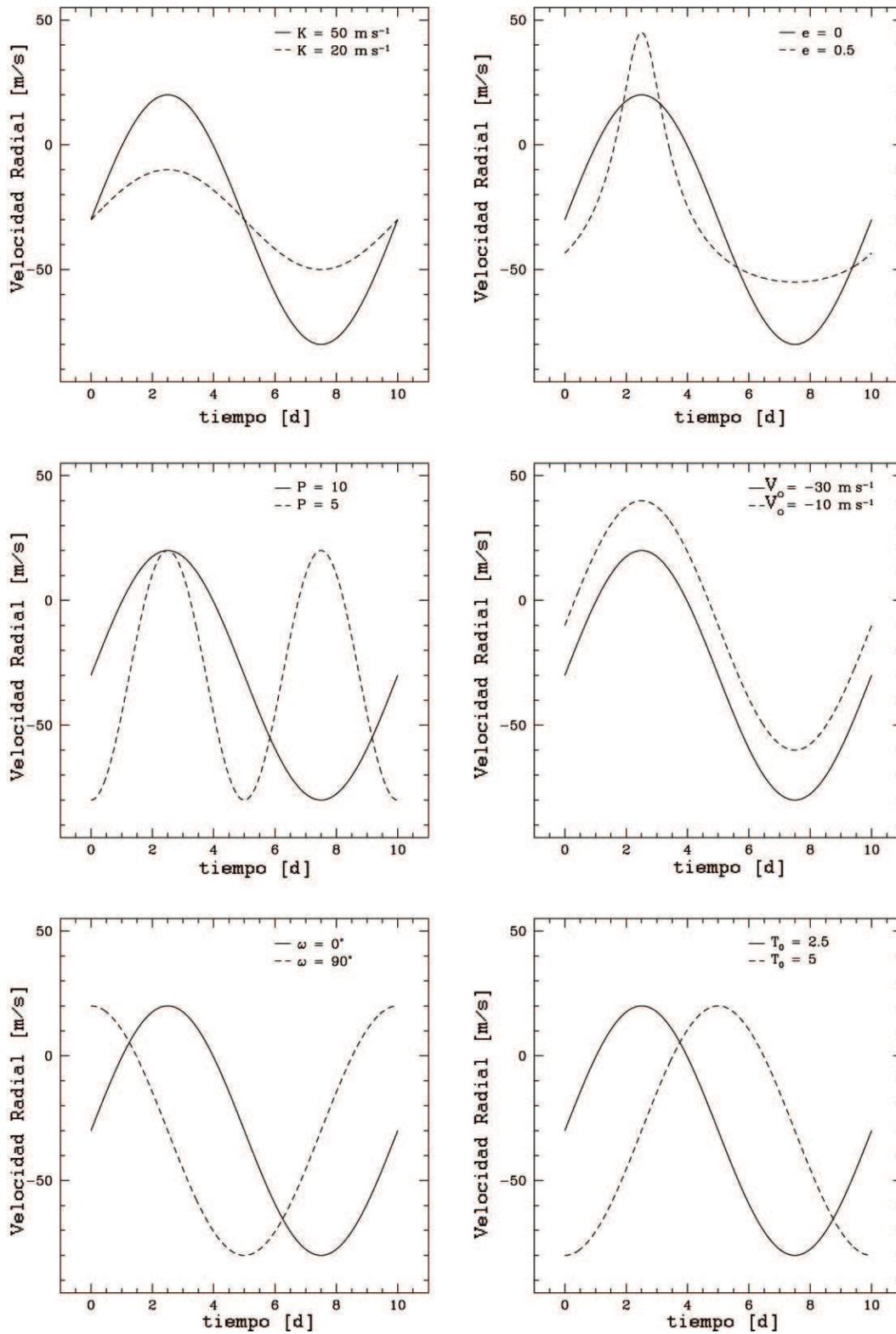


Figura 1.14: Ejemplos de dependencia del modelo -Ecuación (1.34)- con cada uno de los parámetros primarios. La señal representada por la línea continua tiene $P=10$ días, $e=0$, $K=50 \text{ m/s}$, $\omega = 0^\circ$, $\tau=2.5$ días, y $V_0=-30 \text{ m/s}$. Los parámetros primarios fueron cambiados de a uno por vez y en línea a trazos se dibuja la velocidad radial resultante. Adaptado de Eggenberger y Udry (2010).

1.3.6.5. Expansiones Elípticas

Podemos aproximar la expresión exacta de la Ecuación (1.34), escribiéndola en función de elementos orbitales osculadores y expandiendo las funciones trigonométricas:

$$V_r = K \left(e \cos(\omega) + \cos(f) \cos(\omega) - \sin(f) \sin(\omega) \right) + \mathbf{V}_0 \quad (1.46)$$

La anomalía verdadera f puede ser escrita, a través de la ecuación de Kepler y utilizando las expansiones en serie de Bessel como (Murray y Dermott, 1999):

$$\sin(f) = 2\sqrt{1-e^2} \sum_{s=1}^{\infty} \frac{1}{s} \frac{d}{de} J_s(se) \sin(sM) \quad (1.47)$$

$$\cos(f) = -e + \frac{2(1-e^2)}{e} \sum_{s=1}^{\infty} J_s(se) \sin(sM) \quad (1.48)$$

donde $J_s(se)$ es la función de Bessel de primer orden, dada por la expresión:

$$J_s(se) = \frac{1}{\pi} \int_0^{\pi} \cos(sE - se \sin(E)) dE. \quad (1.49)$$

Escribiendo explícitamente hasta orden 4 en excentricidades, obtenemos:

$$\begin{aligned} \sin(f) = & \sin(M) + e \sin(2M) + e^2 \left(\frac{9}{8} \sin(3M) - \frac{7}{8} \sin(M) \right) + \\ & e^3 \left(\frac{4}{3} \sin(4M) - \frac{7}{6} \sin(2M) \right) + \vartheta(e^4) \end{aligned} \quad (1.50)$$

$$\begin{aligned} \cos(f) = & \cos(M) + e(\cos(2M) - 1) + \frac{9}{8}e^2(\cos(3M) - \cos(M)) + \\ & \frac{4}{3}e^3(\cos(4M) - \cos(2M)) + \vartheta(e^4) \end{aligned} \quad (1.51)$$

Esta serie converge rápidamente para pequeños valores de e , aunque es divergente para valores de $e > 0.6627434$. Más adelante volveremos sobre este punto.

1.3.7. Efecto Rossiter-McLaughlin

En 1924 Richard Rossiter y Dean McLaughlin midieron la proyección, sobre el plano del cielo, del ángulo formado por el plano de la órbita y el eje de rotación estelar (ángulo *spin-órbita*= λ , de ahora en más) para las binarias eclipsantes, *beta – Lyrae* y *Algol* respectivamente. El trabajo consistió en modelar las variaciones de la velocidad radial de las estrellas, medidas durante el tránsito. Este efecto actualmente denominado *Rossiter-McLaughlin* (RM de ahora en más), se manifiesta cada vez que un objeto (estrella o planeta) oculta parte de la superficie estelar.

Es de interés particular medir este efecto en Júpiteres calientes⁵, pues no se entiende bien su proceso de formación. Se supone que los Júpiteres calientes se forman lejos de la estrella y luego sufren una

⁵ Planetas del orden de $1 M_{Jup}$ y períodos $P \sim 3$ días.

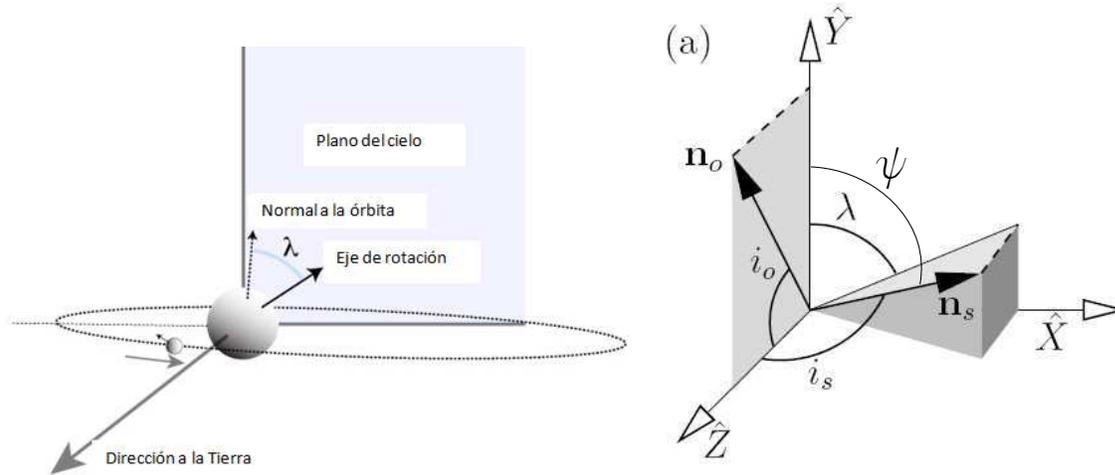


Figura 1.15: **Izquierda.** Esquema de la configuración espacial necesaria para observar el efecto RM. (adaptado de www.oklo.org). **Derecha.** Diagrama de orientación de ejes donde se ve más claramente que λ es una proyección y existe una simetría no resuelta sobre el eje X. Adaptado de Fabrycky y Winn (2009).

migración hacia la estrella hasta su posición actual. Trabajos recientes como Benitez-Llambay et al. (2010) tratan de explicar mejor estos procesos. Los mecanismos de migración no son bien conocidos y algunos de ellos difieren en el alineamiento spin-órbita:

- Existe una basta literatura que considera la interacción disco-planeta (Lin et al. 1996, Papaloizou y Terquem 2006) y, según los resultados de las integraciones numéricas, este mecanismo no afecta el ángulo λ .
- En contraste otras teorías sugieren eventos disruptivos como interacciones planeta-planeta (Rasio y Ford 1996) y colisiones de planetesimales (Murray et al. 1998) que pueden causar una aleatoriedad en la alineación spin-órbita.
- Los planetas en órbitas estables pueden presentar grandes oscilaciones en inclinación y excentricidad debido a perturbaciones seculares de planetas más alejados. Estas oscilaciones se encuentran acopladas y ocurren en un rango específico de valores; este efecto es conocido como Mecanismo de Kozai (Kozai 1962). Una combinación entre la circularización tidal y el mecanismo de Kozai podría explicar la formación de planetas cercanos a la estrella (Nagasawa et al. 2008) y midiendo la alineación spin-órbita se puede identificar este mecanismo en sistemas extrasolares.

La geometría del problema puede ser analizada en la parte derecha de la Figura 1.15. El sistema de coordenadas está orientado de tal forma que el eje \hat{Z} está en dirección al observador y \hat{X} yace sobre la línea de nodos de la órbita planetaria y \hat{Y} completa el sistema ortogonal. El nodo ascendente del planeta (la ubicación donde el planeta cruza el plano del cielo con $\dot{Z} > 0$), es en $X < 0$. En este sistema de coordenadas, n_o (la normal al plano de la órbita) está en el plano YZ y es especificado por el ángulo de inclinación de la órbita:

$$i_o = \cos^{-1}(n_o \cdot \hat{Z}), \quad (1.52)$$

que varía de 0 a $\pi/2$. Para especificar n_s se necesitan 2 ángulos, el ángulo de inclinación del spin de la estrella

$$i_s = \cos^{-1}(n_s \cdot \hat{Z}) \quad (1.53)$$

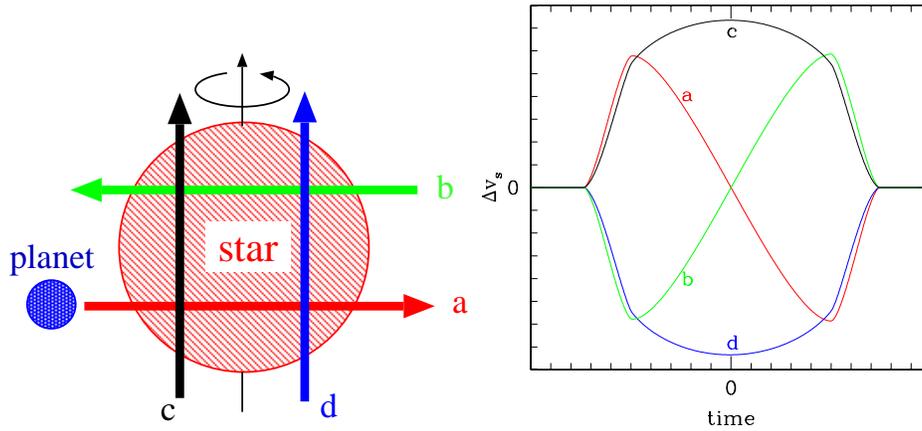


Figura 1.16: Ilustración esquemática de la curva de velocidad anómala que produce el efecto RM. Los 4 distintos caminos que puede seguir un planeta, a-d, en el panel de la izquierda corresponden a las curvas de velocidades en la derecha. Reproducido de Ohta et al. 2005

y el ángulo final λ que se entiende de la figura es en ángulo entre el eje \hat{Y} y la proyección de n_s sobre el plano del cielo XY . Concluyendo,

$$\begin{aligned} n_o &= \hat{Y} \sin i_o + \hat{Z} \cos i_o \\ n_s &= \hat{X} \sin i_s \sin \lambda + \hat{Y} \sin i_s \cos \lambda + \hat{Z} \cos i_s \end{aligned} \quad (1.54)$$

Para un planeta que transita, i_o es medible mediante el análisis de la fotometría del tránsito, y λ (el ángulo **proyectado** spin-órbita) mediante el efecto RM (el ángulo ψ es el ángulo **completo** spin-órbita). Cuando $|\lambda|$ es pequeño, entonces $|\lambda|$ es un límite inferior del valor de ψ . Un valor alto de $|\lambda|$ implica que hay una gran diferencia de ángulo spin-órbita. Por simetría una configuración (i_o, λ) no se puede distinguir de una configuración $(\pi - i_o, -\lambda)$. Valores positivos de λ tienen el spin estelar rotado en el sentido de las agujas del reloj y valores negativos de λ en sentido contrario de las agujas del reloj. Valores de $|\lambda|$ mayores que $\pi - i_o$ corresponden a órbitas retrógradas.

Si en su recorrido el planeta cruza a través del disco estelar de forma no paralela al ecuador de la estrella, el cero de velocidad radial no ocurre en el punto medio del tránsito y la curva de velocidad radial es asimétrica. La Figura 1.16 muestra como este efecto actúa dependiendo de la dirección del tránsito.

Las velocidades radiales observadas durante el tránsito pueden ser combinadas con otros datos, incluidas las V_r fuera del tránsito y la fotometría del tránsito, acotando la ubicación de la órbita del planeta.

Muchos trabajos recientes de exoplanetas realizan mediciones del efecto RM, aunque los límites de precisión restringen su aplicabilidad a casos específicos. Los ángulos medidos para los planetas extrasolares⁶ no evidencian tendencia hacia algún valor y se obtienen incertezas de unos $\pm 10^\circ$.

⁶http://www.aip.de/People/rheller/content/main_spinorbit.html.

1.3.7.1. Ecuaciones para modelar el efecto RM

Presentaremos un desarrollo básico de las ecuaciones en el efecto *RM*, aunque los pormenores y mejoras al modelo pueden ser consultadas en Giménez (2006) y Gaudi y Winn (2007).

Consideremos una estrella de masa M_* y radio R_* que tiene un planeta que transita con masa m y radio R_p . La órbita tiene un período P , excentricidad e y argumento de pericentro ω . Se puede escribir la variación de velocidad radial neta de la estrella (ΔV) como la suma de la velocidad radial de la estrella debido al movimiento orbital (ΔV_O , ver Ecuación (1.34)) y la velocidad radial anómala debido al efecto *RM* (ΔV_R):

$$\Delta V(t) = \Delta V_O(t) + \Delta V_R(t). \quad (1.55)$$

donde “O” es reservado para las contribuciones por el movimiento del planeta en su órbita y la “R” para contribuciones de los efectos Rossiter-McLaughlin y rotación estelar.

Por conveniencia se escribe el efecto de RM como:

$$\Delta V_R(t) = K_R g(t; x_p, y_p, \gamma, \epsilon, \dots), \quad (1.56)$$

separando la amplitud total K_R del efecto RM de la función adimensional $g(t) < 1$. La amplitud es dada por:

$$K_R \equiv V_S \sin I_S \frac{\gamma^2}{1 - \gamma^2}, \quad (1.57)$$

$$= 52.8 \text{ m/s} \left(\frac{V_S \sin I_S}{5 \text{ km/s}} \right) \left(\frac{R_p}{R_{\text{Jup}}} \right)^2 \left(\frac{R_*}{R_{\odot}} \right)^{-2} \quad (1.58)$$

donde $\gamma \equiv R_p/R_*$. En la última igualdad, se asumió que $\gamma \ll 1$. La función adimensional g depende de la posición proyectada (x_p, y_p) del planeta, pero también de γ y del oscurecimiento límbico (ϵ). Por simplicidad, se usa una función parametrizada linealmente con la ley de oscurecimiento límbico, de tal forma que el brillo superficial (no ocultado) de la estrella es:

$$\frac{\mathcal{I}(x, y)}{\mathcal{I}_0} = 1 - \epsilon \left[1 - (1 - x^2 - y^2)^{1/2} \right], \quad (1.59)$$

La Figura 1.17 muestra tres trayectorias de un planeta que transita sobre el disco estelar. Estas trayectorias tienen el mismo parámetro de impacto b , y consecuentemente todos ellos producen la misma señal fotométrica. (recordemos que el parámetro de impacto es dado por $b = a \cos i/R_*$, donde a es el semieje mayor). Sin embargo las trayectorias difieren en el valor de λ , y consecuentemente producen distintas formas de RM, como se dibuja en los paneles inferiores. La sensibilidad de la forma del efecto de RM con λ es lo que permite al observador estimar el alineamiento spin-órbita.

Un caso especial, simple de analizar, ocurre cuando el disco planetario es totalmente contenido dentro del disco estelar y el oscurecimiento límbico es despreciable ($\epsilon = 0$). En este caso, g es la distancia perpendicular desde el eje orbital estelar proyectado, $g(t) = x_p(t)$. Si consideramos una trayectoria rectilínea a través la cara de la estrella, con parámetro de impacto b , podemos escribir la posición del centro del planeta en función del tiempo como:

$$\begin{aligned} x_p(t) &= t_\tau \cos \lambda - b \sin \lambda, \\ y_p(t) &= t_\tau \sin \lambda + b \cos \lambda, \end{aligned} \quad (1.60)$$

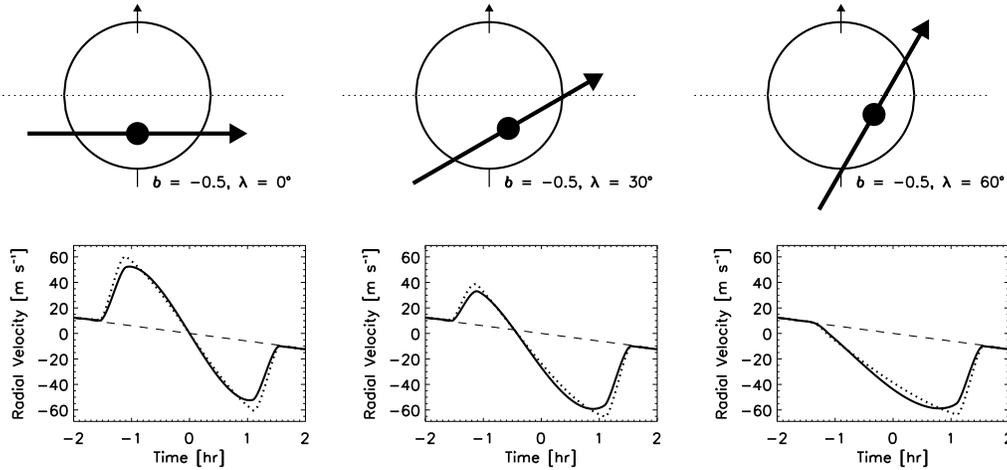


Figura 1.17: La dependencia del efecto RM con respecto a λ . Se dibujaron tres trayectorias posibles de un planeta que transita. Las configuraciones tienen el mismo parámetro de impacto ($b = -0.5$) y producen la misma curva de luz, pero difieren en λ y producen distintas curvas RM. En los paneles inferiores se muestra la velocidad radial resultante, para el caso donde no hay oscurecimiento por limbo ($\epsilon = 0$) con líneas punteadas; mientras que en líneas sólidas se reproduce la velocidad radial cuando $\epsilon = 0.6$. Adaptado de Gaudi y Winn, (2007).

donde $t_\tau \equiv (t - t_{\text{tra}})/T_{\text{tra}}$, t_{tra} es el tiempo del punto medio del tránsito, $T_{\text{tra}} = R/v_{\text{orb}}$ es el tiempo de cruce por el radio correspondiente a la velocidad orbital del planeta en el tiempo de tránsito (debido a esto la duración del tránsito es aproximadamente $2T_{\text{tra}}\sqrt{1-b^2}$) y λ es el ángulo de la trayectoria con respecto a el ecuador estelar aparente. Se define λ entre -180° y $+180^\circ$; Si $\lambda > 0$, el planeta se mueve hacia el norte estelar cuando cruza a través del disco. La indeterminación del parámetro de impacto $b \leftrightarrow -b$ presente en la solución de la señal fotométrica, es resuelta cuando se considera el efecto RM.

En el caso de rotación uniforme, sin oscurecimiento límbico y cuando el planeta está totalmente contenido en el disco estelar, la curva de RM es simplemente una función lineal del tiempo (ecuación 1.60). Si asumimos que el planeta es pequeño ($\gamma \ll 1$), y no se tienen en cuenta las fases de ingreso y egreso⁷, se usa la aproximación simple dada por:

$$g(t) = \begin{cases} x_p(t) & \text{si } \sqrt{t_\tau^2 + b^2} \leq 1, \\ 0 & \text{caso contrario.} \end{cases} \quad (1.61)$$

En el caso de considerar el oscurecimiento límbico, la expresión de g es más complicada y necesita ser evaluada numéricamente. Sin embargo se pueden consultar los trabajos de Ohta et al. (2005) y Giménez (2006) para obtener algunas aproximaciones analíticas.

Para conocer la importancia del efecto RM podemos comparar la amplitud de la velocidad radial, K_O (Ec. 1.35), con la amplitud de la velocidad radial anómala K_R (Ec. 1.62) que produce el efecto RM. Asumiendo órbitas circulares ($e = 0$), con ángulos proyectados pequeños $\gamma \ll 1$, y con

⁷ya que constituyen una pequeña fracción de la duración total del tránsito ($\sim 2\gamma$)

inclinación $i = 90^\circ$, se obtiene (Gaudi y Winn, 2007):

$$\frac{K_R}{K_O} = \left(\frac{PV^3}{2\pi Gm} \right)^{1/3} \left(\frac{\rho_*}{\rho_p} \right)^{2/3}, \quad (1.62)$$

donde ρ_p y ρ_* son las densidades medias del planeta y la estrella, respectivamente. Suponiendo que ρ_*/ρ_p no varían demasiado entre sistemas estelares, el orden de magnitud de K_R/K_O depende principalmente de: el período orbital, la masa del planeta y la velocidad proyectada de rotación de la estrella:

$$\frac{K_R}{K_O} \sim 0.3 \left(\frac{m}{M_{\text{Jup}}} \right)^{-1/3} \left(\frac{P}{3 \text{ días}} \right)^{1/3} \left(\frac{V}{5 \text{ km/s}} \right) \quad (1.63)$$

Para sistemas extrasolares comparables a la masa de Júpiter y con períodos de 1–4 días (*Júpiters Calientes*) esta relación es menor que 1. Sin embargo las propiedades de los sistemas conocidos tienen efectos de selección ya mencionados: los tránsitos de planetas grandes y de corto período son mucho más fáciles de detectar que planetas pequeños de largo período (Gaudi, 2005). Una aplicación interesante de la Ec. (1.63) es que para los planetas más difíciles de detectar, la amplitud del efecto de RM **excede** la amplitud de la velocidad orbital. En particular para un planeta de masa terrestre con un período de 1 año, $K_R/K_O \sim 3$ para $V = 5 \text{ km/s}$ (Gaudi y Winn, 2007).

Podemos hacer una estimación simple para determinar los órdenes de magnitud de λ que produce el efecto RM. Para esto, es útil considerar los tiempos relativos de 3 eventos observables:

- I. el instante t_{tra} de la mayor profundidad del tránsito. En $t = t_{\text{tra}}$, la distancia proyectada estrella-planeta es mínima.
- II. el instante t_{orb} cuando la variación de velocidad radial es nula. En $t = t_{\text{orb}}$, la estrella se mueve en el plano del cielo.
- III. el instante t_{rot} cuando la velocidad anómala que produce la variación RM es nula. En $t = t_{\text{rot}}$, el planeta yace directamente en frente del eje de rotación estelar.

Para una órbita circular con $\lambda = 0$, estos 3 eventos son simultáneos. Si la órbita es circular, pero $\lambda \neq 0$, entonces $t_{\text{tra}} = t_{\text{orb}}$ pero t_{rot} ocurre antes o después (utilizando la Ec. 1.60):

$$\frac{t_{\text{rot}} - t_{\text{tra}}}{T_{\text{tra}}} \approx b \tan \lambda. \quad (1.64)$$

Para órbitas no circulares, t_{tra} y t_{orb} no son más simultáneos. A primer orden en excentricidades se obtiene,

$$t_{\text{orb}} - t_{\text{tra}} = \frac{P}{2\pi} (e \cos \omega), \quad (1.65)$$

donde ω es el argumento del pericentro. Al mismo orden, la diferencia de tiempo $t_{\text{rot}} - t_{\text{tra}}$ y la duración del tránsito T_{tra} son multiplicadas por el mismo factor $1 - e \sin \omega$, y así la Ec. (1.64) continúa siendo válida. Estas expresiones dan una idea de la precisión en tiempo necesaria para poder medir λ . Por ejemplo, para un tránsito en latitud media en $b = 0.5$ que dure 2.5 horas, una inclinación spin-órbita de $\lambda = 1^\circ$ corresponde a un desplazamiento de tiempo de 45 segundos entre el tiempo medio de tránsito y la anulación de la forma del evento RM.

De la misma forma se puede derivar la incerteza en λ , pero basados en las mediciones espectroscópicas. Consideremos una serie de N velocidades radiales tomadas en tiempos t_k durante el

Tabla 1.1: Valores de observables para detectar planetas tipo Júpiter y tipo terrestres. La estrella central se asume que es idéntica al Sol y que se encuentra a 10 pc del observador.

Técnica	Observable necesario	Júpiter	Tierra
Imagen directa	contraste de brillo en $\lambda_{visible}$	6×10^{-7}	1.5×10^{-10}
Imagen directa	contraste de brillo en $\lambda = 10\mu\text{m}$	1.5×10^{-7}	1.2×10^{-7}
Imagen directa	separación angular	0".50	0".1
Astrometría	amplitud astrométrica	500 μas	0.3 μas
Velocidad radial	amplitud de V_r	13 m/s	0.1 m/s
Tránsito	profundidad del tránsito ΔF	1 %	0.01 %
Timing	residuos del timing TOAs	2.5 s	1.5 ms
RM	amplitud RV dentro tránsito	52.8 m/s	0.4 m/s

tiempo de tránsito (entre el primer y cuarto contacto, ver Figura 1.9), cada una de las cuales tiene un error asociado σ (el mismo para todas). Consideremos también el caso donde las mediciones RM son la única fuente de error (asumimos que ambos, el tránsito fotométrico y la órbita determinada por espectroscopia han sido medidas precisamente). De esta forma las incertezas en los tiempos de contacto, t_{tra} , y los parámetros γ y b tienen incertezas despreciables y la velocidad orbital $\Delta V_O(t)$ puede ser precisamente extraída de la variación total de velocidad observada durante el tránsito para aislar la forma de RM. Los únicos parámetros a determinar cuando se ajusta el tránsito son λ y V , que se pueden combinar en un vector bidimensional \vec{a} . Las incertezas esperadas en λ y V para un conjunto de datos equiespaciados y grandes son:

$$\sigma_\lambda = Q_R^{-1} \left[\frac{(1 - b^2) \sin^2 \lambda + 3b^2 \cos^2 \lambda}{b^2(1 - b^2)} \right]^{1/2}, \quad (1.66)$$

$$\frac{\sigma_V}{V} = Q_R^{-1} \left[\frac{(1 - b^2) \cos^2 \lambda + 3b^2 \sin^2 \lambda}{b^2(1 - b^2)} \right]^{1/2}, \quad (1.67)$$

donde se definió

$$Q_R \equiv \sqrt{N} \frac{K_R}{\sigma}. \quad (1.68)$$

El factor Q_R es proporcional a la razón señal sobre ruido. Es destacable la relación existente entre las Ecs. (1.66) y (1.67) es $\sigma_V(\lambda)/V = \sigma_\lambda(\pi - \lambda)$. De todas formas se puede deducir que la incerteza depende de la geometría del tránsito b y la geometría de la órbita λ .

1.4. Efectos de Selección y Límites de detección

El amplio rango de propiedades de los exoplanetas implica grandes diferencias a la hora de detectarlos. Muchas veces se usan las referencias de la Tierra y Júpiter como comparación de los potenciales observables provenientes de los métodos de detección. La Tabla 1.1 muestra los valores típicos de señal que deberían ser detectados por una técnica para asegurar detecciones positivas.

Podemos resumir la información de las secciones precedentes en el plano *Masa mínima* en función del semieje como se ve en la Figura 1.18. Los planetas del Sistema Solar se encuentran marcados como referencia. Las líneas verticales indican períodos típicos de referencia. Las líneas horizontales están

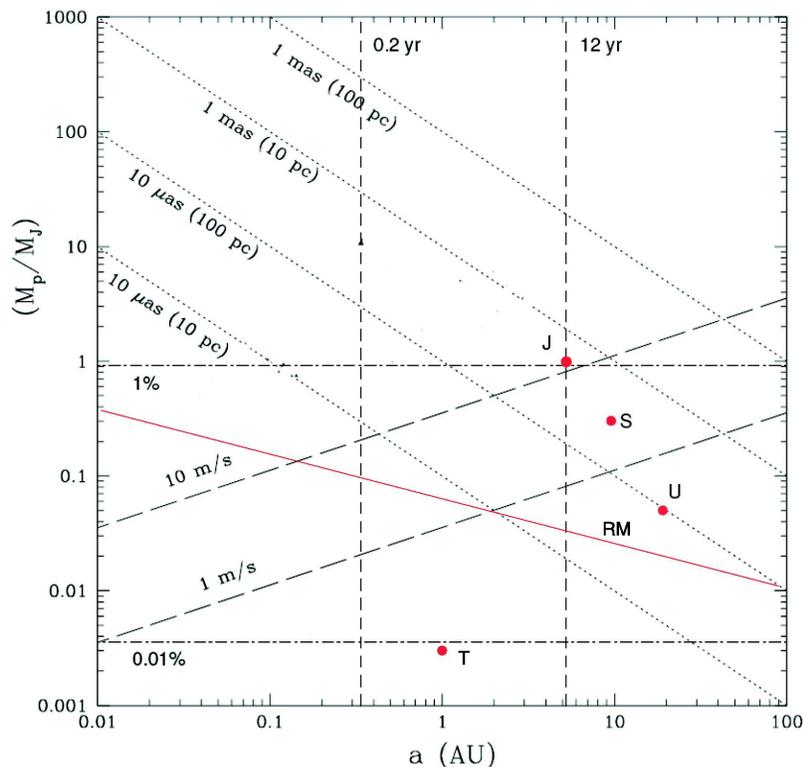


Figura 1.18: Límites de detección según las distintas técnicas. Los puntos rojos indican la posición de los planetas del Sistema Solar nombrados con sus iniciales.

asociadas a la profundidad del tránsito necesarias para la detección (1 % y 0.01 %), las líneas a trazos con pendiente positiva indican los límites para la amplitud detectada por velocidad radial (10 m/s y 1 m/s), mientras que las líneas a trazos con pendiente negativa indican la precisión astrométrica necesaria en las posiciones para estrellas a distintas distancias. Por ejemplo para detectar un planeta tipo Saturno necesitamos medir profundidades de tránsito menores que $\sim 0.1\%$; o una amplitud de velocidad radial de 5 m/s; o una precisión astrométrica de $10\ \mu\text{as}$ para una estrella a 100 pc. La línea roja negativa marca el límite de detección del efecto RM para las precisiones actuales; se observa que gran cantidad de los actuales exoplanetas son plausibles de medir y así brindar más pistas acerca del pasado evolutivo de cada sistema (ver Subsección 1.3.7).

Ahora bien, si sobre la Figura 1.18 dibujamos los elementos orbitales de los exoplanetas descubiertos hasta abril de este año podemos fijar aún más ideas (Figura 1.19, adaptada con la Figura de http://en.wikipedia.org/wiki/File:Exoplanet_Period-Mass_Scatter.png). En primer lugar notamos que la gran mayoría de los sistemas extrasolares tienen obviamente grandes masas, mucho más que decenas de veces la masa terrestre. Esta tendencia hacen el análisis estadístico complicado, sin embargo estas altas masas son en parte a un efecto de selección: todos los métodos de detección son más sensibles a descubrir planetas masivos. A pesar que los planetas tipo terrestres son más difíciles de detectar, aparentemente son mucho más comunes (Butler et al. 2006) y es así que las técnicas de detección más sensibles como microlente gravitacional y pulsar timing tienen su peso estadístico.

Muchos exoplanetas orbitan alrededor de su estrella a distancias mucho más cercanas que en nuestro propio Sistema Solar. Esto nuevamente es debido a efectos de selección, ya que la velocidad

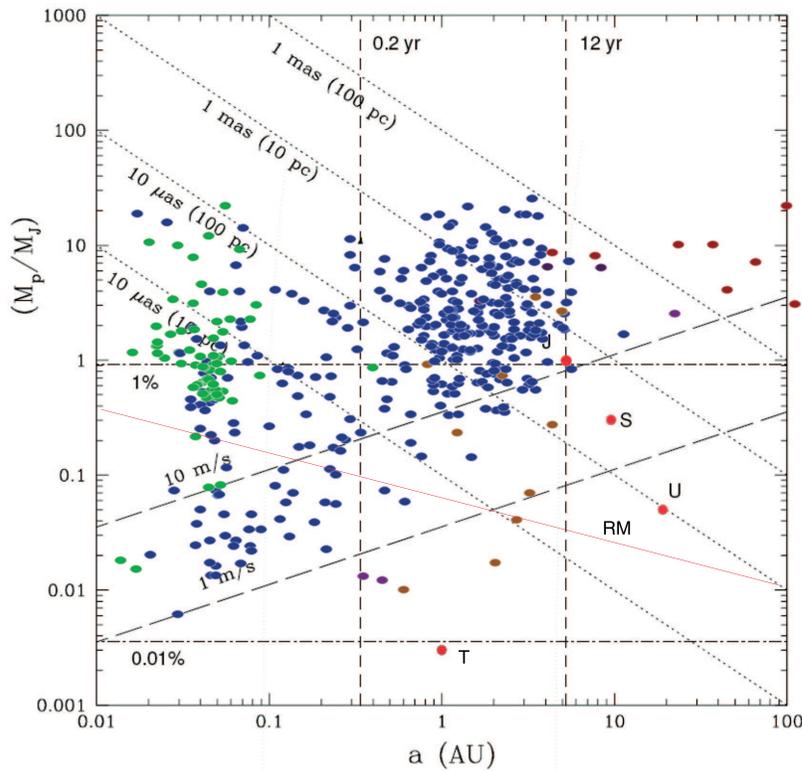


Figura 1.19: Límites de detección según las distintas técnicas. Los puntos rojos son referentes a la posición de los planetas del Sistema Solar. Los distintos puntos de colores son exoplanetas catalogados según el método de detección: verde: tránsitos, azul: velocidad radial, violeta: pulsar timing, marrón: imagen directa y naranja: microlente gravitacional. Existen 2 grupos bien diferenciados: los puntos verdes, planetas predominantemente cercanos a la estrella y los azules un poco más dispersos, pero exhibiendo cierta concentración alrededor de ~ 1 [UA].

radial es un método más sensible a detectar planetas con órbitas más cercanas. Los astrónomos, inicialmente se sorprendieron por estos planetas gigantes cercanos a las estrellas que denominaron *hot Jupiters* (Júpiteres Calientes). Según los más recientes trabajos acerca de la formación de exoplanetas, aparentemente al menos uno o dos exoplanetas deberían existir en cada sistema, con órbitas y tamaños comparables a Júpiter y Saturno (Crida et al., 2009).

Los planetas detectados por Tránsito presentan períodos más cortos debido principalmente al método de detección: las probabilidades de detección de tránsito de un planeta que está lejos de la estrella son muy bajas porque su posibilidad de transitar sobre el disco estelar es inversamente proporcional al tamaño de la órbita. Sumado a ello, la probabilidad de realizar la observación justo durante el tránsito es un factor importante.

Gran parte de los exoplanetas tienen órbitas excéntricas. Esto, en principio, no es un efecto observacional ya que un planeta puede ser detectado igual, sea o no excéntrico. La prevalencia de órbitas elípticas es una incógnita, ya que algunas teorías de formación sugieren que deberían ser circulares. Sin embargo veremos que parte de este sesgo puede ser un efecto de selección (Capítulo 5).

Capítulo 2

Obtención de los parámetros planetarios

A pesar que en el capítulo anterior discutimos varias técnicas de detección, a lo largo del resto de la tesis nos restringiremos a las detecciones por velocidad radial.

La determinación de los parámetros orbitales de sistemas exoplanetarios a partir de los datos de velocidad radial, es un proceso complejo. Aunque los primeros sistemas exoplanetarios eran rutinariamente observados por varios períodos antes de su publicación, en la actualidad los candidatos planetarios son publicados cuando las observaciones cubren apenas un período orbital e inclusive con regiones de la órbita no uniformemente observadas. Existen también algunos sistemas donde se dispone de varios conjuntos observacionales, como para *Epsilon Eridani* y *Upsilon Andromedae*, aumentando la información disponible, pero a su vez haciendo más difícil la tarea de la interpretación de los datos. Por estos motivos es fácil interpretar erróneamente los parámetros orbitales y el incremento del número de observaciones a lo largo del tiempo puede modificar los ajustes obtenidos. Las ecuaciones que relacionan las observaciones con los elementos orbitales (y masas planetarias mínimas) son altamente no lineales, permitiendo en principio la existencia de varios mínimos locales en el espacio de parámetros.

La mayoría de las veces las observaciones de velocidad radial (y_i) se encuentran disponibles de manera libre. En algunos casos los autores no proveen los datos y se deben emplear métodos de extracción de la información contenida en las imágenes publicadas, en forma de archivos postscript (en la bibliografía este tipo de datos es denominado *dexterated data*).

Dado el conjunto de observaciones se necesita resumir e interpretar los datos ajustándolos a un “modelo” que depende de parámetros ajustables. Esto se conoce como: **Modelización de Datos**. Disponemos de las N velocidades radiales observadas (y_i) para cada tiempo (t_i), y por otro lado para cada tiempo (t_i) podemos calcular la señal de la velocidad $V_r(t_i)$ con la Ec. (1.34). El objetivo que se plantea es encontrar el conjunto de parámetros planetarios que reproduzcan la función $V_r(t_i)$ de tal forma que su comportamiento sea lo más parecido posible a los datos observacionales (y_i). Existen distintos enfoques para determinar los parámetros planetarios $A_i = A(K, T, e, \omega, \tau, \mathbf{V}_0)$:

- Los **métodos de Fourier**. Se utiliza una descomposición armónica de la señal en un conjunto de frecuencias. Las frecuencias pueden ser preestablecidas (como el caso de la transformada discreta de Fourier, aplicable sólo en el caso de datos equiespaciados) o se puede utilizar un rango cualquiera de frecuencias de interés (como en el caso de la transformada de Fourier para datos compensados).

- Los **métodos de Cuadrados Mínimos**. Dada una función modelo se calculan simultáneamente frecuencias y parámetros planetarios, tal que ellos sirvan para reproducir el comportamiento global de la función que queremos modelar. Debido a la forma no-lineal de la función, existen dos enfoques que utilizaremos:

Los **métodos determinísticos**. Consisten en un método iterativo, en los cuales se comienza con una solución (potencialmente pobre) y se van introduciendo cambios en la solución, cada uno mejorando la anterior. Estos métodos buscan la dirección en la cual la función residuo \mathcal{R} decrece, ya sea utilizando evaluaciones de la función (por ejemplo Simplex), e inclusive sus derivadas (Métodos de Gradiente, por ejemplo Levenberg-Marquardt). Son métodos muy sensibles a las condiciones iniciales donde comienzan las iteraciones.

Los **métodos estocásticos**. Se incorpora la selección probabilística (aleatoria) en el proceso de minimización. Esta aleatoriedad puede estar presente en la forma de la selección de los elementos de la función (por ejemplo Simulated Annealing) o bien en el algoritmo mismo (como por ejemplo métodos evolutivos como Algoritmos Genéticos). La aleatoriedad introducida puede proveer la “iniciativa” necesaria de posicionarse lejos de la solución local cuando se busca un óptimo global.

También es necesario definir una *función de mérito* que mida el acuerdo entre los datos y el modelo. La función de mérito es convencionalmente diseñada de tal forma que pequeños valores de ella representen el mejor acuerdo con los datos. Los parámetros del modelo son entonces ajustados para lograr un mínimo en la función de mérito, obteniendo así los parámetros del mejor ajuste. Dicho de otra manera, el proceso de ajuste es un problema de minimización en varias dimensiones.

La obtención de los parámetros no es el final del problema, para ser genuinamente útil el proceso de ajuste debe:

- a) brindar los parámetros del mejor ajuste
- b) estimar los errores en los mismos
- c) medir estadísticamente la bondad del ajuste

Cuando el tercer punto sugiere que el modelo no es representado por los datos, entonces los puntos (a) y (b) no tienen importancia. Ellos nos indican si un ajuste es aceptable cuando el gráfico de los datos y el modelo son parecidos. La base de los siguientes desarrollos pueden ser consultadas en Press et al. (1992).

Este tipo de problemas muchas veces utiliza los denominados métodos globales, ya que realizan un barrido general de todas las soluciones posibles en el espacio de parámetros, su costo computacional es elevado y nunca se tiene la certeza de haber encontrado el mínimo global de la función. De alguna manera todos los métodos pueden modificarse para que su búsqueda sea global dentro del espacio de parámetros, con mayor o menor eficiencia, recordando tres reglas:

- **Nunca sentirse afortunado.** Puede ser que sólo tengamos un mínimo local en lugar del mínimo global.
- **Todo tiene un costo.** Si queremos encontrar el mínimo global debemos trabajar para obtenerlo. Asociado a esta regla se deben considerar tres aspectos relacionados con la eficiencia:

1. Eficiencia Absoluta. ¿Cuán precisa es numéricamente la solución dada por mi método?
2. Eficiencia Global. ¿Cuán seguro estoy que la solución es el verdadero mínimo de mi solución?
3. Eficiencia Relativa. ¿Cuánto tiempo de cómputo es necesario para hallar la solución?

Los métodos más elegantes son designados para cumplir tanto como sea posible los puntos (1) y (3). Sin embargo el punto (2) es bien resuelto sólo si se tiene una condición inicial conveniente.

- La evolución es más inteligente que uno! Si bien por el momento esta frase parece no tener sentido, veremos con la introducción de **Algoritmos Genéticos** la importancia de la misma.

2.1. Métodos de Fourier

Si bien los métodos de Fourier constituyen una de las herramientas numéricas más comunes, también resultan ser una de las más complicadas. Los posibles problemas no solo se restringen a dificultades numéricas, sino también teóricas. El tratamiento erróneo de los datos no solo puede dar resultados imprecisos sino también pueden ser completamente equivocados. En cierto sentido, el principal propósito del Análisis de Fourier es el siguiente: dada una función $f(t)$, encontrar una aproximación $\hat{f}(t)$ tal que:

$$\hat{f}(t) = \sum_{\nu} F_{\nu} e^{-i2\pi\nu t}. \quad (2.1)$$

En otras palabras, deseamos encontrar una descomposición *armónica* de la función. Si ésta es naturalmente periódica o casi-periódica (también llamada de multiperiódica), lo que buscamos esencialmente es determinar cuales son las frecuencias ν involucradas en la variación de la función V_r con respecto de t . Por el momento vamos a suponer que V_r es definida sobre un cierto intervalo finito de la variable independiente $t \in [0, T]$. Notemos que, en la sumatoria de la Ec. (2.1), ν es una variable y no un índice entero¹.

No obstante este método tiene la importante limitación de que únicamente se evalúa la potencia del espectro en determinadas frecuencias (las frecuencias de Fourier) que pueden no ser en las que estamos interesados. Veamos a continuación como funcionan los métodos de Fourier.

2.1.1. Principios de Fourier

Para representar una V_r cualquiera en el espacio de funciones $L_2[0, T]$ (cuadrado integrables), primero debemos elegir una base para el espacio. Recordemos que el espacio posee dimensión infinita, y por lo tanto precisamos infinitas funciones base para representar cualquier elemento. Se puede denotar una base genérica de la forma:

$$\{\phi_{\nu}\} \quad \text{donde} \quad \phi_{\nu}(t) \in L_2[0, T] \quad (2.2)$$

y donde necesitamos infinitos (y adecuados) valores de ν tal de generar todo el espacio. El parámetro ν puede variar en forma discreta o continua pero, cualquiera sea el caso, si solo consideramos un numero **finito** de ellos, el conjunto $\{\phi_{\nu}\}$ generara un sub-espacio (del espacio total).

¹A fin de evitar confundir terminología denominaremos **e** a la base del logaritmo natural, “e=2.7182..”; alternativamente trabajaremos con desarrollos planares y reservamos $i = \sqrt{-1}$.

Si consideramos un conjunto discreto de valores de ν , la aproximación de la función V_r en este espacio modelo es representada por la combinación lineal:

$$\hat{f} = \sum_{\nu} F_{\nu} \phi_{\nu} \quad (2.3)$$

donde los coeficientes F_{ν} son incógnita.

Debido a que \hat{f} puede ser considerada como una proyección de V_r en un sub-espacio, podemos definir la *función residuo* R tal que:

$$V_r = \hat{f} + R \quad (2.4)$$

que se corresponde con las componentes de V_r que no fueron proyectadas; o sea, aquellas componentes que son normales al sub-espacio modelo. En general, los coeficientes F_{ν} serán elegidos de forma que $|R|^2$ sea el menor posible (i.e. la mejor aproximación).

Un caso particular de gran interés, dentro de la Mejor Aproximación, consiste en elegir las funciones bases como:

$$\phi_{\nu}(t) = \mathbf{e}^{-i2\pi\nu t} \quad (2.5)$$

para ciertos valores de ν que pueden ser tanto discretos como continuos. Estas funciones (2.5) forman lo que se llama de *base armónica* del espacio de funciones. Existen infinitas bases posibles, ya que podemos generar distintas secuencias:

$$\{\phi_{\nu}\} = (\mathbf{e}^{-i2\pi\nu_1 t}, \mathbf{e}^{-i2\pi\nu_2 t}, \mathbf{e}^{-i2\pi\nu_3 t}, \dots) \quad (2.6)$$

simplemente eligiendo los ν_i de forma diferente. Si reemplazamos (2.5) en (2.3), siempre que podamos suponer el residuo muy pequeño, obtenemos que los coeficientes de la mejor aproximación armónica son dados por:

$$F_{\nu} = \frac{1}{T} \sum_{\omega} \mathbf{B}_{\omega, \nu}^{-1} \int_0^T f(t) \mathbf{e}^{i2\pi\omega t} dt \quad (2.7)$$

donde el tensor métrico es dado por:

$$\begin{aligned} \mathbf{B}_{\nu, \omega} &= \frac{1}{T} \int_0^T e^{i2\pi(\nu - \omega)t} dt \\ &= \mathbf{e}^{i\pi(\nu - \omega)t} \frac{\sin(\pi(\nu - \omega)T)}{\pi(\nu - \omega)T}. \end{aligned} \quad (2.8)$$

Una consecuencia muy importante de esta ecuación es que las funciones base $\mathbf{e}^{-i2\pi\nu t}$ en general no definen una base ortogonal, por lo tanto resulta complicado calcular la matriz inversa de $\mathbf{B}_{\nu, \omega}$. En otras palabras, aunque el producto interno $\langle \mathbf{e}^{-i2\pi\nu t}, \mathbf{e}^{-i2\pi\omega t} \rangle = 1$ si $\nu = \omega$, en general no se cumple que el producto interno sea nulo cuando $\nu \neq \omega$.

Independientemente de las dificultades puramente numéricas, esto implica que el coeficiente correspondiente a la frecuencia ν se verá afectado (o contaminado) por las contribuciones de otras frecuencias diferentes. Esto es inevitable al trabajar con bases no-ortogonales, y es una característica que será importante a la hora de definir y trabajar con las Transformadas de Fourier.

2.1.2. Transformada discreta de Fourier

La definición clásica de la *Transformada de Fourier* de una función f en una frecuencia ν arbitraria esta dada por:

$$F(\nu) \equiv \langle \mathbf{e}^{-i2\pi\nu t} | f \rangle = \frac{1}{T} \int_0^T f(t) \mathbf{e}^{i2\pi\nu t} dt. \quad (2.9)$$

En otras palabras, la transformada de Fourier es la *proyección* de la función original f sobre la función $\phi_\nu(t) = \mathbf{e}^{-i2\pi\nu t}$.

Si disponemos, por ejemplo de una integración numérica, podemos muestrear con un conjunto equiespaciado de N puntos, de forma tal que $f_n = f(t_n)$ donde

$$t_n = (\Delta, 2\Delta, \dots, N\Delta) \quad \rightarrow \quad t_n = n\Delta. \quad (2.10)$$

La cantidad Δ se denomina *escala del muestreo* (i.e. *sampling rate*) y denota el intervalo de t entre datos. Tomando la función peso como $w_n \equiv 1/N \forall n$, tenemos que la Transformada Discreta de Fourier (i.e. *Discrete Fourier Transform - DFT*) esta definida como:

$$\hat{F}(\nu) = \frac{1}{N} \sum_{n=1}^N f_n \mathbf{e}^{i2\pi\nu n\Delta}. \quad (2.11)$$

Ahora bien, para cada valor de Δ existe un valor critico de ν , llamada *frecuencia critica de Nyquist* denotada por ν_c , que es dada por:

$$\nu_c = \frac{1}{2\Delta}. \quad (2.12)$$

Esta cantidad es muy importante porque denota cual es la máxima frecuencia que es identificada por el muestreo de datos.

En las aplicaciones prácticas, no es posible evaluar la transformada de Fourier de una función discreta en *todos* los valores continuos de frecuencias, sino que se trabaja con un cierto conjunto finito de frecuencias específicas. Estas frecuencias son elegidas de tal forma que el tensor métrico (Ec. 2.8) sea la matriz identidad. De hecho, lo que comúnmente se entiende como *Transformada de Fourier Discreta* (DFT) se refiere no sólo a la discretización de los datos, sino también de las frecuencias del espectro.

El final del intervalo de la variable independiente se corresponde con $T = N\Delta$, de forma tal que tenemos en total N datos. También vamos a suponer (y esto es importante) que el numero total de datos es par.

Con estas especificaciones, vamos a considerar un espacio modelo generado por las funciones base $\mathbf{e}^{-i2\pi\nu_k t_n}$, donde las frecuencias cumplen con la condición (**frecuencias de Fourier**):

$$\nu_k = \frac{k}{T} = \frac{k}{N\Delta} \quad ; \quad k = -\frac{N}{2}, \dots, 0, \dots, \frac{N}{2}. \quad (2.13)$$

Con esta elección, el tensor métrico es la matriz identidad (ver Ec. 2.8), y la transformada de Fourier (ecuación (2.11)) es dada por:

$$F_{\nu_k} = \langle \mathbf{e}^{-i2\pi\nu_k t_n} | f \rangle = \frac{1}{N} \sum_{n=1}^N f_n \mathbf{e}^{i2\pi k n \Delta / T}. \quad (2.14)$$

Notemos que, para los efectos de cálculo, hicimos variar el índice k de 0 a N , siendo que los valores de F_k para $k \geq \frac{N}{2}$ van a corresponderse a las frecuencias negativas ν_{k-N} .

La *Anti-Transformada Discreta* posee la forma análoga:

$$f_n = \frac{1}{N} \sum_{k=1}^N F_k e^{-i2\pi kn\Delta/T} \quad (2.15)$$

Esta función discreta reconstituida se denomina *Serie de Fourier Discreta*.

La DFT posee otra desventaja que tiene que ver con el número de operaciones numéricas que son necesarias para su cálculo. Aunque esta limitación es de índole numérica y no teórica, puede ser muy importante.

Una forma de reducir el número de operaciones es lo que se conoce como la *Transformada Rápida de Fourier - FFT*. No obstante este método tiene la importante limitación de que únicamente se evalúa la potencia del espectro en determinadas frecuencias (las frecuencias de Fourier) que pueden no ser las que interesan. Más precisamente, es una descomposición de la serie temporal en una base ortonormal definida por un número finito $N/2$ de frecuencias, donde N es el número de datos.

El método además es definitivamente inútil para evaluar bajas frecuencias. A pesar de que a menudo es utilizada, la FFT no debe aplicarse a observaciones no equiespaciadas puesto que no se cumple la hipótesis fundamental (datos equiespaciados). Peor aún, suele utilizarse la FFT para evaluar la potencia en frecuencias que no son las de Fourier, alegando de que se está trabajando con la transformada continua. Todo esto hace que la potencia evaluada mediante FFT sea incorrecta y por lo tanto se identifiquen mal las frecuencias asociadas a la señal periódica.

La idea de la FFT consiste en aprovechar el caso cuando N es un número par, y reagrupamos los términos de la ecuación (2.14) en la forma:

$$\begin{aligned} F_{\nu_k} &= \sum_{n=1}^N f_n e^{i2\pi kn/N} = \sum_{n=1}^{N/2} f_n e^{i2\pi kn/N} + \sum_{n=N/2+1}^N f_n e^{i2\pi kn/N} \\ &= \sum_{n=1}^{N/2} f_n e^{i2\pi kn/N} + \sum_{n=1}^{N/2} f_{n+N/2} e^{i2\pi k(n+N/2)/N} \\ &= \sum_{n=1}^{N/2} (f_n + f_{n+N/2} e^{i\pi k}) e^{i2\pi kn/N}. \end{aligned} \quad (2.16)$$

En conclusión, tenemos que:

$$F_k = \sum_{n=1}^N f_n e^{i2\pi kn/N} \longrightarrow \sum_{n=1}^{N/2} f_n^{(1)} e^{i2\pi kn/N}, \quad (2.17)$$

donde los nuevos coeficientes son $f_n^{(1)} = (f_n + f_{n+N/2} e^{i\pi k})$. Si ahora $N/2$ es nuevamente un número par (i.e. N divisible por 2^2), el proceso puede ser repetido, subdividiendo nuevamente la última sumatoria en dos términos, reagrupando los sumandos, definiendo la función $f_n^{(2)}$, etc.

Este proceso puede ser continuado m veces, siempre que el límite de la sumatoria resultante $N/2^m$ sea par. Así, el caso más favorable ocurre cuando el número de datos es en realidad una potencia de dos. En tal caso, la transformada de Fourier se traduce en un único término donde su coeficiente $f_0^{(m)}$

es una suma de $(m - 1)$ evaluaciones de la función. El número total de operaciones a ser realizadas se reduce desde N^2 a $N \log_2 N$. Según Press et al. (1992), si tomamos $N = 10^6$ el cálculo de la FFT es 40000 veces más rápido que la DFT.

2.1.3. Determinación de Parámetros orbitales con análisis de Fourier

Recordemos la forma de la función periódica que queremos analizar:

$$V_r = K[\cos(f + \omega) + e \cos(\omega)] + \mathbf{V}_0, \quad (2.18)$$

donde vemos que la dependencia temporal se encuentra en la anomalía verdadera f . Haciendo uso de la ecuación de Kepler podemos expandir f^2 en series de excentricidades e_j tales que:

$$\mathbf{e}^{if_j} = \sum_{k=-\infty}^{k=\infty} C_k(e_j) \mathbf{e}^{ikM_j} \quad (2.19)$$

donde $M_j = n_j(t - \tau)$ es la anomalía media, τ el tiempo de pasaje por el perihelio y $C_k(e_j)$:

$$C_k(e_j) = \frac{1}{2\pi} \int_0^{2\pi} \left(\cos E - e_j + i\sqrt{1 - e_j^2} \sin E \right) \mathbf{e}^{-ik(E - e_j \sin E)} \quad (2.20)$$

Reemplazando en la expresión general se obtiene:

$$V_r = \sum_{j=1}^N K_j \mathbf{e}^{i\omega_j} \sum_{k=1}^{\infty} C_k(e_j) \mathbf{e}^{ik \cdot n_j(t - \tau)} + \mathbf{V}_0, \quad (2.21)$$

A quinto orden en excentricidades, los coeficientes $C_k(e_j)$ pueden escribirse como³:

$$C_1(e_j) = \left(1 - \frac{9}{8}e_j^2 + \frac{25}{192}e_j^4 \right) + i \left(1 - \frac{7}{8}e_j^2 + \frac{17}{192}e_j^4 \right) \quad (2.22)$$

$$C_2(e_j) = \left(1 - \frac{4}{3}e_j^2 + \frac{3}{8}e_j^4 \right) e_j + i \left(1 - \frac{7}{6}e_j^2 + \frac{1}{3}e_j^4 \right) e_j \quad (2.23)$$

Según Correia (2008) podemos escribir la transformada de Fourier para un número discreto de observaciones N , no equispaciadas, en un intervalo de tiempo $[0, T]$:

$$F(\nu) \approx \frac{1}{T} \sum_{k=2}^N V_k \mathbf{e}^{-i\nu t_k} (t_k - t_{k-1}), \quad (2.24)$$

donde V_k es la velocidad radial medida en la fecha t_k . Los parámetros orbitales pueden ser determinados como se detalla a continuación:

■ Determinación de V_0

²haciendo uso de los desarrollos de la Sección (1.3.6.5).

³solo escribiremos $k=1$ y $k=2$, ya que es suficiente considerar sólo dos frecuencias para este análisis.

Reemplazando la Ec. (2.18) en la transformada de Fourier dada por la Ec. (2.24) y eligiendo $\nu = 0$, siempre que los datos sean más o menos uniformemente distribuidos en el intervalo de observación, se obtiene:

$$F(0) = V_0, \quad (2.25)$$

y una vez obtenido este corrimiento es sugerible sustraer este valor de los datos V_k y continuar el análisis de Fourier sin esa parte en la Ec. 2.18.

■ Determinación de n_j

La frecuencia n_1 corresponde al compañero de la estrella con mayor amplitud K_1 , dado por la frecuencia ν asociada al pico más alto en el espectro de potencias, es decir:

$$|F(n_1)| \geq |F(\nu)|, \forall \nu \quad (2.26)$$

Luego de localizar n_1 es relativamente sencillo encontrar los otros parámetros. Una vez que la órbita del primer compañero es completamente establecida, se recomienda sustraer su contribución y seguir evaluando para encontrar otros posibles compañeros de la estrella.

■ Determinación de parámetros orbitales restantes (e_j, K_j, τ_j y ω_j)

Reemplazando la expresión (2.21) en (2.24) y eligiendo $\nu = n_j$ y $\nu = 2n_j$

$$F(n_j) = K_j e^{i\omega_j} C_1(e_j) e^{-in_j\tau_j} \quad (2.27)$$

$$F(2n_j) = K_j e^{i\omega_j} C_2(e_j) e^{-i2n_j\tau_j} \quad (2.28)$$

donde las cantidades $F(n_j)$ y $F(2n_j)$ se determinan haciendo las transformadas de Fourier con la Ec. (2.24). Tomando módulo de las expresiones de arriba se obtiene:

$$|F(n_j)| = K_j |C_1(e_j)| \text{ y } |F(2n_j)| = K_j |C_2(e_j)| \quad (2.29)$$

de la que se obtiene una expresión implícita para la excentricidad que puede ser resuelta con algún método iterativo tal que se satisfaga la ecuación:

$$\frac{|F(2n_j)|}{|F(n_j)|} = \frac{|C_2(e_j)|}{|C_1(e_j)|} \quad (2.30)$$

Una vez determinado e_j se puede deducir directamente el valor de K_j de la Ec. (2.29):

$$K_j = \frac{|F(2n_j)|}{|C_2(e_j)|} = \frac{|F(n_j)|}{|C_1(e_j)|}, \quad (2.31)$$

luego tomando el cociente entre 2.27 y 2.28:

$$e^{in_j\tau_j} = \frac{F(n_j)C_2(e_j)}{F(2n_j)C_1(e_j)}, \quad (2.32)$$

para finalmente obtener:

$$e^{i\omega_j} = \frac{F^2(n_j)C_2(e_j)}{K_j F(2n_j)C_1^2(e_j)} \quad (2.33)$$

Concluyendo entonces, es posible determinar un compañero estelar a partir de los datos observacionales, haciendo la correspondiente transformada de Fourier para 3 frecuencias, a saber $F(0)$, $F(n)$ y $F(2n)$. En esta deducción se empleó n y $2n$ pero se podría haber elegido cualquier frecuencia múltiple de n . Sin embargo, a no ser que la excentricidad sea muy elevada, las frecuencias consideradas corresponden a los picos más altos y por lo tanto son fácilmente reconocibles en el espectro de potencias.

Más aún, si la excentricidad es cercana a 0, (que es en teoría el caso para los “Júpiteres Calientes”, $F(kn) \approx 0$ excepto para $k = 0, 1$; En este caso, w_j y T_j no pueden ser determinados, pero aún así es posible establecer la posición del planeta en la órbita, $\lambda_j = \omega_j - n_j\tau$ como:

$$e^{i\lambda_j} = \frac{F(n_j)}{K_j C_1(e_j)}. \quad (2.34)$$

Los parámetros orbitales determinados con este método, obviamente tienen error y son proporcionales a la precisión de las observaciones e inversamente proporcionales al número de datos disponibles. Si se dispone de un gran número de datos se incrementa la convergencia entre la transformada discreta y la continua.

Como ya mencionáramos el acuerdo entre los parámetros de Fourier y los parámetros verdaderos se puede mejorar trabajando con la función de mérito χ^2_ν . La utilización de estos métodos es mucho más rápida que métodos de cuadrados mínimos, ya que los parámetros determinados ya están asociados a la periodicidad de la señal y por lo tanto muy cerca de la función real. Inclusive cuando las interacciones entre los planetas son débiles, los resultados son satisfactorios. Este método ha sido probado en tres sistemas planetarios distintos (Correia et al. 2005, Lovis et al. 2006 y Pepe et al. 2007) dando los mismos resultados que otros métodos alternativos.

El problema de este método es que si tenemos, por ejemplo, 2 planetas con períodos conmensurables, es probable que el planeta de menor señal no sea detectado. Esto se debe a que la señal débil puede ser enmascarada por un armónico de la primer señal. Otra desventaja es que las frecuencias de Fourier pueden no estar bien determinadas, debido a que las observaciones son no equiespaciadas y a veces la señal no cubre suficientemente varios períodos orbitales.

Existen alternativamente otros métodos de análisis de Fourier, como el propuesto por Konacki y Maciejewski (1999), donde se analiza la relación entre los picos del periodograma. El método ha sido desarrollado utilizando series convergentes de elementos orbitales y demostrando que cada uno de los elementos de la serie tiene amplitudes más pequeñas a medida que el orden de la serie aumenta.

2.1.4. Análisis de Fourier en datos no equiespaciados

La estimación de frecuencias (y del espectro de potencia) de una serie temporal formada por datos no-equiespaciados es un problema común. En astronomía, la disponibilidad del telescopio, el clima y las estaciones son los principales responsables de que las series de observaciones consecutivas, varíen desde horas a semanas e inclusive meses.

Como hemos visto en las secciones precedentes, a veces se usa la expresión Ec. (2.11) para datos no equiespaciados con la excusa que los resultados son comparables. De hecho, el análisis de datos simulados muestra que los resultados están en buen acuerdo con la frecuencia real en una amplia variedad de situaciones, excepto algunos casos que mencionaremos más adelante. Gran parte de los

resultados y métodos explicados a continuación fueron desarrollados por Ferraz-Mello (1981). Teóricamente el uso de una transformada de Fourier finita para datos no equiespaciados es inválida, salvo en el caso de un proceso Gaussiano (ruido blanco) en cuyo caso encontrar el espectro no tiene sentido.

Varias técnicas se propusieron para estimar el espectro, algunos métodos consideran métodos de ordenamiento de fase, que son derivados de la idea empírica de probar con distintos períodos hasta que una curva suave es obtenida. Con respecto a este estadístico podemos citar los trabajos de Pel't (1975), el método de Minimización de Dispersión de Fase (e.g. Stellingwerf, 1978) o el Método de Mínima Entropía (e.g. Cincotta et al. 1995 y 1999).

Los métodos actuales de transformada de Fourier para tratar con datos no-equiespaciados consisten en el análisis de frecuencia por cuadrados mínimos, donde la función $\hat{f} = H_0 + H_1 \cos(2\pi\nu t) + H_2 \sin(2\pi\nu t)$. Para valores tentativos de la frecuencia ν se calculan las constantes H_i y el Coeficiente de Correlación Espectral (Ferraz-Mello 1981, Foster 1996) que es un número menor a 1 y está asociado a la probabilidad de que la frecuencia ν sea real y no producto del azar.

Existe una amplia mención en la literatura al método Lomb (1976) y Scargle (1982) (*Lomb-Scargle*). Sin embargo la base ortogonalizada utilizada es incompleta ($H_0 = 0$) y no considera los pesos individuales de los datos. Como consecuencia, no es posible realizar un filtrado de los armónicos en una serie temporal. El método fue mejorado por Ferraz-Mello (1981) (seguidamente por Foster (1996)) utilizando la base completa y considerando los pesos individuales.

Consideremos el proceso periódico $g(t) = y(t) + x(t)$, donde $y(t)$ es una función periódica, y $x(t)$ un proceso Gaussiano representado por una variable normal cuya media es nula y su varianza es σ^2 ($\equiv N(0, \sigma^2)$).

Asumamos la serie temporal $f(t_1), f(t_2), \dots, f(t_N)$, formada por muestreo del proceso periódico, a los cuales se sumó una constante tal que:

$$\sum_{l=1}^N f(t_l) = 0 \quad (2.35)$$

Para una frecuencia de prueba ν , el ajuste de la curva de regresión sinusoidal de máxima similitud tiene una base:

$$\begin{aligned} H_0(t) &= 1 \\ H_1(t) &= \cos(2\pi\nu t) \\ H_2(t) &= \sin(2\pi\nu t) \end{aligned} \quad (2.36)$$

Podemos construir una base $\{h_0, h_1, h_2\}$ ortonormal por medio del método Gram-Schmidt,

$$\begin{aligned} h_0 &= a_0 H_0, \\ h_1 &= a_1 H_1 - a_1 h_0(h_0, H_1) \\ h_2 &= a_2 H_2 - a_2 h_0(h_0, H_2) - a_2 h_1(h_1, H_2) \end{aligned} \quad (2.37)$$

En estas fórmulas, los paréntesis indican el producto interno,

$$(g_1, g_2) = \sum_l^N g_1(t_l) \cdot g_2(t_l) \quad (2.38)$$

Los coeficientes a_0 , a_1 y a_2 en la Ec. (2.37) son determinados de tal forma que:

$$(h_0, h_0) = (h_1, h_1) = (h_2, h_2) = 1 \quad (2.39)$$

así finalmente se obtiene:

$$\begin{aligned} a_0^{-2} &= N, \\ a_1^{-2} &= (H_1, H_1) - a_0^2(H_0, H_1)^2 \\ a_2^{-2} &= (H_2, H_2) - a_0^2(H_0, H_2)^2 - a_1^2(H_1, H_2)^2 - a_1^2 a_0^4 (H_0, H_1)^2 (H_0, H_2)^2 \\ &\quad + 2 a_0^2 a_1^2 (H_0, H_1)(H_0, H_2)(H_1, H_2) \end{aligned} \quad (2.40)$$

donde los coeficientes de la regresión son:

$$\begin{aligned} c_0 &= (f, h_0) = 0, \\ c_1 &= (f, h_1) = a_1(f, H_1) - a_0^2(H_0, H_1)^2 \\ c_2 &= (f, h_2) = a_2(f, H_2) - a_1 a_2 c_1 [(H_1, H_2) - a_0^2(H_0, H_1)(H_0, H_2)] \end{aligned} \quad (2.41)$$

La función en la nueva base ortonormal puede ser entonces expresada como:

$$f(t) = c_0 h_0(t) + c_1 h_1(t) + c_2 h_2(t) \quad (2.42)$$

Para datos equiespaciados este procedimiento equivale a una aproximación de ajuste a una curva utilizando tres parámetros: una “sinusoide” más una constante. Si el período no es conmensurable con el espaciado de los datos y la serie es suficientemente larga para cubrir toda la fase uniformemente, uno tiene:

$$(H_0, H_1) \approx (H_0, H_2) \approx (H_1, H_2) \approx 0 \quad (2.43)$$

$$(H_1, H_1) \approx (H_2, H_2) \approx \frac{1}{2} a_0^2 = \frac{N}{2} \quad (2.44)$$

con

$$\begin{aligned} c_1 &\approx a_0 \sqrt{2} (f, H_1) \\ c_2 &\approx a_0 \sqrt{2} (f, H_2) \end{aligned} \quad (2.45)$$

De esta forma, el cálculo del *espectro de potencia* es dado por:

$$I(\nu) = 2a_0^2 \|F(\nu)\|^2 = c_1^2 + c_2^2 \quad (2.46)$$

donde $F(\nu)$ es la transformada discreta de Fourier (**DFT**) dada por la Ec. (2.11). Para series temporales con datos no-equiespaciados, la medición del espectro de potencia es equivalente, pero ahora $F(\nu)$ es la **transformada de Fourier discreta de datos compensados (DCDFT)**:

$$F(\nu) = (f, h_1 + ih_2) / (a_0 \sqrt{2}). \quad (2.47)$$

El producto interno dado por la ecuación (2.38), debe ser modificado de acuerdo a los datos que se consideran. Para considerar la diferencia de precisiones entre distintas observaciones, se usa la función de peso $w_i = w(t_i)$ (asociada al inverso cuadrado del error en las mismas)

$$(g_1, g_2) = \sum_l^N w_{t_l} g_1(t_l) \cdot g_2(t_l), \quad (2.48)$$

debido a esta nueva definición del producto interno, se deduce que:

$$\begin{aligned} a_0^{-2} &= \sum w_i, \\ a_1^{-2} &= \sum w_i \cos^2 x_i - a_0^2 \left(\sum w_i \cos x_i \right)^2 \\ a_2^{-2} &= \sum w_i \sin^2 x_i - a_0^2 \left(\sum w_i \sin x_i \right)^2 - a_1^2 \left[\sum w_i \cos x_i \sin x_i - a_0^2 \left(\sum w_i \sin x_i \right) \left(\sum w_i \cos x_i \right) \right]^2, \end{aligned} \quad (2.49)$$

donde $x_i = 2\pi\nu t_i$ y todas las sumas desde $i = 1$ hasta $i = N$. Los coeficientes de la regresión son:

$$\begin{aligned} c_0 &= 0 \\ c_1(\nu) &= a_1 \sum w_i f(t_i) \cos x_i \\ c_2(\nu) &= a_2 \sum w_i f(t_i) \sin x_i - a_1 a_2 c_1 \left[\sum w_i \cos x_i \sin x_i - a_0^2 \left(\sum w_i \sin x_i \right) \left(\sum w_i \cos x_i \right) \right]^2, \end{aligned} \quad (2.50)$$

el cálculo del espectro de potencia es la intensidad dada por:

$$I(\nu) = c_1^2 + c_2^2 \quad (2.51)$$

La teoría de la regresión lineal pide que $0 \leq I(\nu) \leq Q$, donde Q :

$$Q = (f, f) = \sum w_i f^2(t_i) \quad (2.52)$$

Esta propiedad es utilizada para introducir una escala de normalización el estadístico:

$$S(\nu) = I(\nu)/Q \quad (2.53)$$

que se denomina *coeficiente de correlación espectral* y es tal que $0 \leq S(\nu) \leq 1$ para todo ν . En un caso degenerado, cuando $f(t_i)$ es un ruido blanco (es decir no tiene ningún tipo de periodicidad), los coeficientes de regresión c_1 y c_2 son variables normales con la misma varianza que $f(t_i)$ y $I(\nu)/\sigma^2$ sigue una distribución chi-cuadrado. La cantidad $(N-1)S(\nu)$ puede ser considerada como una variable de prueba. Para un dado resultado $S(\nu)$ permite calcular la probabilidad de obtener un resultado cuando se trabaja con ruido blanco calculando:

$$\text{Prob}(S(\nu)) \geq A = \text{Prob}(\chi_\nu^2) \leq (N-1)A \quad (2.54)$$

es decir para un dado ν , la probabilidad que $S(\nu)$ sea mayor que un dado límite de detección preestablecido (A) puede ser expresado en función de la distribución de una variable chi-cuadrado.

2.1.5. Diferencias entre DFT y DCDF

La principal diferencia entre ambos métodos es que DFT utiliza frecuencias de Fourier ($\nu_k = \frac{k}{T}$), mientras que DCDF realiza una regresión en base sinusoidal para cualquier valor de frecuencia ν_i deseado. Además, las bases y justificaciones de la DFT son basadas en el equiespaciado de los datos, mientras que en la DCDF esta condición no es necesaria.

Existen trabajos donde se utiliza la DFT (Ec. 2.11) para datos no equiespaciados, con el pretexto que los resultados son comparables (Deeming, 1975). Ferraz-Mello (1981) realizó una amplia variedad de pruebas, estimando el espectro para datos que simulaban observaciones no-equiespaciadas de un proceso periódico muestreado en los tiempos t_i :

$$x_i = \text{sen}(\nu t_i + A) + B\epsilon_i \quad (2.55)$$

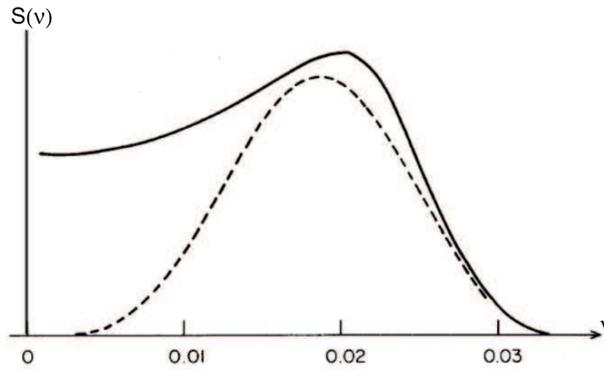


Figura 2.1: Espectro obtenido usando DCDFT (curva sólida) y DFT (curva a trazos) de un proceso periódico simulado para una frecuencia muy baja ($\nu = 0.02$) y relación $S/N=3.5$.

donde ϵ_i es una variable de distribución normal $N(0, 1)$, siendo A una constante arbitraria y B un factor relacionado a la razón señal-ruido (S/N).

A continuación expondremos algunos de sus resultados. En primer lugar se confirmó que se puede utilizar la DFT cuando solamente se está interesado en la estimación del **período** y cuando la relación señal/ruido es pequeña ($S/N \leq 2.5$). Cuando la relación S/N es alta, la estimación no compensada dada por la DFT lleva a *errores en la determinación del período*, que dependen del tamaño del muestreo y del número de períodos cubiertos. Sin embargo los cálculos usando (2.41) son mucho más rápidos que con (2.50) por lo que es conveniente, por lo menos en primera aproximación, calcular usando DFT y luego mejorar con DCDFT. La justificación del uso de la DFT estándar, reside en el hecho que los coeficientes a_1 y a_2 (de la DCDFT) oscilan alrededor del valor medio $a_0\sqrt{2}$.

Ferraz Mello mostró que, aún cuando la DCDFT no sea mejor que la DFT en la determinación del período en una serie de puntos (cuando $S/N \leq 2.5$), tiene la ventaja de dar una mejor estimación de la altura del pico. Este último punto es crucial si se quiere interpretar correctamente la señal, más que solamente determinar un período. Las discrepancias entre los espectros determinados con la DCDFT y con la DFT se vuelven notables para frecuencias pequeñas. De hecho para pequeñas frecuencias (períodos grandes) no hay suficiente cobertura en el espacio de fase y los valores de a_1 y a_2 de la DCDFT pueden tomar valores muy distintos de $a_0\sqrt{2}$. La Figura 2.1 muestra el espectro estimado para un proceso periódico de muy baja frecuencia ($\nu = 0.02$) y relación $S/N=3.5$, con un muestreo de 60 datos distribuidos sobre un período de la señal (2.55). Los resultados muestran la DCDFT en curva continua y la DFT curva a trazos: la DFT muestra un error del 5% mientras que la DCDFT da el valor correcto para el período.

2.1.6. Filtrado de armónicos

El proceso periódico estudiado puede estar conformado por varias señales. A pesar que se puede seleccionar un período para la inspección del espectro, nada se puede decir de las otras periodicidades a través del análisis del mismo espectro: otros picos presentes pueden ser alias del pico principal. Si uno desea buscar otras periodicidades aparte de la seleccionada, es fácil filtrar la serie temporal y eliminar la *onda* cuyo período ha sido identificado. Denominemos $\bar{\nu}$ al período cuya onda está definida

por:

$$d(\bar{\nu}, t) = c_1(\bar{\nu})h_1(\bar{\nu}, t) + c_2(\bar{\nu})h_2(\bar{\nu}, t). \quad (2.56)$$

La serie temporal filtrada será:

$$\bar{f}(t) = f(t) - d(\bar{\nu}, t) \quad (2.57)$$

y por lo tanto obtenemos:

$$d(t_j) = d_0 + d_1 \cos(\bar{x}_j) + d_2 \sin(\bar{x}_j), \quad (2.58)$$

con $\bar{x}_j = 2\pi\bar{\nu}t_j$. Según Ferraz-Mello (1981) puede obtenerse que:

$$\begin{aligned} d_2 &= a_2 c_2 \\ d_1 &= a_1 c_1 + d_2 a_1^2 \left[a_0^2 \left(\sum w_j \cos \bar{x}_j \right) \left(\sum w_j \sin \bar{x}_j \right) \left(\sum w_j \cos \bar{x}_j \sin \bar{x}_j \right) \right] \\ d_0 &= -a_0^2 \left[d_1 \sum w_j \cos \bar{x}_j + d_2 \sum w_j \sin \bar{x}_j \right] \end{aligned} \quad (2.59)$$

En adición, se pide que:

$$(\bar{f}, h_0) = 0 \quad (2.60)$$

Finalmente calculando \hat{Q} :

$$\hat{Q} = (\bar{f}, \bar{f}) = Q[1 - S(\bar{\nu})] \quad (2.61)$$

que permite que nos demos cuenta que $S(\bar{\nu})$ es una estimación de la información relacionada con la frecuencia seleccionada. Los coeficientes de correlación espectral de la serie filtrada, evaluados en $\bar{\nu}$ son cero, es decir:

$$\hat{S}(\bar{\nu}) = 0 \quad (2.62)$$

Vale la pena mencionar que en el caso donde la base no incluye $H_0(t)$ (como en el periodograma Lomb-Scargle), se puede construir el filtrado armónico pero no se cumple la Ec. (2.62) debido a la incompletitud de la base (es realizado un filtrado incompleto). La calidad del filtrado depende de la calidad del estimador de la intensidad del espectro de potencias en la frecuencia seleccionada ν y eso justifica el uso de DCDFE.

2.1.7. Probabilidades de detección y límites de confianza

Cuando uno construye la función $S(\nu)$ en un rango de frecuencias y con un determinado paso de muestreo, se puede graficar la función $S(\nu)$ en función de ν , denominado **Periodograma**. Si uno quisiera evaluar de forma aproximada cuando es importante una señal $S(\nu)$, se puede establecer un nivel de ruido y considerar importantes todos los picos dos o tres veces mayores que dicho límite. Una forma más rigurosa de evaluar la importancia de un pico en el periodograma es calcular el FAP⁴, definido como la probabilidad de obtener un pico por azar con alta señal, cuando se trabaja con ruido blanco. Algunas formas utilizadas para calcular el FAP de una señal son:

⁴del inglés *False Alarm Probability*, i.e. Probabilidad de Falsa Alarma.

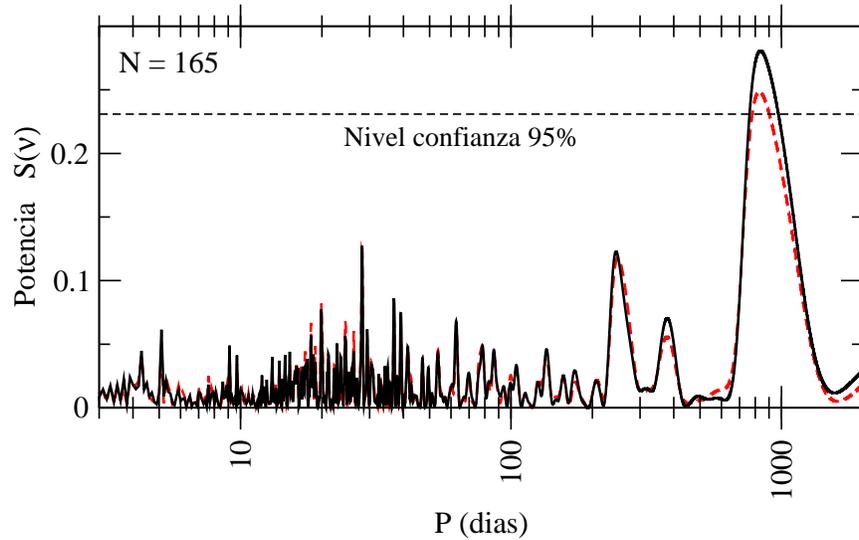


Figura 2.2: Espectro obtenido usando DCDFT con pesos individuales todos iguales (curva sólida) y pesos individuales proporcionales a los errores en cada medición (curva a trazos). Los datos corresponden al residuo de un ajuste con dos planetas en el sistema HD82943. La relación S/N se puede estimar en ~ 2.5 y el límite de confianza del 95 % fue calculado con las aproximaciones de las Ecs. 2.63. Dado el alto nivel de confianza del pico adicional, podemos asumir como real la existencia del tercer planeta en el sistema.

- Cálculos Analíticos. Según se menciona en Ferraz-Mello y Quast (1987), se puede deducir una fórmula general y calcular las siguientes cantidades:

$$\begin{aligned}
 G &= -\frac{N-3}{2} \ln(1-S) \\
 H &= \frac{N-4}{N-3} (G + e^{-G} - 1) \\
 \alpha &= \frac{2(N-3)\Delta t \cdot \Delta\nu}{3(N-4)} \\
 C &= (1 - e^{-H})^\alpha,
 \end{aligned} \tag{2.63}$$

donde Δt es el intervalo de tiempo cubierto por las observaciones N (con $N \gg 10$), $\Delta\nu$ es el rango de las frecuencias muestreadas que integran el periodograma construido. C es la confianza del resultado y $(1 - C) \equiv FAP$ debe ser interpretado como la probabilidad de tener la altura del pico más alto por azar. Por ejemplo $C = 0.05 \equiv 5\%$ es considerado un número grande para esta probabilidad. En los trabajos de descubrimientos de exoplanetas es normal encontrar valores de $FAP = 10^{-3}$ para considerar una buena detección.

- Trabajos del grupo del Observatorio de Ginebra proponen determinar el FAP numéricamente (Lovis et al. 2010). La idea es realizar permutaciones aleatorias de los residuos de un ajuste, calculando el pico de potencia de cada intento. Se construye la distribución de estas potencias (que puede ser modelada con una distribución gaussiana). La estimativa del FAP sería entonces la probabilidad de obtener un valor igual a la potencia de la señal original. La Figura 2.3 ilustra como funciona este método.
- Otros autores (e.g. Anglada-Escudé et al. 2009) proponen una forma numérica alternativa. Los autores generan un conjunto de datos sintéticos utilizando la mejor solución del ajuste Keplero. Cada conjunto sintético se genera sumando a la velocidad radial nominal un ruido gaus-

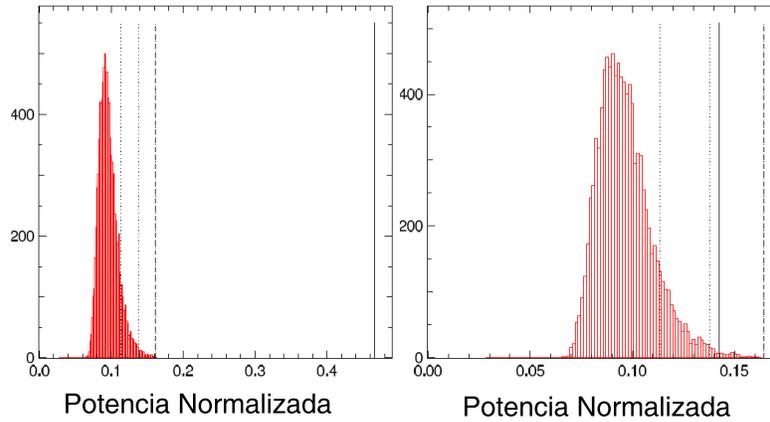


Figura 2.3: Distribuciones de picos de potencia para los planetas del sistema HD10180. Izquierda una señal que es significativa y derecha una señal que tiene un FAP de 0.6 %.

siano⁵ y se redetermina el mejor ajuste orbital. Este proceso se repite varias veces. Se realizan 1000 conjuntos sintéticos y se evalúa el porcentaje de ajustes orbitales que mejora la solución original. Ese porcentaje es el *FAP* buscado. Este es un proceso computacionalmente muy lento y engorroso y si el FAP es pequeño, muchas veces se deben generar hasta 10^5 conjuntos sintéticos.

La Figura (2.2) muestra el periodograma sobre los residuos del ajuste orbital kepleriano de las velocidades radiales (para el sistema HD82943 con 2 planetas). En otras palabras se busca la presencia de una señal periódica adicional. Ambas curvas fueron calculados con DCDFT pero uno considera todos los datos con el mismo peso (curva sólida, $P \sim 836$ d) y el otro dando los pesos individuales a cada observación (curva a trazos, $P \sim 829$ d). En la figura se muestra el nivel de confianza $C=95$ % (marcado como una línea horizontal) y calculado con la aproximación dada por la Ec. (2.63). El valor del pico del espectro (S_ν) y la frecuencia asociada (ν) a este pico difieren levemente.

Veamos un ejemplo de la determinación numérica del FAP. La Figura 2.3 muestra la distribución de potencia de picos (Peak power distribution) generando permutaciones aleatorias de los residuos del ajuste. A la izquierda se muestra una señal con gran probabilidad de ser planeta, mientras que a la derecha se muestra una señal con FAP del 0.6 % (= 0.006), es decir 94 % probable de ser planeta. La línea continua muestra la potencia del pico en los datos reales, mientras que las líneas a trazos muestran los límites de detección con FAPs del 10 %, 1 % y 0.1 % basados en estas distribuciones. En el gráfico de la izquierda se dibuja la Potencia de la señal de ~ 0.45 (línea negra vertical). La distribución de los picos en cada conjunto de residuo generado posee la distribución mostrada en rojo. Se ajusta una distribución Gaussiana y se averigua el área normalizada bajo la curva que queda a la derecha de la señal, esta área corresponde al FAP de la señal, que en este caso es de 10^{-4} . En el gráfico de la derecha, el período a analizar tiene una Potencia Normalizada de ~ 0.14 (marcado nuevamente como una línea vertical continua). La distribución de potencias obtenidas, sobre los conjuntos de residuos artificiales, muestran una distribución gaussiana que se encuentra visiblemente sobre la potencia de la señal. El área normalizada a la derecha de esta potencia es 0.6 %, es decir la señal posee un FAP del 0.6 %.

⁵con desviación estándar igual al RMS del ajuste kepleriano.

2.2. Métodos Cuadrados Mínimos

El método de cuadrados mínimos puede pensarse como un caso particular de lo que se conoce como *Estimadores de Máxima Similitud* (Maximum Likelihood Estimators).

Habiendo especificado el tipo de base que vamos a adoptar (i.e. las funciones ϕ_k), el problema de modelización se reduce a la determinación de los coeficientes A_k . Como siempre, estos son los parámetros del ajuste y son tratados como las variables del sistema. Cada conjunto de valores del espacio de parámetros $\{A_k\}$ denota una aproximación distinta, algunas parecidas con los datos originales, otras muy diferentes.

La idea básica es partir de la premisa que los datos (observacionales o experimentales) pueden tener errores. O sea, nosotros podemos medir y_i , pero el valor “real” puede de hecho ser $y_i \pm \varepsilon_i$.

La minimización de residuos o maximización de la función *Likelihood* (\mathcal{L}) es un proceso equivalente. Para un dado conjunto de datos medidos D y dado un modelo M con parámetros A , la función \mathcal{L} es definida mediante un modelo de ruido (Barnes, 2007):

$$\mathcal{L}(D; M, A) = \text{prob}(D|M, A) \quad (2.64)$$

En palabras, \mathcal{L} es la probabilidad de medir los datos D , dado el modelo M mediante los parámetros A . La estimación del máximo de \mathcal{L} implica obtener los valores de los parámetros A de tal forma que los datos **medidos** son los correspondientes a las mediciones más probables de obtener⁶.

Veamos un ejemplo: trabajemos con una hipotética distribución Gaussiana de errores, la probabilidad de obtener un dato es,

$$\mathcal{L}(D; M, A) = P_i(y_i|A, \sigma_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp \left[-\frac{1}{2} \left(\frac{y_i - f(x_i, A)}{\sigma_i} \right)^2 \right] \quad (2.65)$$

Mientras mayor sea la distancia entre el dato y_i (cuyo error asociado es σ_i) y el ajuste $f(x_i)$ (calculado con los parámetros A), menor será su probabilidad⁷. Haciendo lo mismo para los $N + 1$ datos, tendremos que la probabilidad compuesta será dada por el producto de las probabilidades individuales, o sea:

$$P = \prod_{i=0}^N \exp \left[-\frac{1}{2} \left(\frac{y_i - f(x_i)}{\sigma_i} \right)^2 \right]. \quad (2.66)$$

Dados coeficientes A_k , la expresión (2.66) define la probabilidad de que todos los datos se encuentren en y_i suponiendo que sus valores “verdaderos” son $f(x_i)$. A partir de aquí surge claramente que el mejor ajuste será aquel donde la cantidad (2.66) sea máxima.

⁶En realidad esta no es la pregunta que estamos realizando. Lo que en realidad queremos saber es cuales son los parámetros más probables dados los datos y el modelo. De esta forma deberíamos maximizar $\text{prob}(A|M, D)$ en lugar de $\text{prob}(D|M, A)$, lo que requeriría aplicar el teorema de Bayes (e.g. Barlow 1993).

⁷ En el caso de trabajar con una distribución de Poisson (utilizada para datos discretos.) la probabilidad de medir n fotones en un pixel del CCD dado un valor medio μ de fotones en dicho pixel está dada por:

$$\mathcal{L}(n; \mu) = P_i(n|\mu) = \frac{\mu^n}{n!} e^{-\mu}$$

donde se aprecia que en este caso la estimación del parámetro es independiente del modelo.

Ahora bien, maximizar esta cantidad es equivalente a maximizar el logaritmo natural de P . Existen dos razones para trabajar con logaritmo, primero que hay una multiplicación de números potencialmente pequeños que cuando N es grande puede producir un problema computacional. Segundo se adopta una forma más conveniente en forma matemática. Así maximizaremos:

$$\text{Ln}(P) = \sum_{n=1}^N \text{Ln} \left[e^{-\frac{1}{2} \left(\frac{y_i - f(x_i)}{\sigma_i} \right)^2} \right] + C \quad (2.67)$$

identificando la suma,

$$\chi^2 = \sum_{n=1}^N \left(\frac{y_i - f(x_i)}{\sigma_i} \right)^2 \quad (2.68)$$

donde C son los términos constantes (por construcción). En otras palabras maximizar el \mathcal{L} se reduce a minimizar χ^2 , ya que:

$$\text{Ln}(P) = -\frac{1}{2}\chi^2 + C \quad (2.69)$$

donde χ^2 es igual a la función residuo R^2 tomando la función peso igual a:

$$\boxed{w_i = \frac{1}{\sigma_i^2}} \quad (2.70)$$

Concluyendo, el MCM es un Estimador \mathcal{L} (de Máxima Similitud) de que los datos se correspondan a la función ajuste, sí y solo sí suponemos que los “errores” tienen una distribución normal alrededor del ajuste. Vemos también que *la función peso puede ser asociada a la dispersión o error de cada dato individual*. O sea, aquellos datos que están pobremente determinados, y por ende tendrán un valor grande de σ_i , tendrán un peso pequeño dado por la inversa del cuadrado de la desviación estándar. Si consideramos que todos los datos poseen el mismo error, entonces podemos simplemente obviar la función peso y tomarla igual a uno en todo el intervalo.

Cuadrados mínimos sin las funciones de pesos $w(x)$ son también llamados de Cuadrados Mínimos *per se*. Cuando hacemos uso de la función peso, el método es a veces conocido como *Método de Chi-Cuadrado*. El nombre proviene del hecho que, en estos casos, se designa la función

$$\langle w(x)R^2(x) \rangle \rightarrow \chi^2 = \sum_{i=0}^N \frac{\left(y_i - f(x_i) \right)^2}{\sigma_i^2}, \quad (2.71)$$

con lo cual el mejor ajuste será aquel que minimice χ^2 .

La pregunta que surge inmediatamente es si es correcto considerar una distribución normal para los errores. Es bien sabido que la distribución de probabilidad de la composición de un gran número de desviaciones aleatorias siempre converge a una distribución normal. Esto es lo que se llama de *Teorema Central del Límite*, ver Sección 3.

Sin embargo, a la hora de trabajar con datos observacionales o experimentales, muchas veces esta convergencia no ocurre, especialmente si el número de datos es limitado. El problema más importante

es que, en general, las distribuciones experimentales poseen una “cola” mucho mayor que la prevista por una distribución Gaussiana. En otras palabras, en los casos reales existen muchos más puntos fuera del rango 3σ de lo que uno esperaría si la distribución fuese normal. También es posible que, debido a errores groseros como el funcionamiento inadecuado del equipo, técnica, etc..., existan datos muchísimo más allá del resto de las observaciones. Estos datos a veces se denotan como *outliers* en el sentido que están ubicados afuera del rango esperado, y se corresponden con una estadística diferente respecto de la Gaussiana.

El principal problema de los outliers es que, según un ajuste de cuadrados mínimos, la probabilidad de ocurrencia de estos puntos es tan baja que el método optará por distorsionar todo el ajuste para acomodarlos equivocadamente (en Press et al. 1992, se menciona que en una distribución normal, la probabilidad que un dato esté ubicado a una distancia $\pm 20\sigma$ es del orden de 10^{-88}). Esto es uno de los principales motivos por el cual el MCM es muy sensible, y puede dar resultados erróneos si tenemos (aunque más no sea) un único punto significativamente fuera del resto del conjunto de datos.

Esta sensibilidad de los Cuadrados Mínimos puede ser atenuada (y a veces eliminada) con algunas modificaciones. Estos son los llamados *Métodos Robustos* que se caracterizan por adoptar distribuciones de errores diferentes de la normal para los datos. También existen variantes que permiten un tratamiento especial de los outliers. Si bien estos son métodos válidos para el tratamiento de datos, basan su método en considerar una gran cantidad de puntos, contrario a nuestro caso donde todas las observaciones son necesarias y no podemos prescindir de algunos datos en forma arbitraria.

Asumiendo una distribución Gaussiana de errores, dos formas de \mathcal{R} son usualmente encontradas en la literatura:

$$\begin{aligned} (\text{wrms})^2 &= \frac{1}{S(N-1)} \sum_{i=1}^N \frac{(y_i - V_{ri})^2}{\sigma_i^2} & \text{con } \frac{1}{S} &= \frac{1}{N} \sum_{i=1}^N \frac{1}{\sigma_i^2} \\ \chi_\nu^2 &= \frac{1}{N-M} \sum_{i=1}^N \frac{(y_i - V_{ri})^2}{\sigma_i^2} \end{aligned} \quad (2.72)$$

donde y_i es la velocidad radial al tiempo de observación t_i , y V_{ri} es el valor calculado de la V_r a partir del modelo que tiene los elementos orbitales en el momento de la observación t_i . Como mencionáramos M es el número de parámetros ajustados ($M = 5n_{pla} + n_{obs}$), mientras que $\nu = N - M$ es usualmente referido como el número de grados de libertad de la regresión. El peso que se asigna a cada observación es dado por σ_i^2 y es la suma cuadrática del error en cada la velocidad radial y el jitter estelar (i.e $\sigma_i^2 = \varepsilon_i^2 + \sigma_j^2$, ver Sección 1.3.6.3).

El valor de χ_ν^2 es usualmente utilizado como estimador de buen ajuste entre la comunidad astronómica. Sin embargo, como puede verse en la ecuación, es muy sensible a los valores que se asuman en los errores de los datos y en el estimado del jitter. En este sentido el wrms es mucho más independiente de los valores ε_i^2 y σ_j^2 , consistente con un estimador de buen ajuste. Por esta razón a lo largo del trabajo, se darán ambos valores de las funciones residuo, pero la comparación con otros trabajos debe realizarse utilizando el wrms.

2.2.1. El Principio de Cuadrados Mínimos

Aunque el método de cuadrados mínimos (MCM) es generalmente utilizado para el ajuste (o modelización) de datos, resulta conveniente comenzar a discutirlo como una aproximación de funciones.

Supongamos que tenemos una función $y(x)$ conocida y dada en forma analítica (no es lo mismo que pedir que la función en sí sea analítica). La idea es conseguir representar o aproximar esta función como una combinación lineal de un conjunto de funciones más simples que sean una base del espacio de funciones cuadrado-integrables. El caso típico es aproximar por polinomios, pero esta no es la única opción.

Vamos a definir:

- $y(x)$: la función que queremos aproximar en un intervalo $x \in [a, b]$ de la variable independiente. Ya que x puede ser una variable n -dimensional, vamos a generalizar el intervalo de interés como $x \in D$, con $D \subset \mathbb{R}^n$.
- $\{\phi_k(x)\}$: un conjunto de infinitas funciones analíticas (i.e. continuas y con derivadas de orden arbitrario continuas). Vamos a exigir que estas funciones sean linealmente independientes, de forma tal que no existan ninguna secuencia de reales $\{a_i\}$ tal que $\phi_k(x) = \sum_{i=0}^{k-1} a_i \phi_i(x)$. Decimos que nuestro conjunto $\{\phi_k(x)\}$ forman una *base* para un cierto espacio de funciones.
- $f(x) = \sum_{k=0}^M A_k \phi_k(x)$: la aproximación de orden M de la función $y(x)$ en esta base. Nuestro objetivo será justamente la determinación de estos coeficientes constantes A_k .

A modo de ejemplo, el conjunto de monomios $\phi_k(x) = x^k$ forman una base para el espacio de las funciones analíticas C^∞ . Cualquier función $g(x) \in C^\infty$ podrá entonces escribirse como $g(x) = \sum_{i=0}^{\infty} A_i x^i$. La serie truncada a orden M será una aproximación polinomial de $y(x)$.

- También vamos a definir una *función de peso* definida como $w(x) \geq 0 \forall x \in D$. Esta función puede tener varios significados. Por ejemplo, puede representar una relación del error en la determinación de $y(x)$ en función de x . En el caso mas simple, generalmente se toma $w(x) \equiv 1$.
- Si la aproximación fuese exacta, tendríamos $f(x) = y(x)$ en todo el intervalo. Como generalmente esto no ocurre, definimos la *función Residuo* entre ambas como:

$$R(x) \equiv y(x) - f(x) = y(x) - \sum_{j=0}^M A_j \phi_j(x). \quad (2.73)$$

En términos de esta notación, definiremos la *mejor aproximación* a $y(x)$ (en el sentido de cuadrados mínimos) como el conjunto de coeficientes A_k tal que el *agregado* (“agregate” en inglés: suma o integral) del producto $w(x)R^2(x)$, y denotado por $\langle w(x)R^2(x) \rangle$, sea el menor posible dentro del intervalo D . En el caso de funciones de una variable, esto se traduce en:

$$\langle wR^2 \rangle \equiv \left\langle w \left(y - \sum_{k=0}^M A_k \phi_k \right)^2 \right\rangle = \sum_i^N w(x_i) \left(y(x_i) - \sum_{k=0}^M A_k \phi_k(x_i) \right)^2. \quad (2.74)$$

Si queremos que esta cantidad sea mínima considerando los coeficientes como variables, esto es equivalente a pedir que se cumplan las condiciones:

$$\frac{\partial}{\partial A_r} \langle w(x)R^2(x) \rangle = 0 \quad \forall r = 0, \dots, M. \quad (2.75)$$

Reemplazando R , obtenemos:

$$\begin{aligned} \frac{\partial}{\partial A_r} \langle wR^2 \rangle &= \left\langle w \frac{\partial}{\partial A_r} \left(y - \sum_{k=0}^M A_k \phi_k \right)^2 \right\rangle \\ &= -2 \left\langle w \phi_r \left(y - \sum_{k=0}^M A_k \phi_k \right) \right\rangle = 0 \quad r = 0, \dots, M. \end{aligned} \quad (2.76)$$

Con lo cual,

$$\sum_{k=0}^M A_k \langle w \phi_r \phi_k \rangle = \langle w \phi_r y \rangle \quad r = 0, \dots, M. \quad (2.77)$$

Aplicando estos principios al caso de la Modelización de Datos, donde en vez de $y(x)$ tenemos un conjunto discreto de $(N + 1)$ puntos del tipo (x_i, y_i) con $i = 0, \dots, N$. En tal caso, el principio de cuadrados mínimos queda definido por el producto interno para datos discretos y las ecuaciones quedan:

$$\boxed{\sum_{k=0}^M A_k \sum_{i=0}^N w(x_i) \phi_r(x_i) \phi_k(x_i) = \sum_{i=0}^N w(x_i) \phi_r(x_i) y(x_i) \quad r = 0, \dots, M} \quad (2.78)$$

En otras palabras, tenemos un conjunto de $(M + 1)$ ecuaciones lineales con el mismo numero de incógnitas, dadas por los coeficientes A_k . El sistema (2.78) es conocido como las *ecuaciones normales* del proceso de cuadrados mínimos. Resolviendo este sistema tendremos explícitamente los valores de los coeficientes que minimizan el residuo:

$$\begin{pmatrix} \langle w \phi_0 \phi_0 \rangle & \langle w \phi_0 \phi_1 \rangle & \langle w \phi_0 \phi_2 \rangle & \dots & \langle w \phi_0 \phi_M \rangle \\ \langle w \phi_1 \phi_0 \rangle & \langle w \phi_1 \phi_1 \rangle & \langle w \phi_1 \phi_2 \rangle & \dots & \langle w \phi_1 \phi_M \rangle \\ \vdots & & & & \vdots \\ \langle w \phi_M \phi_0 \rangle & \langle w \phi_M \phi_1 \rangle & \langle w \phi_M \phi_2 \rangle & \dots & \langle w \phi_M \phi_M \rangle \end{pmatrix} \cdot \begin{pmatrix} A_0 \\ A_1 \\ \vdots \\ A_M \end{pmatrix} = \begin{pmatrix} \langle w \phi_0 y \rangle \\ \langle w \phi_1 y \rangle \\ \vdots \\ \langle w \phi_M y \rangle \end{pmatrix}. \quad (2.79)$$

Una característica desfavorable de las ecuaciones normales es su dificultad en determinar explícitamente los coeficientes. Dependiendo de la base $\{\phi_k\}$, esto puede ser sumamente complicado, no solo algebraicamente sino numéricamente. Sin embargo, son raros los casos donde el valor de M sea tan grande como para introducir imprecisiones numéricas significativas. Así, la principal desventaja del Método Directo para la resolución de (2.79) es simplemente su complejidad y uso extensivo de recursos computacionales.

Una forma muy utilizada para simplificar las ecuaciones normales es adoptar una base ortogonal (respecto de $w(x)$) para las funciones $\{\phi_k\}$. En este caso, y por definición de ortogonalidad, se cumple que:

$$\langle w \phi_i \phi_j \rangle = \delta_{i,j} \quad (2.80)$$

donde $\delta_{i,j} = 0 \forall j \neq i$. Cumpliéndose esto, la matriz de la ecuación (2.79) se reduce a una expresión diagonal, y cada ecuación algebraica correspondiente queda desacoplada del resto del sistema. Así obtenemos explícitamente:

$$A_r \langle w \phi_r^2 \rangle = \langle w \phi_r y \rangle \quad r = 0, \dots, M \quad (2.81)$$

como lo cual

$$A_r = \frac{\langle w\phi_r y \rangle}{\langle w\phi_r^2 \rangle} \quad r = 0, \dots, M. \quad (2.82)$$

Si, además de ortogonal, la base es ortonormal, entonces se cumple que $\langle w\phi_r^2 \rangle = 1$ para todo r , y la expresión para los coeficientes se simplifica aun más, quedando:

$$A_r = \langle w\phi_r y \rangle \quad r = 0, \dots, M. \quad (2.83)$$

Cuadrados mínimos sólo tiene solución si $M \leq N$. Es más, en el caso donde $M = N$, tendremos que la solución de aproximación pasa exactamente por todos los puntos del conjunto de datos. En otras palabras, nos reducimos al caso de una interpolación. Así vemos que la interpolación polinomial (cualquiera sea la base) es un caso particular de aproximación polinomial de funciones o modelización de datos.

El valor $\nu = N - M$ se conoce como el *numero de grados de libertad* del ajuste. El caso de interpolación, donde no hay parámetros a definir en el ajuste, se corresponde con cero grados de libertad.

Cuando trabajamos con el método de velocidad radial los coeficientes incógnita A_k no tienen una expresión explícita, puesto que trabajamos con un sistema no-lineal. En general debemos apelar a métodos iterativos utilizando la derivada y Hessiana de χ^2 .

Si escribimos $\chi^2 = \langle wR^2 \rangle$ como:

$$\chi^2(\mathbf{A}) = \sum_{i=0}^N w_i \left(y_i - f(x_i; \mathbf{A}) \right)^2 \quad (2.84)$$

donde dejamos explícita la dependencia con el vector incógnita $\mathbf{A} = (A_0, A_1, \dots, A_M)$. Las componentes del gradiente de esta función con respecto de A_k será:

$$\frac{\partial \chi^2}{\partial A_r} = -2 \sum_{i=0}^N w_i \left(y_i - f(x_i; \mathbf{A}) \right) \frac{\partial f}{\partial A_r} \quad r = 0, \dots, M. \quad (2.85)$$

En el caso lineal, donde $f(x_i; \mathbf{A}) = \sum_{k=0}^M A_k \phi_k(x_i)$, tendremos que $\partial f / \partial A_r = \phi_r(x_i)$ es independiente de los coeficientes. En tal caso, las ecuaciones normales quedan lineales y fácilmente resolubles.

En nuestro caso genérico (i.e. no-lineal), debemos calcular las derivadas segundas (i.e. las componentes de la Hessiana). Estas son dadas por:

$$\frac{\partial^2 \chi^2}{\partial A_r \partial A_s} = 2 \sum_{i=0}^N w_i \left(\frac{\partial f}{\partial A_r} \frac{\partial f}{\partial A_s} - (y_i - f(x_i; \mathbf{A})) \frac{\partial^2 f}{\partial A_r \partial A_s} \right) \quad r, s = 0, \dots, M. \quad (2.86)$$

Según las bases de los métodos en todos los MCM no-lineales, el segundo término de la Ec. (2.86) es despreciable, justificado por el hecho que si el ajuste es bueno, entonces el término $(y_i - f(x_i; \mathbf{A}))$ es pequeño. De hecho, según se menciona en Press et al. (1992), la inclusión de este término es, en la práctica, poco aconsejable sea cual fuese la precisión del ajuste, ya que la presencia de outliers puede

dar origen a una desestabilización del modelo numérico. Por lo tanto, podemos redefinir ad-hoc la ecuación anterior aproximándola por:

$$\frac{\partial^2 \chi^2}{\partial A_r \partial A_s} = 2 \sum_{i=0}^N w_i \left(\frac{\partial f}{\partial A_r} \frac{\partial f}{\partial A_s} \right) \quad r, s = 0, \dots, M. \quad (2.87)$$

y trabajar con esta expresión en vez de la Ec. (2.86).

2.2.2. Método del Gradiente Conjugado

Existe otro tipo de métodos donde se analiza la dirección en la cual el residuo R decrece. Consideramos los coeficientes \mathbf{A} como variables, podemos expandir la función residuo en una serie de Taylor alrededor del origen (truncada a orden dos), como:

$$\chi^2(\mathbf{A}) = \chi_0^2 + \mathbf{A}^T \cdot \mathbf{d} + \frac{1}{2} \mathbf{A}^T \cdot \mathbf{D} \cdot \mathbf{A} \quad (2.88)$$

donde \mathbf{d} es un vector de componentes $\{\partial f / \partial A_r\}$ y \mathbf{D} es una matriz $(M + 1) \times (M + 1)$ cuyas componentes son dadas por la Ec. (2.87). En este grado de aproximación, el gradiente de $\chi^2(\mathbf{A})$ desarrollado en Taylor será:

$$\nabla \chi^2(\mathbf{A}) = \mathbf{d} + \mathbf{D} \cdot \mathbf{A}. \quad (2.89)$$

Denominamos ahora como \mathbf{A}_{min} el conjunto de coeficientes que dan el mejor ajuste, o sea: $\nabla \chi^2(\mathbf{A}_{min}) = 0$. Reemplazando en la expresión general del gradiente, y agrupando, tenemos que:

$$\nabla \chi^2(\mathbf{A}) = -\mathbf{D} \cdot \mathbf{A}_{min} + \mathbf{D} \cdot \mathbf{A} = \mathbf{D} \cdot \delta \mathbf{A} \quad (2.90)$$

donde $\delta \mathbf{A} = \mathbf{A} - \mathbf{A}_{min}$ es la diferencia entre los coeficientes actuales (cualquiera que sean; por ejemplo, determinados vía una aproximación lineal) y los exactos. Con esto podemos construir un procedimiento de determinación iterativa de la solución vía la expresión:

$$\mathbf{A}^{(i)} = \mathbf{A}^{(i-1)} - \mathbf{D}^{-1} \cdot \nabla \chi^2(\mathbf{A}^{(i-1)}) \quad (2.91)$$

donde $\mathbf{A}^{(i)}$ es el valor de la i -ésima iteración.

En otras palabras, cada iteración corrige el valor anterior según la dirección del gradiente de la función residuo (pero en sentido contrario) y con una magnitud que esta pesada por la Hessiana. Existen casos donde la serie de Taylor truncada a segundo orden no es una buena aproximación del sistema no-lineal, especialmente si $\delta \mathbf{A}$ no es pequeño. En tal caso, podemos modificar levemente el esquema anterior, definiendo la próxima iteración como:

$$\mathbf{A}^{(i)} = \mathbf{A}^{(i-1)} - C \nabla \chi^2(\mathbf{A}^{(i-1)}) \quad (2.92)$$

donde C es una constante (generalmente pequeña). Esta variación, donde se elige la dirección del gradiente pero no el modulo de la aproximación cuadrática, se conoce como el *Método del "Steepest Descent"*, ver Sección 15.5, Press et al. (1992). El cambio de $\mathbf{D}^{-1} \rightarrow C$ tiene como objetivo evitar oscilaciones alrededor de la solución exacta.

Expandiendo la ecuación (2.92) en sus componentes, podemos escribirlo como un sistema de $(M + 1)$ ecuaciones algebraicas de la forma:

$$A_r^{(i)} = A_r^{(i-1)} + C \frac{\partial \chi^2}{\partial A_r^{(i-1)}}. \quad r = 0, \dots, M. \quad (2.93)$$

que brindan en forma explícita la solución de cada coeficiente por aproximaciones sucesivas.

Recapitulando un poco antes de continuar, explícitamente en nuestras variables tenemos la función χ_ν^2 :

$$\chi_\nu^2(\mathbf{A}) = \frac{1}{N - M} \sum_{i=1}^N \frac{(y_i - V_r(t_i, \mathbf{A}))^2}{\sigma_i^2} \quad (2.94)$$

Un desarrollo en Taylor hasta orden 2 para esta función multiparamétrica nos da la siguiente aproximación:

$$\chi_\nu^2(\mathbf{A}) \approx \chi_\nu^2(\mathbf{A}_0) + (\mathbf{A} - \mathbf{A}_0)^T \cdot \mathbf{d} + \frac{1}{2} (\mathbf{A} - \mathbf{A}_0)^T \cdot \mathbf{D} \cdot (\mathbf{A} - \mathbf{A}_0), \quad (2.95)$$

donde su gradiente \mathbf{d} y Hessiana \mathbf{D} están dadas explícitamente por:

$$\beta_k = \frac{\partial \chi_\nu^2}{\partial A_k} = \frac{-2}{N - M} \sum_{i=1}^N (y_i - V_{ri}) \frac{1}{\sigma_i^2} \frac{\partial V_r(i)}{\partial A_k} \quad (2.96)$$

$$\alpha_{kl} = \frac{\partial^2 \chi_\nu^2}{\partial A_k \partial A_l} = \frac{2}{N - M} \sum_{i=1}^N \frac{1}{\sigma_i^2} \left[\frac{\partial V_r(i)}{\partial A_k} \frac{\partial V_r(i)}{\partial A_l} - (y_i - V_{ri}) \frac{\partial^2 V_r(i)}{\partial A_k \partial A_l} \right], \quad (2.97)$$

La segunda derivada en la Ec. (2.97) puede ignorarse cuando es cero (el caso lineal por ejemplo) o bien cuando es suficientemente pequeña comparada con el término que involucra la primer derivada. También como argumentáramos antes, existe la posibilidad que el término $(y_i - V_{ri})$ sea suficientemente pequeño. Más aún si el modelo es correcto, este término debería ser aleatorio y oscilar alrededor de cada punto. El error puede tener signo y debe no debe estar correlacionado con el modelo. De esta forma los términos de segunda derivada pueden descartarse y trabajar con una ecuación reducida.

2.2.3. Método de Levenberg-Marquardt

El método del Steepest Descent es el procedimiento más simple para la modelización de datos para una función de ajuste no-lineal. Sin embargo, posee ciertos inconvenientes. El primero tiene que ver con la ambigüedad en la especificación del coeficiente constante C (Ec. (2.92)). No se sabe su escala, que valor se toma para él ni como depende de las unidades que elegimos para los datos. Para aclarar un poco este punto, y fijar un poco más rigurosamente el valor de C , podemos usar parte de la matriz Hessiana, aún sin utilizarla como un todo.

Vamos a adoptar la siguiente notación:

$$\beta_k = -\frac{\partial \chi^2}{\partial A_k} \quad ; \quad \alpha_{k,l} = \frac{\partial^2 \chi^2}{\partial A_k \partial A_l} \quad (2.98)$$

con lo cual tenemos que los Métodos del Gradiente Conjugado y del Steepest Descent adquieren la forma expandida (no-vectorial):

$$\sum_{i=0}^M \alpha_{k,i} \delta A_i = \beta_k \quad (\text{Gradiente Conjugado}) \quad (2.99)$$

$$\delta A_i = C \beta_i \quad (\text{Steepest Descent}).$$

En el primer caso, debemos resolver un sistema lineal para obtener cada δA_i ; en el segundo estos datos son obtenidos de forma explícita.

Veamos como elegir el tamaño de C . Comenzamos recordando que la cantidad χ^2 es un número no-dimensional, lo que es bastante claro cuando relacionamos w_i con $1/\sigma_i^2$. Por otro lado, a partir de la definición de gradiente, tenemos que β_k tendrá unidades de $(A_k)^{-1}$. Peor aún, dependiendo del problema, cada β_k puede tener unidades completamente distintas para diferentes índices k . Reemplazando en la segunda de las ecuaciones (2.99), vemos que las unidades de C deben ser $(A_k)^2$.

En la matriz $\{\alpha_{k,l}\}$, notamos que, dado que son componentes de la Hessiana, cada elemento tiene dimensión igual a $(A_k)^{-2}$. En particular, podemos considerar que la inversa del cuadrado de los elementos de la diagonal, $1/\alpha_{k,k}^2$ son una escala natural para C . Así podemos introducir una nueva constante $\lambda = (C\alpha_{k,k})^{-1}$, y reescribir la ecuación para el Steepest Descent como:

$$\delta A_i = \frac{1}{\lambda\alpha_{i,i}} \beta_i. \quad (2.100)$$

Si λ es muy grande, estamos disminuyendo el “paso” de la iteración que vendría de considerar una matriz Hessiana diagonal.

La gran ventaja de esta variación es que nos permite independizarnos de las unidades y escalas propias de nuestro problema físico. Sin embargo, aún podemos mejorar más las cosas. El Método de Levenberg-Marquardt esta basado en reemplazar la segunda ecuación (2.99) por (2.100), obteniendo:

$$\begin{aligned} \sum_{i=0}^M \alpha_{k,i} \delta A_i &= \beta_k && \text{(Gradiente Conjugado)} && (2.101) \\ \lambda\alpha_{k,i} \delta A_i &= \beta_i && \text{(Steepest Descent Modificado)}. \end{aligned}$$

Así podemos *agrupar* ambos métodos si definimos una nueva matriz

$$\begin{aligned} \alpha'_{i,i} &= (1 + \lambda)\alpha_{i,i} && (2.102) \\ \alpha'_{i,j} &= \alpha_{i,j} && (j \neq i) \end{aligned}$$

y conseguimos reemplazar ambas ecuaciones de (2.101) por un *único* sistema:

$$\sum_{i=0}^M \alpha'_{k,i} \delta A_i = \beta_k \quad k = 0, \dots, M. \quad (2.103)$$

Vemos que cuando $\lambda \gg 1$ estamos sobrevaluando los elementos diagonales del Hessiano, lo cual es equivalente a considerar el método del Steepest Descent modificado. Recordemos que dijimos que ésta es la mejor opción si nuestros coeficientes A_k iniciales están muy lejos de los óptimos. Por otro lado si $\lambda = 0$, retornamos al método del Gradiente Conjugado original, el cual es más recomendable si estamos cerca de la solución no-lineal completa. Por lo tanto, el coeficiente λ nos permite tener un parámetro ajustable para optar por uno u otro método, dependiendo si estamos lejos o cerca de la solución deseada. Esta es la base del **Método de Levenberg-Marquardt**. Esquemáticamente, funciona con los siguientes pasos:

1. Elegimos valores iniciales para $A_k^{(i)}$ (con $i = 0$) y calculamos la función residuo χ^2 , más su gradiente y Hessiana.

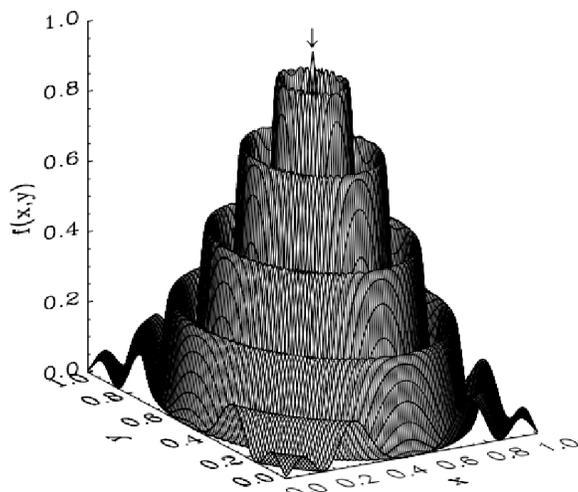


Figura 2.4: Superficie bidimensional $f(x,y)$ con $(x,y) \in [0,1]$, definiendo un problema complicado de resolver para maximización. El máximo global está marcado con una flecha en $(x,y)=(0.5,0.5)$.

2. Elegimos un valor modesto para λ , por ejemplo: $\lambda = 0.001$. Con esto le damos la oportunidad al sistema para que las iteraciones sean generadas con la parte cuadrática incluida (Gradiente Conjugado).
3. Resolvemos las ecuaciones algebraicas lineales (2.103), calculamos $\delta A_k = A_k^{(i+1)} - A_k^{(i)}$ y a partir de allí tenemos una primera estimación de la próxima iteración $A_k^{(i+1)}$.
4. Calculamos el residuo $\chi^2(\mathbf{A}^{(i+1)})$.
5. Chequeamos si $\chi^2(\mathbf{A}^{(i+1)}) < \chi^2(\mathbf{A}^{(i)})$, o sea, si nos estamos aproximando al mínimo.
6. En caso positivo, el método del Gradiente Conjugado esta funcionando. Disminuimos λ en un factor 10, confirmamos $A_k^{(i+1)}$ como la siguiente iteración, y volvemos al punto 3.
7. En caso negativo, la iteración no funciona. En tal caso, eliminamos $A_k^{(i+1)}$, *aumentamos* λ en un factor 10 y volvemos al punto 3.

También es necesario una condición para detener el algoritmo. Iterar hasta la convergencia (precisión de máquina o límite de redondeo) es generalmente una pérdida de tiempo innecesaria ya que el mínimo es a lo sumo la **mejor** estimación estadística de los parámetros \mathbf{A} . Como veremos en la sección de errores (3.1) un cambio en los parámetros que cambie $\chi^2 \ll 1$ nunca es estadísticamente significativo. En la práctica, uno puede suponer una salida del ciclo de iteración cuando χ^2 decrece una cantidad despreciable (por ejemplo un error absoluto menor que 0.01). No es recomendable detenerlo después de un paso donde χ^2 se incrementa, pues esto demuestra que λ no fue ajustado correctamente. Una vez que se encontró un mínimo aceptable, uno quiere fijar $\lambda = 0$ y calcular la matriz

$$[C] \equiv [\alpha]^{-1} \quad (2.104)$$

que es la matriz de covarianza de los errores estándar en los parámetros ajustados \mathbf{A} (más detalles en 3.2).

Este tipo de enfoque sin embargo es determinístico, ya que la solución final depende del punto inicial tomado en el espacio de parámetros. En la Fig. 2.4 se presenta un ejemplo complicado para los

métodos de maximización/minimización locales, donde se observa un pico central rodeado de anillos concéntricos de máximos secundarios.

Por supuesto si se logra iniciar con condiciones iniciales lo suficientemente cercanas al pico central, los métodos de Gradiente Conjugado convergen rápidamente. Si elegimos una estrategia Monte-Carlo, es decir tomamos condiciones iniciales al azar para nuestro punto de partida y luego resolvemos el problema de minimización, dependiendo de la cantidad de condiciones iniciales algunas de nuestras soluciones seguramente lograrán identificar en el máximo. El problema es que dentro del espacio de parámetros por ejemplo para la Fig. 2.4 la región central del pico representa sólo el 1 % del espacio, requiriendo entonces $\simeq 10^2 - 10^3$ condiciones iniciales distintas para encontrar el mínimo de la función. El dibujo puede ser generado con la función:

$$\begin{aligned} f(x, y) &= \cos^2(n\pi r) e^{-r^2/\sigma^2}, \\ r^2 &= (x - 0.5)^2 + (y - 0.5)^2; \quad x, y \in [0, 1], \end{aligned} \quad (2.105)$$

Cuando se requiere aplicar este método en la búsqueda de soluciones planetarias la función es más compleja. Se requiere minimizar la función residuo R , multiparamétrica y muy rugosa. En particular, en nuestro trabajo buscamos soluciones para sistemas de dos exoplanetas, por lo que tenemos una hipersuperficie a explorar de al menos 11 parámetros ($M=11$). Para asegurarnos la convergencia al mínimo global debemos tener al menos unas $10^6 \sim 10^7$ condiciones iniciales. Esto convierte el método en ineficiente, pudiendo tardar varios días de cómputo en encontrar el mínimo (Ferraz-Mello et al. 2003).

2.2.4. Algoritmo Genético

Este método de minimización basa sus principios en las ideas generales de evolución y adaptación que fueron introducidas por Charles Darwin (1859) en “On the Origin of Species” por medio de la selección natural. De esta forma, mucha terminología asociada a biología será usada para describir sus funciones. El mecanismo que rige la evolución es la selección natural, que es el proceso a través del cual los individuos mejor **adaptados** a su ambiente tienden a producir en promedio, más descendencia que sus competidores menos adaptados. Dos componentes adicionales son requeridas para la selección natural:

- En primer lugar, la “herencia”, ya que un descendiente debe heredar de alguna manera algunas de las características que hacen que sus padres sean “aptos”.
- El segundo ingrediente es la variabilidad: en cualquier momento debe haber un espectro de “adaptados” dentro de los miembros de la población, de otra forma la selección natural no puede operar.

¿Como es regida la evolución? Básicamente se sabe que cada célula de cada individuo (denominadas **fenotipos**) contiene información codificada en una secuencia que constituye el **genotipo**. Sin embargo debido a que fenotipos de alta adaptación se traducen, en promedio, en altas probabilidades de reproducción, el fenotipo tiene influencia directa en el pool de genes (conjunto de genotipos) en la próxima generación. Los fenotipos con mejor adaptación “copian” su fenotipo más frecuentemente en la próxima generación, que lleva a una porción creciente de fenotipos con buen ajuste en la población. Para condiciones de medioambiente fijas, la población naturalmente converge como un todo hacia los individuos mejor adaptados.

La velocidad de la convergencia es una función de la *presión de selección*, que es la relación funcional entre el ajuste y el éxito de la reproducción. Una vez que la convergencia es completada, cambios severos en las condiciones del ambiente pueden drásticamente alterar las reglas de adaptación y hacer que todos los individuos no sean aptos con consecuencias desastrosas para una población.

La reproducción involucra la combinación del genotipo de los padres. No es una copia exacta de los genotipos de los padres en la descendencia. Errores en las copias y/o eventos aleatorios pueden introducir mutaciones al genotipo.

El conjunto de todos los genotipos disponibles en la población se denomina “**pool de genes**”. Resumiendo, la evolución puede ser pensada y matemáticamente modelada en términos de cambios temporales a través del “pool de genes”.

Los algoritmos genéticos son una clase de técnicas de búsqueda heurística que incorporan estas ideas en sentencias computacionales y constituyen métodos extremadamente robustos de minimización. Utilizamos la rutina PIKAIA (Charbonneau, 1995) en nuestros programas, que se encuentra disponible de manera libre en <http://download.hao.ucar.edu/archive/pikaia> e introduce todas estas ideas de evolución genética.

2.2.4.1. Principios del Algoritmo Genético

En primer lugar vale aclarar que el “pool de genes” y su población de genotipos evoluciona en respuesta a:

- El éxito de la reproducción diferencial en la población
- La recombinación genética (crossover) ocurrida en los descendientes
- Las Mutaciones aleatorias que afectan a un porcentaje de la descendencia, eligiendo al azar un descendiente (offspring) y un número de gen donde aplicar la misma.

En nuestro caso el parámetro que define la **adaptación o ajuste** de un individuo está asociado al valor de χ_ν^2 . Debido a que PIKAIA está desarrollado para maximizar una función y requiere que la adaptación se defina como un valor positivo. Entonces se sugiere χ_ν^{-1} . Debido a que PIKAIA internamente rankea (ordena) para definir la probabilidad de selección, no necesitamos preocuparnos sobre la forma funcional entre el *ajuste* y el valor de χ_ν ; hacer el *ajuste* proporcional a $\chi_\nu^{-1/2}$ conducirá a la misma distribución de orden y por lo tanto la misma probabilidad de selección. Se recomienda especialmente no utilizar $-\chi_\nu$ ya que el algoritmo para seleccionar los padres llamado *Roulette Wheel* requiere que el parámetro *ajuste* sea definido positivo (e.g. Charbonneau, 2002b).

Podemos pensar el genotipo como una secuencia numérica donde los parámetros se encuentren ordenados de la manera que especificamos en un principio. El programa está optimizado para trabajar con una función $f(\mathbf{x})$ en un espacio acotado m-dimensional, donde

$$\mathbf{x}_j \equiv (x_1, x_2, \dots, x_m) \quad (2.106)$$

Siendo con $j \in [1, m]$ cada uno de los individuos dentro de la población de n individuos. Cada $x_j \in [0, 1]$ es uno de los parámetros de cada individuo; debido a que en el caso de velocidades radiales,

la señal está dada por $V_r = V_r[K, T, e, M, \varpi, V_0]$ y podemos asociar cada parámetro a un x_i , es decir:

$$\begin{aligned} x_1 &\rightarrow K, \\ x_2 &\rightarrow T, \\ &\vdots \\ x_m &\rightarrow V_0 \end{aligned} \quad (2.107)$$

Entrada y Salida de Datos

Como primer paso, se debe definir el rango válido de parámetros que puede tomar cada elemento del individuo, es decir los parámetros orbitales A_j del planeta se deben transformar al espacio de parámetros x_j entre $[0,1]$ (cada x_j es un fenotipo y el conjunto de x_j ordenados uno a continuación del otro es el genotipo, $x_1 x_2 \dots x_M$).

A modo de ejemplo podemos mencionar que la amplitud que modelemos de un planeta K estará acotada entre el mínimo valor que esperamos detectar como señal (al menos 2 veces el valor del jitter estelar, σ_s) y el máximo valor que un planeta podría representar (en teoría, menor que la diferencia entre el máximo y mínimo valor de las velocidades radiales). Es decir si tenemos un valor de jitter de 4.2 m/s y amplitud de velocidades de 108.4 m/s, una señal de exoplaneta de $K=54.2$ m/s significa un valor $x_k = 0.5000$.⁸

La codificación en fenotipos es para cada parámetro A_i dada por:

$$x_i = \frac{(A_i - A_{min})}{\Delta A_i}, \quad (2.108)$$

donde llamaremos a esta transformación Γ ; inversamente Γ^{-1} será la decodificación, que implica obtener los parámetros A_i originales, con

$$A_i = A_{min} + x_i \cdot \Delta A_i \quad (2.109)$$

Obviamente debemos ser muy cuidadosos de los límites establecidos ya que el algoritmo busca dentro del espacio especificado y descartaremos posibles soluciones si acotamos demasiado la región donde viven los parámetros. Por otro lado si la elección es muy amplia corremos el riesgo de perder mucho tiempo en la exploración global del espacio.

Nuestra elección general y que funciona para una amplia variedad de casos toma los siguientes límites:

- $K \in [2 \cdot \sigma_{jitter} ; \Delta V]$
- $P \in [\min\{t_{i+1} - t_i\}/2 ; (\Delta t)/2]$
- $e \in [0 ; 0.6]$ ($e > 0.6$ ralentizan mucho la búsqueda)
- $\omega \in [0 ; 2\pi]$
- $\tau \in [t_{inicial} - \frac{1}{3}\Delta t ; t_{final} + \frac{1}{3}\Delta t]$, el tiempo de pasaje por el pericentro puede ocurrir antes de la primera observación o después de la última.

⁸En este ejemplo se tiene entonces: $K_{min}=8.4$ m/s, $K_{max}=108.4$ m/s y $\Delta K = 100$ m/s.

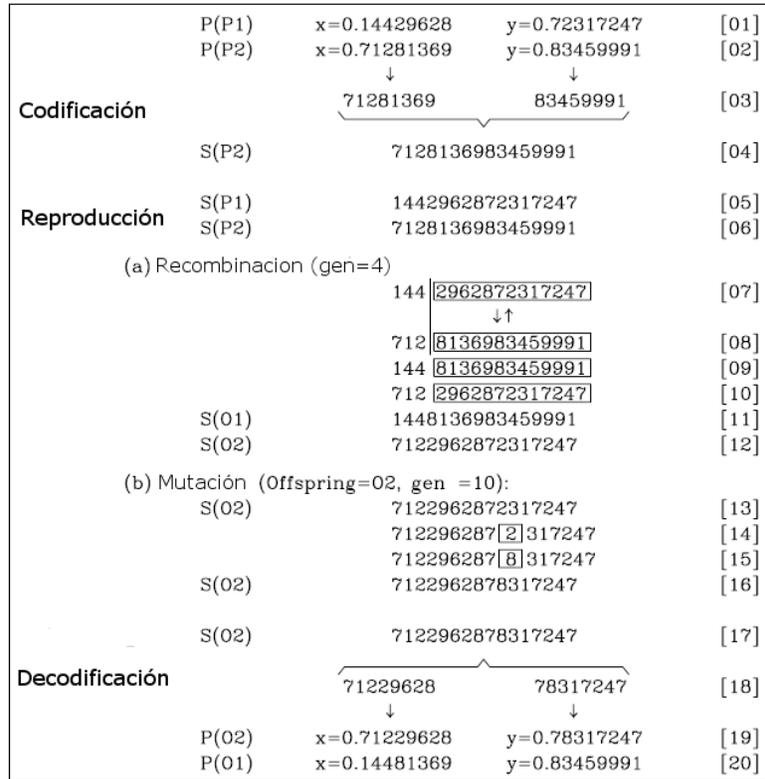


Figura 2.5: Esquema de operaciones básicas que realiza el algoritmo genético, en un problema de minimización 2D. Un individuo es dado por un punto (x,y) y dos individuos que denominamos padres son necesarios para la reproducción denominados P(P1) y P(P2). El operador recombinación actúa sobre las cadenas que representan los padres (ver texto para más detalle). Adaptado de Charbonneau y Knapp (1995).

$$\blacksquare V_0 \in [\min\{V_i\} ; \max\{V_i\}]$$

Donde ΔV está dado por la diferencia entre el valor máximo y mínimo de V_r disponibles, es decir ($\Delta V = \max\{V_i\} - \min\{V_i\}$); mientras que Δt está dado por la diferencia de tiempo entre la observación final y la inicial, es decir $\Delta t = t_{final} - t_{inicial}$.

2.2.4.2. Algoritmo genético esquemáticamente

La rutina PIKAIA busca el máximo de la función trabajando con $x_{j,j=1}^{j=M} \in [0, 1]$. Cada individuo \mathbf{x} puede ser transformado en \mathbf{A} con la Ec. (2.109). La llamada a la rutina es:

$$\text{pikaia}(\text{function}, \mathbf{n}, \text{ctrl}, \mathbf{x}, \mathbf{f}, \text{status}) \quad (2.110)$$

donde **function** es la función que se evaluará para buscar el máximo, **n** es la dimensión del vector \mathbf{x} , **ctrl** es un vector que define parámetros para el algoritmo genético, **f** es el valor de **function** en el vector \mathbf{x} y finalmente **status** es un entero cuyo resultado es **status**=0 si se ejecutó bien la rutina mientras que es **status**≠0 cuando falló la implementación de la rutina (para más detalles ver Charbonneau, 2002ab).

El programa codifica el genotipo internamente para obtener los fenotipos. Para fijar un poco las ideas veamos la Figura (2.5). En la figura solo se muestran los pasos básicos de reproducción y mutación en los algoritmos genéticos para el caso de un problema en un espacio bidimensional. A continuación se describen los pasos y se dan las consideraciones para el caso de los ajustes orbitales:

- Se genera aleatoriamente un conjunto de individuos, con elementos orbitales dentro del espacio de parámetros permitido. En nuestros cálculos este número oscila entre 100 a 200 individuos. En el caso de la figura cada par (x,y) es un individuo con fenotipos (x,y) . Luego cada individuo es codificado (línea 3) para producir un genotipo que caracteriza el individuo, siendo un escalar continuo que denominamos xy y tiene toda la información previa.
- Se forman pares de individuos $(P1_{xy})$ y $(P2_{xy})$ dada una tasa de aceptación de la recombinación genética, para mezclar la información de los genotipos. En este caso la mezcla se da en el $gen=4$ y se intercambia la información. Así producimos 2 individuos de descendencia ($O1$ y $O2$) de manera de no disminuir el número total de individuos disponibles (línea 7 y 8, Fig 2.5).
- Se especifica una tasa de mutación que introduce el factor de aleatoriedad a la descendencia en un cierto gen elegido aleatorio, en el ejemplo la mutación actúa sobre el descendiente $O2$ y el $gen=10$ cambiando el valor 2 por 8 (ver líneas 14 y 15 en la Fig 2.5).
- Cada información de individuo es decodificada para obtener los $(x,y)_1$ y luego transformada al espacio de parámetros A_1, A_2 . Con los valores de A_i se calcula un valor de χ^2_ν (línea 18 y 19, Fig 2.5).
- Cuando se ha realizado este proceso con todos los individuos se dice que se ha aumentado una generación.
- Se renueva parte del pool de individuos (la i -ésima generación) dejando solo algún porcentaje de aquellos que sean los mejor adaptados, es decir los de menor valor χ^2_ν .
- Se repite el proceso volviendo al inicio a lo largo de 10^4 a 10^5 generaciones.

Finalmente se obtiene una solución que resulta estar muy cerca del mínimo global en el espacio de parámetros explorado y en un tiempo que para 2 planetas en órbitas conmensurables rara vez supera los 10 minutos en una computadora Pentium Dual Core de 2GHz utilizando uno solo de sus núcleos.

Características evolutivas incluidas en estos esquemas:

1. la descendencia (O, offspring) incorpora porciones de material genético “intacto” que proviene de ambos padres, hecho que es necesario para la herencia. Sin embargo tanto los operadores reproducción (crossover) como la mutación involucran componentes estocásticos puros, como por ejemplo el punto de corte, el lugar de la mutación y el efecto de la mutación.
2. el proceso de codificado/decodificado es sólo uno de los posibles esquemas utilizados, los algoritmos genéticos tradicionales usan sistemas binario, aunque no es particularmente ventajoso para la optimización numérica.
3. los operadores reproducción y mutación operan en conjunto con el proceso de codificación / decodificación, preservando el rango total dentro del espacio de parámetros, es decir las soluciones están restringidas dentro del espacio $[0, 1]$ y la descendencia también estará restringida al espacio de parámetros $[0, 1]$.
4. el operador mutación tiene una interesante consecuencia. Dependiendo de cuál de los dígitos sea afectado la exploración de parámetros es paralelamente: global o pequeña (fine tuning), lo que permite sortear muy bien algunos problemas donde existen bordes pronunciados en la función residuo R .

En la introducción de muchos libros de texto se enumeran siete características básicas para considerar que un método adoptado es bueno: exactitud, consistencia, justificación, certeza, orden, ponderación y resolución. La lectura de esta sección nos guía a pensar que los Algoritmos Genéticos no son métodos adecuados. Sin embargo desde un punto de vista práctico surge una sugerencia práctica mencionada por Charbonneau (1995) y con la que concordamos luego de realizar exhaustivas pruebas:

USAR CUALQUIER MÉTODO QUE FUNCIONE

2.2.4.3. Parámetros de inicialización

En la rutina PIKAIA encontramos un conjunto de 12 parámetros de control, denominados **ctrl**, que permiten que adecuemos la forma en la que trabaja nuestro algoritmo genético (e.g. Charbonneau, et al., 2002). A continuación enumeraremos los parámetros y especificaremos los valores típicos que hallamos útiles en nuestros estudios de determinaciones orbitales:

- $\text{ctrl}(1) = 100 \sim 200$ Número de individuos en la población
- $\text{ctrl}(2) = 10^3 \sim 10^4$ Número de generaciones a evolucionar
- $\text{ctrl}(3) = 6$ Número de dígitos (número de genes).

Es el número de dígitos retenidos en el codificado del fenotipo en el genotipo. Recordemos que el genotipo termina siendo un número entero y aunque sea ineficiente en término de almacenamiento es útil a la hora de codificar, entender y poder rastrear lugares donde surjan errores.

- $\text{ctrl}(4) = 0.85$ Probabilidad de reproducción; debe ser ≤ 1.0 .

Una vez que los padres son seleccionados se genera un número aleatorio y la operación de reproducción (ver Fig. 2.5) se aplica solamente si el número aleatorio es menor que este valor.

- $\text{ctrl}(5) = 2$ Modo de Mutación; 1/2/3/4/5 (por defecto 2)

1=mutación de un punto, a tasa fija

2=mutación de un punto, ajustable basándose en el valor del ajuste.

3=mutación de un punto, ajustable basándose en la distancia $\delta\Gamma$

4=mutación de un punto+creep, a tasa fija

5=mutación de un punto+creep, ajustable basándose en el valor del ajuste

6=mutación de un punto+creep, ajustable basándose en la distancia

Las 3 primeras opciones dan mutación en un gen ya sea a una tasa fija (1) o variable, utilizando el valor del ajuste R (2) o alternativamente basado en la distancia métrica entre el mejor individuo y la media que llamaremos $\delta\Gamma$ (3).

Desde el punto de vista de la exploración del espacio de parámetros es necesaria una tasa variable, ya que como sabemos cuando la población es distribuida de forma amplia, el operador crossover realiza una búsqueda eficiente y recombina información de las soluciones existentes. Si la población converge (ya sea a un mínimo absoluto o local) la reproducción NO mejora mucho porque es realizada intercambiando fragmentos de padres casi idénticos; obviamente en estos casos se necesita una tasa de mutación alta para lograr inyectar variabilidad en la población.

La mutación varía entre el mínimo y máximo $\in [\text{ctrl}(7), \text{ctrl}(8)]$ durante la evolución con un valor inicial $\text{ctrl}(6)$. En cualquier momento se puede calcular $\delta\Gamma$ que es la diferencia entre el ajuste del más apto de los individuos y el valor del ajuste de la mediana de individuos (obtenido a partir del ordenamiento de los mismos). Dicho parámetro es entonces una medida de la convergencia de la población; si $\delta\Gamma$ es grande entonces la población es presumiblemente distribuida de manera más extensa en el espacio de parámetros que si $\delta\Gamma$ fuera pequeña. Entonces cuando $\delta\Gamma$ es pequeña se incrementa la tasa de mutación y cuando es grande se disminuye.

La mutación **creep** intenta definir una mutación *suave*. Veamos lo que significa con un ejemplo. Consideremos un problema donde buscamos una solución óptima que contiene la subcadena producida por el proceso de evolución:

.....21000.....

decodificada en un punto flotante (número real) 2.1000. Previamente imaginemos que tenemos el número

.....19123.....

que produce un ajuste mejor que la media, de forma que su material genético es dispersado en la población. Después de cierta cantidad de generaciones que den reproducciones y mutaciones favorables tenemos la subcadena que se ve por ejemplo como:

.....19994.....

que evidentemente está cerca de **21000**. Sin embargo dos mutaciones bien coordinadas deben actuar para producir este número: el “1” debe cambiar a “2” y el primer “9” a “0”, para producir **20994**, ie. 2.0994. Cualquier otra mutación que ocurra aisladamente o mutación por otro dígito nos llevará lejos del número que buscábamos. Como ya mencionamos la mutación es un proceso lento y la probabilidad que el par de mutaciones ocurra simultáneamente es relativamente ínfima, significando que necesitaremos muchas generaciones para que ocurra. La población comenzará a **amontonarse** en bordes internos del sistema codificado. Estos bordes son llamados *Paredes de Hamming*. Ellas pueden ser sorteadas eligiendo un esquema de codificación tal que las mutaciones pueda dirigirse a variaciones continuas en el parámetro decodificado. La mutación **creep** justamente actúa de la forma que necesitamos: una vez que el dígito de la cadena codificada fue seleccionado para mutar, en lugar de reemplazar el dígito existente por uno aleatorio, solamente se agrega +1 o -1 con igual probabilidad y si el resultado es < 0 (porque cambiamos un “0” sumándole “-1”) o > 9 (porque cambiamos un “9” sumándole “+1”), se lleva la adición al próximo dígito de la izquierda (como si realmente estuviéramos sumando). Así por ejemplo si el dígito a cambiar es el “9” central,

.....19994.....

con +1 obtendremos

.....20094.....

que tiene el efecto deseado de saltar la pared.

El problema sin embargo que se presenta con este tipo de mutación es que no se permite dar grandes saltos en el espacio de parámetros. Como argumentamos previamente los saltos son una capacidad necesaria, consecuentemente en la práctica

un test de probabilidad debe decidir si aplica mutación de un punto o creep con iguales probabilidades.

- ctrl(6) = 0.005 Tasa de mutación inicial; debe ser pequeña
- ctrl(7) = 0.025 Tasa de mutación mínima; debe ser ≥ 0.0
- ctrl(8) = 0.25 Tasa de mutación máxima; debe ser ≤ 1.0
- ctrl(9) = 1 Ajuste relativo a la función residual; desde 0 a 1
 PIKAIA realiza una clasificación para asignar ajuste. Los individuos son inicialmente clasificados como $[1, 2, \dots, n_p]$ de acuerdo al ajuste verdadero (dado por la función residuo R), donde el individuo de mejor ajuste tiene "ranking" 1 y el peor tiene ranking n_p . PIKAIA define internamente una probabilidad de reproducción para cada elemento basado en este ranking. Si este parámetro es ctrl(9)=0, no hay presión de selección y todos los individuos tienen la misma probabilidad de ser seleccionados para la reproducción. Si se elige una alta presión de selección, ctrl(9)>0.6, se obtienen mejores resultados ya que los mejores individuos son los que participan en la evolución del pool de genes (Charbonneau, 1995).
- ctrl(10) = 3 Plan de reproducción; 1/2/3.
 (1) Reemplazo generacional completo. Se calculan todos los individuos descendientes (Offspring) en un almacenamiento temporario y una vez que se tienen suficientes descendientes se reemplaza completamente la población original de padres.
 (2) Estado estacionario reemplazo aleatorio.
 (3) Estado estacionario reemplazando los peores (por defecto 3)
 El estado estacionario es un plan de reproducción que inserta los descendientes en la población a medida que son calculados. Para mantener los individuos totales se reemplazan los padres de forma aleatoria (2) o bien reemplazando los peores individuos (3).
 En nuestros cálculos hemos corroborado la conclusión mencionada por Charbonneau (1995): ninguno de los planes de reproducción es aparentemente mejor a otro. En casos donde el número de grados de libertad es alto, $\nu \sim 100$, los resultados son mejores si ctrl(10)=3. Mientras que si ν es bajo, ctrl(10)=1 funciona bien.
- ctrl(11) = 1 Condición de elitismo; 0/1=off/on.
 Consiste en asegurar que el genotipo de mejor ajuste de la generación n es copiado al menos 1 vez sin alteraciones en la generación $n + 1$.
- ctrl(12) = 0 Tipo de salida de datos 0/1/2=Nada/Mínimo/Explícito

Se podría suponer que un algoritmo genético debería funcionar bien en una población grande, manteniendo alta presión de selección y altas tasas de mutación y crossover a través de la evolución. Sin embargo como mencionamos los algoritmos genéticos imitan los procesos de evolución y muchas veces estos son complejos e impredecibles. Crossover, mutación, presión de selección y el tamaño de la población interactúan en forma extremadamente no lineal y no pueden ser claramente separados como superficialmente se podría suponer. Por ejemplo una población grande en un principio del algoritmo converge rápidamente a un espacio grande, porque la probabilidad que un individuo en la población inicial aleatoria llegue cerca del mínimo global es directamente proporcional al tamaño de la población. Sin embargo durante los estadios posteriores una población grande tiene mucha inercia cuando la probabilidad surge de una mutación favorable.

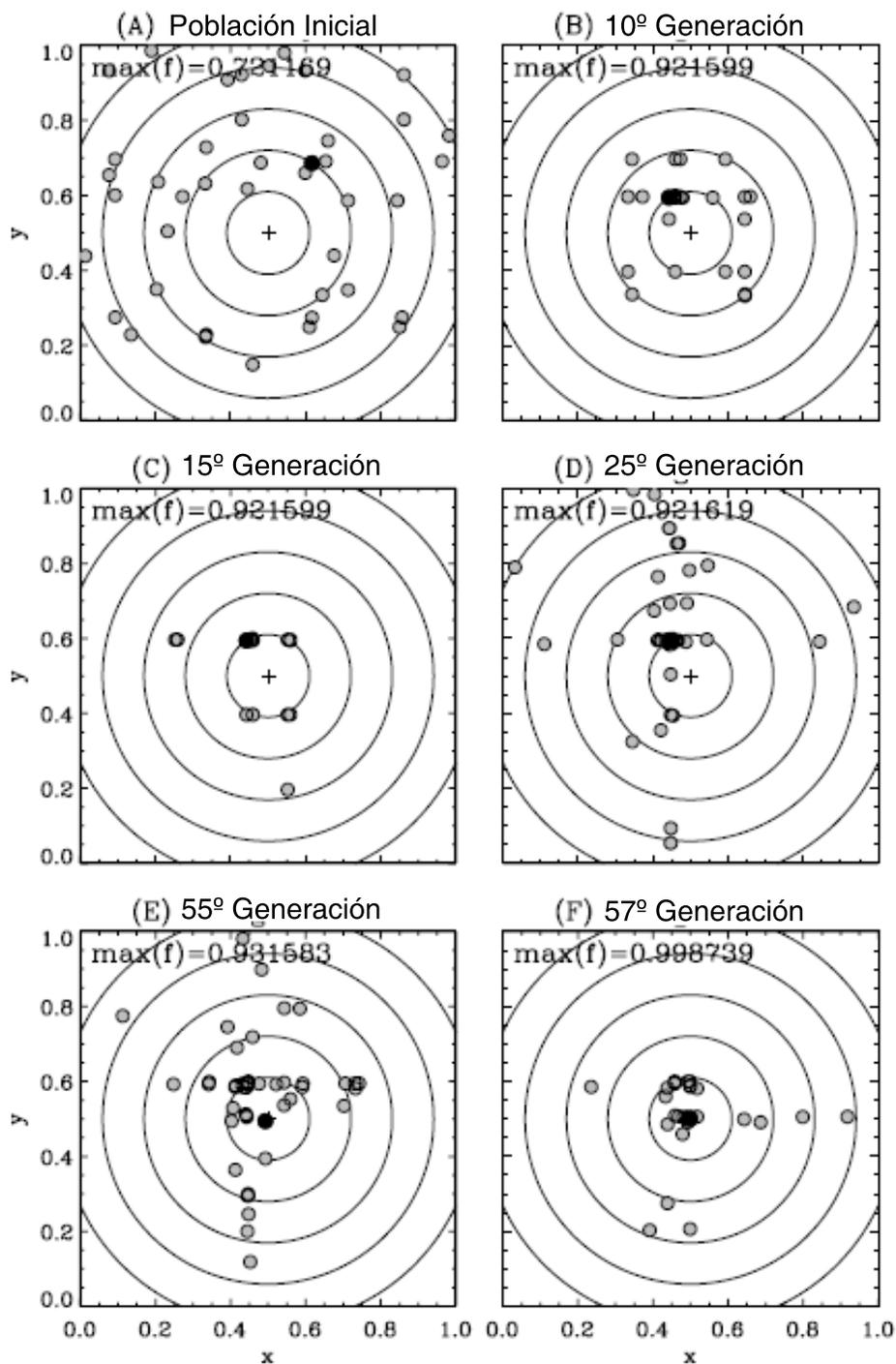


Figura 2.6: Evolución de la población de soluciones de prueba para el problema definido por la Fig. (2.4) utilizando algoritmo genético ($f(x, y) = \cos^2(n\pi r) e^{-r^2/\sigma^2}$). Los círculos concéntricos indican los anillos de mínimos secundarios, mientras que el círculo negro es el individuo con mejor ajuste en la generación actual. Adaptado de Charbonneau (1995).

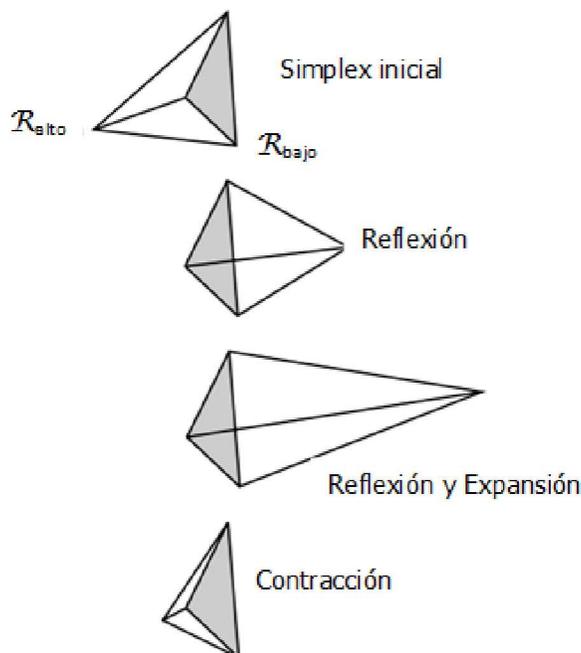


Figura 2.7: Esquema de operaciones básicas que realiza el algoritmo simplex para minimizar la función.

La Figura 2.6 muestra la distribución de la población en el espacio bidimensional definido por la función dada en la Fig. 2.4 para distintas generaciones. En la Fig 2.6.(A) se muestra que ningún individuo en la población inicial aleatoria se encuentra cercano al pico de la función (0.5,0.5). Las primeras iteraciones generacionales muestran un conjunto poblacional cercano al centro de la figura (Fig. 2.6.(B)), aunque la diferencia de ajuste entre el mejor y la mediana es todavía grande. La mutación decrece levemente desde su valor inicial pero se mantiene casi constante. En la 15^o generación, la mayoría de la población ha convergido hacia algunos lugares del anillo interno (Fig. 2.6.(C)) y extremo secundario de la función ($f=0.9216$). De esta forma, el ajuste entre el mejor y la mediana son comparables, esto conduce al algoritmo a un incremento abrupto de la tasa de mutación (entre la 12^o y 20^o generación). La alta tasa de mutación da como resultado una descendencia que se esparce sobre todo el espacio de parámetros en el curso de la reproducción (Fig. 2.6.(D)). Mientras que algunos individuos se posan regularmente en la pendiente del pico central, recién en la 55^o generación tal mutación catapulta suficientemente cerca del pico central a un individuo para convertirse en el mejor elemento de la población (Fig. 2.6.(E)). Subsecuentes reproducciones durante generaciones generaran más individuos cercanos al pico central. Notar como el elitismo (**ctrl(11)**) aquí es esencial, de otra forma las mutaciones podrían alejarnos de los buenos individuos ya identificados. De estos últimos pasos también se entiende la razón por la cual el algoritmo genético es ineficiente para hallar el valor del máximo con la precisión necesaria.

2.2.5. Método Simplex

A pesar que la rutina de Algoritmo genético, es extremadamente útil, para hallar el mínimo global, la solución que encuentra debe ser mejorada. Para ello es posible utilizar el Método Simplex (Simplex de ahora en más). El Simplex o *método downhill simplex* construye una figura geométrica llamada simplex de $N + 1$ vértices y todos ellos interconectándose con segmentos de líneas que se encuentra

en el espacio de \mathbf{M} dimensiones. Cada vértice es un conjunto de parámetros al cual se le puede calcular su valor asociado de la función residuo R (Press et al. 1992).

En el espacio 2-D un simplex es un triángulo mientras que en el espacio 3-D es un tetraedro. En cada iteración este simplex actualiza el aplicando transformaciones específicas a uno de sus vértices, encerrando la solución, hasta que se encuentra el mínimo. El método requiere solo evaluaciones de la función residuo R , y aunque en términos computacionales sea menos eficiente que usar derivadas de la función para encontrar el mínimo, la rugosidad de la función es bien sorteada. Cada uno de los $N+1$ vértices es un vector denominado \mathbf{M}_i con componentes x_i y es construido como:

$$\mathbf{M}_i = \mathbf{M}_0 + \beta \hat{e}_i \quad (2.111)$$

donde β es un parámetro de escala (alrededor del 0.1 % del rango válido para parámetro asociado) y \hat{e}_i es el vector unidad y tiene todas las direcciones posibles del espacio de parámetros.

El método realiza una serie de acciones sobre el vértice de mayor residuo R . La mayoría de las acciones es solo moviendo el vértice del simplex donde la función es menor atravesando la cara opuesta del mismo. Estos pasos son llamados reflexiones y son construidos para conservar el volumen del simplex. Cuando se puede el método expande el simplex en una u otra dirección para tomar pasos mas grandes. Cuando se llega a una región plana, es decir todos los vértices tienen el mismo valor de la función residuo R , el método contrae el simplex y encuentra el mínimo que denominamos \mathbf{M}_0 . A pesar de su definición “global” de búsqueda de mínimos, ciertamente la elección de la configuración inicial tiene influencia en la determinación del mínimo global. Una forma de visualizar como funciona el método es esquematizado en la Figura 2.7.

Existe una mejora al algoritmo Simplex, denominada **Simplex Iterativo**, donde se realizan repetidamente varios simplex. En la primer iteración se toman todos los otros vértices del simplex con los parámetros orbitales de forma aleatoria. A partir de esta solución e iterativamente se aplica simplex, tomando como condición inicial para uno de los vértices la mejor solución y los restantes vértices nuevamente elegidos aleatoriamente. En nuestro enfoque ya partimos desde una muy buena solución (proveniente del algoritmo genético) y a esta altura del código ya exploramos el espacio de parámetros. El trabajo de Charbonneau (1995) muestra las evidentes ventajas del Algoritmo Genético frente al Simplex en problemas de minimización globales.

El número de iteraciones del simplex utilizado es del orden de 10^3 , y es muy rápido tardando solamente 1 minuto para realizarlas. Si bien puede aparentar que es un método muy eficiente, esto se debe al valor acotado que tenemos para el simplex inicial.

2.2.6. Comparación Simplex Iterativo y Algoritmo Genético

Resulta interesante comparar las capacidades exploratorias entre **Simplex Iterativo** y **Algoritmo genético**. Para ello se puede definir el número de pruebas de cada uno de los métodos de acuerdo al número de evaluaciones de la función N_f es:

$$N_f = N_p \times N_g \quad \text{Algoritmo Genético} \quad (2.112)$$

$$N_f = N_t \times N_s \quad \text{Simplex Iterativo} \quad (2.113)$$

donde N_p y N_g son el tamaño de la población y longitud generacional, N_t es el número de intentos y N_s el número promedio de evaluaciones requeridos por un simplex único (es una cantidad dependiente

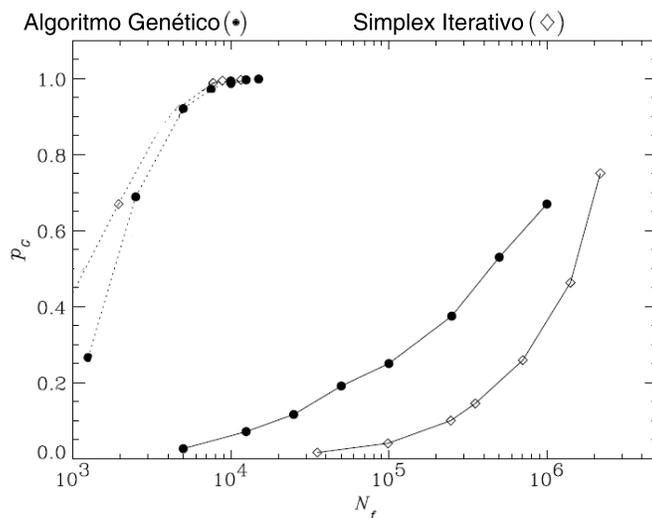


Figura 2.8: Probabilidad de convergencia global en función del número de evaluación de funciones N_f requerido por Simplex Iterativo (diamantes) y Algoritmo Genético (círculos negros), en el problema descrito por la Fig. 2.4 con línea punteada. El problema es bidimensional, ya que $r = r(x, y)$. Si modificamos el problema de tal forma de hacerlo 4-dimensional, i.e. $r = r(x, y, z, w)$, el problema es más complejo y los resultados de convergencia global son mostrados con la línea continua. Adaptado de Charbonneau (2002).

del problema y de los parámetros aleatorios iniciales). Se aplicó la minimización al problema descrito por la Fig. 2.4, incrementando el número de generaciones/iteraciones y se comprobó la convergencia global. Para obtener un valor estadístico se realizó 1000 veces para cada método. Esto permitió definir empíricamente la probabilidad de convergencia global ($p_G, \in [0, 1]$) como función de la cantidad de generaciones/iteraciones. Para decidir si se obtuvo una convergencia global se definió $f(x, y) > 0.95$ y los resultados se muestran en la Figura 2.8.

Este tipo de comparación fue realizado para varios problemas complicados de minimización global por Charbonneau (2002) y en todos se observa el mismo resultado general: los algoritmos genéticos no son tan buenos como podría esperarse respecto del simplex iterativo. Sin embargo, el trabajo de Charbonneau muestra un aspecto interesante que hemos corroborado con nuestros experimentos numéricos: en espacios de baja dimensión el simplex es muy competitivo, sin embargo a medida que crece la dimensión del problema y/o complejidad del mismo surge la ventaja de utilizar Algoritmos Genéticos. Esto se observa en la Fig 2.8 cuando usamos el problema cuatridimensional para la misma función f pero con r definido en un espacio 4-dimensional como $r = (x - 0.5)^2 + (y - 0.5)^2 + (z - 0.5)^2 + (w - 0.5)^2$.

2.2.7. Simulated Annealing

El método usualmente mencionado como *Simulated Annealing* (*SA*) en la literatura es una herramienta de exploración global cuyo nombre e inspiración viene del proceso de recocido del acero, una técnica que consiste en calentar y luego enfriar controladamente un material para aumentar el tamaño de sus cristales y reducir sus defectos. El calor causa que los átomos se muevan de sus posiciones iniciales (un mínimo local de energía) aleatoriamente; el enfriamiento lento les da mayores probabilidades de encontrar configuraciones con menor energía que la inicial.

En cada iteración, el *SA* considera algunos vecinos del estado actual s , y probabilísticamente de-

cide entre cambiar el sistema al estado s' o quedarse en el estado s . Las probabilidades se escogen para que el sistema tienda finalmente a estados de menor energía. La probabilidad de hacer la transición al nuevo estado s' es una función $P(\delta E, T)$ de la diferencia de energía $\delta E = E(s') - E(s)$ entre los dos estados, y de la variable T , llamada temperatura.

Una cualidad importante del método es que la probabilidad de transición P es siempre distinta de cero, aún cuando δE sea positivo, es decir, el sistema puede pasar a un estado de mayor energía (peor solución) que el estado actual. Esta cualidad impide que el sistema se quede atrapado en un óptimo local. Cuando la temperatura tiende al mínimo, la probabilidad tiende a cero asintóticamente. Así, cada vez el algoritmo acepta menos movimientos que aumenten la energía. Si δE es negativo, es decir, la transición disminuye la energía, el movimiento es aceptado con probabilidad $P=1$. Otra cualidad del Simulated Annealing es que la temperatura va disminuyendo gradualmente conforme avanza la simulación. Hay muchas maneras de disminuir la temperatura (T), siendo la más usual una ley exponencial, donde T disminuye por un factor < 1 en cada paso. Es por ello que en este método las primeras iteraciones cambian bruscamente los parámetros y van disminuyendo lentamente los cambios a medida que T decrece. También es posible incorporar saltos aleatorios y de esta forma evitar quedar atascados en algún mínimo/máximo local. En el caso de minimización, las energías representan los valores de χ^2_ν . Además la temperatura está relacionada con el parámetro β_i que define la escala característica de cada parámetro.

Si bien este enfoque es útil, en nuestro conjunto de rutinas resultó provechoso modificar esta idea, debido a que el mínimo global es eficientemente identificado por la rutina de Algoritmo Genético. Nuestra rutina *Annealing* trabaja con un enfoque MonteCarlo no determinístico. Cada conjunto de soluciones está formada por el vector de M parámetros: $\mathbf{x}_{1,2,3,\dots,M}$ y se eligen dos parámetros al azar dentro de los M parámetros disponibles (x_{i1} y x_{i2}). Luego se varía su magnitud en forma aleatoria, de manera que no pueda exceder el valor dado del parámetro de escala β_i . Finalmente se calcula el valor del residuo con el vector de componentes nuevas $\mathbf{x}'_{1,2,3,\dots,M}$. Esquemáticamente el algoritmo es:

$$\begin{aligned}
 \chi^2_\nu(\mathbf{x}) & \quad \# \text{ calcula el valor del residuo con los } M \text{ parámetros originales.} \\
 i1 & = M \cdot \text{ran}0 \quad \# \text{ elige un parámetro al azar dentro del conjunto de } M \text{ parámetros.} \\
 i2 & = M \cdot \text{ran}0 \quad \# \text{ elige un segundo parámetro al azar.} \\
 x'_{i1} & = x_{i1} + (\text{ran}0 - 0.5) \cdot \beta_1 \quad \# \text{ varía } x_{i1}, \text{ una cantidad aleatoria menor que } \beta_1. \\
 x'_{i2} & = x_{i2} + (\text{ran}0 - 0.5) \cdot \beta_2 \quad \# \text{ varía } x_{i2}, \text{ una cantidad aleatoria menor que } \beta_2. \\
 \chi^2_\nu(\mathbf{x}') & \quad \# \text{ calcula el valor del residuo con los } M \text{ parámetros modificados}
 \end{aligned} \tag{2.114}$$

Se utiliza el generador de números aleatorios de distribución uniforme *ran0* (Prees et al. 1992). Los valores óptimos para β_i , son de alrededor del 1% de la región permitida para cada parámetro. Si el cálculo de la función R utilizando los nuevos valores de los parámetros orbitales, es menor que el valor previo se aceptan los nuevos valores, de otra forma se vuelve a tomar la condición original, tomando unas 10^5 pruebas. Si la solución no mejora con esta cantidad de pruebas la rutina es terminada, en caso contrario se incrementa el número de iteración del Annealing. Normalmente se itera el Annealing unas 10^3 a 10^4 veces. Aunque la convergencia al mínimo se logra antes de llegar a estos valores.

La Figura (2.9) muestra el funcionamiento del método en un espacio biparamétrico. Se aprecia como puede quedar atrapada la solución en un mínimo local. La mejora del método con saltos aleatorios es ilustrada con la Figura 2.10, donde un salto aleatorio (un valor arbitrariamente grande para β_i) convenientemente dirige la búsqueda del máximo a una región que de otra forma sería inalcanzable.

En nuestros problemas de minimización de velocidades radiales hemos observado que las pequeñas rugosidades de la función residuo R no son bien sorteadas por el Simplex, y se adaptó una

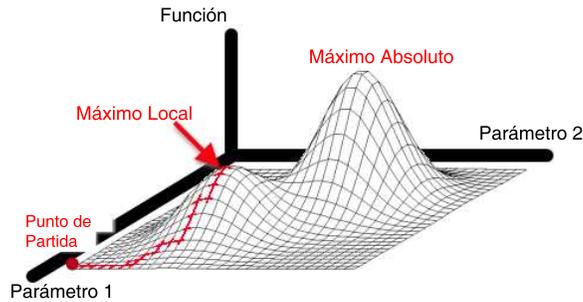


Figura 2.9: Si bien en un principio de iteraciones el algoritmo puede producir resultados equivalentes a un **método de gradiente** este método sólo incorpora evaluaciones de la función, aunque también puede quedar atrapado en un máximo local como muestra la figura.

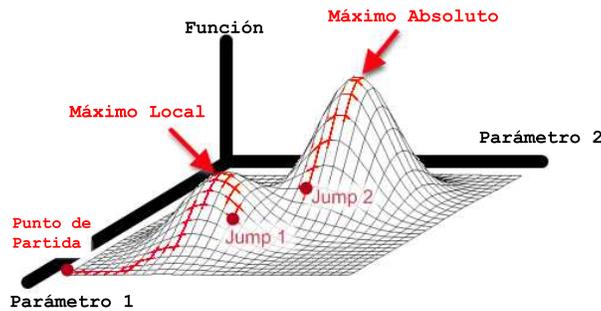


Figura 2.10: El algoritmo puede ser mejorado con un salto aleatorio en el espacio de parámetros. El salto indicado como **jump 1** → **jump 2** es dado gracias un valor arbitrariamente grande del escalar de exploración β .

rutina de Annealing a fin de detectar fehacientemente el mínimo en la función.

2.3. Ajustes Dinámicos

En sistema de varios exoplanetas, el modelo kepleriano es una aproximación. Esta aproximación es constituida por la suma de las contribuciones individuales de la señal de cada planeta, sin tener en cuenta los otros cuerpos, y la interacción que puede existir entre los mismos. Deben chequearse entonces si existen estas interacciones en las observaciones.

Encontrar cuál es la velocidad radial de la estrella entorno a su baricentro, debido a la presencia de los otros cuerpos es una tarea simple. Sólo se deben integrar las ecuaciones de movimiento de los N-cuerpos involucrados:

$$\frac{d^2 \mathbf{x}_i}{dt^2} = - \sum_{j=1}^N \frac{Gm_j (\mathbf{x}_i - \mathbf{x}_j)}{|\mathbf{x}_i - \mathbf{x}_j|^3}, \quad (2.115)$$

Luego, para cada tiempo de observación (t_i), donde se tiene la velocidad radial $V_r(i)$, se puede obtener la velocidad en la dirección z que es por definición la velocidad radial.

La idea es comenzar con una solución kepleriana (que en principio es una buena aproximación del problema) e intentar mejorar el ajuste cambiando levemente algunos de los parámetros planetarios. Con estos parámetros planetarios modificados se integran las ecuaciones de movimiento en el intervalo de tiempo que cubren las velocidades radiales, tomando como elementos osculadores iniciales del sistema planetario la solución a probar.

La estrategia para cambiar levemente los parámetros y mejorar el ajuste es idéntica a la de nuestra rutina Annealing. Utilizamos un integrador Runge Kutta de orden 8 (RK8) y con precisión 10^{-12} , siendo los resultados indistinguibles comparados con un integrador Bulirsch-Stoer de la misma precisión, aunque los tiempos de cómputo con RK8 son apreciablemente más cortos.

Existen dos puntos a destacar en este tipo de ajustes:

- El gasto computacional es extremadamente alto ya que se integra CADA condición de prueba, siendo necesario para la convergencia del método desde unas horas a días dependiendo de la condición inicial kepleriana de partida y del sistema analizado.
- El intervalo de las observaciones rara vez supera los 5 a 10 años, tiempo extremadamente corto para que se evidencien interacciones entre planetas y más aún, si existiesen difícilmente sean acusados debido a los errores de las mediciones.

Si bien podríamos mejorar la velocidad del programa utilizando un integrador simpléctico, los sistemas ya fueron analizados al momento de escribir la tesis y no se encontró razón para creer que pueda ser útil re-analizar todas nuestros ajustes de N-cuerpos.

Para comenzar con los ajustes dinámicos, se realizaron pruebas asumiendo configuraciones coplanares y con inclinación 90° . Sin embargo sabemos que este tipo de integración también permite modelar la posible inclinación mutua entre los cuerpos. Es decir tenemos 4 parámetros adicionales para modelar: ambas inclinaciones i y ambas longitudes de nodo Ω . Sin embargo como exponremos en nuestros ejemplos, el valor del residuo R no disminuye demasiado, a pesar de estar agregando más parámetros al modelo. Por otra parte, los valores de inclinación resultantes son muy cercanos a valores de órbitas coplanares, por lo cual este tipo de ajuste **3-D** no resultó fructífero en los sistemas expuestos.

Entre los pocos sistemas extrasolares en los cuales se encontraron interacciones y, por ende, diferencias cuantificables con el modelo kepleriano podemos citar: GJ 876 (Rivera et al. 2005), HD73526 (Tinney et al. 2003), HD202206 (Correia et al. 2005) y HD60532 (Laskar y Correia 2009).

2.4. Obtención de parámetros planetarios para el sistema HD82943

2.4.1. Reseña histórica

Nuestro método de minimización sirve para cualquier sistema exoplanetario y a lo largo de la tesis fue testeado en otros sistemas de exoplanetas. Sin embargo aquí se expondrán solo los resultados del sistema exoplanetario HD82943 por dos razones: primero porque está demostrado que los mejores ajustes orbitales de este sistema muestran un comportamiento dinámicamente inestable, implicando

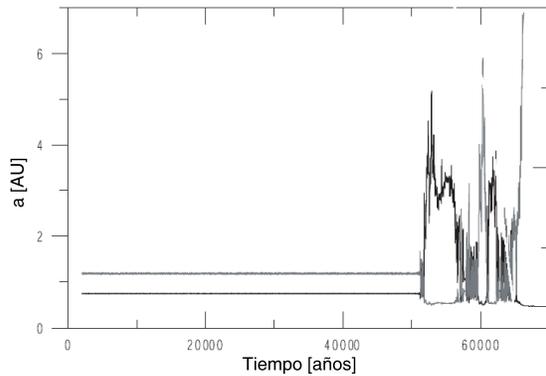


Figura 2.11: Integración de la solución dada por Mayor et al. (2004) donde se muestra la evolución de los semiejes de los planetas considerando masas: $M_* = 1.15M_\odot$, $m_1 = 1.85M_J$, $m_2 = 1.84M_J$). Reproducido de Ferraz Mello et al. (2005).

que no es consistente con la configuración real de los planetas; segundo, porque ya fue estudiado por varios autores proveyendo una buena base con la cual comparar nuestros resultados.

Actualmente existen dos planetas conocidos en el sistema HD82943 (Mayor et al. 2004), ambos descubiertos por el grupo de Ginebra, el primero en 2000 y el segundo en 2001. A pesar que en 2001 se tenía bastante certeza de la existencia de dos planetas que se encontraban en la vecindad de la resonancia de movimientos medios (MMR) 2/1, recién en 2002 hubo suficiente cantidad de observaciones disponibles para determinar un ajuste orbital confiable. Estos parámetros fueron puestos a disposición en la página del grupo de Ginebra pero luego actualizados con nuevos valores (los valores originales se pueden observar en la Tabla de Ji et al. 2003). Estos parámetros podían ser integrados y mostraban que los planetas estaban capturados en una resonancia de corotación apsidal (ACR) del tipo (π, π) , es decir, los ángulos que muestran el movimiento resonante y secular oscilan en torno a los valores entre paréntesis, con una amplitud de oscilación pequeña (ver Beaugé et al. 2003 para más detalles). A pesar que esta órbita es dinámicamente estable por al menos 10^6 años, este tipo de ACR es incompatible con un escenario de migración planetaria suave desde órbitas casi circulares (Beaugé et al. 2006).

En un artículo siguiente, Mayor et al. (2004), se presentó un nuevo conjunto de elementos orbitales con más datos de velocidad radial, tomados con el instrumento CORALIE y totalizando $N=142$ observaciones. A pesar que la solución mostraba ambos planetas aún en MMR 2/1, existían diferencias significativas en las masas y elementos orbitales. La razón de masas m_2/m_1 varió desde 1.9 hasta ~ 1 , y las excentricidades variaron significativamente. Con este nuevo conjunto de parámetros orbitales, la configuración aparentaba corresponder a un ACR tipo (0,0), más compatible con escenarios de migración planetaria.

A partir de las integraciones numéricas de los ajustes orbitales Ferraz-Mello et al. (2005) encontraron que la solución publicada por Mayor et al. (2004) era dinámicamente inestable tan sólo en 6×10^4 años (ver Figura 2.11). Debido a que la estrella central entró en secuencia principal hace unos 3 Gy resulta improbable que sus planetas sean estables en intervalos de tiempo mucho menores que esto. Así, aparentemente el mejor ajuste orbital no se corresponde con el sistema real y se debe encontrar otra solución compatible con los datos observacionales. Previo a este trabajo, raramente se comprobaba que el sistema extrasolar obtenido sea estable. Con los datos de velocidad radial, Ferraz-Mello et al. (2005) encontró que el residuo R alrededor del mínimo era una función muy suave de los parámetros y muchos ajustes orbitales distintos (algunos dinámicamente estables) daban origen a residuos similares. En otras palabras el mejor ajuste no necesariamente correspondía a los parámetros reales del sistema.

Lee et al. (2006) presentaron observaciones adicionales con el instrumento HIRES del observato-

rio KECK, que junto con los datos CORALIE totalizan $N=165$ observaciones. Nuevamente el ajuste orbital con estos datos condujo a una solución dinámicamente inestable, esta vez en un intervalo de tiempo del orden de los 10^3 años. Más aún, el valor del residuo del ajuste orbital multikepleriano se incrementó significativamente, indicando que un mayor intervalo de observación no necesariamente disminuye la diferencia entre los valores observados y los valores calculados (o-c de ahora en más). Sin embargo, los autores encontraron también soluciones con residuos similares al del mínimo: algunas corresponden a ACR tipo (0,0), mientras que algunas sólo muestran una libración del ángulo resonante y circulación para la diferencia de las longitudes de pericentro.

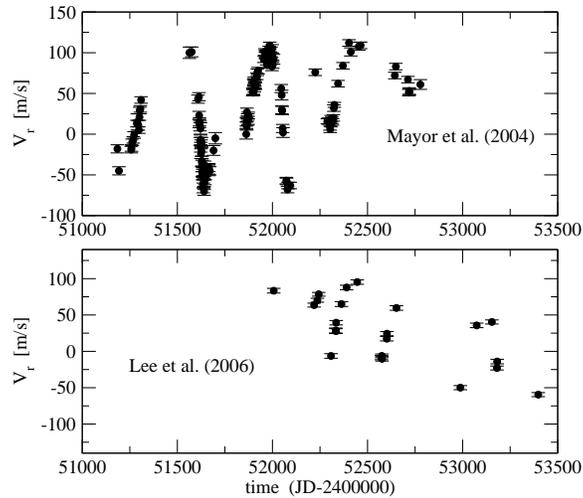


Figura 2.12: **Arriba.** Velocidades radiales del artículo de Mayor et al. (2004). **Abajo.** Velocidades radiales del artículo de Lee et al. (2006). Ambos conjuntos totalizan $N=165$ datos observacionales y son la base para nuestra minimización.

Más recientemente, un análisis distinto sobre el mismo conjunto de datos fue realizado por Goździewski y Konacki (2006), utilizando un algoritmo híbrido que incluye un análisis de estabilidad dentro del ajuste por cuadrados mínimos. En su trabajo se muestra la existencia de dos islas de movimiento estables asociadas a la resonancia MMR 2/1, en la cercanía del mejor ajuste orbital. Soluciones totalmente diferentes fueron analizadas y en particular los autores mostraron que dos planetas en movimiento coorbital y no coplanares pueden tener curvas de velocidad radial similares, con un error en el ajuste orbital un tanto mayor ($wrms \sim 8.4$). Sin embargo, la existencia de planetas gigantes coorbitales está lejos de ser establecida, aunque hay evidencia que planetas con masas mayores que $\sim 0.7M_{\oplus}$ no pueden acretarse en los puntos equiláteros lagrangianos de planetas extrasolares gigantes (Beaugé et al. 2007a). Sin embargo, en un trabajo que recientemente publicamos, hemos observado que existen amplias regiones de movimiento estable para planetas del tamaño de Júpiter que se encuentren en movimiento coorbital (Giuppone et al. 2010).

En el mismo trabajo Goździewski y Konacki (2006) también analizaron la posible existencia de un tercer planeta en HD82943 con un período orbital de aproximadamente 1000 días, exterior al más externo de los planetas conocidos. El ajuste orbital con 3 planetas mostró un decrecimiento significativo del residuo ($wrms \sim 6.3$ m/s), sin embargo la solución fue una vez más inestable en muy cortos tiempos de integración. Alternativamente se encontró una solución que es marginalmente estable con valores más pequeños en la excentricidad del hipotético nuevo planeta. A pesar que los residuos fueron mayores ($wrms \sim 7.4$ m/s), constituye un escenario interesante de estudio.

2.4.2. Velocidades radiales y mejores ajustes

Tabla 2.1: Mejores ajustes orbitales keplerianos para 2 planetas en el sistema HD82943. El tiempo del pasaje por el pericentro es dado en unidades a partir de [JD-2400000]. Todos los elementos orbitales son astrocéntricos y los períodos osculadores.

Parámetro	CORALIE=142 V_r		CORALIE+Keck=165 V_r	
	HD82943c	HD82943b	HD82943c	HD82943b
K [m/s]	61.65	45.75	65.77	43.69
P [d]	219.48	435.09	219.29	441.26
e	0.383	0.183	0.356	0.222
ω [°]	124.29	239.51	127.32	282.86
τ	50747.86	51325.84	50751.02	52695.54
m [M_{Jup}]	1.860	1.845	2.006	1.756
a [UA]	0.746	1.178	0.746	1.189
$\lambda - \varpi$ [°]	356.41	239.51	351.83	207.33
V_{0c} [m/s]	43.79		43.87	
V_{0k} [m/s]	-		33.69	
wrms [m/s]	6.964		7.954	
$\sqrt{\chi^2_\nu}$	1.165		1.369	

Analizamos el sistema HD82943 con el conjunto disponible de 165 velocidades radiales. La Figura 2.12 muestra los conjuntos disponibles. El marco superior muestra los datos CORALIE (no son públicos y fueron extraídos de una imagen postcript) y el marco inferior los datos de Keck (disponibles para toda la comunidad científica).

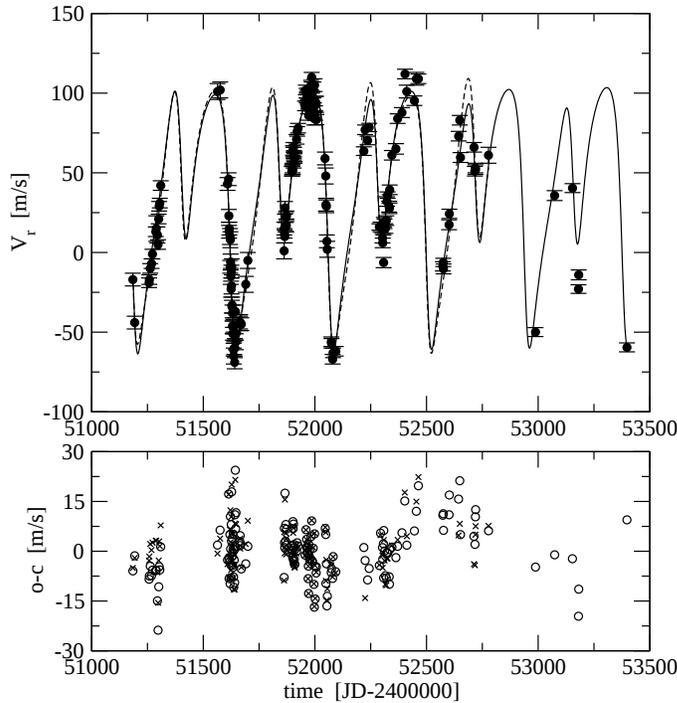
Para todos los cálculos a continuación utilizaremos la masa estelar central igual a $M_* = 1.15M_\odot$ masas solares, adoptadas por Ferraz-Mello (2005) y Lee et al. (2006). También se asumirá jitter estelar igual a $\sigma_j = 4.2$ m/s, deducido por Lee et al. (2006) a partir de modelos empíricos de interiores estelares. De esta forma la incertidumbre en las observaciones está dada según la Ec. (1.33), que introduce pequeños cambios en los ajustes orbitales cuando se toma un valor de jitter muy grande.

Con los datos de velocidad radial procedimos a calcular el mejor ajuste orbital kepleriano en un marco de referencia astrocéntrico, asumiendo la presencia de 2 cuerpos planetarios. Basados en nuestros resultados se minimizó la función residuo R , utilizando primero el Algoritmo genético con una población de 200 miembros, que evolucionan 10^4 generaciones. Luego debido a la característica exploratoria mejoramos la solución con Simplex (Sección 2.2.5) empleando unas 10^3 iteraciones. Finalmente las pequeñas irregularidades del espacio fueron sorteadas aplicando la rutina annealing (Sección 2.2.7) durante 10^3 iteraciones. Si todo el proceso convergió al mínimo, la última rutina termina por sí sola antes de acabar el número total de iteraciones. Este proceso completo no demora más de 10 minutos en una computadora con procesador Pentium Dual Core, de 1.8 Mhz. Se trabajó básicamente con 2 conjuntos de datos, el primero que consta de las 142 V_r correspondientes a CORALIE y el segundo conjunto utilizando los 165 datos que incluyen las observaciones CORALIE y las de Keck. Los resultados son mostrados en la Tabla 2.1 y coinciden con los calculados por los otros autores, permitiendo avalar nuestras rutinas de búsquedas de mínimos.

La Tabla 2.2 muestra para comparación los resultados del **ajuste dinámico** con los 2 conjuntos de V_r , donde se aprecian que ambos ajustes son casi idénticos. Si bien algunos parámetros tienen valores diferentes entre ambos ajustes, los mismos provienen de la forma poco profunda de la función residuo respecto a la excentricidad del planeta exterior. La Figura 2.13 compara los datos de velocidad radial junto con las curvas que se generan a partir de los ajustes keplerianos con los distintos conjuntos

Tabla 2.2: Mejores ajustes orbitales dinámicos para 2 planetas en el sistema HD82943. Todos los elementos orbitales son astrocéntricos y los períodos son osculadores.

Parámetros	CORALIE=142 V_r		CORALIE+Keck=165 V_r	
	HD82943c	HD82943b	HD82943c	HD82943b
P [d]	218.50	444.04	218.40	452.79
e	0.384	0.125	0.355	0.213
ϖ [°]	119.35	247.15	123.55	295.57
λ [°]	115.80	123.60	115.64	134.95
V_{0C} [m/s]	43.46		43.74	
V_{0K} [m/s]			34.29	
m [M_{Jup}]	1.866	1.809	2.025	1.749
a [UA]	0.744	1.194	0.744	1.210
wrms [m/s]	7.010		7.952	
$\sqrt{\chi^2_\nu}$	1.173		1.369	


 Figura 2.13: **Arriba.** Curva de velocidad radial sintética obtenidas con ajustes multi-keplerianos. La curva continua corresponde al conjunto de datos completo CORALIE+Keck, mientras que la curva a trazos corresponde a la de datos CORALIE solamente. **Abajo.** Residuos $o-c$ para los ajustes, con círculos abiertos para el conjunto completo y con cruces solo utilizando las V_r CORALIE.

de V_r . Se observa que hay una leve diferencia entre ambas curvas, especialmente en los máximos locales, donde no hay demasiadas observaciones. Además del cambio en las masas planetarias, una de las razones fundamentales para las diferencias yace en el valor de ϖ_2 . En el 2004 el mejor ajuste obtenido resultaba en $\varpi_2 \sim 238^\circ$, mientras que actualmente es $\varpi_2 \sim 280^\circ$. A pesar de esta diferencia en las señales que producirían ambas configuraciones, existe sólo una pequeña diferencia entre los puntos $o-c$.

La conclusión más importante del repaso de los trabajos previos es, la falta concreta de conocimiento de la estructura dinámica del sistema de HD82943. No solamente es probable una cantidad enorme de configuraciones posibles, sino que además los ajustes orbitales parecen ser sensibles al conjunto de datos observacionales. En cuanto al mejor ajuste, comienza a notarse la necesidad de describir un espacio de regiones donde tengamos soluciones compatibles con los datos observacionales y además estimar de alguna forma el error en cada parámetro determinado.

Capítulo 3

Estimación de Incertezas

Como dijéramos en un principio, no basta con estimar los parámetros que mejor satisfacen los datos, también se deben conocer los errores en los mismos. Los errores formales publicados para los parámetros de exoplanetas son, muchas veces, subestimados. Esta es una clara razón por la cual, cuando aparecen nuevos ajustes orbitales, los parámetros cambian sus valores mucho más allá de los errores formales. Existen tratamientos que demandan mucho tiempo de cómputo y que brindan una idea clara sobre las indeterminaciones y sobre los cuales basaremos este capítulo.

Como fuera mencionado en Ford (2005), a pesar de que las determinaciones de velocidades radiales provienen de una distribución Gaussiana, los errores en los parámetros orbitales no son Gaussianos. Resulta así necesario tener una visión global de los métodos para estimar errores en sistemas no-lineales y multiparamétricos. Como punto de referencia para aplicabilidad de los métodos y ejemplos podemos citar Brown (2004), Ford (2005) y en Beaugé et al. (2008).

En este capítulo primero recordaremos algunos puntos importantes acerca del **Teorema Central del Límite** y luego discutiremos como se determina el **Intervalo de Confianza** alrededor de la mejor solución.

En las secciones subsiguientes se discuten distintos enfoques para determinar incertezas en los parámetros planetarios. Como punto de partida se presenta la **matriz de Fisher** para determinar errores en los ajustes, justificando su escasa validez en el tipo de problema que enfrentamos. Luego los métodos que consideramos más útiles a la hora de determinar intervalos de confianza y soluciones compatibles con los datos: **Cadenas MonteCarlo de Markov** y **Generación de Datos Sintéticos**. Alternativamente y como método más rápido se discuten los **Ajustes en una Grilla**. Finalmente, para evaluar la dependencia de los ajustes con el intervalo de observación se presenta el método **Jackknife**. Muchos métodos son aplicados al sistema HD82943 y se analiza la compatibilidad de las soluciones con la factibilidad de existencia de las configuraciones (i.e, si integramos las soluciones, éstas deben ser estables pues representan un sistema planetario observado).

El Teorema Central del Límite: Consideremos un conjunto de variables aleatorias X_1, X_2, \dots, X_n , muestreadas de forma independiente, pero provenientes de una misma distribución con media μ y varianza σ^2 . Entonces, si n es suficientemente grande, su promedio es la variable aleatoria:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (3.1)$$

y tiene una aproximación normal, con media $\mu_{\bar{x}} = \mu$ y varianza $\sigma_{\bar{x}}^2 = \sigma^2/n$.

Es importante remarcar que este teorema no dice nada acerca de la distribución de X_i , excepto la existencia de media y varianza. La aproximación entre las dos distribuciones es, en general, mayor en el centro de las mismas que en sus extremos o colas, motivo por el cual se prefiere el nombre “teorema del límite central” (“central” califica al límite, más que al teorema).

En palabras simples, dada cualquier distribución $\mathcal{L}(\theta)$, es *asintóticamente* Gaussiana cerca de su máximo (valor más probable), tal como lo muestra la Figura 3.1. Consideremos la función $\ln \mathcal{L}(\theta)$ que posee M parámetros $\theta = \{\theta_1, \dots, \theta_M\}$ y realicemos una expansión de Taylor hasta segundo orden alrededor del valor más probable θ_{\max} ,

$$\ln \mathcal{L}(\theta) \approx \ln \mathcal{L}(\theta_{\max}) + \sum_{i,j} \frac{1}{2} \left. \frac{\partial^2 \ln \mathcal{L}}{\partial \theta_i \partial \theta_j} \right|_{\theta_{\max}} (\theta - \theta_{\max})_i (\theta - \theta_{\max})_j. \quad (3.2)$$

donde los términos lineales se anulan en el máximo. Consecuentemente, $\ln \mathcal{L}(\theta)$ es aproximada de forma cuadrática en θ cerca del máximo, siendo evidente que $\mathcal{L}(\theta) = e^{\ln \mathcal{L}(\theta)}$ tiene la forma de una Gaussiana. La bondad de la aproximación Gaussiana depende de la situación específica, es decir de los datos medidos, su ruido y el modelo utilizado. El ancho de la Gaussiana no es siempre una buena representación del error. La Figura 3.1 muestra un ejemplo donde esto ocurre. Consecuentemente si se invoca el Teorema Central del Límite, uno debería chequear la bondad de la aproximación Gaussiana.

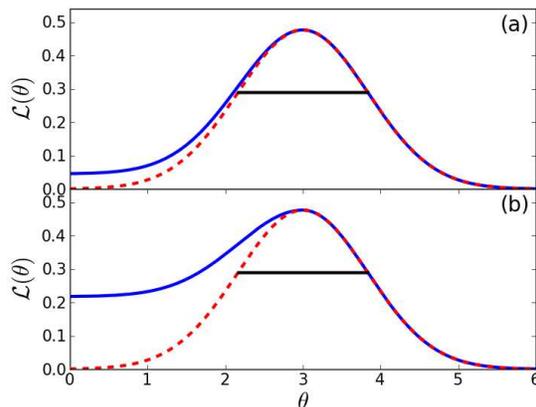


Figura 3.1: Posibles fallas del Teorema del Límite Central. Esta figura muestra un ejemplo de la función distribución \mathcal{L} como función de algún parámetro θ (curvas sólidas azules). En curva roja a trazos se muestra su aproximación Gaussiana y el ancho de la Gaussiana con una línea horizontal negra. En el panel (a) la aproximación Gaussiana es buena y la estimación del error es confiable. En el panel (b) la aproximación Gaussiana es de hecho pobre y el **error** está sustancialmente subestimado. Reproducido de Andrae (2010).

3.1. Intervalo de Confianza

Consideremos una variable θ que es aleatoria y tiene una distribución Gaussiana con media $\langle \theta \rangle$ y desviación estándar σ . Si tomamos un valor al azar de θ , la diferencia $|\theta - \langle \theta \rangle|$ será menor que σ con el 68.3 % de probabilidad, y será menor que 2σ con el 95.5 % de probabilidad. Los intervalos de confianza para una distribución Gaussiana (monoparamétrica), se muestran en la Figura 3.2.

La distribución Gaussiana es un ejemplo casi trivial debido, en parte, a su simetría alrededor de la media. En general, las funciones \mathcal{L} pueden no ser simétricas y se pueden definir los intervalos de confianza tales que $\theta_- \leq \hat{\theta} \leq \theta_+$ para una dada función de distribución que se denomina $\text{prob}(\theta)$

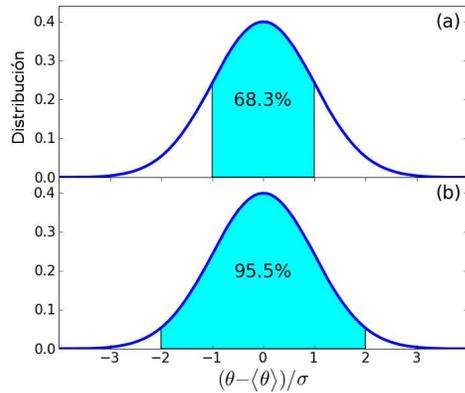


Figura 3.2: Intervalos de confianza para una distribución Gaussiana de media $\langle\theta\rangle$ y desviación estándar σ . Si tomamos al azar N valores de θ de la distribución Gaussiana, el 68.3 % de estos valores estarán dentro del intervalo $[\langle\theta\rangle - \sigma, \langle\theta\rangle + \sigma]$ como se muestra en el panel (a), mientras que el 95.5 % de estos valores se encontrarán dentro del intervalo $[\langle\theta\rangle - 2\sigma, \langle\theta\rangle + 2\sigma]$ como se muestra en el panel (b). Reproducido de Andrae (2010).

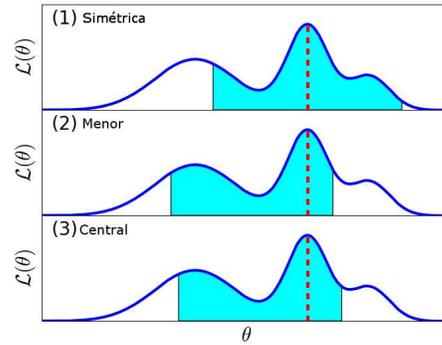


Figura 3.3: Diferentes tipos de intervalos de confianza de 68.3 % para funciones de distribución \mathcal{L} . La línea vertical (roja) muestra la estimación del máximo $\hat{\theta}$. Los paneles son numerados de acuerdo a las definiciones del texto. Reproducido de Andrae (2010).

mediante¹ (Andrae, 2010):

$$\text{prob}(\theta_- \leq \hat{\theta} \leq \theta_+) = \int_{\theta_-}^{\theta_+} d\theta \text{prob}(\theta) = C, \quad (3.4)$$

donde usualmente $C = 0.683$ es el intervalo Gaussiano de 1σ .

En la práctica, la función de distribución $\text{prob}(\theta)$ es desconocida y solamente se dispone de un histograma de muestras de θ . En este caso la integral de la ecuación (3.4) se reduce a determinar la fracción de todas las muestras de θ que se encuentran entre θ_- y θ_+ (similar al método utilizado para determinar las señales FAP en 2.1.7). Sin embargo la Ecuación (3.4) no está *unívocamente* definida dentro del intervalo de confianza, y un criterio adicional es requerido (Barlow, 1993):

1. Intervalo simétrico: θ_- y θ_+ son simétricas alrededor del parámetro estimado, i.e., $\hat{\theta} - \theta_- = \theta_+ - \hat{\theta}$.
2. Intervalo menor: $\theta_+ - \theta_-$ es el menor de los intervalos que satisfaga la Ecuación (3.4).
3. Intervalo central: Las probabilidades antes y después del parámetro estimado son iguales i.e., $\int_{-\infty}^{\theta_-} d\theta \text{prob}(\theta) = \int_{\theta_+}^{\infty} d\theta \text{prob}(\theta) = (1 - C)/2$.

Cuando la distribución es Gaussiana, las tres definiciones son iguales. En los casos generales, los intervalos son distintos, como se muestra en Figura 3.3. En la figura, el intervalo de confianza del 68.3 % no coincide con el “intervalo 1σ ” de una Gaussiana que ajuste en el máximo. En general no hay preferencia acerca de ninguna de las definiciones aunque se debe aclarar con cual se está trabajando.

¹Recordemos que la función de distribución está normalizada, es decir:

$$\text{prob}(-\infty \leq \hat{\theta} \leq +\infty) = \int_{-\infty}^{+\infty} d\theta \text{prob}(\theta) = 1, \quad (3.3)$$

Podemos calcular el límite de confianza, teniendo en cuenta el comportamiento estadístico de la función de distribución \mathcal{L} . En astronomía generalmente se utiliza el estadístico χ^2_ν , que se comporta como una variable chi-cuadrado, y existen distintas aproximaciones para evaluar los límites de confianza del mejor ajuste (al que denominamos $\chi^2_{\nu \min}$). Por ejemplo:

- Según Press et al. (1992) y Andrae (2010) el intervalo de confianza del 68.3 % cuando la variable posee un grado de libertad, se sitúa en algún lugar del parámetro para el cual:

$$\chi^2 = \chi^2_{min} + 1 \quad (3.5)$$

- Si consideramos que tenemos más grados de libertad, esta aproximación no es más válida. Suponiendo errores Gaussianos, con un valor grande de datos N , el **Teorema Central del Límite** nos asegura que cualquier función de distribución \mathcal{L} (multidimensional) es aproximada por una Gaussiana en su posición de máximo (alternativamente mínimo). Tales Gaussianas (multiparamétricas) tienen asociado un vector $\vec{\mu}$ (M -dimensional con valor medio de cada parámetro estimado) y una matriz de covarianza Σ ($M \times M$ -dimensional, que contiene las varianzas de cada Gaussiana). Podemos calcular la probabilidad de obtener un vector \vec{x} como:

$$\text{prob}(\vec{x}|\vec{\mu}, \Sigma) = \frac{1}{\sqrt{(2\pi)^M \det \Sigma}} \exp \left[-\frac{1}{2} (\vec{x} - \vec{\mu})^T \cdot \Sigma^{-1} \cdot (\vec{x} - \vec{\mu}) \right] \quad (3.6)$$

Si calculamos esta probabilidad sobre cualquier proyección en un subespacio bidimensional del $\vec{\mu}$, se obtienen **contornos de error elipsoidales**, es decir podemos describir las correlaciones *lineales* en los parámetros. Sin embargo, el contorno de 1σ no identifica más el 68.3 % del nivel de confianza como se mostraba en la Fig. 3.2. Veamos cómo calcular directamente el valor de confianza de 1σ utilizando una Gaussiana multiparamétrica según la cantidad de parámetros libres:

- Cuando tenemos 1 grado de libertad:

$$\text{prob}(x|\mu_1, \Sigma) = \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \Rightarrow \int_{-\sigma_1}^{\sigma_1} \text{prob}(x|\mu_1, \Sigma) dx = 68.3\%$$

- En el caso de tener una Gaussiana bi-dimensional se tiene:

$$\begin{aligned} \text{prob}(x, y|\mu_1, \mu_2, \Sigma) &= \frac{1}{\sqrt{2\pi\sigma_1^2}} \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2} - \frac{(y-\mu_2)^2}{2\sigma_2^2}} \\ &\Rightarrow \int_{-\sigma_1}^{\sigma_1} \int_{-\sigma_2}^{\sigma_2} \text{prob}(x, y|\mu_1, \mu_2, \Sigma) dx dy = 46.6\% \end{aligned} \quad (3.7)$$

- mientras que en tres dimensiones es:

$$\int_{-\sigma_1}^{\sigma_1} \int_{-\sigma_2}^{\sigma_2} \int_{-\sigma_3}^{\sigma_3} \text{prob}(x, y, z|\mu_1, \mu_2, \mu_3, \Sigma) dx dy dz = 31.9\% \quad (3.8)$$

- Se puede generalizar para más grados de libertad y obtener una expresión más simple, recordando que la función χ^2 se comporta como una variable tipo chi-cuadrado (con valor medio ν y varianza $\sigma^2 = 2\nu$). Por el **Teorema Central del Límite**, la distribución

chi-cuadrado puede ser aproximada por una Gaussiana con la misma media y varianza, entonces:

$$\begin{aligned}\chi_{1\sigma}^2 &= \langle \chi_\nu^2 \rangle \pm \sigma \\ \chi_{1\sigma}^2 &= \nu \pm \sqrt{2\nu} \\ \chi_{1\sigma}^2/\nu &\sim 1 \pm \sqrt{2/\nu} \\ \chi_{\nu 1\sigma} &\sim 1 \pm \sqrt{\frac{1}{2\nu}}\end{aligned}\quad (3.9)$$

Debido a que el mínimo de los ajustes se encuentra en $\chi_{\nu \min}$, obtenemos:

$$\chi_{\nu 1\sigma} = \chi_{\nu \min} \left(1 \pm \sqrt{\frac{1}{2\nu}} \right) \quad (3.10)$$

Finalmente, recordamos que $(\text{wrms})^2$ tiene una dependencia lineal con χ_ν^2 , entonces el valor correspondiente al nivel de confianza de 1σ debería ser aproximadamente dado por:

$$\text{wrms}_{1\sigma} \simeq \text{wrms}_{\min} \left(1 \pm \sqrt{\frac{1}{2\nu}} \right). \quad (3.11)$$

Una vez que tenemos establecidos los límites de confianza, el siguiente problema consiste en encarar una estrategia para determinar las curvas de nivel en distintos planos. Una idea inicial sería que en nuestro esquema de minimización (Algoritmo Genético + Simplex + Annealing) se acumule información acerca de las isosuperficies de χ_ν^2 en el curso de cada generación del algoritmo genético (o cualquiera de los métodos globales exploratorios utilizados). Recordemos que mientras la población evoluciona a través de varios cientos (o miles) de generaciones (algoritmos genéticos), se obtiene un muestreo significativo (pero no homogéneo) del espacio de parámetros en el curso de la evolución. Sin embargo una de las cosas que no se debe hacer es usar la población de prueba de la última generación para establecer los límites de error, pues la población final será evidentemente distribuida alrededor de la mejor solución (en el caso que el algoritmo converja). La forma en la que los miembros de la población son distribuidos en el espacio de parámetros es muy influenciada por características del algoritmo genético, como por ejemplo la mutación. Dichos factores extraños, son claramente no relacionados a la estructura del espacio de parámetros en la vecindad de la mejor solución. Existen métodos estadísticos que son más eficientes y que veremos en las próximas secciones, pero primero resulta conveniente mencionar un ejemplo citado por Andrae (2010), que constituye un incorrecto uso de la estadística y conduce a posibles resultados erróneos:

Reescaleo de errores, σ_i Cuando se ajusta un modelo con M parámetros a N datos, el valor de χ^2 obtenido luego la minimización debería satisfacer:

$$\chi^2 = N - M \quad \Leftrightarrow \quad \chi_\nu^2 = \frac{\chi^2}{N - M} = 1, \quad (3.12)$$

donde χ_ν^2 se denomina “ χ^2 reducido”. Relajadamente hablando, esto significa que cada grado de libertad contribuye **una** desviación estándar. Muchas veces se observan casos en la literatura donde, los autores tratan de corregir los errores potencialmente malos de σ_i reescaleándolos, de tal forma de lograr que la minimización tenga como resultado $\chi_\nu^2 = 1$ (por ejemplo, aumentando el tamaño del jitter estelar).

Si bien la distribución de errores es Gaussiana en cada dato, existen objeciones a este tipo de enfoque (Andrae, 2010):

1. Se asume que el modelo es *lineal* en *todos* los parámetros M del ajuste. Sin embargo, cuando el modelo no es lineal no podemos decir que $\chi_\nu^2 = 1$, porque la desviación de $\chi^2 = N - M$ implícitamente asume linealidad en todos los parámetros.
2. Se asume que el modelo utilizado es *correcto* (subyacente de exigir que $\chi_\nu^2 = 1$). Sin embargo es una afirmación optimista, por ejemplo porque podemos estar escondiendo compañeros adicionales o errores sistemáticos en los datos.

Incluso si las suposiciones anteriores fueran ciertas, el método de reescalar errores es de hecho aplicable solo si el número de grados de libertad (ν) es grande. La razón es que la incerteza en las mediciones no solamente causa incerteza en los parámetros del modelo, sino también en el valor mismo de χ^2 . El valor de χ^2 está sujeto a la llamada distribución chi-cuadrado, cuyo valor esperado es de hecho $\nu = N - M$. Sin embargo esta distribución no posee un ancho infinitesimal, sino que tiene una varianza de $\sigma^2 = 2(N - M) = 2\nu$. Consecuentemente si $N - M$ es pequeña, existe una incerteza relativa grande en el valor de χ^2 . Esto significa que χ^2 puede desviarse sustancialmente de $N - M$ incluso si el modelo es lineal y correcto (Andrae, 2010).

3.2. Matriz de Fisher

Un método tradicional para estimar el error en un ajuste es determinarlos “localmente” mediante la matriz Hessiana (\mathbf{D}^{-1}), evaluándola en la mejor solución (Press et al. 1992, 15.5 y Andrae 2010). Recapitulemos un poco y recordemos que nuestro objetivo es estimar el mínimo de χ_ν^2 , cuyo desarrollo de Taylor hasta segundo orden alrededor de la mejor solución \mathbf{A}_0 es:

$$\chi_\nu^2(\mathbf{A}) \approx \chi_\nu^2(\mathbf{A}_0) + (\mathbf{A} - \mathbf{A}_0)^T \cdot \mathbf{d} + \frac{1}{2}(\mathbf{A} - \mathbf{A}_0)^T \cdot \mathbf{D}^{-1} \cdot (\mathbf{A} - \mathbf{A}_0), \quad (3.13)$$

donde \mathbf{d} es el gradiente. Por construcción, queremos evaluar los errores alrededor del mínimo de χ_ν^2 , donde su gradiente es nulo ($\mathbf{d}=0$), obteniendo así que el error estará dado por la aproximación \mathbf{D}^{-1} .

La matriz Hessiana, \mathbf{D}^{-1} , contiene en sus diagonales, la estimación de la varianza de cada parámetro (A_M) individual y fuera de la diagonal están las estimaciones de las covarianzas. Para entender esto volveremos a utilizar el Teorema Central del Límite y la aproximación Gaussiana (ver Barlow, 1993 y Heavens, 2010 para más detalles):

$$\mathcal{L}(\theta) \propto \exp \left[-\frac{1}{2}(\theta - \theta_0)^T \cdot \Sigma^{-1} \cdot (\theta - \theta_0) \right], \quad (3.14)$$

donde θ es una variable M -dimensional Gaussiana, con media θ_0 y matriz de covarianza Σ . Esta matriz de covarianza es la estimación de error deseada. Comparando Ecs. (3.2) y (3.14), se identifican las componentes y la matriz de segundas derivadas de $\ln \mathcal{L}$ es denominada “Matriz Fisher”:

$$\hat{\Sigma} = \left(-\frac{\partial^2 \ln \mathcal{L}}{\partial \theta_i \partial \theta_j} \right)^{-1}. \quad (3.15)$$

En el caso ideal, las derivadas segundas del ($\ln \mathcal{L}$) se pueden evaluar analíticamente. Sin embargo si es muy complicado, como ocurre en los sistemas no-lineales, se pueden hallar numéricamente. Por construcción este método sólo describe contornos de error elípticos e involucra dos suposiciones que se deben verificar:

- La distribución de errores de las medidas es conocida, i.e. la función \mathcal{L} está definida correctamente.
- La expansión a segundo orden en Taylor es una buena aproximación.

La segunda suposición es la más problemática. A pesar que el Teorema Central del Límite asegura esto asintóticamente, la Figura 3.1 es un ejemplo donde la matriz Fisher falla. Existen dos pruebas simples para verificar si la matriz covariante resultante puede describir correctamente los errores (Heavens, 2010):

- Calcular el determinante $\det \hat{\Sigma}$. Si $\det \hat{\Sigma} < 0$, entonces $\hat{\Sigma}$ no describe correctamente los errores.
- Diagonalizar la matriz $\hat{\Sigma}$ para determinar sus autovalores. Si algún autovalor es negativo o cero, nuevamente $\hat{\Sigma}$ no describe correctamente los errores.

Sin embargo los errores que se obtienen con esta información local de la curvatura, no son particularmente útiles cuando trabajamos en problemas de optimización global. Lo que necesitamos es información acerca de la forma y extensión de la región donde la minimización es $\chi^2_\nu < \chi^2_\nu(\text{cota})$ (que es definida por un rango de parámetros). Esta región puede ser identificada con métodos MonteCarlo, perturbando la mejor solución y calculando el residuo (χ^2_ν) de esas soluciones perturbadas. Esto es indudablemente el camino más confiable para tener estimación de los errores. Si se halló la mejor solución en espacios de parámetros relativamente pequeños, esta tarea es relativamente simple, como construir un hipercubo centrado en ella y calcular $\chi^2_\nu(\mathbf{A})$ en lugares previamente seleccionados con la resolución espacial requerida. De esta forma se pueden construir isocontornos de χ^2_ν en las regiones de interés. La estimación del error es directa, aunque consume bastante tiempo de cómputo. Ciertamente funciona si los costos computacionales asociados con el modelo son lo suficientemente bajos para permitir el cálculo de varias soluciones adicionales.

Veremos a continuación algunos métodos para definir regiones de validez de los parámetros.

3.3. Cadenas Markov con MonteCarlo

Los métodos MonteCarlo directamente toman muestras a partir de la función de distribución *likelihood* (\mathcal{L}), es decir, suponen valores de los parámetros del ajuste y los aceptan con la probabilidad definida por el valor correspondiente de la función \mathcal{L} . A pesar que parezcan complicados, en primer aproximación el muestreo MonteCarlo es la aproximación más intuitiva para la estimación de errores. La razón es su similitud con medir errores en los datos, repitiendo el proceso de medición y monitoreando los resultados. La fortaleza de los métodos MonteCarlo es que usan la mínima cantidad de suposiciones. La única suposición es que la distribución de los errores en los datos se conoce correctamente, sin requerir específicamente que sean Gaussianos, lo que hace general su aplicación.

Hay distintos tipos de métodos MonteCarlo dependiendo, por ejemplo, del número de parámetros del modelo \mathcal{M} . Cuando \mathcal{M} es bajo (< 3) se usan: **muestreo uniforme**, **muestreo de importancia** y **muestreo de eliminación** (ver MacKay, 2003 para una introducción a los mismos). Sin embargo estos métodos se hacen ineficientes o computacionalmente intensivos para modelos con muchos parámetros, donde se utiliza **Cadenas de Markov con MonteCarlo** (MCMC).

Los métodos MCMC son una clase de algoritmos que muestrean distribuciones a partir de una probabilidad, basándose en construir una cadena de iteraciones (denominada cadena Markov) que tiene una distribución de equilibrio deseada. El análisis de esta distribución de equilibrio permite estimar

las cantidades de interés como la media, varianza, correlaciones, etc. El estado de la cadena de iteraciones después de un largo número de pasos es usado para tomar muestras de la distribución deseada y la calidad de la muestra mejora en función del número de pasos tomados. Es sencillo construir un algoritmo MCMC con las propiedades deseadas, el problema reside en determinar cuantos pasos son necesarios para la convergencia a la distribución estacionaria dentro de un error aceptable.

El muestreo MCMC solo puede llegar a ser próximo a la distribución deseada, debido a que siempre queda un efecto residual dependiendo de la posición inicial. Los métodos MCMC se mueven alrededor de la posición de la distribución de equilibrio en pasos relativamente pequeños, sin tendencia a ir en una dada dirección. El problema es que demandan mucho tiempo de cómputo debido al aspecto MonteCarlo que involucra técnicas de *Camino Aleatorio* (Random Walk). Existen 2 tipos de camino aleatorios utilizados en MCMC:

- Algoritmo Metropolis-Hastings: genera el camino aleatorio utilizando una densidad de probabilidad², con el objetivo de descartar algunos movimientos en el espacio de parámetros.
- Muestreo Gibbs: es requerido para que la distribución condicional de la distribución deseada sea muestreada exactamente. Consideremos el vector A cuyas componentes son los M parámetros del modelo en nuestro ajuste. El Muestreo Gibbs genera un estado de prueba A' alterando un subconjunto de parámetros de A en cada paso. Luego se combina el muestreo con una función de transición de candidatos que es Gaussiana. La función de probabilidad definida bajo estas condiciones es:

$$q(A'_\mu | A_\mu) = \frac{1}{\sqrt{2\pi\beta_\mu^2}} \exp \left[-\frac{(A'_\mu - A_\mu)^2}{2\beta_\mu^2} \right] \quad (3.16)$$

donde utilizamos el índice μ para distinguir los elementos del vector de parámetros, y se utilizará el índice n para indicar el número de paso de la cadena Markov. Cada β_μ es un parámetro que controla el tamaño de los pasos para el parámetro μ .

La probabilidad de transición entre dos configuraciones subsecuentes A_n y A_{n+1} es dada por la expresión:

$$\begin{aligned} \alpha(A_{n+1} | A_n) &= \min \left(\exp \left[\frac{\chi_n^2 - \chi_{n+1}^2}{2} \right], 1 \right) && \text{para valores de parámetros válidos} \\ \alpha(A_{n+1} | A_n) &= 0 && \text{para valores de parámetros inválidos.} \end{aligned} \quad (3.17)$$

donde χ_{n+1}^2 es el valor del residuo calculado con el vector A_{n+1} y χ_n^2 es el residuo calculado con el vector A_n . Asumiendo que los errores observacionales son exactamente conocidos y aproximadamente gaussianos, esta transición entre estados asegura que (después de descartar una porción inicial de la cadena) la distribución de las configuraciones generadas va a muestrear la distribución desconocida de errores (ver por ejemplo Ford (2005) y Gregory (2005)).

Una vez que se obtiene convergencia de la cadena (arbitrariamente determinada), se pueden representar los valores de χ_ν^2 en superficies de nivel. Estas superficies muestran la distribución real de errores en los parámetros seleccionados.

Esquemáticamente el algoritmo MCMC con su base Metropolis-Hasting es definido por los siguientes pasos:

²la elección natural para esta densidad de probabilidad se basa en una distribución uniforme.

1. Inicializar la cadena en algún A_0 en $n = 0$. Si iniciamos el algoritmo con las componentes del vector A en el **mejor ajuste**, se tiene la ventaja de no necesitar un diagnóstico de convergencia (Andrae, 2010).
2. Generar un estado de prueba, A' , de acuerdo a $q(A'|A_n)$.
3. Calcular el residuo del estado de prueba $\chi^2(A')$ y el residuo $\chi^2(A_n)$ para el estado actual.
4. Determinar $\alpha(A'|A_n)$
5. Escoger un número aleatorio, u , de una distribución uniforme entre 0 y 1.
6. Si $u \leq \alpha(A'|A_n)$, como se define en la Ec. (3.17), entonces el nuevo valor del vector es: $A_{n+1} = A'$.
Si $u > \alpha(A'|A_n)$, entonces el nuevo valor es igual que el anterior (es decir no actualizamos el vector): $A_{n+1} = A_n$.
7. Avanzar un paso en la iteración, $n = n + 1$.
8. ir al paso #2.

El monitoreo de la fracción de estados de prueba que son aceptados es una manera de verificar que la escala elegida para $q(A'|A)$ no sea demasiado ineficiente. Los valores óptimos para la tasa de aceptación pueden ser determinados para los casos más simples. Por ejemplo cuando A tiene una sola dimensión, la tasa de aceptación de ~ 0.44 es óptima. Para parámetros de espacios con varias dimensiones una tasa de aceptación óptima es de ~ 0.25 (Gelman et al. 2003).

Si después de varios pasos de la cadena de Markov se calcula que la tasa de aceptación es distinta que la mencionada arriba, entonces los parámetros de escala deben ser reajustados y reinicializarla. Puede ser necesario repetir este procedimiento varias veces para determinar un conjunto aceptable de parámetros de escala.

Este método fue aplicado en el sistema HD82943. En la Figura 3.4 se muestran las superficies de nivel para los elementos orbitales ($e \cos(\varpi)$, $e \sin(\varpi)$) del planeta interno construidos a partir de 100 cadenas MCMC de 10^5 iteraciones. El mínimo que marcamos en el segundo cuadrante es el que corresponde a la mejor solución del sistema (inestable según muestran las integraciones dinámicas), mientras que marcamos también con un punto rojo un segundo mínimo que aparece en el tercer cuadrante. Esta cadena fue insuficiente para muestrear superficies de nivel en los elementos del segundo planeta.

Fue necesario construir 10 cadenas MCMC, de 10^7 iteraciones, para cubrir un área considerable alrededor de la mejor solución orbital del segundo planeta y mostramos los resultados de interés en la Figura 3.5. En negro se dibujaron las superficies de nivel para el planeta interno y en azul para el externo. Los puntos rojos indican la posición de los mínimos del planeta interno (1) y en azul los del planeta externo (2). Las superficies de nivel permiten estudiar el comportamiento alrededor de la solución mínima, es decir $\varpi_1 \sim 120^\circ$ y $\varpi_2 \sim 280^\circ$. Es interesante ver como se extienden las superficies de nivel del planeta externo: existe un área que conecta soluciones entre el segundo y tercer cuadrante. Si la solución fuera en el segundo cuadrante, ambos planetas se encontrarían alineados en $\varpi_1 \sim \varpi_2 \sim 120^\circ$. Las superficies de nivel en los otros planos no son expuestas pues se encontraban muestreadas pobremente alrededor del mínimo y carecían de interés.

Entre las desventajas del método se encuentra la demanda de tiempo extremadamente larga (para este sistema una cadena Markov de 10^7 pasos tomó 4 días de cómputo), siendo que los resultados

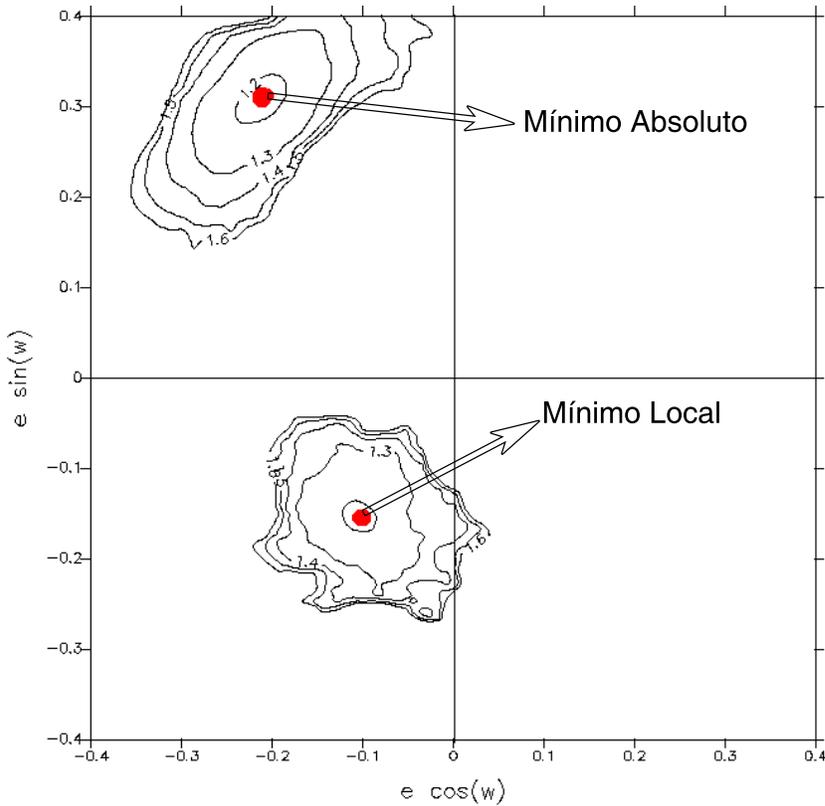


Figura 3.4: Isosuperficies de χ_ν construidas a partir de cadenas MCMC sobre el conjunto total de datos de HD82943, sobre el plano $e \cos(\varpi), e \sin(\varpi)$ con los elementos orbitales del planeta interno. Los datos fueron obtenidos con 100 cadenas MCMC de 10^5 pasos cada una. Se observan dos mínimos: el que se ubica en el tercer cuadrante es un mínimo absoluto y el que está en el cuarto cuadrante es un mínimo local.

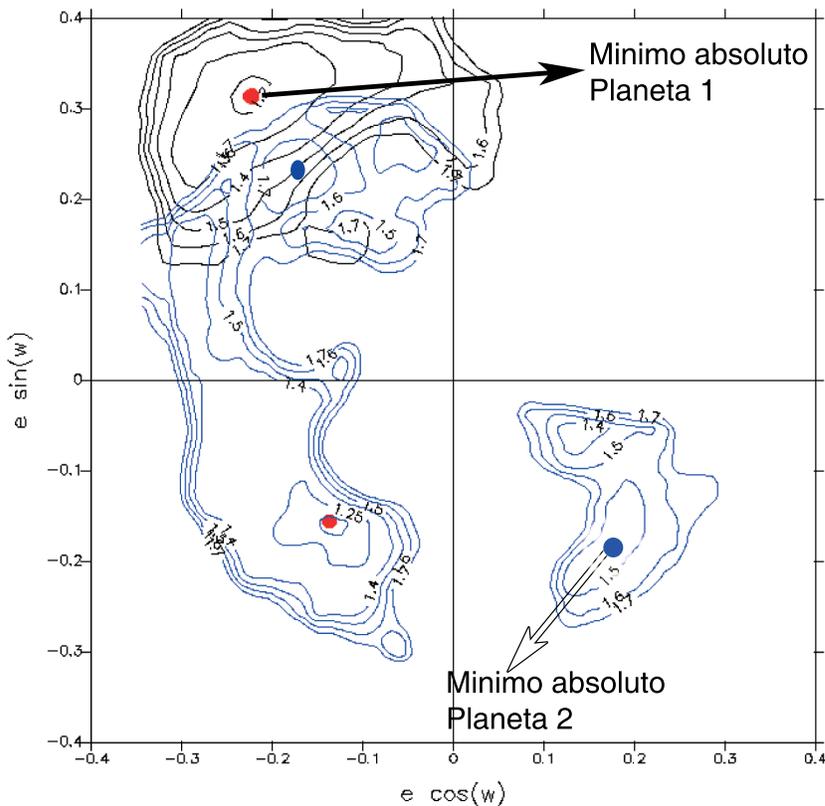


Figura 3.5: Isosuperficies de χ_ν construidas a partir de cadenas MCMC en HD82943 sobre el plano $e \cos(\varpi), e \sin(\varpi)$. Las superficies de nivel del planeta interno se muestran en negro y para el externo en azul. Los datos fueron obtenidos con 10 cadenas de 10^7 pasos cada una. Se observan tres mínimos para el planeta exterior: el que se ubica en el cuarto cuadrante es un mínimo absoluto (configuración inestable), mientras que el que está en el segundo cuadrante es un mínimo local (configuración estable).

no siempre son los deseados (no se mapean homogéneamente las regiones alrededor de todos los parámetros planetarios). Como mencionamos anteriormente, el tamaño del paso β_μ y la probabilidad de salto entre estados posibles q son parámetros ajustables. El trabajo artesanal requerido a fin de encontrar el estado de equilibrio (que mapea los errores que buscamos) no se justifica y existen métodos alternativos más eficaces como veremos a continuación.

3.4. Generación de Datos Sintéticos: Resampling

Un método alternativo para estimar incertezas en los parámetros orbitales es aplicar la técnica de minimización de residuos a varios conjuntos de datos simulados (Brown et al. 2001 y Ford, 2005). Debido a que los errores observacionales son muy cercanos a errores Gaussianos y cada velocidad observada ($V_r(i)$) tiene su correspondiente error (σ_i) asociado, es directo construir conjuntos simulados añadiendo errores Gaussianos a los datos nominales:

$$V_r^G(i) = V_r(i) + N(0, \sigma) \quad (3.18)$$

donde $N(0, \sigma)$ es una distribución Gaussiana de media cero y dispersión σ .

Cada conjunto de velocidades radiales simuladas $V_r^G(i)$ (también llamado conjunto sintético) representa un posible conjunto de valores reales en las medidas. Si se aplica el mismo procedimiento de ajuste orbital a cada conjunto sintético, entonces se obtiene una distribución de los mejores parámetros orbitales. Este método fue aplicado anteriormente a sistemas como HD72659 (Ford, E. 2005 y Brown, R. 2004).

Aplicamos esta técnica sobre el sistema HD82943, construyendo un total de 1000 conjuntos de datos sintéticos. Los resultados se muestran en la Figura 3.6, donde cada solución al conjunto sintético es representada por un círculo. Cada marco representa la distribución entre pares de elementos planetarios que consideramos más importantes. En el plano $(e_2 \cos(\varpi_2), e_2 \sin(\varpi_2))$ (marco inferior derecho) se aprecian claramente tres regiones de solución disconexas, cada una de ellas identificada con un color diferente para facilitar su correspondencia en los otros planos:

- Región (I) - Negro. Es la región alrededor de la mejor solución de los datos originales, con $e_2 \sim 0.2$ y $\varpi_2 \sim 280^\circ$
- Región (II) - Azul. Región con $e_2 \sim 0$ y por lo tanto ϖ_2 pobremente definido.
- Región (III) - Rojo. Región con $e_2 \sim 0.4$ y $\varpi_2 \sim 120^\circ$ que define una solución dinámicamente estable como veremos más adelante.

En el marco superior izquierdo se muestra el plano de las amplitudes de velocidad radial (K_1, K_2) y se observa que a pesar que K_2 es casi el mismo para todos los conjuntos, el valor de K_1 puede variar en un 50%; K_1 para la región (I) vale ~ 65 m/s, mientras que en la región (III) vale ~ 47 m/s. En la misma figura, marco superior derecho se observa el plano de períodos (P_1, P_2), donde el período del planeta P_1 está muy bien definido ~ 219 días, mientras que P_2 varía unos 10 días dependiendo la región donde se encuentre, desde 430 a 440 días. En el plano $(e_1 \cos \omega_1, e_1 \sin \omega_1)$ mostrado abajo a la izquierda no se aprecian grandes diferencias entre las regiones (I), (II) y (III) y parece que están muy bien definidos (el método MCMC permite mapear un poco mejor este espacio de parámetros).

Los grupos en el plano $(e_2 \cos \omega_2, e_2 \sin \omega_2)$ son muy distintos, y corresponden aproximadamente a los mismos favorecidos con el método que veremos más adelante en *Jackknife*. La distribución de

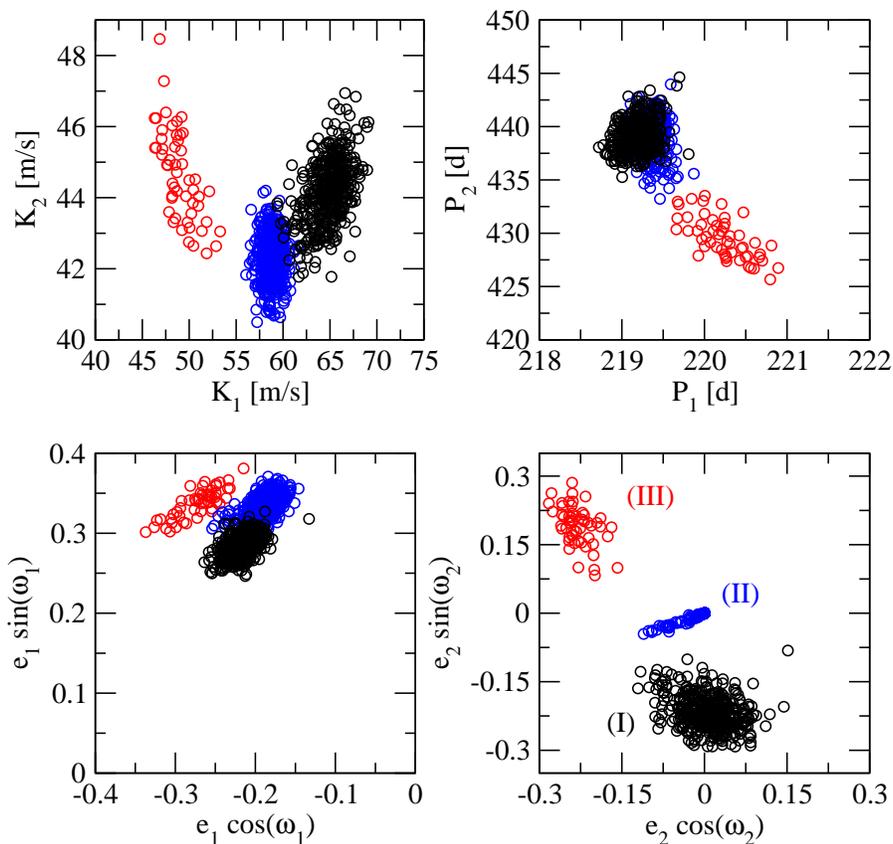


Figura 3.6: Resultados de 1000 resamplings de los datos originales. Cada círculo representa el mejor ajuste del correspondiente conjunto de datos. Cabe destacar la presencia de 3 regiones distintas, cada una identificada con distinto color.

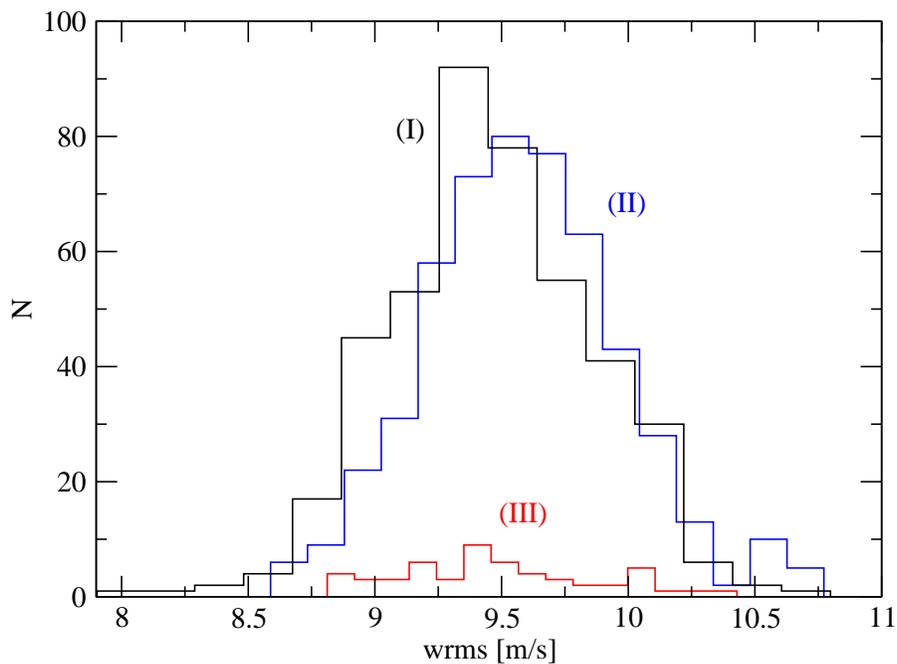


Figura 3.7: Histogramas de los mejores ajustes encontrados (escritos en función del wrms) que muestran la distribución de los miembros de cada grupo identificado en la Fig. 3.6 con la misma identificación de colores.

los miembros de cada grupo de acuerdo al valor de $wrms$ se muestra en la Figura 3.7. A pesar que los tres histogramas tienen aproximadamente la misma media y dispersión, hay varias diferencias que podemos enumerar:

$$\begin{aligned}
 \text{(I)} \quad & \text{miembros} = 42.8\% \quad wrms_{min} = 7.902 \text{ m/s} \\
 \text{(II)} \quad & \text{miembros} = 52.0\% \quad wrms_{min} = 8.588 \text{ m/s} \\
 \text{(III)} \quad & \text{miembros} = 5.3\% \quad wrms_{min} = 8.812 \text{ m/s}
 \end{aligned} \tag{3.19}$$

Así las regiones (I) y (II) tienen casi el 95 por ciento de todas las soluciones encontradas con Resampling, mientras que la región (III) sólo un 5%. Esta última región es caracterizada también por valores de $wrms$ más grandes y tiene un valor de señal más pequeña (K_1).

El mínimo global se encuentra en la región (I) que como se mencionara es inestable; por otro lado la región (II) tiene valores de $wrms$ similares y como veremos más adelante está dentro de la región de confianza de la mejor solución (ver 3.5). La región (III) es la que proporciona valores muy cercanos entre $\varpi_1 \sim \varpi_2$, en otras palabras, hacia una *corotación apsidal* del tipo (0,0) (ver Beaugé et al. 2006).

Una de las desventajas en esta técnica de reajustes es que se deben identificar los mejores parámetros orbitales para cada conjunto sintético de velocidades radiales, que en la práctica requiere muchos recursos computacionales. Por ejemplo para un sistema de un planeta (seis parámetros), es necesario generar 10^3 a 10^4 conjuntos sintéticos para tener niveles de confianza dentro de los 3σ . Asumiendo que la rutina de minimización requiere unas 10^4 evaluaciones para converger, esto hace un total de 10^8 evaluaciones de χ^2_ν . Incluso el uso de minimización local, buscando soluciones cercanas al mínimo global, demandaría al menos unas 10^6 evaluaciones del residuo χ^2_ν .

Finalmente no olvidemos que la distribución de errores estimada por medio del *Resampling* no sirve para construir superficies de nivel de error constante, sino para identificar posibles regiones de mínimos.

3.4.1. Análisis dinámicos sobre las regiones

En cada una de las regiones mencionadas previamente se identificó el ajuste orbital con menor $wrms$, teniendo en cuenta que esa configuración resulta compatible con los datos originales de V_r . La región (I) es representada por la solución mostrada en la Tabla 2.2. Para las regiones (II) y (III) se realizaron ajustes dinámicos coplanares a partir de la mejor solución multikepleriana y los resultados se muestran en la Tabla 3.1.

Ambas soluciones dinámicas en la Tabla 3.1 muestran valores similares de $wrms$, del orden de 8.03 m/s, apenas mayor que el mínimo global (i.e. 7.92 m/s) pero no muy diferente. Ambas fueron integradas numéricamente con un código de N-cuerpos por 10^6 años, mostrando estabilidad dinámica y corotación apsidal (ACR) tipo (0,0) (Figura 3.8). Sin embargo, la amplitud de libración es grande, particularmente en el caso de la Región (II). Para la Región (III), la semi-amplitud en $\sigma_1 = 2\lambda_2 - \lambda_1 - \varpi_1$ es cercana a 50° , mientras que en $\Delta\varpi = \varpi_2 - \varpi_1$ es cercana a 90° .

No se muestra la evolución de la Región (I), debido a que como es bien conocido es inestable en cortos períodos de tiempo. A pesar que pueda ser posible encontrar algunos miembros marginales en la región que sean estables, no es posible decir que son representativos de la región general, ni siquiera robustos (es decir, pequeños cambios en los elementos orbitales pueden desestabilizar el movimiento).

El comportamiento extraño de la solución de la Región (II) en $(k, h) = (e_1 \cos \sigma_1, e_1 \sin \sigma_1)$ puede ser explicado y deducido con la Figura 3.9, donde se ha dibujado la evolución en el plano

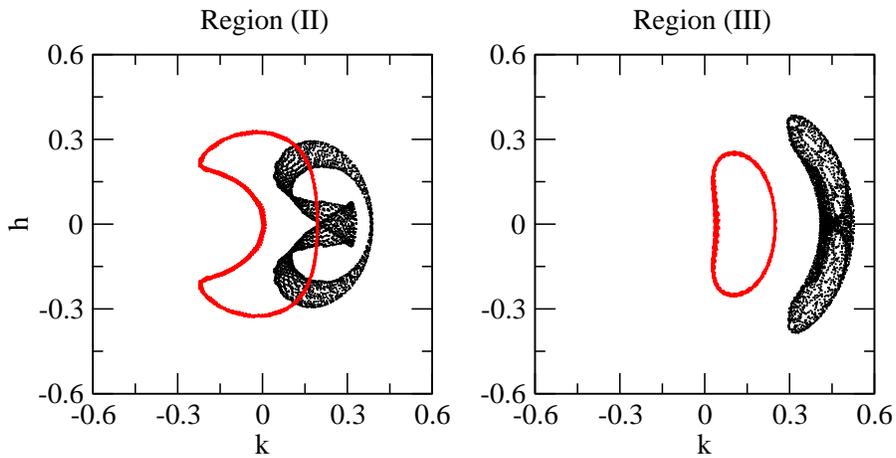


Figura 3.8: Evolución de los mejores ajustes orbitales representativos de las soluciones en la Región (II) (izquierda) y la Región (III) (derecha). Los puntos negros corresponden a las variables $(k, h) = (e_1 \cos \sigma_1, e_1 \sin \sigma_1)$, y los puntos rojos a las variables $(k, h) = (e_2 \cos \Delta\varpi, e_1 \sin \Delta\varpi)$. Aquí σ_1 es el ángulo resonante $\sigma_1 = 2\lambda_2 - \lambda_1 - \varpi_1$. En ambos casos se ven ACR de amplitud de oscilación grande tipo del $(0, 0)$.

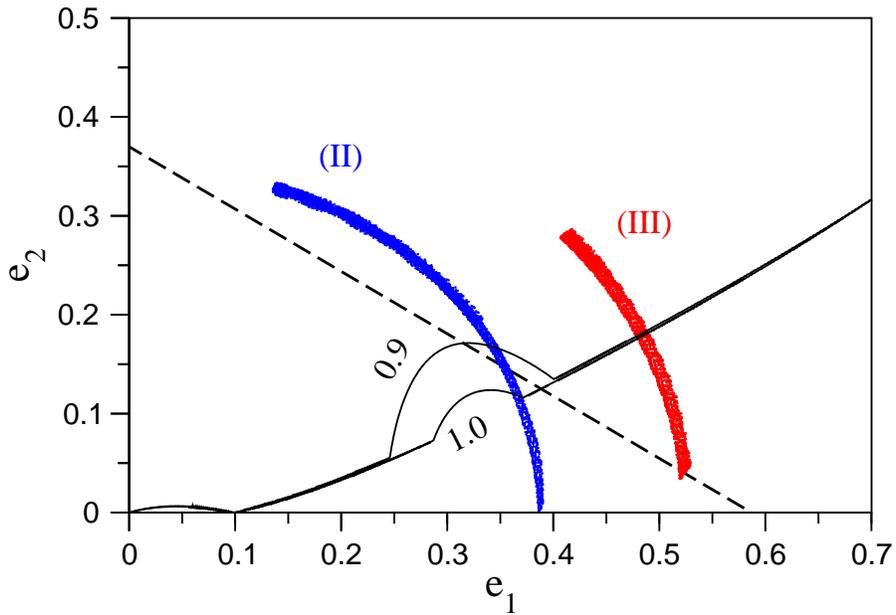


Figura 3.9: Evolución de ambas soluciones en el plano (e_1, e_2) , junto con la familia de ACR estable del tipo $(0, 0)$ para cada una de las razones de masa. La línea punteada muestra la curva de colisión.

Tabla 3.1: Mejores ajustes orbitales para soluciones representativas de las regiones (II) y (III) para un modelo de 2 planetas del sistema HD82943.

Parámetro	HD82943c	HD82943b
Ajuste dinámico coplanar - Región (II)		
$m [M_{Jup}]$	1.7681	1.7396
P [d]	219.890	442.879
a [UA]	0.7473	1.1919
e	0.3863	0.0166
ω [°]	120.915	201.379
M [°]	356.842	281.828
V_{01} [m/s]	43.497	
V_{02} [m/s]	33.754	
wrms [m/s]	8.037	
χ_ν	1.529	
Ajuste dinámico coplanar - Región (III)		
$m [M_{Jup}]$	1.4605	1.7442
P [d]	220.081	439.375
a [UA]	0.7477	1.1856
e	0.4369	0.2548
ω [°]	126.211	137.232
M [°]	358.181	337.541
V_{01} [m/s]	42.427	
V_{02} [m/s]	32.305	
wrms [m/s]	8.031	
χ_ν	1.528	

(e_1, e_2) juntamente con las familias ACR de distintas razones de masa. Para la razón de masa de esta solución ($m_2/m_1 \sim 1$), deberíamos tener ACR asimétrica. Sin embargo lo que vemos es una amplitud de libración tan grande que la órbita abarca ambos lóbulos asimétricos y rodea la solución simétrica cuyo centro es inestable.

3.5. Ajustes en una Grilla

Hasta el momento hemos visto que tanto *Resampling* como cadenas *MCMC* brindan mucha información acerca de cuales son los parámetros compatibles con el conjunto de datos. Sin embargo el tiempo computacional es elevado y finalmente no es sencillo construir curvas de nivel sobre alguna proyección de plano que elijamos. Para solucionar este último punto se puede seguir la idea de Lee et al. (2006), realizando *Ajustes en una Grilla*.

La idea es realizar una serie de ajustes multikeplerianos (o dinámicos) sobre una grilla de variables de interés (A_1, A_2). Para cada punto de la grilla los valores numéricos de estos parámetros son fijados, y se realiza un ajuste sobre los parámetros restantes (que pueden variar sin restricciones). El valor resultante de residuo R corresponde al mejor ajuste para los valores prefijados de (A_1, A_2). Luego se dibujan las curvas de nivel de *wrms* en esta grilla, que da información de la forma y la profundidad relativa de la función residuo en el plano seleccionado.

Para interpretar mejor los resultados, podemos intentar estimar los niveles de confianza de 1σ del

Tabla 3.2: Principales parámetros de los ajustes orbitales para $N = 164$ observaciones. Todos los elementos orbitales son astrocéntricos y los períodos osculadores.

Parameter	Multi-Kepleriano		Dinámico	
	HD82943c	HD82943b	HD82943c	HD82943b
$m [M_{Jup}]$	1.435	1.765	1.397	1.782
$a [UA]$	0.748	1.169	0.744	1.188
e	0.447	0.331	0.460	0.360
$\omega [^\circ]$	133.19	144.25	134.79	139.82
$\lambda - \varpi [^\circ]$	359.57	326.21	357.45	331.45
wrms [m/s]	7.835		7.838	
$\sqrt{\chi^2}$	1.348		1.504	

mejor ajuste en las superficies dentro de la grilla. Para valores grandes de N , la función normalizada chi-cuadrado χ_ν^2 se comporta como una variable tipo χ^2 , y debido a que $(wrms)^2$ tiene una dependencia lineal con χ_ν^2 , entonces el valor correspondiente al nivel de confianza de 1σ es dado por (ver Sección 3.1):

$$wrms_{1\sigma} \simeq wrms_{\min} \left(1 \pm \sqrt{\frac{1}{2\nu}} \right). \quad (3.20)$$

Aplicamos este método al sistema HD82943 y describiremos algunos resultados a continuación. En primer lugar, el grado de inestabilidad estadística (analizado en la Sección *Jackknife*) es tal que la eliminación de sólo un dato observacional hace que obtengamos distintos parámetros orbitales (ver Figura 3.11 y 3.12). La Tabla 3.2 muestra los mejores ajustes multikeplerianos y dinámicos para $N = 164$ (el último dato eliminado) que puede ser comparado con los obtenidos para $N = 165$ mostrados en la Tablas 2.1 y 2.2. A pesar que existen cambios importantes en la solución, ellos prácticamente no causan efectos en el residuo ($wrms$ o $\sqrt{\chi_\nu^2}$), ni siquiera en la curva de velocidad radial, indicando que ambas soluciones son compatibles con los datos.

Para analizar más globalmente esta dependencia con el número de puntos, estudiamos el plano $(k_2, h_2) = (e_2 \cos \omega_2, e_2 \sin \omega_2)$. El procedimiento fue llevado a cabo con 4 conjuntos de observaciones correspondientes a $N = 100$, $N = 142$, $N = 164$ y $N = 165$. Los resultados se muestran en la Figura 3.10 solamente en el plano de mayor interés (k_2, h_2) . En cada caso el mínimo global $wrms_{\min}$ es identificado con un círculo negro relleno. A pesar que cerca del mínimo las curvas de nivel son casi elípticas, aparecen distorsiones para valores más grandes, indicando que los niveles de confianza obtenidos del análisis de la matriz de correlación no son válidos más allá de la inmediata vecindad del $wrms_{\min}$ (ver Sección 3.2).

Es interesante notar los cambios en las curvas de nivel en función de N . Este cambio sin embargo no es gradual: para $N = 164$ la figura muestra dos mínimos globales localizados en $\omega_2 \sim 140^\circ$, mientras la otra solución está en $(\omega_2 \sim 270^\circ)$ asemejándose al mínimo global de la solución obtenida para $N = 165$. De esta forma no solamente la función residual es muy poco profunda alrededor de la mejor solución, sino que a veces se pueden observar dos mínimos locales. Los pequeños cambios en el conjunto de datos pueden favorecer a uno u otro, identificándolos como el mejor ajuste. Este resultado muestra concordancia con los obtenidos en la Sección 3.4.

Para el caso del mínimo con 165 puntos, el $wrms_{\min} = 7.952$ y $\nu = 153$, obteniendo $wrms_{1\sigma} \simeq$

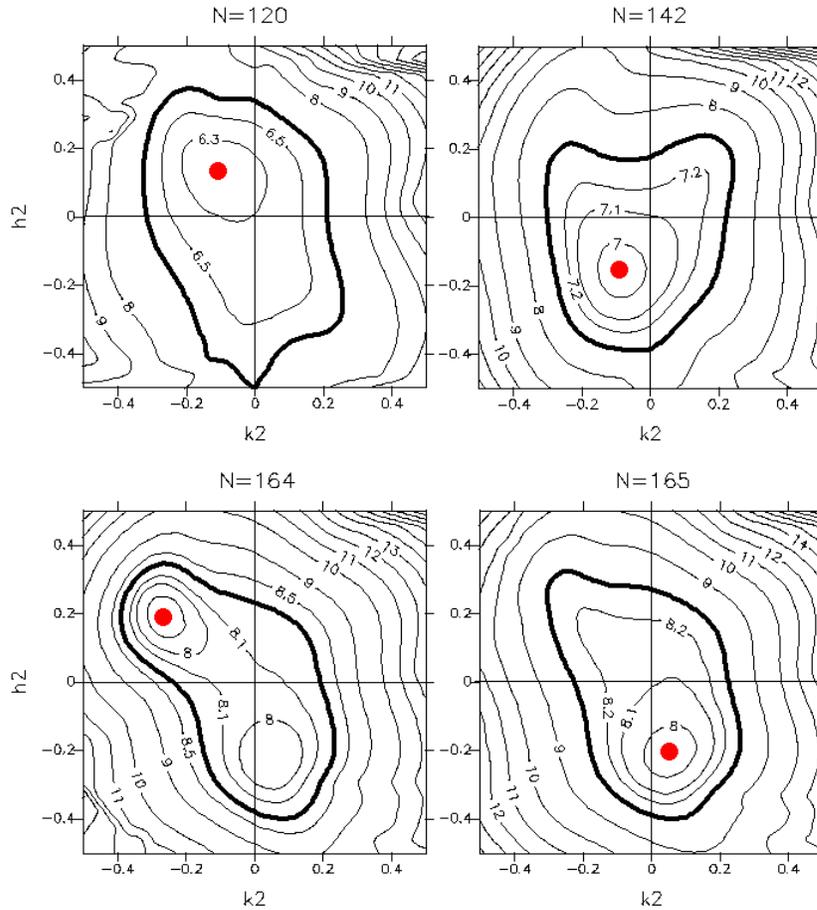


Figura 3.10: Curvas de nivel de $wrms$ constante, para ajustes multikeplerianos considerando valores fijos de $(k_2, h_2) = (e_2 \cos \omega_2, e_2 \sin \omega_2)$ en una grilla regular. Cada gráfico fue realizado usando un número distinto de datos N , simulando la información disponible en el momento de la publicación de cada solución orbital. El mínimo global es identificado con un círculo grande rojo. Las curvas de nivel más oscuras indican el nivel de confianza de 1σ .

$7.952(1 \pm 0.057)$, lo que se corresponde con:

$$wrms_{1\sigma} \simeq 8.41. \quad (3.21)$$

Es interesante remarcar que para $N=142$ la solución es la misma que ya habíamos observado como mínimo secundario con el método de MCMC. Inclusive, eliminando gran cantidad de datos, con $N=120$, se observa que la región de confianza de 1σ engloba los cuatro cuadrantes: tenemos acotada la excentricidad máxima del planeta exterior, aunque su argumento de pericentro no es bien establecido.

En conclusión, los *Ajustes en una Grilla* son un método útil para estimar los niveles de confianza y correlaciones en un plano preseleccionado. Recordemos que de esta forma muestreamos homogéneamente el espacio de interés y es más fácil de construir las superficies de nivel. Sin embargo, esta herramienta es un complemento para *Resampling* o *MCMC*, pues no tenemos forma de saber *a priori* si los valores de los parámetros que quedaron libres para el ajuste son similares en todo el plano.

3.6. Ajustes Orbitales con conjuntos de datos reducidos: Jackknife

En la Sección 2.4 vimos que el sistema HD82943 cambia sus soluciones a medida que se van agregando más observaciones. El primer ajuste orbital publicado para ambos planetas tenía sólo unos 100 puntos (ver Ji et al. 2003), dando soluciones estables consistentes con una pequeña oscilación de los ángulos resonantes. Las observaciones subsiguientes llevaron a cambios significativos en las masas

y la mayoría de los parámetros orbitales, así como también a soluciones dinámicamente inestables. Este hecho también se observa en otros sistemas exoplanetarios. Es por ello que empleamos una modificación de un análisis usualmente llamado *Jackknife*, con el objetivo de evaluar las variaciones de los resultados de los ajustes orbitales con el número de observaciones.

El método consiste en realizar una serie de ajustes orbitales, cada uno usando un número reducido de datos donde se elimina la última observación por vez. En otras palabras, se realiza un ajuste con N puntos, luego un ajuste con $N-1$ puntos (eliminando el último punto cronológicamente hablando), etc. Si la solución actual es robusta, entonces deberíamos esperar solo pequeños y suaves cambios en los parámetros en función de N . Más importante aún es que podemos esperar que la incorporación de observaciones futuras no cambiarán significativamente nuestro conocimiento del sistema.

De la Ec. (2.72), si asumimos valores constantes de σ_i^2 y $(y_i - V_{r,i})$ para todos los datos, puede deducirse que :

$$wrms \propto \sqrt{\frac{N}{N-1}} \quad (3.22)$$

$$\sqrt{\chi_v^2} \propto 1 + M/N \quad (3.23)$$

Entonces cuando el número de observaciones N es mucho mayor que el número de parámetros libres M , ambos indicadores deberían ser prácticamente constantes, siempre que la solución orbital sea adecuada y el sistema satisfaga las condiciones estadísticas implícitas en el cálculo de cuadrados mínimos. Si se observa un incremento en el valor de $wrms$ con el número de datos, podría indicar que el modelo de n -planetas se convierte crecientemente inadecuado para representar las observaciones.

Los resultados de los ajustes multikeplerianos son mostrados en la Figura 3.11 donde los valores de V_r fueron ordenados cronológicamente, es decir mezclando los datos CORALIE y Keck. Mostramos en cada cuadro los parámetros de interés (aunque todos fueron analizados): masas (m_i), períodos (P_i), excentricidad (e_i), argumento de pericentro (ω_i), residuos ($\sqrt{\chi_v^2}$ y $wrms$) y número de datos (N). Se reserva el subíndice 1 para el planeta interno (en negro) y 2 para el externo (en rojo).

Estos gráficos brindan bastante información. Primero y más importante, no parece haber convergencia de los parámetros hacia valores específicos. Desde este punto de vista, las nuevas observaciones pueden cambiar nuevamente las soluciones para ambos planetas y por lo tanto el ajuste de HD82943 no es confiable.

A pesar de ello, hay parámetros que cambian menos que otros. Los períodos orbitales P_i , parecen ser bien determinados y por lo tanto, prácticamente no existe duda que se encuentran en la vecindad de la MMR 2/1. Incluso pocos datos son suficientes para establecer que existe conmensurabilidad entre ambos planetas.

Las masas son parámetros más sensibles, a pesar que la masa del planeta m_2 en principio es más estable por un tiempo más largo. La masa interna m_1 salta entre distintos valores especialmente cuando $N > 100$. Finalmente los parámetros menos robustos son la excentricidad y la longitud de pericentro. No es un hecho sorprendente, ya que como mencionamos son determinados a partir de las asimetrías en las curvas, aunque los cambios son significativos. Como ocurre con m_1 , existe una tendencia y los valores se alternan repetidamente. Por ejemplo para los últimos 60 datos el valor de excentricidad e_2 es cercano a 0.18 o 0.30. Los valores más comunes para ω_2 son ~ 120 o ~ 270 grados. Más aún las variaciones de e_1 y e_2 parecen estar relacionadas manteniendo una diferencia casi constante. La variación de e_1 también parece estar opuesta a la de m_1 .

Es interesante ver como los valores de los residuos evolucionan con el número de observaciones, mostrado abajo a la izquierda ($wrms$ y $\sqrt{\chi^2_{\nu}}$). Particularmente en las últimas decenas de observaciones se incrementan los residuos: esto significa que más observaciones no necesariamente implican mejores ajustes orbitales. El marco inferior derecho muestra el número de puntos en función del intervalo de observación. Con este gráfico podemos identificar en cada uno de los otros marcos los valores de los parámetros en cada fecha de publicación.

Dos conclusiones adicionales pueden extraerse de esta figura. Primero, que los ajustes orbitales con menos que $N \sim 100$ puntos son poco confiables. Esto ayuda a explicar porque la diferencia entre las soluciones de Ji et al. (2003) y Mayor (2004). Segundo, existe una notoria diferencia en los ajustes con respecto al último punto observacional. El pasaje de $N=164$ a $N=165$ puntos implica cambios drásticos en algunos elementos, particularmente excentricidad y longitud de pericentro. Este cambio no es acusado en los valores mínimos de $wrms$. Esto implica que ambas soluciones aunque diferentes son consistentes con los datos observacionales. Una vez más esto nos ayuda a concluir que el sistema no es robusto y sólo un dato observacional puede cambiar nuestras soluciones. El incremento en $wrms$ como función de N no era esperado (ver Ecuación 3.22).

La Figura 3.12 presenta un análisis similar, esta vez sólo considerando los $N=142$ datos CORALIE. Debido a que el intervalo de observación de los datos Keck intersectan parcialmente los puntos CORALIE, esta figura ayuda a mostrar la evolución de los parámetros desde 2000 a 2004. Comparando con la Fig 3.11 encontramos una evidente disminución en la dispersión de las soluciones. En este caso los parámetros planetarios, especialmente las excentricidades, muestran una tendencia más regular. Se observa que e_2 es más estable y e_1 mostrando un suave descenso en su valor. Se observa también que no hay un salto de e_2 alrededor del instante $T=52500$. Las longitudes de pericentro ϖ ,

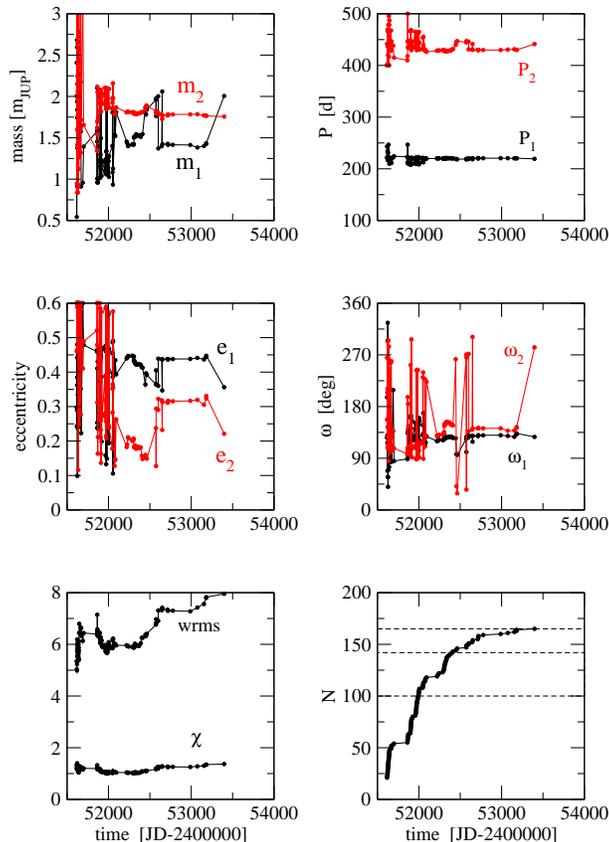


Figura 3.11: Variación de los parámetros orbitales de los mejores ajustes multi-keplerianos como función del conjunto de datos observacionales. Los datos corresponden al conjunto completo CORALIE-Keck ($N=165$).

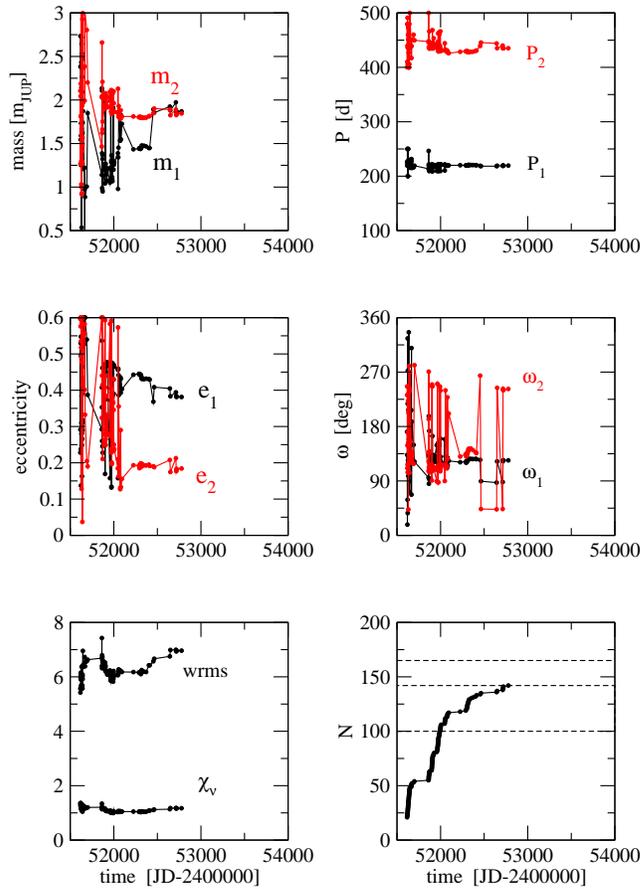


Figura 3.12: Variación de los parámetros orbitales de los mejores ajustes multikeplerianos como función del conjunto de datos observacionales. Los datos corresponden al conjunto reducido CORALIE ($N=142$).

sin embargo todavía presentan repentinos y grandes saltos. Una vez más la eliminación de solo un par de puntos finales causa gran variación en el ángulo, a pesar que no se observen grandes variaciones en los residuos (ver el marco inferior izquierdo).

En conclusión, este análisis muestra que las soluciones se mueven en el espacio de parámetros. Afortunadamente su movimiento ocurre dentro de la región de 1σ , que se observa por ejemplo en la Figura 3.10. Además, los resultados coinciden perfectamente con los obtenidos en el método de *Resampling* y *Análisis en Grilla*. Debido a la diferencia de tiempo en el cómputo entre estos enfoques, parece más sensato utilizar Jackknife para conocer las ubicaciones de los mínimos secundarios.

3.6.1. Análisis dinámicos en soluciones Jackknife y Grilla

Analizaremos la compatibilidad de las soluciones previas desde el punto de vista dinámico. Los resultados nuevamente son obtenidos con un integrador Runge Kutta de orden 8 y con precisión 10^{-12} . Los mejores ajustes para $N = 164$ y $N = 165$ son dinámicamente inestables en escalas de tiempo del orden de 10^3 años. De hecho, como fuera mencionado previamente, es poco común que los mejores ajustes orbitales para este sistema y con distintos números de observaciones (para distintos N) correspondan a soluciones estables. Sin embargo los resultados de las secciones previas han mostrado que existe una región amplia de posibles soluciones orbitales dentro del nivel de confianza de 1σ , todos consistentes con los datos observacionales. Podemos de alguna forma extender el análisis dinámico a todos los puntos en la región y estimar cuales soluciones son estables.

Cada punto de la grilla mostrada en la Figura 3.10 no solo corresponde con valores específicos de $(k_2, h_2) = (e_2 \cos \omega_2, e_2 \sin \omega_2)$, sino también valores distintos de los otros elementos orbitales y masas

planetarias mínimas que son los que minimizan el residuo $wrms$ en ese punto.

Esto es importante pues existen otros valores posibles de soluciones con el mismo (k_2, h_2) que tengan distinto (pero más grande) $wrms$, puede pensarse como la proyección de las mejores soluciones sobre ese plano. En principio es posible que algunas de estas soluciones también tengan $wrms$ dentro del nivel de confianza de 1σ . En otras palabras, a pesar que los gráficos son una buena representación existen posibles soluciones tan “buenas” como las encontradas.

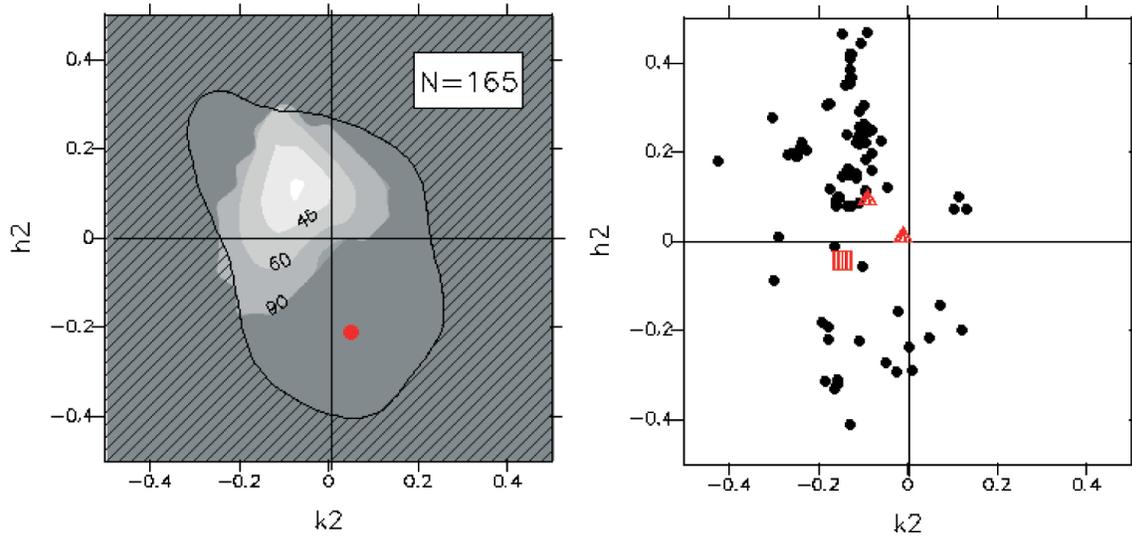


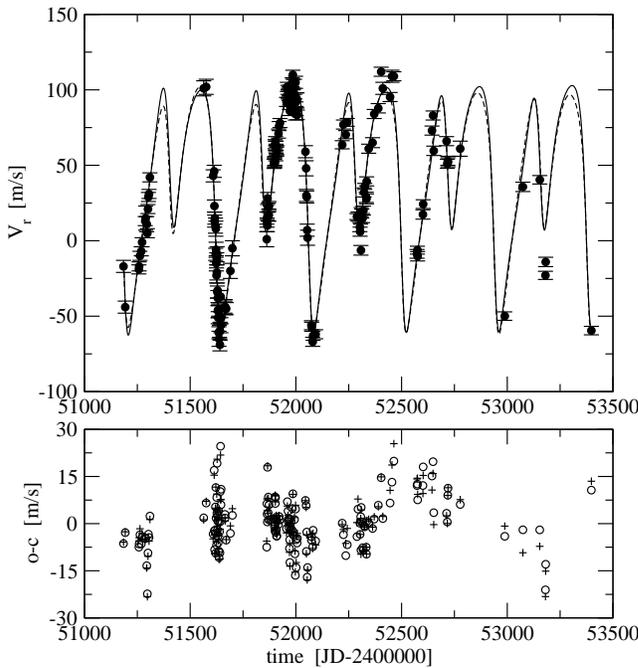
Figura 3.13: **Izquierda:** En el plano $(k_2, h_2) = (e_2 \cos \omega_2, e_2 \sin \omega_2)$ se aprecia el mejor ajuste (menor $wrms$) para $N = 165$ como un círculo rojo. La curva continua cerrada es el nivel de confianza de 1σ . En escala de grises se muestra el comportamiento dinámico en la región. Blanco corresponde a ACR estables tipo $(0, 0)$ con amplitudes menores que 30° . La sombra más oscura indica regiones inestables. **Derecha:** En círculos negros se muestran los mejores ajustes obtenidos para distintos valores de N (Sección 3.6. Los triángulos rojos muestran los mejores ajustes de Lee et al. (2006) correspondientes a *FIT II* y *FIT III*, mientras que el cuadrado rojo es la solución B de Ferraz-Mello et al. (2005).

Los parámetros de cada punto en la grilla para $N = 165$ fueron integrados numéricamente por 10^6 años. La mayoría de ellos tienden a eyectar uno de los planetas dentro del intervalo de tiempo, aquellos que sobreviven se encuentran en configuración ACR del tipo $(0, 0)$ en la resonancia $2/1$ (ver Beaugé et al. 2006) y estos fueron considerados estables. Se calculó la amplitud de la libración de ambos ángulos, el resonante $\theta_1 = 2\lambda_2 - \lambda_1 - \varpi_1$ y la diferencia de longitudes de pericentro $\Delta\varpi = \varpi_2 - \varpi_1$. Aquí λ_i son las longitudes medias de los planetas y ϖ_i son las longitudes medias del pericentro. Los resultados se muestran en la Figura 3.13, una vez más en el plano (k_2, h_2) . En la izquierda de la figura se observan dos gráficos superpuestos: uno que indica niveles de confianza conteniendo las soluciones del *Ajuste en Grilla* y el otro el análisis dinámico de la región. El círculo rojo muestra la posición del mejor ajuste para $N = 165$. La curva negra continua alrededor del punto muestra el nivel de confianza de 1σ (ver Figura 3.10). Todos los puntos con $wrms > wrms_{1\sigma}$ se ubican en la región rayada fuera de esta curva. En escala de grises se muestra el análisis dinámico. La pequeña región blanca en la parte superior izquierda del gráfico corresponde a condiciones iniciales que conducen a soluciones ACR estables con amplitudes de libración menores que 30° . El sombreado gris indica amplitudes crecientes (los valores numéricos se identifican en la figura), mientras que las regiones oscuras ubican la localización que regiones dinámicamente inestables.

Todas las soluciones iniciales correspondientes a ACR estables en la Figura 3.13 muestran buen acuerdo con los datos observacionales. A pesar de que no somos capaces de especificar completamente los parámetros para los planetas de HD82943, parece probable que los valores de los mismos sean

Tabla 3.3: Soluciones orbitales multikeplerianas para $N = 165$ correspondientes a la menor amplitud de co-rotación con $wrms$ dentro del nivel de confianza de 1σ .

Parámetro	HD82943c	HD82943b
$m [M_{Jup}]$	1.569	1.728
$P [d]$	219.24	440.28
$a [UA]$	0.746	1.180
e	0.417	0.150
$\omega [^\circ]$	119.17	123.69
$\lambda - \varpi [^\circ]$	356.74	357.35
$V_{01} [m/s]$	43.90	
$V_{02} [m/s]$	34.28	
$wrms [m/s]$	8.037	
$\sqrt{\chi^2}$	1.537	


 Figura 3.14: **Arriba:** Dos curvas sintéticas de velocidad radial obtenidas de ajustes multikeplerianos. La línea continua corresponde al mejor ajuste orbital con $N = 165$, y la curva punteada a la solución de menor amplitud de ACR tipo $(0, 0)$ presentada en la Tabla 3.3. **Abajo:** residuos $o - c$, con círculos abiertos el mejor ajuste orbital y con cruces la solución estable.

próximos a los dados en la Tabla 3.3 y que los planetas se encuentren atrapados en una ACR estable del tipo $(0, 0)$.

En la parte derecha de la Figura 3.13 se dibujaron con círculos negros la localización en el plano (k_2, h_2) de las soluciones obtenidas con el método (Jackknife). Los triángulos rojos corresponden a dos soluciones estables (Fit II y Fit III) de Lee et al. (2006), y el cuadrado rojo marca la posición de la solución estable B de Ferraz-Mello et al. (2005). Una vez más la gran diversidad de configuraciones, cuya distribución en este plano está más o menos acotada por la curva de nivel de 1σ . Es interesante notar que la mayor parte de los mejores ajustes orbitales corresponden a órbitas alineadas (i.e. $\varpi_2 \sim 130^\circ$), bastante diferente a los mejores ajustes orbitales publicados.

La Tabla 3.3 muestra los parámetros del punto en el plano con la mínima amplitud de co-rotación posible. La razón de masa $m_2/m_1 = 1.1$, un poco mayor que las estimaciones previas incluidas en los mejores ajustes. La curva sintética de velocidad radial resultante y el residuo ($o - c$) de esta solución fueron incluidos en la Figura 3.14, donde también se incluye para comparar la mejor solución. No

existe prácticamente alteración entre las curvas a pesar que obviamente hay un incremento en los valores $o - c$ para el caso de la solución ACR. La estabilidad a largo plazo de esta configuración fue comprobada con una simulación numérica por 10^9 años. No se detectaron cambios en las amplitudes de libración o en los valores máximo/mínimo de los semiejes o excentricidades.

En el próximo capítulo analizaremos cómo el hecho de que el mejor ajuste sea inestable conduce a la posibilidad de que existan más planetas.

Capítulo 4

Búsqueda de Planetas Adicionales

Una solución alternativa para explicar por qué el mejor ajuste de un sistema extrasolar puede ser inestable, es sugerir que pueda existir un planeta adicional que no ha sido modelado. Esta idea también es favorecida cuando el incremento del número de observaciones aumenta los residuos. La inclusión de un nuevo cuerpo en el sistema puede cambiar drásticamente los parámetros de los planetas ya detectados y arrojar nueva luz sobre el sistema real observado. La nueva señal debe tener significado físico razonable (una aceptable relación señal-ruido, excentricidad relativamente baja, etc.) y obviamente presentar una baja probabilidad de ser un falso positivo.

Cuando se intenta agregar nuevos planetas a un sistema, debemos ser cuidadosos en la correcta identificación de la señal. Existen períodos relacionados con el muestreo temporal de las observaciones, que producen picos de señal en el periodograma (los más importantes son los de ~ 1 día y sus correspondientes alias). También es posible identificar en los análisis de velocidad radial falsos períodos, relacionados con el período de rotación de la estrella (ver Howard et al. 2010 y Hatzes et al. 2003 para algunos ejemplos). Sin embargo, ambos procesos tienen una probabilidad de ser falsos positivos (FAP¹) muy alta (es decir 10^{-1} a 10^{-2}).

Volviendo a la idea de la adición de un nuevo planeta para un sistema que tiene ajuste inestable, Gozdziewski y Konacki (2006) propusieron un tercer planeta para HD82943 que podría estar presente en el sistema, utilizando el conjunto total de datos observacionales del sistema ($N=165$). Ellos notaron que los residuos o-c del ajuste de dos planetas, parecía contener una señal periódica con período aproximadamente $P = 1000$ días. Lee et al. (2006) también mencionaron este punto pero no indagaron en el tema. Gozdziewski y Konacki buscaron un ajuste con tres planetas, en el cual el nuevo planeta exterior tendría un período orbital de ~ 930 días. Como resultado el $wrms$ decrece significativamente de ~ 8 m/s hasta aproximadamente 6.4 m/s, y el correspondiente valor de χ^2_ν alcanza valores cercanos a la unidad. Sin embargo este ajuste corresponde a configuraciones dinámicamente inestables. De hecho, las únicas soluciones estables encontradas tienen residuo $wrms=7.4$ m/s, que no representan una disminución satisfactoria debido al hecho que se incrementó el número de parámetros libres. La solución encontrada corresponde a una resonancia 2/1 MMR entre los dos cuerpos interiores, mientras que el tercer planeta hipotético tiene una órbita de muy baja excentricidad en una órbita no resonante. Una simulación numérica de tiempos largos mostró indicaciones de movimiento caótico, a pesar que no se detectaron inestabilidades dinámicas en escalas de tiempo de 10^8 años.

Gozdziewski y Konacki (2006) siguieron su discusión proponiendo un caso hipotético distinto, en

¹un valor de FAP de 10^{-3} es un valor aceptable para detectar un exoplaneta.

el cual se proponía un ajuste orbital con un planeta interior y dos planetas exteriores en configuración coorbital. El problema es que nuevamente no se encontraron soluciones dinámicamente inestables.

4.1. Búsqueda de un tercer planeta en HD82943

En el trabajo que realizamos en 2008 (Beugé et al. 2008) indagamos más profundamente en la búsqueda del tercer planeta, tratando de buscar soluciones más compatibles con escenarios de formación planetaria y cuyos *wrms* no sean tan altos.

En primer lugar utilizamos nuestro algoritmo de minimización para buscar soluciones con 3 planetas en el sistema HD82943 (con $N=165$ de ahora en más), sin restricción alguna en los parámetros. En nuestros resultados, al igual que los mencionados por otros autores, encontramos varias soluciones en las cuales el tercer planeta tiene período $P \sim 900$ días, algunos de ellos tan bajos como 840 días. Sin embargo todas las soluciones encontradas son inestables.

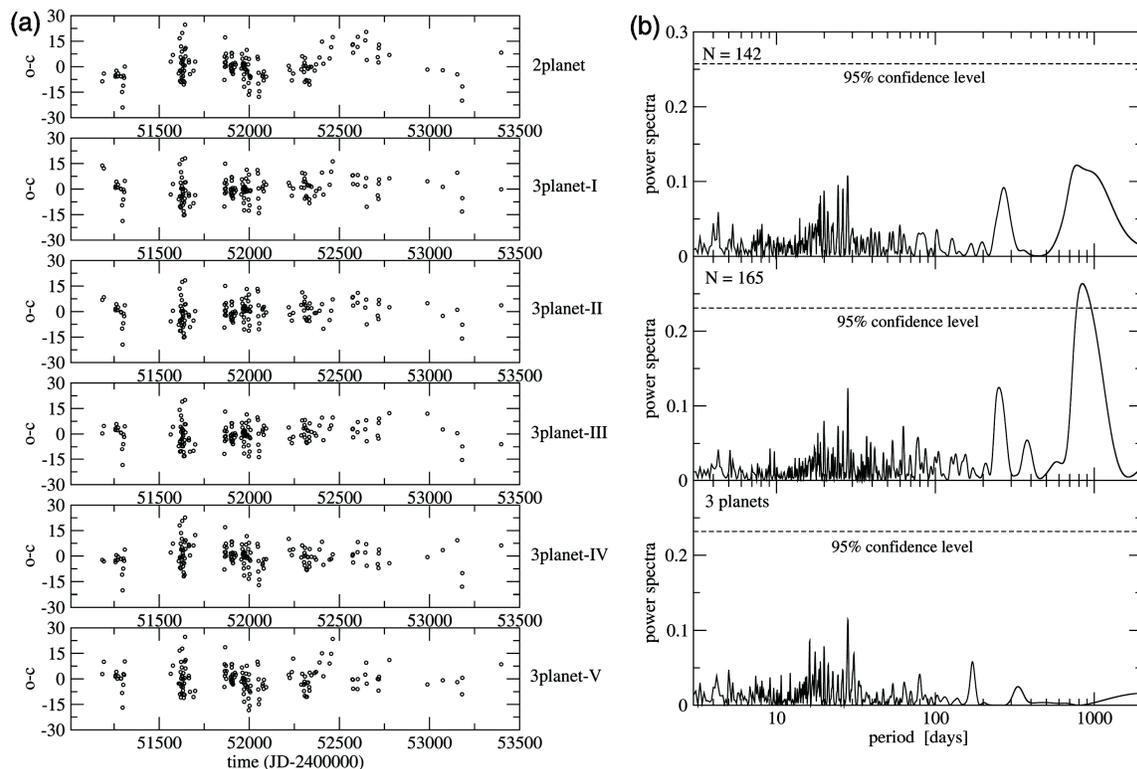


Figura 4.1: (a). Arriba: $o-c$ para el mejor ajuste orbital con 2 planetas. Se puede observar una señal periódica con $P \sim 1000$ días. Los dibujos inferiores corresponden a diferentes ajustes con 3 planetas (ver detalles en el texto). En gran parte de ellos se aprecia la reducción del *wrms* como así también la desaparición de la tendencia periódica. (b). Periodogramas de los residuos $o - c$ de varios ajustes orbitales. El gráfico de arriba corresponde a datos CORALIE ($N = 142$) una vez que la mejor solución orbital con 2 planetas fue eliminada. El gráfico del medio muestra el mismo resultado pero utilizando los datos completos CORALIE+Keck ($N = 165$). Existe un pico pronunciado con nivel de confianza mayor que el 95% con un período entre 800 a 1000 días. El gráfico inferior muestra el periodograma pero ahora asumiendo la solución con 3 planetas mostrada en la solución *3planetas-III*.

Aparte de la reducción del residuo, una ventaja importante de este tipo de soluciones puede verse en la Figura 4.1a. La parte superior reproduce los residuos $o-c$ del ajuste de dos planetas y los otros

paneles muestran diferentes ajustes con tres planetas (período para el externo mayor que 800 días). La tendencia periódica observada en el panel superior desaparece en todos ellos y en la mayoría de los casos el máximo de o-c decrece. Los principales elementos orbitales de estos cinco ajustes orbitales son dados en la Tabla 4.1.

Tabla 4.1: Ajustes orbitales de 3 planetas para el sistema HD82943.

Parámetro	HD82943c	HD82943b	HD82943d?
3planetas-I			
$m [M_{Jup}]$	1.7730	1.7590	0.5654
$P [d]$	220.683	429.863	988.806
$a [UA]$	0.7491	1.1684	2.0354
e	0.3967	0.0553	0.6073
$\omega [^\circ]$	116.989	95.393	315.078
$M [^\circ]$	3.559	11.898	267.314
wrms [m/s]		6.643	
χ_ν		1.286	
3planetas-II			
$m [M_{Jup}]$	1.8060	1.7236	0.4254
$P [d]$	220.273	437.426	993.770
$a [UA]$	0.7482	1.1821	2.0421
e	0.3806	0.0251	0.5676
$\omega [^\circ]$	117.967	142.676	294.073
$M [^\circ]$	0.778	338.029	280.876
wrms [m/s]		6.618	
χ_ν		1.281	
3planetas-III (Mejor Solución)			
$m [M_{Jup}]$	2.045	1.693	0.508
$P [d]$	217.88	456.59	934.85
$a [UA]$	0.743	1.216	1.961
e	0.353	0.198	0.579
$\omega [^\circ]$	124.80	288.67	213.56
$M [^\circ]$	350.15	298.13	27.63
wrms [m/s]		6.005	
χ_ν		1.051	
3planetas-IV			
$m [M_{Jup}]$	1.7630	1.8968	0.8176
$P [d]$	220.520	443.405	1959.682
$a [UA]$	0.7488	1.1929	3.2116
e	0.3722	0.0283	0.7000
$\omega [^\circ]$	113.556	8.696	325.099
$M [^\circ]$	1.645	114.232	115.272
wrms [m/s]		6.449	
χ_ν		1.248	
3planetas-V			
$m [M_{Jup}]$	1.2547	1.6202	0.6236
$P [d]$	220.191	436.868	2150.213
$a [UA]$	0.7479	1.1810	3.4164
e	0.4357	0.4089	0.6864
$\omega [^\circ]$	144.879	124.480	153.728
$M [^\circ]$	346.763	348.759	358.286
wrms [m/s]		7.095	
χ_ν		1.373	

En el caso de los 3 primeros ajustes todos los parámetros fueron determinados sin imposición de límites. Todos los períodos para el nuevo planeta se encuentran entre los 930 días y 993 días, siendo el menor período correspondiente al menor $wrms$. Sin embargo todos son dinámicamente inestables debido a la alta excentricidad de HD82943d. Gozdziewski y Konacki (2006) encontraron que sólo órbitas con muy baja excentricidad se corresponden con configuraciones estables, a pesar que ellas tienen residuos mayores.

En *3planetas-IV* pusimos un límite inferior de 1600 días en los posibles períodos para HD82943d. Para *3planetas-V* este límite inferior es de 2000 días. La idea en ambos casos busca identificar compañeros más distantes que puedan ayudar al ajuste siempre que no introduzcan inestabilidades dinámicas en los compañeros. Sin embargo esto no fue posible debido a la alta excentricidad necesaria que minimiza el residuo, pero una vez más el sistema es inestable dinámicamente.

Como punto final, vale mencionar que en todos los casos la tendencia periódica observada desaparece. Puede inferirse que esta tendencia se debe a la falta de robustez de la solución del ajuste con 2 planetas mientras el número de observaciones aumenta. Sin embargo no está muy bien definido el período probable y existe una amplia gama de configuraciones posibles para el tercer planeta.

Con el objetivo de establecer el período del probable tercer planeta, se puede realizar un análisis de Fourier (con DCDFt) y calcular la importancia de los picos presentes en los espectros de potencia (ver Sección 2.1.7). Los gráficos superior y medio de la Figura 4.1.b muestran el periodograma de los residuos $o - c$ del mejor ajuste dinámico usando los datos CORALIE ($N = 142$) y el conjunto completo de observaciones CORALIE+Keck ($N = 165$). Ambos espectros de potencia fueron normalizados de tal forma que el área total bajo la curva sea la misma y las curvas horizontales muestran el nivel de confianza del 95 %. Si bien con $N=142$ no existe un pico significativo en la señal, cuando se usan $N=165$ existe una señal marcada en $P = 900$ días (gráfico del medio).

La solución *3planetas-III* muestra el mejor ajuste encontrado del conjunto completo de datos asumiendo los 3 planetas. El correspondiente periodograma de los residuos es mostrado en la parte inferior de la Figura 4.1.b. El pico presente en períodos orbitales grandes ha desaparecido y no queda señal significativa en el espectro de potencias nuevo. Más aún, el valor de $\sqrt{\chi^2_\nu}$ se redujo de 1.369 a 1.051 mientras que el $wrms$ cambió de 7.952 a 6.005. A pesar que la reducción es importante, el número de parámetros de la regresión se incrementó desde $M=12$ a $M=17$.

Debemos emplear algún estadístico que cuantifique esta mejora entre ambos modelos. Comencemos recordando que χ^2_ν es una variable aleatoria que puede ser representada por una distribución chi-cuadrado con ν grados de libertad. El mejor ajuste, representado por el valor encontrado de $\chi^2_{\nu_i}$, es el valor más probable de dicha distribución. La mejora relativa entre los ajustes asociados a dos modelos independientes de ν_1 y ν_2 grados de libertad es dada por:

$$\hat{z} \equiv \left(\frac{\chi^2_{\nu_1} \cdot \nu_1 - \chi^2_{\nu_2} \cdot \nu_2}{\nu_1 - \nu_2} \right) \left(\frac{\nu_2}{\chi^2_{\nu_2} \cdot \nu_2} \right) \quad (4.1)$$

donde la multiplicación por el número de grados de libertad (ν_i) debe realizarse para poder comparar los valores de χ_{ν_i} . Según teoría estadística, ésta división variables aleatorias independientes chi-cuadrado produce como resultado una variable aleatoria del tipo Fisher-Snedecor, $F_{\nu_1-\nu_2, \nu_2}$ (Wackerly et al. 2003, p-340 y Cumming, 2004). En la Figura 4.2 mostramos la Función de Densidad de Probabilidad (PDF) de la variable $F_{\nu_1-\nu_2, \nu_2}$.²

²con $\nu_1 = (165 - 12)$ y $\nu_2 = (165 - 17)$ grados de libertad, correspondientes al modelo con dos y tres planetas en el sistema HD82943.

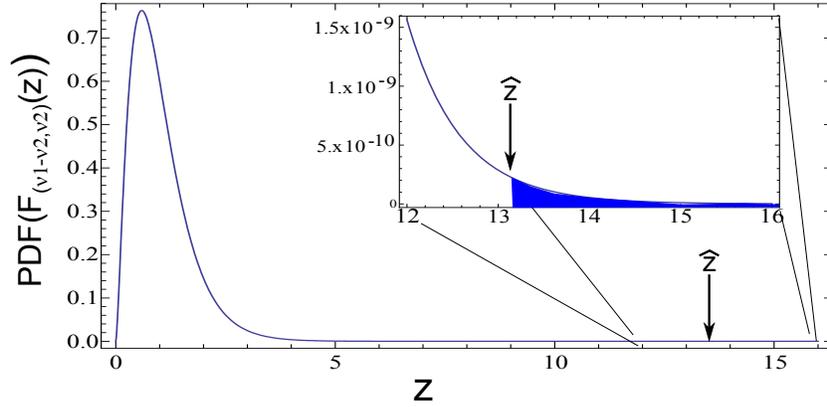


Figura 4.2: Función de densidad de probabilidad de $F_{\nu_1-\nu_2, \nu_2}$. Cuando se averigua un valor \hat{z} , el FAP de esa señal es la integral desde ese punto hasta ∞ (la región que sombreamos en azul).

Mientras mayor sea la diferencia entre los dos ajustes, más fácil optar por el modelo con mayor número de parámetros M . Se puede estimar cual es la probabilidad de obtener un valor z mayor o igual al calculado³, resolviendo:

$$P(z \geq \hat{z}) = \int_{\hat{z}}^{\infty} PDF\left(F_{\nu_1-\nu_2, \nu_2}(z)\right) dz \equiv FAP \quad (4.3)$$

Utilizando nuestros resultados obtenemos un valor de $\hat{z} = 13.06$ y la probabilidad de obtener un valor mayor o igual a este valor es 1×10^{-10} (es decir un $FAP = 1 \times 10^{-10}$). Esto indica que la mejora en el ajuste orbital tiene significado estadístico y no se debe solamente al incremento de parámetros libres. Si a este hecho, añadimos que el pico de la señal periódica es eliminado del periodograma (gráfico interior en la Figura 4.1.b), es muy probable que los datos observacionales apunten en la dirección de un tercer planeta ubicado en las cercanías de un período $P \sim 930$ días.

Nuestro mejor ajuste orbital con 3 planetas se asocia con movimiento inestable, al igual que los resultados de Gozdziewski y Konacki (2006), aunque con una disminución considerable del residuo (obtuvimos $wrms = 6.0$ en lugar de 6.5 m/s). El problema es que la excentricidad del planeta hipotético ($e_3 \sim 0.6$) provoca que el sistema sea muy susceptible a encuentros próximos.

4.2. Soluciones de Laplace en HD82943

A pesar que el ajuste orbital de 3 planetas que analizamos anteriormente tiene la ventaja de eliminar la señal de ~ 1000 días en el residuo, no está claro que sea consistente con un escenario de migración. Recordemos que la relación resonante entre HD82943b y HD82943c es una consecuencia de la migración en gran escala, es decir de varias unidades astronómicas (ver Beaugé et al. 2006), el mismo efecto que debería afectar la órbita primordial de un hipotético tercer planeta. Durante nuestro trabajo notamos que podríamos tener soluciones estables con tres planetas, si suponemos que se

³Por definición de PDF:

$$\int_0^{\infty} F(z) dz = 1. \quad (4.2)$$

encuentra en una resonancia de movimientos medios del tipo *Laplace*; en otras palabras una conmensurabilidad entre los tres movimientos medios de tal forma que:

$$n_1/n_2 \approx n_2/n_3 \approx 2/1 \quad (4.4)$$

El espacio de fase asociado a la resonancia de Laplace es complejo, y parece estar poblado con un número pequeño de islas de movimiento estable, rodeadas por grandes regiones de inestabilidad. Esto hace que la identificación de los mejores ajustes estables, con estas características, sea un proceso complejo.

Modificamos el código de minimización para incluir un análisis dinámico simultáneamente con las soluciones. Aparte de incorporar un chequeo de estabilidad del sistema, también se pusieron restricciones en la rutina de *Annealing* para buscar soluciones solo en aquellas superficies del espacio de fase compatibles con la resonancia de Laplace. Debido a las altas excentricidades, las islas de estabilidad de tal resonancia deberían ser relativamente pequeñas, quizás muy difíciles de encontrar con la búsqueda aleatoria con *Annealing*, más difícil aún si poseen libraciones asimétricas (debido a la libertad de elección de ángulos).

Con este código de ajuste orbital, buscamos el conjunto de tres planetas dinámicamente estables que mejor reproducen los datos completos de velocidad radial. La Tabla 4.2 muestra la mejor solución encontrada y aunque la masa del tercer planeta es un poco menor que la solución de la Tabla 4.1 (*3planetas-III*), los residuos son un poco mayores. Nuestra búsqueda de soluciones con tres planetas no fue exhaustiva y pueden existir mejores configuraciones.

Tabla 4.2: Mejores ajustes orbitales dinámicos de tres planetas correspondientes a una resonancia de Laplace ($N = 165$).

Parámetro	HD82943c	HD82943b	HD82943d(?)
$m [M_{Jup}]$	1.703	1.747	0.351
$P [d]$	218.70	447.50	900.22
$a [UA]$	0.745	1.200	1.912
e	0.361	0.190	0.078
$\omega [^\circ]$	132.32	192.01	116.19
$M [^\circ]$	350.15	298.13	27.63
$V_{01} [m/s]$		42.93	
$V_{02} [m/s]$		31.74	
wrms [m/s]		7.521	
$\sqrt{\chi^2}$		1.316	

Para estudiar la dinámica del sistema presentado en la Tabla 4.2 y su compatibilidad con escenarios de migración, analizamos la evolución dinámica de los tres planetas bajo el efecto adicional de una fuerza exterior disipativa que simula los efectos de la interacción entre el planeta y el disco. Las masas de los dos planetas interiores se eligieron como $m_1 = 1.70M_{Jup}$ y $m_2 = 1.75M_{Jup}$ (resultado del ajuste en configuración de Laplace, Tabla 4.2), mientras que la masa del tercer planeta ficticio fue variada. Todas las órbitas fueron inicialmente localizadas con semiejes mayores más allá de la MMR 2/1 y con excentricidad igual a cero (órbitas circulares).

Como en trabajos previos (e.g. Beaugé et al. 2006) se eligió una fuerza de rozamiento tipo Stokes con valores fijos de tiempos característicos para el semieje y las excentricidades. En todas las simulaciones la fuerza exterior solo afecta los dos planetas exteriores y los parámetros de rozamiento se eligieron para garantizar decaimiento orbital con órbitas convergentes (m_1 no está sujeto a estas fuerzas, pues las simulaciones hidrodinámicas muestran que el disco gaseoso en la región interior al

planeta más interno debió ser agotado). La migración se asume más rápida para el planeta exterior debido a que su masa es menor. Esto se hizo de acuerdo a evidencia de simulaciones hidrodinámicas, pero también para garantizar la migración convergente de todos los planetas involucrados.

Se trabajó bajo un amplio rango de condiciones iniciales y parámetros de rozamientos. Siempre encontramos que los tres planetas evolucionan hacia una resonancia de movimientos medios doble, en los cuales $n_1/n_2 \simeq 2/1$ y la razón n_2/n_3 también corresponde a razón entre enteros. En la mayoría de las simulaciones, los dos planetas exteriores quedan atrapados en una resonancia de movimientos medios (MMR) 2/1 (i.e. $n_2/n_3 \simeq 2/1$). En tales casos, las órbitas de los tres planetas muestran libraciones simultáneas de los correspondientes ángulos resonantes

$$\begin{aligned}\theta_{21} &= 2\lambda_2 - \lambda_1 - \varpi_1 \\ \theta_{32} &= 2\lambda_3 - \lambda_2 - \varpi_2\end{aligned}\quad (4.5)$$

como así también las libraciones de la diferencia en longitudes de pericentro $\varpi_2 - \varpi_1$ y $\varpi_3 - \varpi_2$. En otras palabras, el sistema evoluciona hacia una ACR doble. Un ejemplo de tal simulación se muestra en la Figura 4.3, donde la masa del planeta exterior fue elegida como $m_3 = 0.35M_{Jup}$ (solución de la Tabla 4.2). Las condiciones iniciales para los tres exoplanetas consideran órbitas circulares y los tiempos característicos fueron elegidos para lograr reproducir los valores de la Tabla 4.2 (ver Beaugé et

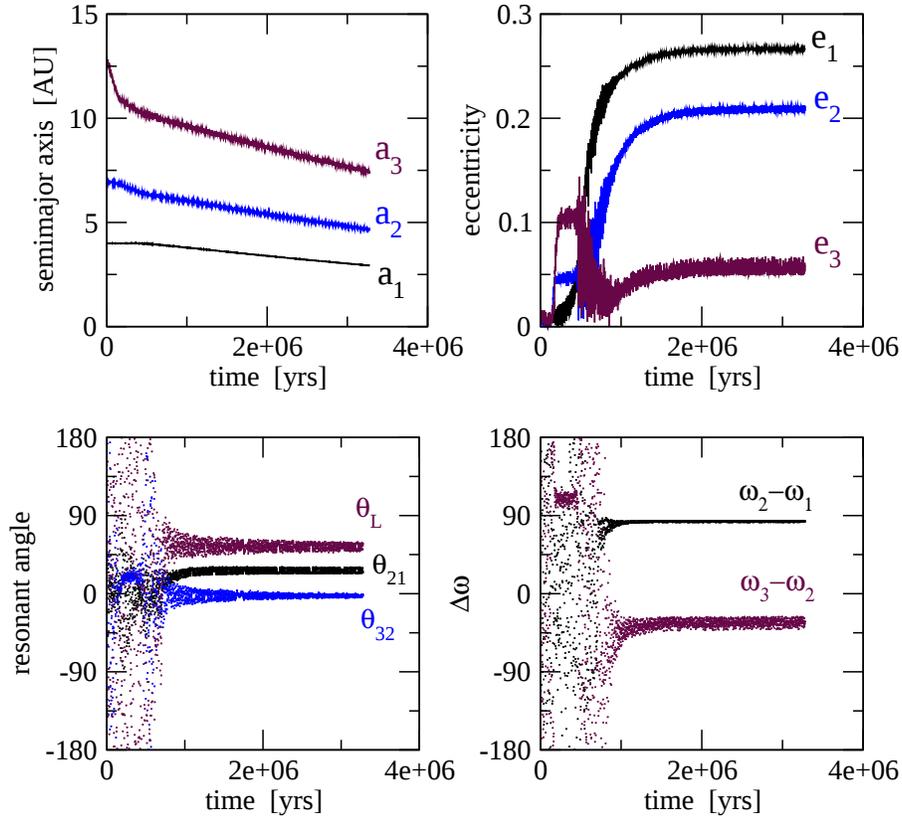


Figura 4.3: Migración de tres planetas ($m_1 = 1.70$, $m_2 = 1.75$, y $m_3 = 0.35$, en masas de júpiter), inicialmente en órbitas circulares con semiejes mayores $a_1 = 4$, $a_2 = 7$ y $a_3 = 13$ UA. La migración convergente es modelada con una fuerza de rozamiento tipo Stokes actuando sólo sobre los dos cuerpos exteriores. Después de $\sim 10^6$ años, todos los planetas quedan atrapados en MMR 2/1, mostrando pequeñas amplitudes de libración de los ángulos resonantes θ_{21} , θ_{32} y la diferencia en longitudes de pericentro. El ángulo resonante de la resonancia Laplace $\theta_L = \lambda_1 - 3\lambda_2 + 2\lambda_3$ también libra. La configuración final corresponde a una ACR doble asimétrica.

al. 2006, para una introducción a teoría de invariantes adiabáticos)

El panel superior izquierdo de la Fig. 4.3 muestra la evolución de los semiejes, que se ubican inicialmente fuera de la resonancia y terminan en la commensurabilidad deseada. El panel superior derecho muestra como las excentricidades evolucionan, obteniendo finalmente órbitas elípticas con valores medios muy parecidos al mejor ajuste. El panel inferior izquierdo muestra las libraciones asimétricas de los ángulos resonantes θ_{21} y θ_{32} , y también se observa la libración del ángulo de Laplace:

$$\theta_L = \lambda_1 - 3\lambda_2 + 2\lambda_3. \quad (4.6)$$

Finalmente en el panel inferior derecho, se observa que las diferencias de longitudes de pericentro $\varpi_2 - \varpi_1$ y $\varpi_3 - \varpi_2$ muestran una pequeña amplitud de oscilación, completando la configuración de ACR asimétrica.

Debido a que nuestras simulaciones de migración planetaria parecen favorecer la evolución orbital de los tres planetas en resonancia de Laplace, es posible esperar que cualquier hipotético tercer planeta en el sistema HD82943 se encuentre en dicha configuración.

Una simulación numérica de largo tiempo muestra que la solución encontrada es dinámicamente estable para tiempos de escala de al menos 1 Gyr, sin cambio secular apreciable en las excentricidades o amplitudes de oscilación. La Figura 4.4 presenta dos mapas dinámicos en el plano $(k_3, h_3) = (e_3 \cos \omega_3, e_3 \sin \omega_3)$. Cada uno fue construido con una grilla de 401×401 condiciones iniciales, y la solución estable de la Tabla 4.2 es identificada con una estrella roja. Las masas planetarias mínimas así como todos los otros parámetros fueron elegidos iguales a los de la configuración estable. El coloreado gris en el mapa izquierdo está relacionado con los tiempos de escape; colores oscuros corresponden a condiciones iniciales inestables que son eyectadas del sistema en pocos miles de años (excentricidades mayores que 1 o distancias mayores que 10 UA), mientras que órbitas más estables son marcadas en blanco. En el mapa derecho se colorea con el número espectral a partir de un análisis de Fourier de las integraciones (Michtchenko y Ferraz-Mello, 2001). Las órbitas regulares se identifican en blanco y las soluciones más caóticas en tonos más oscuros. Las líneas continuas negras muestran las curvas de nivel constante de $wrms$ para cada punto del plano.

La isla de movimiento estable no es muy amplia y está limitada a excentricidades moderadamente bajas para el hipotético tercer planeta. Así, la región asociada al residuo más bajo (ver Tabla 4.1, *3planetas-III*) yace en el medio de una región inestable. El mejor ajuste orbital estable (estrella roja) se localiza en un costado de la isla de estabilidad, a pesar que todas las órbitas regulares tienen aparentemente el mismo valor de residuos.

Finalmente, la Figura 4.5 compara los periodogramas de los residuos de los mejores ajustes orbitales de 3 planetas (línea a trazos) con el mejor ajuste estable en Resonancia de Laplace (línea azul) dado en la Tabla 4.2. El gráfico original corresponde al mejor ajuste con 2 planetas y es mostrado en línea continua. A pesar que la mejor solución estable tenga valor de residuo más alto que el mejor ajuste sin restricciones, es útil a la hora de eliminar el pico en aproximadamente 900 días, y no quedan señales estadísticamente significativas en los datos.

4.3. Otros ejemplos de Resonancia de Laplace

La pregunta pendiente es si ¿existen otros sistemas que se encuentren en esta resonancia?, ya que esta resonancia parece ser un mecanismo de estabilización muy eficiente cuando tenemos tres cuerpos masivos.

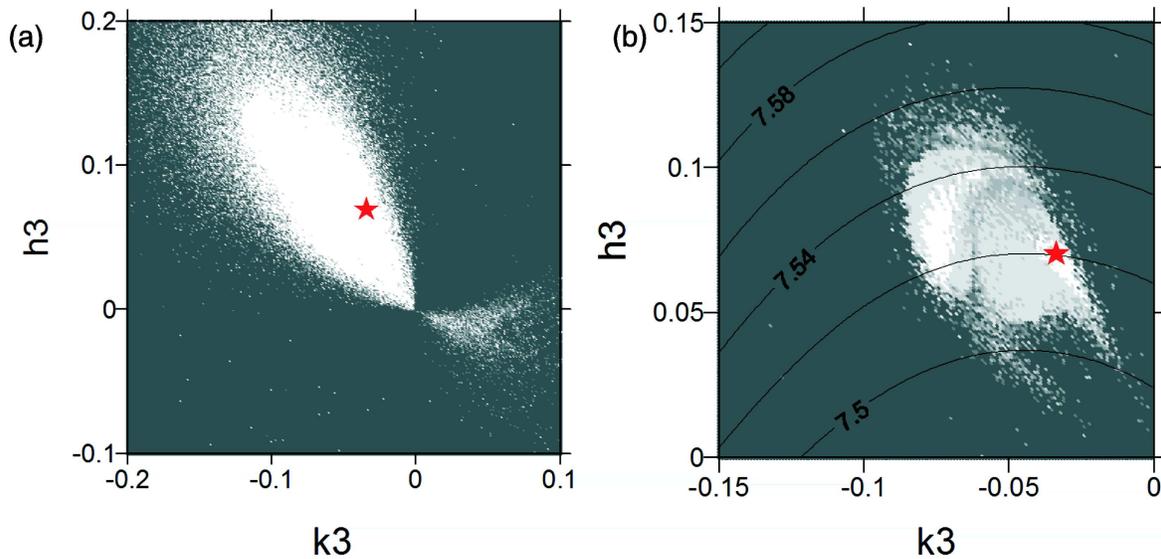


Figura 4.4: **a:** Tiempo de escape para condiciones iniciales en el plano $(k_3, h_3) = (e_3 \cos \omega_3, e_3 \sin \omega_3)$ en la vecindad de la solución estable de tres planetas en la resonancia de Laplace. El blanco indica que no hay escape, mientras que oscuro muestra eyecciones rápidas de un planeta. **b:** Número espectral de un análisis de Fourier de las integraciones numéricas de cada condición inicial. Las órbitas regulares se muestran en blanco mientras que el movimiento caótico en sombreados más oscuro. Las curvas de nivel corresponden a valores constantes de $wrms$ de las condiciones iniciales en dicho plano.

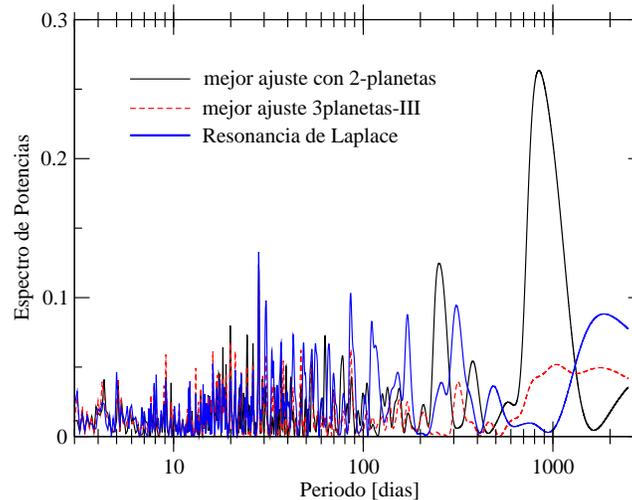


Figura 4.5: Periodograma de los residuos de tres ajustes orbitales distintos considerando el conjunto completo de datos. A pesar que la configuración orbital estable en la vecindad de la resonancia de Laplace tiene $wrms$ más grande que el mejor ajuste orbital de 3 planetas, el pico de alrededor de 900 días es prácticamente eliminado, y no queda ninguna señal estadísticamente significativa.

A mediados del siglo XVIII, ya existían datos adecuados para demostrar que existía una relación muy “curiosa” entre las órbitas de las lunas de Júpiter: Io, Europa y Ganímedes. En 1743, el astrónomo sueco Pehr Wilhelm Wargentin (el primer director del Observatorio de Estocolmo) publicó tablas que permitieron establecer la razón 1:2:4 entre los períodos de las lunas Jovianas. Sus resultados predecían que el ángulo:

$$\phi_L = \lambda_{IO} - 3 \cdot \lambda_{Eu} + 2 \cdot \lambda_{Ga} = 180^\circ \quad (4.7)$$

que relaciona las longitudes de las lunas, era mantenido con un extraordinario grado de exactitud. Geométricamente esto significa que los satélites están acoplados en un ciclo de seis conjunciones sucesivas luna-luna durante el curso de una órbita de Ganímedes.

Laplace descubrió que un mecanismo dinámico debía ser responsable de mantener el ciclo de conjunciones y en 1784 pudo demostrar que el ángulo ϕ_L estaba sujeto a oscilaciones tipo pendulares. Si los satélites son perturbados levemente, entonces con el tiempo las interacciones satélite-satélite colaboran para causar que ϕ_L oscile alrededor de su valor de equilibrio de 180° . Su teoría le permitió derivar las masas de las lunas y también predecir que el período de oscilación del ángulo ϕ_L sería de 2270d 18h (Murray y Dermott, 1999)

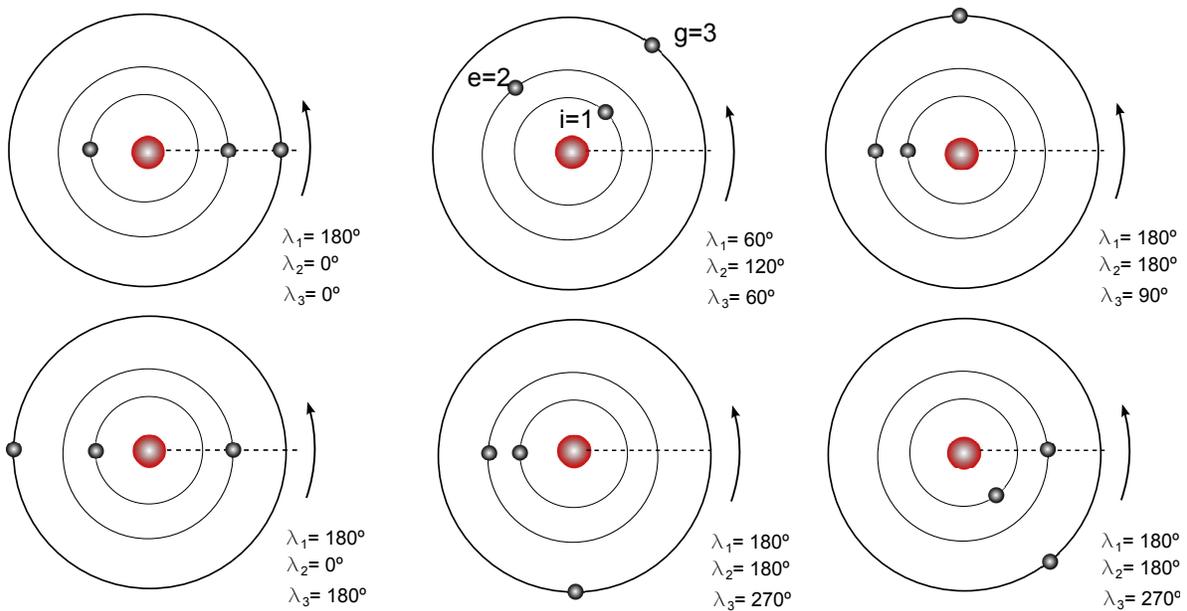


Figura 4.6: Resonancia de Laplace en Satélites de Júpiter. De interior a exterior se muestran las lunas Io (1), Europa (2) y Ganímedes (3). Se esquematizan las seis conjunciones consecutivas que ayudaron a deducir la forma del ángulo resonante.

En los tiempos de Laplace, las observaciones no eran lo suficientemente precisas para medir una amplitud de libración, sino que parecían estar exactamente en la condición resonante. Sin embargo hoy en día se sabe que el ángulo ϕ_L libra con una pequeña amplitud de 0.064° , y que el período de oscilación es 2071d, bastante cercano al predicho por Laplace. Yoder y Peale (1981) mostraron que la libración observada, puede ser entendida teniendo en cuenta la presencia de fuerzas disipativas balanceadas que incluyen disipación tidal de Júpiter y la disipación tidal de Io. La Figura 4.6 muestra una secuencia de posiciones consecutivas que clarifican como actúa la resonancia de Laplace en los satélites de Júpiter.

¿Podríamos esperar entonces que algún tipo de configuración similar exista en otros sistemas extrasolares? -aparte de las soluciones compatibles en el sistema extrasolar HD82943 (Beaugé et al. 2008)-.

Recientemente, una resonancia de Laplace fue propuesta para el sistema extrasolar Gliese 876 (Rivera et al. 2010). En su trabajo mostraron que los conocidos planetas gigantes con períodos $P \sim 30$ días y $P \sim 61$ días son acompañados por un planeta adicional de masa similar a la de Urano con período $P \sim 124$ días. Este sistema muestra que el mejor ajuste se encuentra en una relación de Laplace con el ángulo ϕ_L oscilando alrededor de $\sim 0^\circ$, con una amplitud $\Delta\phi_L = 40 \pm 13^\circ$.

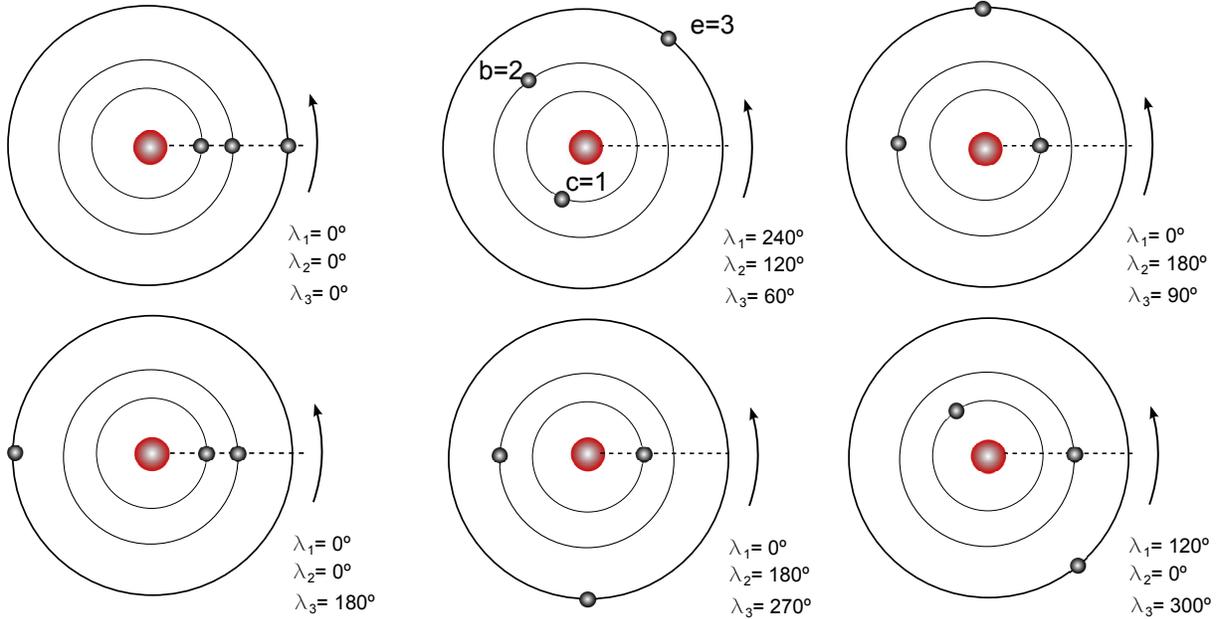


Figura 4.7: Esquema de la resonancia de Laplace para el sistema *GJ 876*. El ángulo resonante es $\phi_L = 0^\circ$. Los períodos son $P \sim 30$ días y $P \sim 61$ días y $P \sim 124$ días.

Tabla 4.3: Ajuste coplanar de 4 planetas para GJ876 con $i=59^\circ$

Parámetro	Planeta d	Planeta c	Planeta b	Planeta e
P (días)	1.94	30.09	61.12	124.26
m	$6.8 M_\oplus$	$0.714 M_{\text{Jup}}$	$2.27 M_{\text{Jup}}$	$14.6 M_\oplus$
a (UA)	0.0208	0.1295	0.2083	0.3343
K (m s^{-1}) 6.56	88.34	214.00	3.42	
e	0.207	0.256	0.0324	0.055
ω ($^\circ$)	234	48.76	50.3	239
M ($^\circ$)	355	294.59	325.7	335 ± 24
V_0 (m s^{-1})		51.06 ± 0.30		
χ_ν^2		2.6177		
RMS (m s^{-1})		2.9604		

La Figura 4.7 muestra la configuración de *GJ 876* y la Tabla 4.3 los parámetros planetarios de la solución. Las integraciones numéricas muestran que el sistema de cuatro planetas es estable por al menos 1 Gy. Mientras que en el sistema de Júpiter, el cuerpo central es 5000 veces más masivo que su sistema de satélites, en Gliese 876 los planetas representan el 1 % de la masa del cuerpo central. Esto implica interacciones más fuertes con la estrella central.

Rivera et al. (2010), en el mismo trabajo, generaron 1000 conjuntos de velocidades radiales. Cada conjunto sintético fue generado con la señal producida por los cuatro planetas de la Tabla 4.3 añadiendo ruido Gaussiano. Luego cada ajuste orbital obtenido se integró por 300000 días (2500 períodos del planeta externo) y se examinó la amplitud de oscilación del ángulo crítico de la resonancia de Laplace, ϕ_L . Como resultado sólo pudieron encontrar 6 ajustes orbitales con amplitud de libración $< 40^\circ$. Mientras que si la condición para generar las velocidades radiales se elige con un ángulo de oscilación bien pequeño y se disminuye el ruido Gaussiano, un 80 % de los resultados tiene amplitud de oscilación $< 40^\circ$. Además también los autores realizaron una serie pruebas donde corroboraron

que si el ángulo de Laplace no se encuentra librando, los sistemas son inestables en escalas de tiempo < 200 Myr. De esta forma se concluyó que es muy difícil detectar sistemas en resonancia de Laplace con el método de velocidades radiales a menos que se tengan muy bajos errores. Sin embargo su estudio tiene en cuenta que ϕ_L oscila alrededor de 180° , mientras que en HD82943 encontramos que la libración es asimétrica (un ángulo distinto de 0 o 180°). Esto implicaría que en nuestro caso las fases de los ángulos pueden ser identificadas más fácilmente.

Resumiendo la información expuesta en el capítulo, obtenemos que la resonancia de Laplace es un mecanismo que, primero, es compatible con escenarios de formación planetaria. Segundo se encuentra presente en nuestro propio Sistema Solar y es posible encontrar buenos ajustes orbitales (con la presencia de un nuevo planeta) para HD82943 y Gliese 876. Sin embargo, aún queda como trabajo pendiente explicar bajo que condiciones la resonancia de Laplace produce libraciones del ángulo resonante ϕ_L que sean simétricas (alrededor de 0° o 180°) o asimétricas (cualquier otro valor del ángulo) y cuales son las diferencias para su detección con el método de velocidad radial.

Capítulo 5

Efectos de Selección con Velocidades Radiales

En adición a la gran cantidad de trabajo observacional dirigido a la observación de planetas extrasolares, existe un esfuerzo paralelo de astrónomos teóricos para explicar la distribución de elementos orbitales en el contexto de formación planetaria y evolución.

Históricamente los astrónomos dedicados a la dinámica poseían elementos orbitales determinados de forma increíblemente precisa. En el siglo XVIII los elementos orbitales del Sistema Solar eran conocidos mucho mejor de lo que actualmente conocemos los de los planetas extrasolares. Las interpretaciones teóricas de los datos de planetas extrasolares a veces son realizadas sin tomar en cuenta las altas incertidumbres reales en los elementos orbitales.

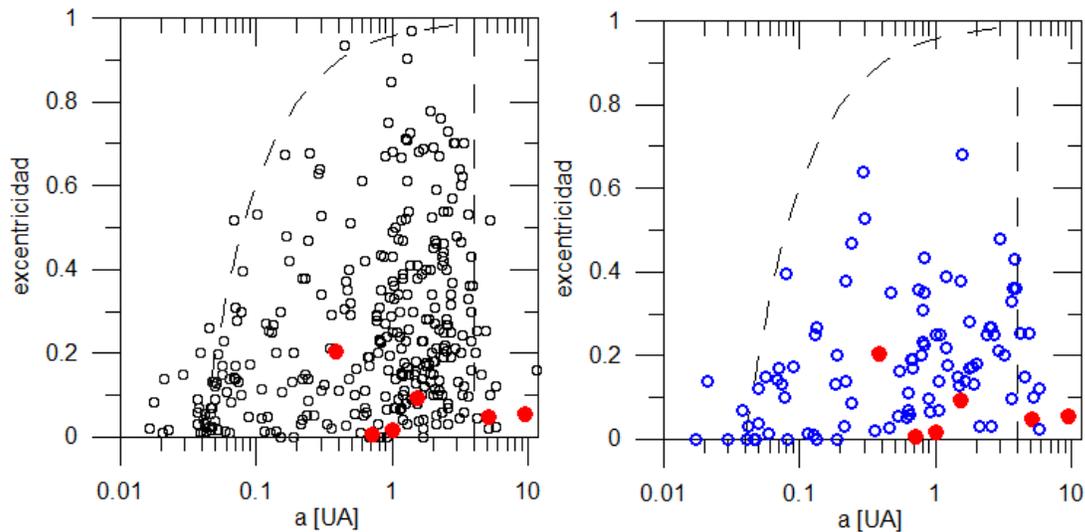


Figura 5.1: **Izquierda.** Distribución de elementos orbitales de todos los planetas extrasolares conocidos al 14 de abril del 2010. **Derecha.** Distribución elementos orbitales de planetas que se encuentran en sistemas de 2 o más planetas. Los puntos rojos representan los elementos orbitales de los planetas del Sistema Solar.

Por ejemplo, en la Figura 5.1 se compara la distribución de elementos orbitales (semieje y excentricidad) entre todos los sistemas conocidos (izquierda) y aquellos sistemas de 2 o más planetas

(derecha). La distribución de elementos orbitales de todos los sistemas muestra valores de excentricidades altos, que no están presentes en sistemas con 2 o más planetas. ¿Eso es evidencia de un pasado dinámico del sistema, una limitación observacional o una falla en el proceso de detección debido a la técnica utilizada?

Desde el punto de vista dinámico los sistemas extrasolares en resonancia de movimientos medios (MMR) son muy importantes porque ellos pueden evidenciar la migración orbital pasada del sistema, y así ayudar a entender los procesos de formación. La compatibilidad de un dado sistema planetario con un escenario de migración depende sensiblemente de los valores calculados de excentricidad e y longitud de pericentro ϖ debido a que especifican que tipo de configuración resonante tenemos (libración, corrotación de apsides, e.g. Beaugé et al. 2006, Michtchenko et al. 2008ab)

En particular, los sistemas multiplanetarios son complicados de analizar y resolver, sobre todo porque la razón entre el número de observaciones N y el número de parámetros a determinar M no es muy grande ($N/M \sim 10-20$) y en algunos casos en número de grados de libertad ν es pequeño (donde $\nu = N-M$). Si el sistema contiene dos o más planetas en resonancia de movimientos medios (MMR) el problema es aún más desafiante. La conmensurabilidad puede causar interacciones entre las componentes, complicando la separación de la señal (Giuppone et al. 2009, Anglada-Escude et al. 2009). Dos de los más importantes elementos orbitales que determinan el comportamiento dinámico son la excentricidad e y la longitud de pericentro ϖ y son justamente los más difíciles de determinar debido a que están dados por las asimetrías en la señal cuasi-periódica.

El propósito de este capítulo es explorar que procesos pueden causar errores sistemáticos en la distribución de los elementos orbitales reales. Primero expondremos algunas pautas acerca del análisis de su señal de velocidad radial y luego los resultados relacionados con la detección de 1 exoplaneta (presentados también en Shen y Turner, 2008). Luego pasaremos a analizar sistemas con 2 planetas y más generalmente al final del capítulo analizaremos los límites de detección para sistemas con 2 planetas en resonancia.

5.1. Detectabilidad de exoplanetas

Imaginemos un exoplaneta orbitando a su estrella y que tenemos un conjunto de N observaciones de velocidades radiales (V_i), todas con el mismo error ($\sigma_i = \sigma$). Obviamente, dependiendo del error en las observaciones es más fácil identificar si la variación en las velocidades radiales se debe a la presencia de un exoplaneta o a una aleatoriedad en la señal. El conjunto de datos puede ser ajustado con distintos modelos W_i y evaluar cual es el que mejor representa los datos. El modelo lineal implica que $W_a = cte$, mientras que el modelo kepleriano con 1 planeta es $W_b = K[\cos(f + \omega) + e \cos(\omega)] + \mathbf{V}_0$ (K es la amplitud de la señal). Debido a que los modelos tienen número distintos de grados de libertad (i.e. número de observaciones menos número de parámetros libres, $\nu = N - M$) no alcanza con decir que un ajuste es mejor que otro solamente orientándonos con el valor de los residuos, sino que debemos cuantificar esta mejora.

La velocidad radial observada debería ser algo con la forma:

$$V_r(i) = W_b(i) + N(0, \sigma) \quad (5.1)$$

donde $W_b(i)$ es la velocidad radial que proviene del sistema físico real y $N(0, \sigma)$ es una variable Gaussiana con dispersión σ . Cuanto mayor es la relación señal sobre ruido ($S/N \equiv K/\sigma$), los datos pueden ser mejor modelados y encontrar la configuración del sistema observado.

A fin de evaluar si somos capaces de identificar bien los ajustes, podemos generar conjuntos de velocidades sintéticas y analizar los resultados de la minimización en función de la S/N, del número de grados de libertad y del número de datos por ejemplo. La Figura 5.2 muestra una síntesis del problema que enfrentamos. Se eligió un planeta con período 365 días, $m_p = 0.5M_{jup}$ y $e = 0.3$. En el panel de la izquierda los datos generados con una relación $K/\sigma=2$ y tres curvas que representan distintos ajustes: una línea horizontal a trazos que representa el modelo lineal, una curva roja que se corresponde con la curva de velocidad radial del sistema real y la curva en azul que muestra la curva que produciría la solución con menor residuo. Es interesante notar que cuando la relación K/σ es baja, el mejor ajuste orbital puede identificar erróneamente un planeta con el mismo período pero una excentricidad más alta. En el panel de la derecha en los mismos tiempos aleatorios que a la izquierda, pero con datos generados con una relación $K/\sigma=4$. Esto permite identificar mucho mejor el sistema (i.e. la curva roja y azul se encuentran casi superpuestas).

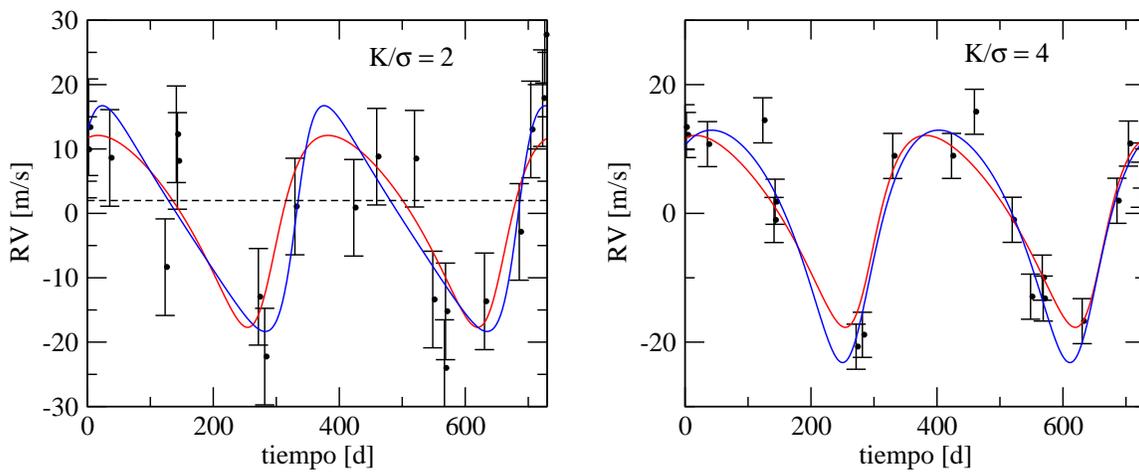


Figura 5.2: **Izquierda.** Conjunto sintético de 20 velocidades radiales generadas por un planeta de $0.5 M_{jup}$ y $e = 0.3$, que gira en torno a una estrella de masa solar. Se cubren dos períodos orbitales, cuyos tiempos de “observación” fueron elegidos aleatorios. En rojo se muestra la señal nominal que produce el planeta. Debido a que la razón K/σ es baja, la curva generada con los datos del mejor ajuste (línea en azul) no coincide con la curva que genera el sistema real. Tampoco es claro si los datos no se comportan solo de forma aleatoria alrededor de su valor medio, marcado como una línea horizontal a trazos. **Derecha.** El mismo conjunto de velocidades radiales pero con menor error (σ) (i.e. K/σ es mayor), permite una mejor identificación del sistema.

Veamos nuevamente el desarrollo de cómo evaluar la mejora de un ajuste orbital respecto de otro. A fin de simplificar la notación trabajaremos con la suma cuadrática de residuos:

$$Q \equiv \chi^2_\nu \cdot \sigma^2 \cdot \nu = \sum_{i=1}^N (V_i - W(t_i))^2 \quad (5.2)$$

donde $W(t)$ es el modelo adoptado. Como es usual, las V_i son asumidas variables normales, estadísticamente independientes $V_i = N(W(t_i), \sigma^2)$. La varianza σ^2 es independiente del tiempo (la misma para todos los datos), pero de ser necesario la definición de Q puede introducir el peso de cada variable y sus diferentes varianzas (ver 2.1.4 y Ferraz-Mello, 1981).

Adoptaremos los subíndices a y b para el primer y segundo ajuste respectivamente. Además, asumiremos que el modelo a está incluido en el modelo b ($M_a < M_b$). La mejora relativa entre los ajustes orbitales puede ser caracterizada por la potencia z , definida como (ver Cumming 2004):

$$z = \frac{\nu_b}{\nu_a - \nu_b} \frac{Q_a - Q_b}{Q_b} \quad (5.3)$$

Cuando nos referimos al resultado en la solución del problema de comparación de modelos, cada $Q = Q_{min}$, debemos tener en cuenta que no se refiere a una elección al azar de parámetros sino al **máximo** de z que se puede obtener (\hat{z}). La probabilidad de falsa alarma (FAP) es la probabilidad de obtener un resultado con el mismo valor de \hat{z} pero sólo por casualidad.

Si se necesita calcular cual es la probabilidad de obtener un valor z mayor o igual que el obtenido, debemos trabajar con la Función de Distribución de Probabilidad (PDF) de esta variable¹ y deberemos resolver:

$$FAP \equiv P(z \geq \hat{z}) = \int_{\hat{z}}^{\infty} PDF \left(F_{\nu_a - \nu_b, \nu_b}(z) \right) dz \quad (5.5)$$

El resultado de la integral involucra funciones Beta y funciones Hipergeométricas, que solamente tienen solución analítica simple cuando tratamos casos sencillos. Finalmente la detectabilidad del exoplaneta es garantizada siempre y cuando la FAP sea un valor muy pequeño (usualmente menor que 10^{-2}).

Por ejemplo si consideramos que el primer ajuste “a” corresponde a un modelo constante ($\nu_a = N-1$) y el modelo “b” a un modelo circular ($\nu_b = N-3$)

$$P(z \geq \hat{z}) = \left(1 + \frac{2\hat{z}}{\nu_b} \right)^{-\nu_b/2} \quad (5.6)$$

En tanto que, por ejemplo, si queremos ver la mejora del modelo constante ($\nu_b = N-1$) comparada con la del modelo kepleriano de 5 parámetros ($\nu_b = N-5$) tenemos:

$$P(z \geq \hat{z}) = \left(1 + \frac{(\nu_b + 2) 4\hat{z}}{2 \nu_b} \right) \left(1 + \frac{4\hat{z}}{\nu_b} \right)^{-(\nu_b+2)/2} \quad (5.7)$$

Los casos más complejos necesitan ser resueltos numéricamente. Estos dos ejemplos han sido estudiados en Cumming (2004) y Cumming et al. (2008) respectivamente. La importancia de un ajuste orbital depende de cuan seguido se obtenga una potencia tan grande como la observada \hat{z} , pero producida solamente por ruido. Sin embargo, la FAP total depende de cuantas frecuencias independientes son analizadas. Para una búsqueda entre muchas frecuencias (\mathcal{M}), cada frecuencia independiente debe ser contada como si fuera una prueba individual. La FAP es (Cumming, 2004):

$$FAP = 1 - [1 - P(z > \hat{z})]^{\mathcal{M}}, \quad (5.8)$$

que cuando la FAP es pequeña se reduce a:

$$\mathcal{F} \approx \mathcal{M} P(z \geq \hat{z}). \quad (5.9)$$

Una estimación del número de frecuencias independientes es $\mathcal{M} = T\Delta f$, donde $\Delta f = f_2 - f_1$ es el rango de frecuencias buscados y T la duración del conjunto de datos (Cumming et al. 2008). Para datos equiespaciados, el número de frecuencias independientes es $N/2$, comenzando desde $1/T$

¹Por definición de PDF:

$$\int_0^{\infty} PDF \left(F_{\nu_a - \nu_b, \nu_b}(z) \right) dz = 1. \quad (5.4)$$

hasta la frecuencia de Nyquist $f_{Nyq} = N/2T$. Sin embargo, cuando el muestreo es no uniforme admite frecuencias más altas que la de Nyquist (ver la discusión en Scargle 1982) y puede ocurrir que $M \gg N$.

Cuando las observaciones no son equidistantes, y no se dispone de una expresión analítica para la ecuación 5.5 se deben realizar simulaciones MonteCarlo para averiguar \mathcal{M} . En los casos estudiados en este trabajo encontramos que $\mathcal{M} \sim N$ parece dar el mejor acuerdo con las simulaciones MonteCarlo.

La ecuación (5.9) puede ser invertida para dar un límite de detectabilidad z_d correspondiente a un FAP_d especificado por el usuario, i.e. resolviendo:

$$FAP_d = \mathcal{F}(z_d) \quad (5.10)$$

5.1.1. Detectabilidad de 1 exoplaneta

Cumming (2004) presentó un procedimiento simple para estimar la detectabilidad de sistemas planetarios con un único planeta, dado el número de datos N y la razón K/σ como función de la probabilidad de falsa alarma (FAP) deseada. A pesar que la formulación original fue desarrollada para órbitas circulares y para grandes conjuntos de datos, sirve como punto de partida para la extensión que se desarrolla a continuación. A través de esta sección asumiremos que el tiempo observacional cubre al menos un período de las masas planetarias.

Trataremos de establecer la relación que existe entre un determinado FAP y su relación K/σ . Para ello primero vamos a identificar el residuo de la mejor solución ² del problema de ajustes como:

$$\hat{Q} = Q_{\min} \quad (5.11)$$

Este residuo representa la diferencia entre los datos de velocidad radial medida y el modelo,

$$\hat{Q} = \|V\| - \|\widehat{W}\| \quad (5.12)$$

donde $\|V\| \equiv (V, V)^\dagger$. En general si $W(t_i)$ son velocidades radiales computadas en los tiempos t_i , usando el modelo propuesto, tenemos que $\sum_1^N W(t_i) \neq 0$.

Si asumimos que los datos de velocidad radial son la representación de un sistema real \mathcal{W} más una variable normal que describe el error, $\epsilon_i \equiv N(0, \sigma_\epsilon^2)$, tenemos que:

$$V_i = \mathcal{W}(t_i) + \epsilon_i \quad (5.14)$$

donde tenemos como condición que los errores no se correlacionan con el modelo (i.e. $(\mathcal{W}, \epsilon) = 0$), entonces obtenemos $\|V\| = \|\mathcal{W}\| + \|\epsilon\|$ y, por lo tanto, reemplazando en (5.12) obtenemos:

$$\|\widehat{W}\| = \|\mathcal{W}\| + \|\epsilon\| - \hat{Q}; \quad (5.15)$$

²a lo largo de la sección vamos a usar el símbolo $\hat{}$ sobre una dada cantidad para indicar que es la estimación del mejor ajuste.

^{2†} prestar atención a los paréntesis ya que representan el producto interno:

$$(V, V) = \sum_{t_i=1}^N V(t_i)V(t_i). \quad (5.13)$$

Si las observaciones están perfectamente distribuidas,

$$\|\mathcal{W}\| = \sum \mathcal{W}_i^2 = N \langle \mathcal{W}^2 \rangle = \lim_{T \rightarrow \infty} \frac{N}{T} \int_0^T \mathcal{W}(t)^2 dt. \quad (5.16)$$

¿Cómo calcular la potencia de la señal en función de un dado modelo (i.e sus parámetros planetarios)? Recordemos que tenemos el modelo a que es una constante (i.e. $W_a = \beta_0$). Entonces podemos simplemente escribir $\|\widehat{W}_a\| = \alpha_0^2 = \langle V \rangle^2$. Si, además los datos dados son tales que $\langle V \rangle = 0$, entonces la potencia z (ver Ec. (5.3)) correspondiente a la solución del mejor ajuste es:

$$\widehat{z} = \frac{\nu_b}{\nu_a - \nu_b} \frac{\widehat{Q}_a - \widehat{Q}_b}{\widehat{Q}_b} = \frac{N - M_b}{M_b - 1} \frac{\|\widehat{W}_b\|}{\widehat{Q}_b}. \quad (5.17)$$

y utilizando la Ecuación (5.15) obtenemos:

$$\widehat{z} = \frac{N - M_b}{M_b - 1} \left(\frac{\|\mathcal{W}_b\| + \|\epsilon\|}{\widehat{Q}_b} - 1 \right). \quad (5.18)$$

considerando que $\|\epsilon\| = \sum_i \epsilon_i^2$ es un estimador de $(N - 1) \sigma_\epsilon^2$, donde σ_ϵ es la desviación estándar característica de los datos V_r , finalmente

$$\widehat{z} = \frac{N - M_b}{M_b - 1} \left(\frac{\|\mathcal{W}_b\| + (N - 1) \sigma_\epsilon^2}{\widehat{Q}_b} - 1 \right) \quad (5.19)$$

$$= \frac{\nu_b}{\nu_a - \nu_b} \left(\frac{\|\mathcal{W}_b\| + (N - 1) \sigma_\epsilon^2}{\widehat{Q}_b} - 1 \right), \quad (5.20)$$

recordando que \mathcal{W}_b es la señal generada por un sólo planeta. Como modelo a , usamos el promedio de los datos, que se asume han sido promediados a cero previamente. La expresión anterior puede ser simplificada para leer:

$$\widehat{z}_1 = \frac{\Lambda \sigma_\epsilon^2}{\widehat{Q}_b} \quad (5.21)$$

donde se introdujo el estadístico auxiliar

$$\widehat{z}_1 = \widehat{z} + \frac{\nu_b}{\nu_a - \nu_b} \quad (5.22)$$

y la constante

$$\Lambda = \frac{\nu_b}{\nu_a - \nu_b} \left(\frac{N \langle \mathcal{W}_b^2 \rangle}{\sigma_\epsilon^2} + N - 1 \right). \quad (5.23)$$

La única cantidad aleatoria en la Ec. (5.21) es \widehat{Q}_b . Se conoce de la teoría de regresión lineal (ver Kurth, 1967) que $\widehat{Q}_b / \sigma_\epsilon^2$ es una variable χ_ν^2 con $\nu = \nu_b$. Entonces el estadístico z_1 es una distribución chi-cuadrado inversa, y:

$$\text{Prob}(z > z_d) = \text{Prob}(z_1 > z_{1d}) = \text{Prob}\left(\frac{\Lambda}{\chi_\nu^2} > z_{1d}\right). \quad (5.24)$$

A partir de estas expresiones, podemos obtener el valor medio y la varianza de \widehat{z} como (e.g. Wackerly et al. 2003):

$$\begin{aligned} \langle \widehat{z} \rangle &= E(\widehat{z}) = \frac{\Lambda}{\nu_b - 2} - \frac{\nu_b}{\nu_a - \nu_b} \\ \sigma_{\widehat{z}}^2 &= D^2(\widehat{z}) = \frac{2\Lambda^2}{(\nu_b - 2)^2(\nu_b - 4)}. \end{aligned}$$

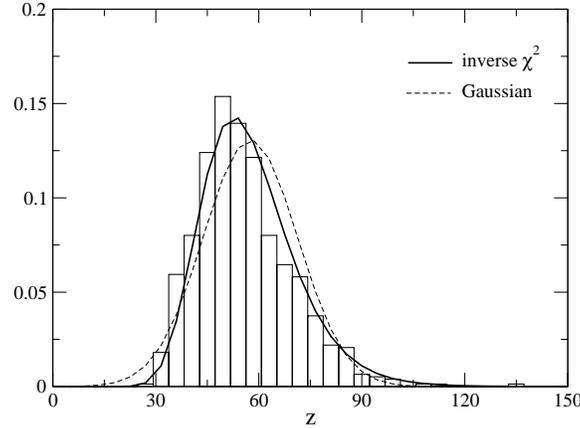


Figura 5.3: Histograma de los valores del mejor ajuste \hat{z} resultante de 774 conjuntos de datos aleatorios contruidos para un planeta único ($m_1 = 1M_{\text{Jup}}$), en una órbita circular con un período orbital de 100 días ($\lambda = [0, 2\pi]$, $N = 60$ y $K/\sigma = 3$). La curva sólida muestra la distribución chi-cuadrado inversa usando la Ec. (5.19), mientras que la curva a trazos corresponde a una distribución Gaussiana con la misma media y varianza. Todos los gráficos están normalizados de forma que el área encerrada por la curva sea uno.

Aproximando para valores grandes $N - M_b$, la desviación estándar de \hat{z} es dada por:

$$\sigma_{\hat{z}} \simeq \frac{\sqrt{2}N}{(M_b - 1)\sqrt{N - M_b}} \left(\frac{\langle \mathcal{W}_b^2 \rangle}{\sigma_\epsilon^2} + 1 \right). \quad (5.25)$$

Estos resultados pueden ser comparados con los dados por Cumming (2004)³. Si adoptamos las aproximaciones $\langle \mathcal{W}_b^2 \rangle = K^2/2$ y $\sigma_\epsilon = \sigma$, los resultados para $\langle \hat{z} \rangle$ no son muy diferentes de aquellos dados previamente y tienden a ser iguales para valores grandes de N . Sin embargo los resultados de Cumming para $\sigma_{\hat{z}}^2$ son significativamente menores que aquellos dados por la ecuación (5.25) (e.g. 9.8 en lugar de 13.7 en el ejemplo considerado en la Fig. 5.3). La Gaussiana correspondiente es entonces mucho más estrecha que la mostrada en la Fig. 5.3 y no reproduce satisfactoriamente la distribución de los valores de \hat{z} .

Para una dada amplitud de señal K , la distribución de \hat{z} nos da la dispersión de posibles potencias obtenidas de distintos conjuntos de datos con el mismo N y relación señal-ruido. Así, para un dado límite de detección z_d (o z_{1d}), podemos buscar el valor de los parámetros correspondientes a una dada probabilidad de detección P_{detect} . Esto es realizado resolviendo la ecuación:

$$\text{Prob} \left(\frac{\Lambda}{\chi_\nu^2} > z_{1d} \right) = P_{\text{detect}} \quad (5.26)$$

con respecto a z_{1d} . Adoptando un valor numérico para P_{detect} (por ejemplo 99%), las ecuaciones (5.24)-(5.25) brindan el valor necesario de Λ . Finalmente, a partir de la aproximación,

$$\langle \mathcal{W}^2 \rangle / \sigma_\epsilon^2 = K^2 / (2\sigma_\epsilon^2) \quad (5.27)$$

³Los resultados de Cumming (2004) predicen:

$$\langle \hat{z} \rangle = (\nu_b/2)(K^2/2\sigma_\epsilon^2) \quad \text{y} \quad \sigma_{\hat{z}}^2 = 2\langle \hat{z} \rangle.$$

podemos obtener una relación entre K/σ_ϵ y N para satisfacer este criterio. Vale la pena mencionar que los resultados así obtenidos son muy robustos y no dependen críticamente de la distribución de los valores de mejores ajustes \hat{z} (Giuppone et al. 2009).

5.1.2. Resultados de detectabilidad de 1 exoplaneta

Shen y Turner (2008) mostraron importantes resultados relacionados con la detección de 1 exoplaneta a partir de la generación de datos sintéticos. La idea básica consistió en suponer un planeta con parámetros planetarios ($K, T, e, \omega, \tau, V_0$) que orbita alrededor de una estrella de masa M_* . Generaron conjuntos de velocidad radial, añadiendo un ruido Gaussiano de dispersión σ como perturbación y definiendo la relación señal sobre ruido S/N como K/σ , donde K es la amplitud de la curva de V_r teórica y como hemos visto antes:

$$K = \left(\frac{2\pi G}{P} \right)^{1/3} \frac{m \cdot \sin(i)}{M_*^{2/3}} \frac{1}{\sqrt{1-e^2}}. \quad (5.28)$$

Por simplicidad σ se consideró igual para todas las observaciones. Además de K/σ , otra cantidad importante es el número de observaciones N . Obviamente mientras mayor sea N y $S/N \equiv K/\sigma$ es mejor la calidad de los datos.

La Figura 5.4 muestra los resultados obtenidos para la relación entre K/σ con los parámetros orbitales determinados en el ajuste orbital. La condición elegida supone una masa estelar $M_* = 1.1M_\odot$ y como elementos planetarios: $P=50$ días, $m_p=1 M_J$ y valores de excentricidad $e = 0.0 - 0.5$ en intervalos de $e = 0.1$. Los tiempos de observación fueron generados aleatoriamente de forma que imitan el efecto general de las observaciones reales. Cada conjunto de órbitas verdaderas fue utilizado para generar 5000 conjuntos artificiales de V_r con ω y τ elegidos aleatoriamente y con ruido Gaussiano determinado a partir del S/N asumido. Se minimizó χ^2_ν con el método Levenberg-Marquardt y se usaron los parámetros de la órbita real como condición inicial para acelerar el proceso. El intervalo de tiempo de las observaciones cubre 2 períodos y se simuló distintos conjuntos temporales desde $N=10$ hasta $N=100$ (aunque mostramos sólo el de $N = 20$). Se utilizó un $FAP = 10^{-2}$, que esencialmente se satisface para valores grandes de K/σ y N . Un resultado muy importante de los autores es que las excentricidades de los planetas pueden ser sobreestimadas incluso si tenemos una buena relación S/N (unos ejemplos son mostrados en la Figura 5.5).

Hemos repetido las simulaciones numéricas con nuestras herramientas, a fines de constatar los resultados expuestos por Shen y Turner (2008). Tomamos inclusive más cantidad de observaciones dentro del intervalo de tiempo a fin de asegurarnos que el efecto de sesgo observado fuera real. Para verificar algunas conclusiones fijamos la excentricidad $e = 0.3$ y $\omega = 0^\circ$, el valor de $K = 54.25$ es el resultante de utilizar la Ec. (5.28) con un planeta de $1 M_{Jup}$. Se generaron 200 conjuntos de V_r y sobre cada conjunto se realizó la minimización utilizando *Annealing*, cuya condición inicial tiene los parámetros orbitales nominales. Se fijaron cuatro valores para $K/\sigma = 2, 4, 6$ y 10 , por lo tanto para un valor fijo de K se obtiene σ . El valor de σ es la dispersión de error que se sumará a cada V_r .

Los resultados obtenidos se muestran en la Figura 5.6. En la izquierda se muestra la diferencia entre los valores obtenidos por el ajuste y el valor de la señal original. A la derecha la dispersión de las diferencias. Como primer punto a destacar es que tanto la media de las excentricidades (panel superior izquierdo) como la media de las amplitudes (y por ende masas individuales) son siempre sobreestimadas, incluso con alto número de observaciones (por ejemplo $N=100$). A medida que se aumenta la relación $S/N = K/\sigma$ los valores medios tienden al valor nominal (i.e. $\Delta_e, \Delta_K, \Delta_\omega \rightarrow 0$).

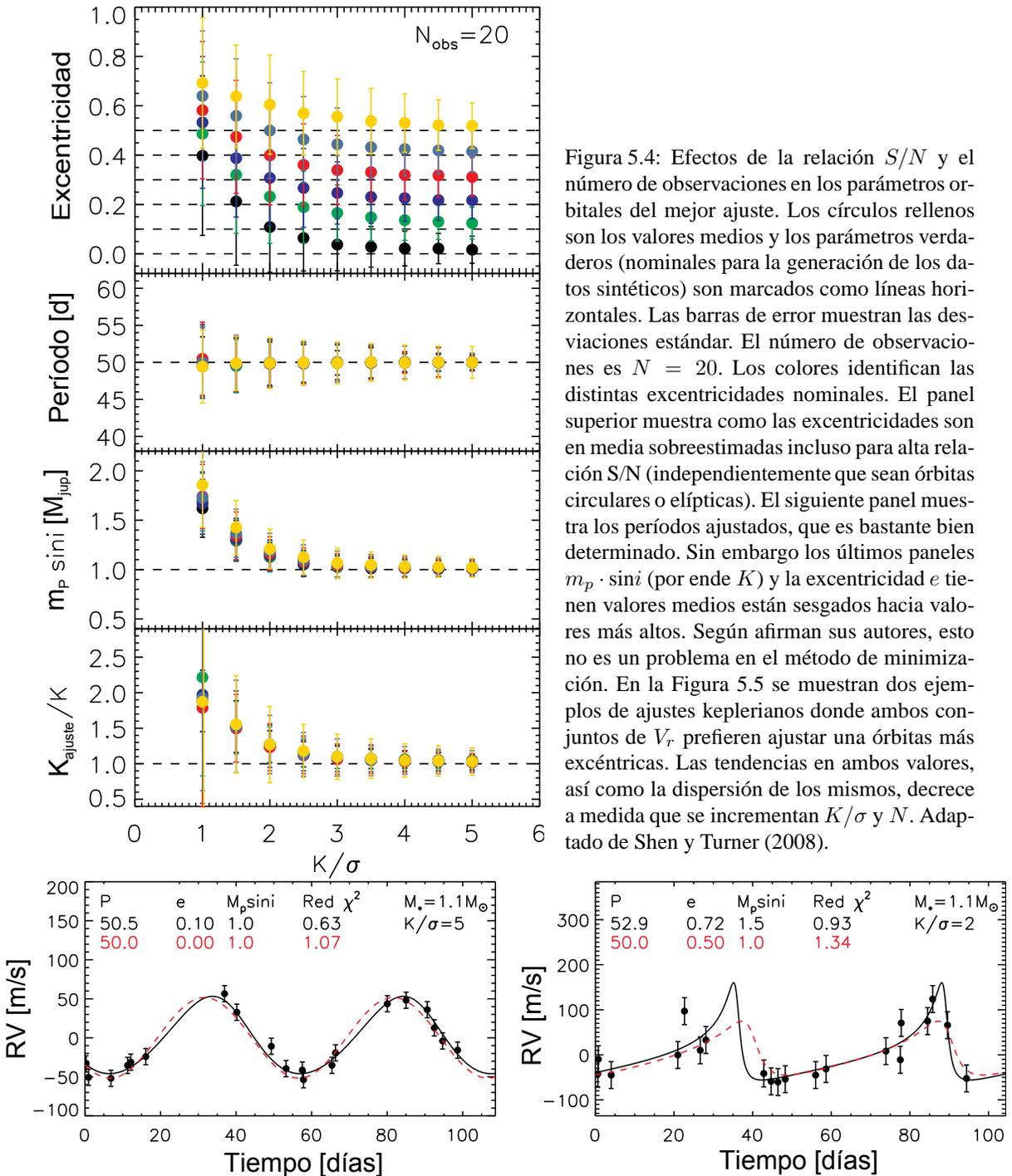


Figura 5.4: Efectos de la relación S/N y el número de observaciones en los parámetros orbitales del mejor ajuste. Los círculos rellenos son los valores medios y los parámetros verdaderos (nominales para la generación de los datos sintéticos) son marcados como líneas horizontales. Las barras de error muestran las desviaciones estándar. El número de observaciones es $N = 20$. Los colores identifican las distintas excentricidades nominales. El panel superior muestra como las excentricidades son en media sobreestimadas incluso para alta relación S/N (independientemente que sean órbitas circulares o elípticas). El siguiente panel muestra los períodos ajustados, que es bastante bien determinado. Sin embargo los últimos paneles $m_p \cdot \sin i$ (por ende K) y la excentricidad e tienen valores medios están sesgados hacia valores más altos. Según afirman sus autores, esto no es un problema en el método de minimización. En la Figura 5.5 se muestran dos ejemplos de ajustes keplerianos donde ambos conjuntos de V_r prefieren ajustar una órbita más excéntrica. Las tendencias en ambos valores, así como la dispersión de los mismos, decrece a medida que se incrementan K/σ y N . Adaptado de Shen y Turner (2008).

Figura 5.5: **Izquierda:** ejemplo de un conjunto de V_r artificial con $K/\sigma = 5$ y el número de observaciones $N = 20$ donde $e_0 = 0$. El ajuste determina una órbita excéntrica en lugar de la órbita circular, el período y la masa parecen bien determinados. **Derecha:** Un ejemplo de sobreestimación de la excentricidad esta vez con $K/\sigma = 2$ y $N = 20$ y $e_0 = 0.5$. Nuevamente la excentricidad es sobreestimada. En ambos casos la curva punteada corresponde a una curva de V_r construida con los parámetros orbitales nominales y las curvas continuas son creadas con el mejor ajuste orbital. Adaptado de Shen y Turner (2008).

También a medida que aumentamos el número de datos N las determinaciones de elementos orbitales mejora (ver Figura 5.7). El ángulo ϖ es bastante bien determinado en media, incluso su dispersión es

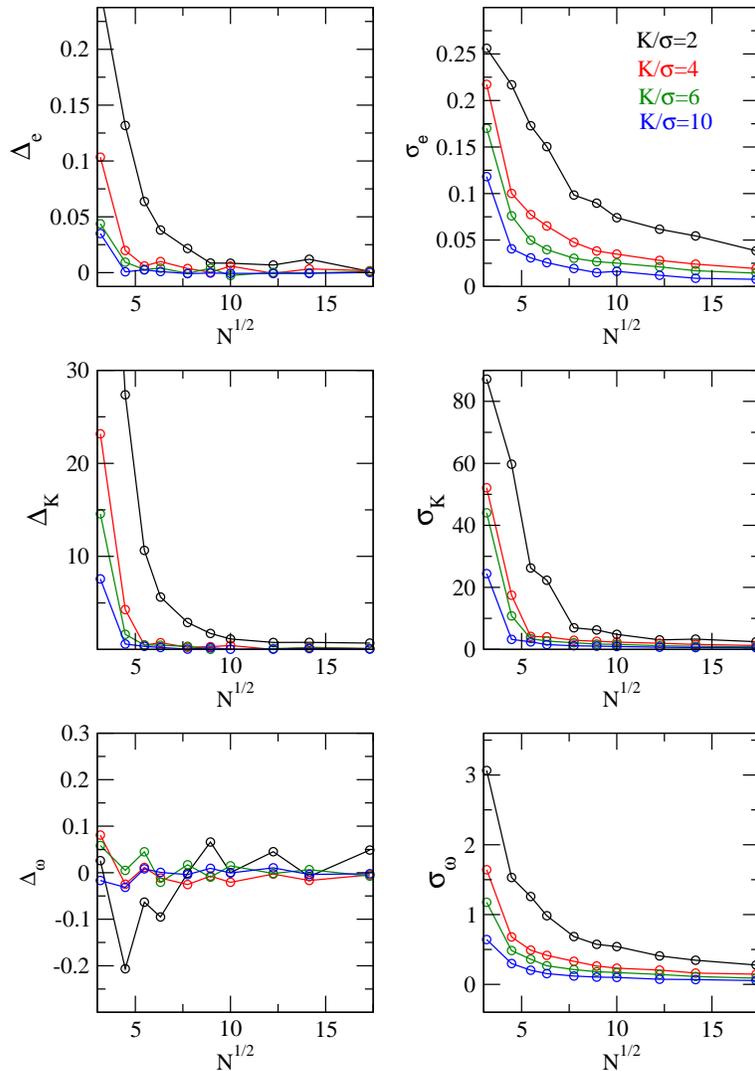


Figura 5.6: Resultados propios que muestran los efectos de la detección de un exoplaneta con la relación S/N y el número de observaciones N en función de los parámetros orbitales del mejor ajuste. Evidentemente existe una sobreestimación en todos los parámetros, más notable en la excentricidad cuando hay baja relación K/σ . El argumento de pericentro ω es el parámetro mejor determinado y pueden existir grandes diferencias en el valor de la masa planetaria (asociada a K).

de sólo 3° en el caso menos favorable (i.e. $K/\sigma = 2$).

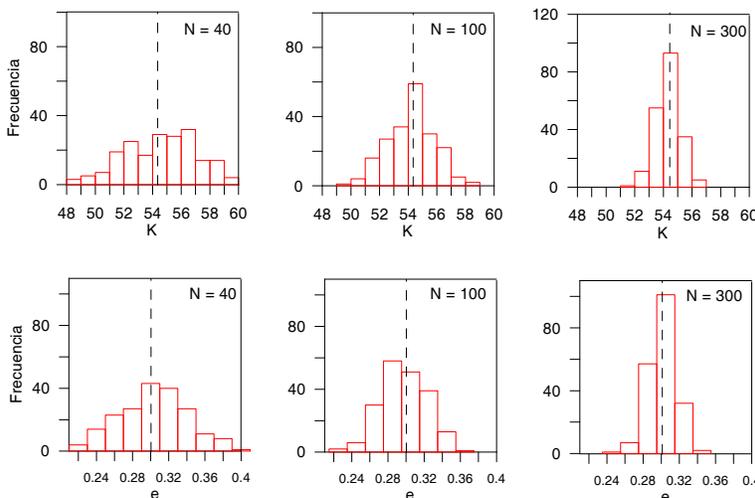


Figura 5.7: Histogramas de las soluciones orbitales encontradas con $K/\sigma = 6$, donde las soluciones nominales se marcaron con líneas punteadas. De izquierda a derecha: $N=40, 100, 300$. En la fila superior se aprecia que la amplitud de la señal (K) es pobremente identificada si se disponen de pocas observaciones ($N=40$). A medida que se aumenta la cantidad de observaciones la dispersión disminuye (lo mismo que se observa en la parte derecha de la Fig.5.6). El mismo efecto se observa en la fila inferior donde se encuentran los histogramas de la excentricidad.

5.2. Sistemas de exoplanetas

Volvamos a nuestro análisis de la Figura 5.1, pero tratando de obtener algo más de información. En la Figura 5.8 hemos identificado cada planeta individualmente de acuerdo a las masas:

	$m_p < 0.02 M_{\text{jup}}$	Supertierras (violeta)
$0.02 < m_p < 0.2 M_{\text{jup}}$		Neptunos (naranja)
$0.2 < m_p < 2 M_{\text{jup}}$		Júpiteres (celeste)
$m_p > 2 M_{\text{jup}}$		SuperJúpiteres (verde)

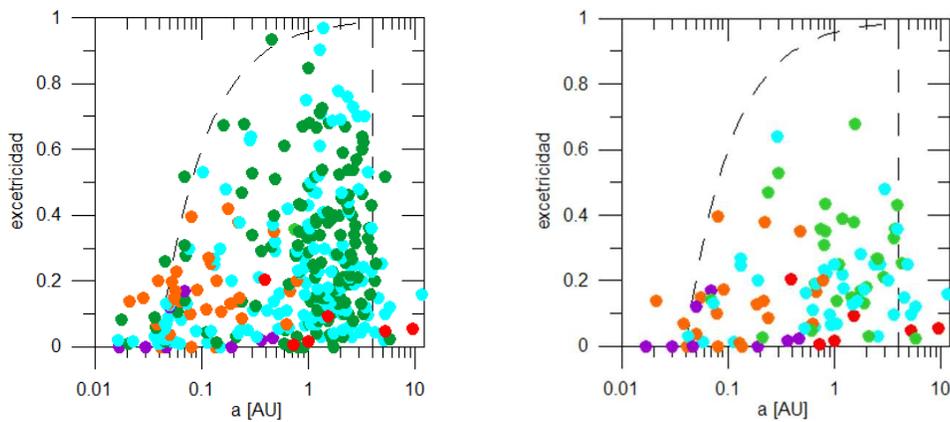


Figura 5.8: Distribución elementos orbitales de planetas detectados con velocidad radial, cada color representa un rango de masas y el color rojo es reservado para los planetas del Sistema Solar. Izquierda. Todos los exoplanetas detectados. Derecha. Exoplanetas en sistemas con 2 o más componentes. Se observa tendencia a descubrir planetas menos masivos más cerca de la estrella central. Aunque en principio no debería ser un efecto observacional ya que los planetas masivos y más cercanos a la estrella son más fáciles de detectar. La curva a trazos indica la línea de pericentro constante para planetas con período de 3 días, región en la cual se supone empiezan a ser más importantes los efectos tidales.

Gracias a la figura, vemos que existe cierto patrón en la distribución de puntos: los planetas menos masivos se ubican a menor distancia de la estrella, mientras que planetas más masivos se encuentran más lejos de la estrella; es decir parece que existe cierta jerarquía que relaciona la masa con la distancia. Más aún, los planetas más masivos y lejanos son en promedio más excéntricos. Esta distribución se hace mucho más evidente en sistemas de 2 o más exoplanetas (marco derecho de la figura) y no existen ideas claras de porque esta segregación en masas.

Otro tipo de análisis se aprecia en la Figura 5.9. En ella muestra la distribución de razones de masa en función de la razón de períodos, para todos los pares planetarios consecutivos. Todos los sistemas planetarios cercanos a la resonancia MMR 2/1 tienen razones de masa m_2/m_1 agrupados cerca del valor de la unidad.

A pesar que la población de exoplanetas se incrementa rápidamente, el número de sistemas con planetas en la resonancia MMR 2/1 parece haberse estancado durante algunos años. Aún cuando la resonancia MMR 2/1, es la más importante dinámicamente hablando y la más poblada de las conmensurabilidades, el número de sistemas planetarios está restringido a cuatro sistemas ampliamente estudiados: GJ876, HD82943, HD73526 y HD128311. Goździewski et al. (2007) propuso que HD160691 puede también tener dos planetas en la MMR 2/1, sin embargo esto no fue confirmado y esto fue

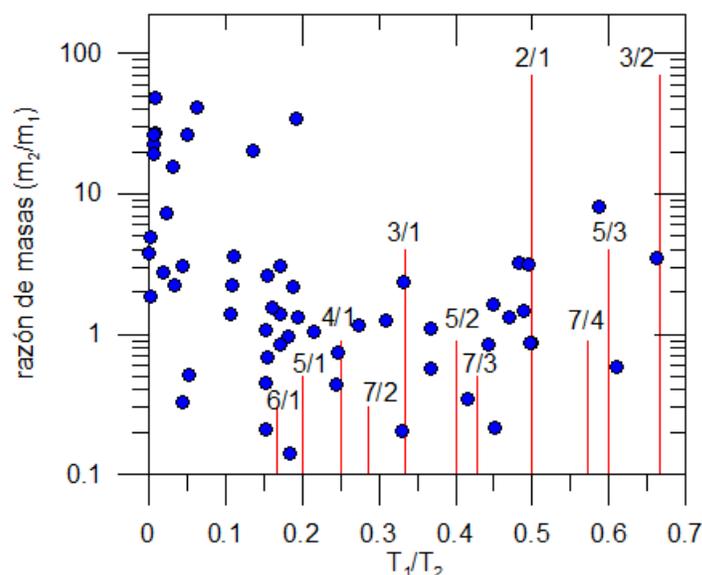


Figura 5.9: Distribución de pares consecutivos de planetas en sistemas extrasolares múltiples. Se muestra la razón de masa con respecto a la razón de períodos orbitales. Las resonancias de movimientos medios están indicadas con líneas rojas verticales cuyo largo es inversamente proporcional al orden de la MMR. Los datos corresponden a los ajustes orbitales hasta el 14 de abril del 2010.

propuesto sólo dos años después que el mismo sistema parecía favorecer una configuración en MMR 5/1.

Durante varios años se esperaba que aparecieran muchos planetas resonantes como consecuencia de migraciones planetarias de larga escala debido a interacciones con el disco de gas. De hecho la existencia de planetas resonantes fue utilizada varias veces como evidencia observacional de que estos tipos de migraciones ocurrieron y que muchos planetas se forman lejos de su actual localización. Las simulaciones hidrodinámicas parecen indicar que la captura en commensurabilidades de bajo orden (particularmente la MMR 2/1) son altamente probables. Esta conclusión es resultado de muchas simulaciones, con amplias condiciones de masas planetarias, condiciones iniciales y parámetros del disco. Varios trabajos centran su atención en el sistema GJ876 (e.g., Kley et al. 2005), concluyendo que la configuración presente de los planetas *b* y *c* puede ser explicada con este escenario.

Recientemente tres explicaciones se propusieron para explicar la actual ausencia de más planetas en esta resonancias:

- Una posibilidad es que no todos los planetas que se aproximen a la resonancia 2/1 puedan ser capturados. Si las razones de masas entre el planeta exterior y el interior son lo suficientemente pequeñas (del orden de la razón entre Saturno y Júpiter), entonces un decaimiento muy rápido del cuerpo exterior (e.g. migración tipo III) puede evitar la captura en 2/1 y producir una posterior captura en la MMR 3/2 (Masset y Snellgrove 2001). Una explicación similar fue propuesta por Morbidelli y Crida (2007) como primer paso para explicar la actual arquitectura de nuestro propio Sistema Solar

Sin embargo no se sabe claramente bajo que circunstancias ocurre una migración rápida (ver D'Angelo y Lubow 2008). Alternativamente, los efectos de turbulencia en los discos gaseosos pueden posiblemente causar perturbaciones orbitales significantes en los cuerpos que van cayendo y así inhibir la captura de la resonancia (Adams et al. 2008).

- Una segunda posibilidad puede recaer en la baja tasa de sobrevivencia de los planetas resonantes durante su evolución dentro de la conmensurabilidad. Como fue mostrado originalmente por Lee y Peale (2002), una vez dentro del dominio resonante, las interacciones tidales con el disco gaseoso incrementarán las excentricidades de los cuerpos planetarios hasta que una colisión o eyección de alguno de los cuerpos ocurra. Esto es debido a la topología de la familias ACR estables en el dominio de excentricidades (e.g., Beaugé et al. 2006, Hadjidemetriou 2008) dentro de la MMR 2/1. La única forma que dos planetas resonantes puedan sobrevivir migraciones orbitales de larga escala es que el mecanismo responsable introduzca un significativo amortiguamiento de las excentricidades orbitales, que produzcan valores de equilibrio de estos elementos comparados con los valores observados (Kley et al. 2005, Beaugé et al. 2006 y aplicado también en la Sección 4.2).

A pesar que se espera este efecto para parámetros de disco normales, especialmente si se asume un disco interno (Crida et al. 2008), no hay evidencia que el efecto sea verdadero en todos los casos. Quizás la mayoría de los sistemas que originariamente estaban atrapados en la resonancia 2/1 pueden haber sido eyectados. Moorhead y Adams (2005) propusieron este mecanismo para explicar la actual distribución de semiejes y excentricidades de los exoplanetas.

- Una tercera posibilidad, no considerada previamente, es que la aparente falta de planetas resonantes podría ser atribuida a limitaciones de detectabilidad. Recordemos que el movimiento resonante causa una casi periódica repetición de la curva de velocidad radial de la estrella, que bajo ciertas circunstancias, puede no permitir una correcta separación de las componentes (Giuppone et al. 2009). Quizás varios sistemas más puedan actualmente encontrarse en dentro de la resonancia 2/1 pero no son discernibles debido a limitaciones en las observaciones o pequeña relación señal/ruido.

Recientemente Anglada-Escudé et al. (2009) discutieron los casos donde dos planetas resonantes con órbitas casi circulares pueden ser confundidos con un único planeta en una órbita elíptica de excentricidad e más alta. Sin embargo, este efecto parece ser posible sólo en casos donde el planeta exterior resonante es mucho más masivo que el compañero interior y la solución de sólo un planeta tiene baja excentricidad (e.g. $e < 0.2$). En nuestro trabajo, construimos un criterio de detectabilidad válido para cualquier relación de masas y que no sea restringido a órbitas cuasi-circulares.

A continuación utilizaremos las herramientas expuestas en la Sección 5.1, mostrando resultados obtenidos en nuestro trabajo (Giuppone et al. 2009).

5.2.1. Detectabilidad de 2 Planetas Resonantes

Consideramos la presencia de un sistema extrasolar, con dos planetas de masa m_1 y m_2 orbitando una estrella de masa M_* con semieje mayor a_i , excentricidades e_i , longitudes medias λ_i y longitudes de pericentro ϖ_i . El índice $i = 1$ será usado para el planeta interno, mientras que el $i = 2$ denotará el externo (i.e. $a_1 < a_2$). Asumimos que ambos planetas se encuentran en el mismo plano orbital y orientado en la dirección del observador (edge-on).

Se eligieron configuraciones de interés y a partir de ellas se generaron curvas sintéticas que imitan el movimiento de la estrella alrededor del baricentro del sistema. Esta curva es resultado de la superposición de dos señales periódicas, cada una de las cuales tiene una amplitud K_j relacionada al j -ésimo planeta. Cada conjunto generado solo cubre pocos períodos orbitales, por lo tanto podemos asumir que no existen efectos significativos provenientes de perturbaciones gravitacionales mutuas.

Planeta	masa $M_j up$	a UA	e e	M \odot	ϖ \odot
1	1.0	0.6298	0.4137	0.0	0.0
2	3.0	1.0	0.0946	0.0	0.0
1	1.0	0.6299	0.3923	0.0	0.0
2	1.0	1.0	0.1271	0.0	0.0
1	1.0	0.6299	0.1105	310.4	115.9
2	0.33	1.0031	0.3840	65.30	0.0

Tabla 5.1: Parámetros orbitales para tres soluciones ACR en la resonancia 2/1 MMR para distintas relaciones de masa m_2/m_1 . La masa central es $M_* = 1M_\odot$.

Una vez que la curva es generada se construyó un muestreo discreto, tomando N tiempos observacionales t_i distribuidos aleatoriamente de acuerdo a una distribución homogénea. En cada punto se calculó el valor de RV como un desplazamiento aleatorio del valor nominal $V_r(t_i)$; este desplazamiento sigue una distribución Gaussiana con varianza constante σ^2 . Los datos sintéticos resultantes constan de tres columnas (t_i, V_{r_i}, σ) y se usaron como entrada en nuestro método de ajuste.

Consideramos 2 planetas con amplitud de oscilación pequeña alrededor de la resonancia MMR 2/1, para estimar como las incertezas en el ajuste orbital afectan el movimiento observado del sistema. Como ejemplo se eligieron tres configuraciones distintas, cada una de ellas correspondientes a una ACR estable con distintas relaciones de masa m_2/m_1 . Las primeras dos corresponden a ACR del tipo $(0,0)^4$, mientras que la última muestra un comportamiento corrotacional asimétrico (el centro de oscilación de σ_1 y $\Delta\varpi$ es cualquiera menos 0° o 180°). Las masas y elementos orbitales son dados en la Tabla 5.1. En todos los casos la integración muestra pequeñas amplitudes de oscilación (aproximadamente 5°) alrededor de su respectivo ACR.

Ejemplos típicos de conjuntos de datos sintéticos con $\sigma = 10$ m/s se muestran en la Figura 5.10. Aunque todos los conjuntos de datos generados son construidos con la misma desviación estándar σ , la razón S/N aparece distinta para cada relación de masa.

Para entender por que ocurre esto escribimos la amplitud total de la curva de velocidad radial $\text{amp}(V_r) = K_1 + K_2$. Expresando K_i en términos de la masa y elementos orbitales (ver, e.g., Beaugé et al. 2007), para un sistema coplanar orientado en dirección al observador:

$$\text{amp}(V_r) = K_1 + K_2 = \frac{m_1}{M_*} n_1 a_1 + \frac{m_2}{M_*} n_2 a_2, \quad (5.29)$$

donde n_i son los movimientos medios, considerando hasta segundo orden en las masas y para órbitas cuasi-circulares. Para planetas en la vecindad de la resonancia MMR 2/1, podemos simplificar la expresión como:

$$\text{amp}(V_r) \simeq \left(1 + \sqrt[3]{\frac{1}{2} \frac{m_2}{m_1}} \right) \frac{m_1}{M_*} n_1 a_1, \quad (5.30)$$

donde $(n_2/n_1)(a_2/a_1) \simeq \sqrt[3]{1/2}$. La cantidad dentro de los corchetes representa el incremento en la amplitud de RV debido a la presencia del planeta exterior (resonante). Este término tiende a la unidad para $m_2 \rightarrow 0$, y muestra dependencia lineal con la razón de masa. Así, para valores grandes

⁴Recordemos que una ACR es una solución periódica estable y que el valor entre paréntesis indica el centro de oscilación de los ángulos resonantes $\sigma_1 = 2\lambda_2 - \lambda_1 - \varpi_1$ y diferencia de longitudes de pericentro $\Delta\varpi = \varpi_2 - \varpi_1$.

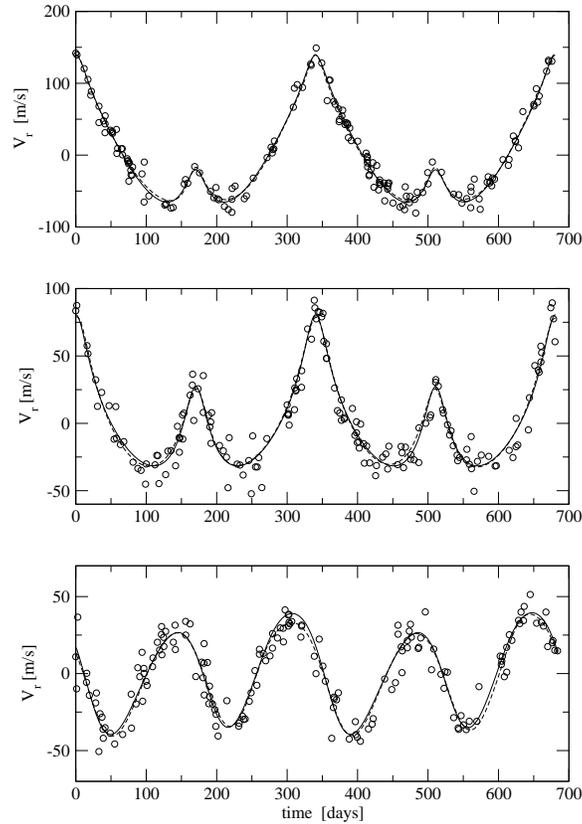


Figura 5.10: La curva continua negra muestra las curvas sintéticas de RV para las tres configuraciones ACR presentadas en la Tabla 5.1. La relación de masa decrece desde arriba ($m_2/m_1=3/1$) hacia abajo ($m_2/m_1=1/3$). Los círculos abiertos son $N = 165$ puntos ficticios con distribución temporal aleatoria y desviación estándar $\sigma = 10$ m/s.

de m_2/m_1 se obtiene una señal RV mayor, resultando una mayor relación S/N, todo esto asumiendo una desviación estándar fija σ igual para todos los datos.

La Ecuación (5.30) también es un indicador preliminar de que planeta domina la señal de RV. La razón de masa crítica es dada por $m_2/m_1 \simeq \sqrt[3]{2} \simeq 1.26$. Para valores menores, la amplitud RV del planeta interior m_1 es mayor, y la curva RV aparece como una señal perturbada con el período primario igual al período del planeta interior (i.e. $2\pi/n_1$). En la parte inferior de la Figura 5.10 se muestra un ejemplo. Por otro lado si la razón de masa es más grande que $\simeq 1.26$, la señal del planeta exterior se vuelve más importante y el período dominante en la curva de RV es dado por $2\pi/n_2$ (ver el gráfico superior en la Fig. 5.10). El gráfico intermedio representa la región de transición en la cual ambas componentes tienen magnitudes similares.

Para estimar la detectabilidad de dos planetas en la resonancia MMR 2/1, asumimos que el cuerpo que genera la mayor amplitud de RV es ya conocido, y queremos calcular bajo cuales condiciones la señal del segundo cuerpo también satisface un determinado límite de detectabilidad. Si volvemos a observar los gráficos en la Figura 5.10 vemos que no existe una separación única de las señales. Cuando la razón de masa m_2/m_1 es grande, la amplitud RV del planeta exterior es también grande (implicando $K_2 > K_1$), y la detectabilidad del par resonante se reduce a discernir la existencia del planeta interior. En el caso opuesto, donde $K_1 > K_2$, el planeta interior es el que domina la curva de RV y la existencia del cuerpo exterior debe ser deducida de la diferencia entre máximos consecutivos.

Extendiendo las ecuaciones de la Sección 5.2.1 para órbitas elípticas, podemos escribir la razón entre las amplitudes de ambos planetas como:

$$\frac{K_2}{K_1} = \sqrt[3]{\frac{1}{2} \frac{m_2}{m_1} \sqrt{\frac{1-e_1^2}{1-e_2^2}}}, \quad (5.31)$$

mostrando que el valor deducido anteriormente $m_2/m_1 \sim \sqrt[3]{2}$ todavía provee un valor crítico cualitativo para $K_1 \sim K_2$, cuando ambos planetas tienen excentricidades similares.

5.2.2. Planeta Interior más masivo

Empezaremos discutiendo el caso para un planeta interior más masivo. Aquí $K_1 > K_2$, y se desea estimar el límite de detectabilidad de hallar m_2 asumiendo m_1 conocido. Podemos estimar exactamente el mismo criterio que para sistemas con 1 planeta, donde el número de parámetros libres ahora es $M_b = 11$ y $M_a = 6$, y reescribir (5.19) como:

$$\hat{z} = \frac{\nu_b}{\nu_a - \nu_b} \left(\frac{\|\mathcal{W}_2\| + (N-1)\sigma_\epsilon'^2}{\hat{Q}_b} - 1 \right). \quad (5.32)$$

Si asumimos que $\sigma'_\epsilon \simeq \sigma$ (para ver la deducción puede consultarse el anexo de Giuppone et al. 2009) y el procedimiento resultante es el mismo deducido en la Sección 5.1.1. Es posible entonces estimar el valor límite de K_2/σ , como función del número de puntos N necesarios para que el pico resultante de potencia z sobrepase un valor preestablecido de FAP, con probabilidad igual a P_{detect} . Debido a que:

$$\frac{K_2}{\sigma} = \left(\frac{K_2}{K_1} \right) \left(\frac{K_1}{\sigma} \right) \quad \text{o} \quad \frac{K_1}{\sigma} = \left(\frac{K_1}{K_2} \right) \left(\frac{K_2}{\sigma} \right), \quad (5.33)$$

podemos encarar este problema también argumentando que la condición de detectabilidad depende de N y las razones K_1/σ y K_2/K_1 . Aplicando la ecuación (5.31) podemos reescribir la condición de K_2/K_1 en una que relacione m_2/m_1 . Debido a que $K_2 < K_1$, el procedimiento se reduce a encontrar que valores de m_2/m_1 (para excentricidades dadas e_i) que reducen el valor de K_2/σ a magnitudes que no son detectables.

La razón K_1/σ depende de la desviación estándar de los datos, la masa estelar así como también los parámetros orbitales y masa del planeta interno. Por ejemplo, cuando se tienen grandes valores de m_1 se incrementa la amplitud de K_1 , permitiendo la detección de pequeños valores de m_2/m_1 para un valor fijo σ .

Llamemos K_{10} a la amplitud RV generada por un planeta del tamaño de Júpiter ($m_1 = M_{Jup}$) ubicado a $a_1 = 1$ UA y con excentricidad e_1 . Debido a que $K_1 \propto n_1 a_1 (m_1/M_*)$, para cualquier otra masa y semieje mayor la razón de K_1/σ puede ser escrita como

$$\frac{K_1}{\sigma} = \frac{K_{10}}{\sigma^*}, \quad (5.34)$$

Donde la desviación estándar “escalada” σ^* es relacionada al valor nominal σ con

$$\sigma = \sigma^* \left(\frac{m_1}{M_{Jup}} \right) \left(\frac{a_1}{1\text{UA}} \right)^{-1/2} \left(\frac{M_*}{M_\odot} \right)^{-1}. \quad (5.35)$$

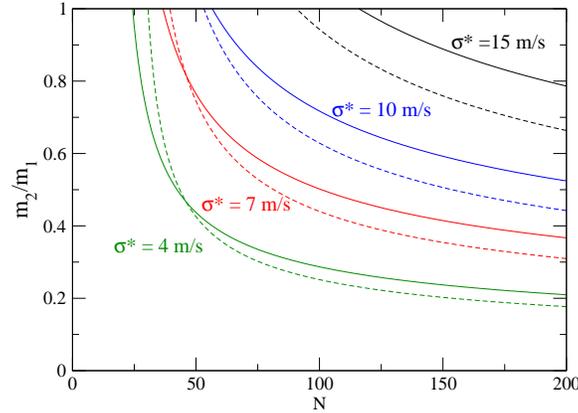


Figura 5.11: Valores mínimos de razón de masa m_2/m_1 necesarios para detectar el planeta resonante exterior menor (con $P_{detect} = 0.99$), como función del número de datos N para cuatro valores distintos de la desviación estándar escalada (σ^*). Ambas excentricidades se escogieron como $e_i = 0.1$. El valor del FAP fue tomado igual a $\mathcal{F} = 10^{-4}$. Las líneas continuas corresponden a predicciones con la distribución **chi-cuadrada inversa** (5.24), mientras que las curvas a trazos es obtenida con la aproximación Gaussiana con iguales valores de media y varianza.

Finalmente, con las Ecuaciones (5.31) , (5.33) y (5.34) podemos escribir

$$\frac{K_2}{\sigma} = \sqrt[3]{\frac{1}{2}} \sqrt{\frac{1 - e_1^2}{1 - e_2^2}} \frac{m_2}{m_1} \left(\frac{K_{10}}{\sigma^*} \right). \quad (5.36)$$

Esta ecuación nos permite estimar el mínimo valor de m_2/m_1 para la detectabilidad del planeta exterior resonante como función de N y σ^* , para una dada excentricidad e_i . Si asumimos que el planeta interno corresponde a un planeta de masa joviana con distancia de 1 UA, σ^* es igual a la desviación estándar de los datos RV σ . Para otras masas internas, σ puede ser calculada con (5.35).

La Figura 5.11 muestra los límites de detectabilidad para $m_2/m_1 < 1$ para varios valores de σ^* en función del número de puntos N . Cada curva brinda un valor mínimo de razón de masa m_2/m_1 necesaria para obtener un FAP de $\mathcal{F} = 10^{-4}$. La masa estelar fue fijada como $M_* = 1M_\odot$ y ambas excentricidades en 0.1. Las líneas continuas fueron construidas usando la distribución **chi-cuadrado inversa** y la (5.24), mientras que la punteada denotan la aproximación con una distribución Gaussiana con la misma media y varianza. Cualitativamente ambos modelos producen resultados similares, sin embargo la expresión (5.24) produce límites de detectabilidad más restrictos.

Para $\sigma^* = 15$ m/s, prácticamente no se detectaron sistemas con razón de masa m_2/m_1 menores que ~ 0.6 , incluso para conjuntos de datos que contenían $N = 200$ observaciones de RV. Para $N < 60$ prácticamente ningún sistema con este σ^* es detectable (para cualquier razón de masa). El panorama mejora con menores valores de la desviación estándar escalada, hasta valores de $\sigma^* = 4$ m/s. Es posible detectar sistemas con razón de masas cercanos a dos décimas, siempre que el número de datos sea suficientemente grandes.

Es importante mencionar que el valor de σ considerado en estos cálculos no es solamente el deducido con las técnicas observacionales, sino el valor de los errores totales que incluye el jitter estelar y la posible incompletitud del modelo de 2 planetas. Los valores típicos de jitter estelar son del orden de 4 – 7 m/s (Wright 2005) y las incertezas de los datos que actualmente son disponibles son de similar magnitud. La suma de los cuadrados de ambos, implica valores típicos para σ en estos

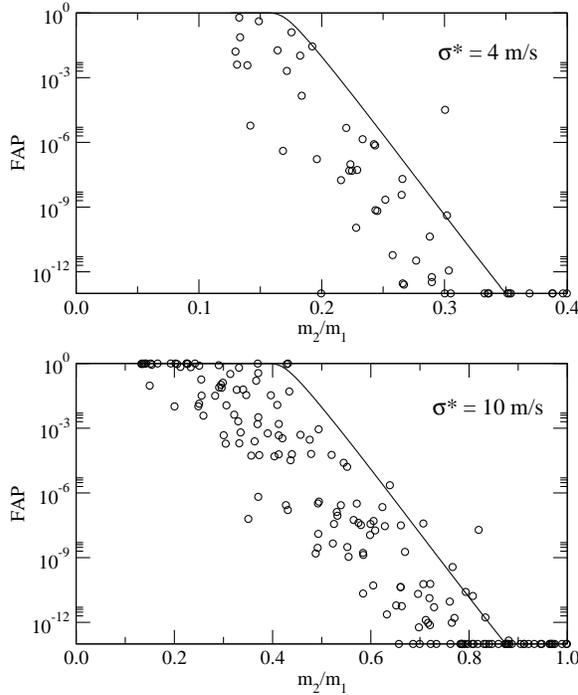


Figura 5.12: Detectabilidad de dos planetas en la resonancia MMR 2/1. Los círculos abiertos muestran simulaciones MonteCarlo para distintos conjuntos sintéticos. El gráfico muestra la probabilidad de falsa alarma (FAP) para valores aleatorios de razón de masas m_2/m_1 . El número total de puntos utilizado fue $N = 100$, distribuidos aleatoriamente sobre cuatro períodos orbitales del cuerpo exterior. La curva continua muestra el valor de predicción usando nuestro modelo con $P_{detect} = 0.99$.

sistemas del orden de 6 – 10 m/s. Una buena estimación de σ puede ser obtenido del valor de rms resultantes del ajuste orbital (Shen y Turner 2008) que proporciona un $\sigma \sim 7 - 10$ para la mayoría de los sistemas multiplanetarios.

Para chequear los límites de precisión obtenidos con este modelo, realizamos una serie de simulaciones MonteCarlo con series de datos sintéticos, cada uno con $N = 100$ observaciones y cubriendo un total de 4 períodos del cuerpo resonante exterior y con razones de masa en el intervalo $m_2/m_1 \in [0.1, 1.0]$. El planeta interior fue fijado con la misma masa y parámetros orbitales usados en la Fig. 5.11 ($m_1 = 1M_{Jup}$, $a_1 = 1$ UA). Una serie de simulaciones fue realizada adoptando $\sigma^* = 4$ m/s, mientras que una segunda corrida fue realizada considerando $\sigma = 10^*$ m/s. Para cada conjunto de datos determinamos el ajuste orbital de dos planetas y calculamos el valor del FAP correspondiente al segundo planeta, utilizando el mismo método que para sistemas de un planeta. Los resultados se muestran en la Figura 5.12, donde los círculos abiertos muestran la relación entre la razón de masa y los respectivos valores de FAP para todos los conjuntos de datos. Los valores de ordenada menores que 10^{-13} fueron igualados a este límite inferior. Finalmente la curva continua muestra los límites de detectabilidad analítica obtenidos a partir del modelo.

Estos gráficos deben ser interpretados de la siguiente manera: para un dado valor de razón de masas m_2/m_1 , todos los conjuntos de datos sintéticos dan ajustes orbitales con FAP menor que la línea continua. Así, de manera de asegurar la detectabilidad, el FAP correspondiente a una dada razón de masas debe ser menor que este límite. El acuerdo entre la curva analítica y las simulaciones numéricas es muy bueno, indicando que el modelo simplificado descrito anteriormente es más que adecuado para predecir la detectabilidad de sistemas múltiples planetarios.

5.2.3. Planeta Exterior más masivo

Se pueden realizar los mismos cálculos que en la sección precedente pero para sistemas con $K_2/K_1 > 1$. En tales casos el criterio de detectabilidad sigue el mismo esquema, pero ahora el pla-

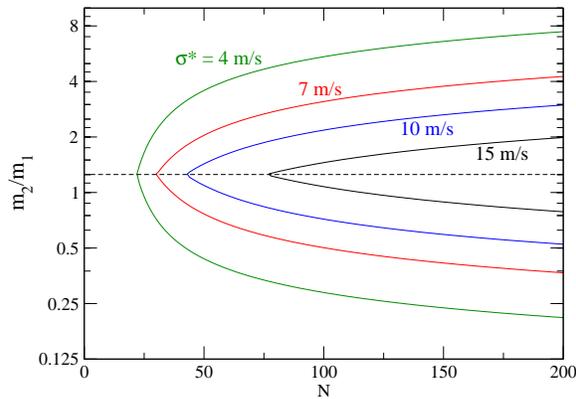


Figura 5.13: Al igual que la Fig. 5.11, pero incluyendo razones de masa $m_2/m_1 > 1$. Se muestran sólo los resultados construidos con la distribución chi-cuadrado inversa (5.24). Para un dado valor de σ^* , la detectabilidad es garantizada sólo dentro de las dos curvas correspondientes.

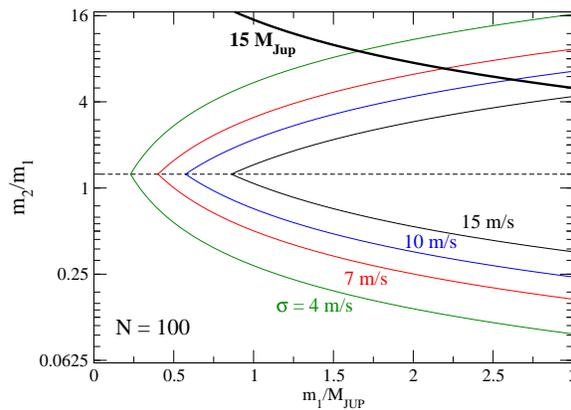


Figura 5.14: Límites de detectabilidad para razones de masa m_2/m_1 (con $P_{detect} = 0.99$ y $N = 100$), como función de la masa del planeta interno m_1 , para cuatro valores distintos de desviación estándar σ . Ambas excentricidades planetarias se eligieron como $e_i = 0.1$. El FAP fue tomado igual a $\mathcal{F} = 10^{-4}$. Como antes, para un dado valor de σ , la detectabilidad es garantizada dentro de las dos curvas correspondientes. La línea negra gruesa marca el límite $m_2 > 15M_{JUP}$, un valor utilizado normalmente para asociar el límite de enanas marrones (ver Sección 1.2).

neta exterior es el responsable del pico dominante, mientras que la señal que define la detectabilidad del sistema es K_1 . Consecuentemente la ecuación (5.33) debe ser invertida y nos brindará el mínimo valor necesario de K_1/K_2 tal que el valor K_1/σ sea todavía detectable. La dependencia con las masas individuales y los semiejes puede ser sorteada nuevamente definiendo la desviación estándar escalada σ^* como en la ecuación (5.35). En este caso hemos mantenido la masa y semieje mayor del planeta interior como parámetros de escala. Esta es una cuestión de elección, pero ayuda a mantener cierta homogeneidad, independientemente de la razón de masas en consideración.

La Figura 5.13 muestra el límite de detectabilidad combinando resultados para todas las razones de masa. Para un dado σ^* , todos los sistemas planetarios detectables se encuentran dentro del par de curvas correspondientes. Se puede apreciar que ambos conjuntos de curvas se intersectan para $m_2/m_1 \sim \sqrt[3]{2}$, como se esperaría de la condición $K_2 = K_1$. Finalmente, existe una variación dependiente de las excentricidades orbitales, originada por la relación entre la amplitud de la señal y la masa planetaria (ver ecuación (5.31)). Sin embargo, cualitativamente, la imagen continúa siendo la misma, excepto para muy bajas desviaciones estándar. En este último caso aparentemente es muy difícil detectar sistemas planetarios resonantes con razones de masa m_2/m_1 mayores que $\simeq 4$ o menores que $\simeq 0.3$. Es importante recalcar que esto no significa de que no hay manera de detectar sistemas fuera de estos límites, sino que la detectabilidad no estará garantizada para cualquier conjunto de datos y dependerá específicamente de la distribución temporal y/o los errores individuales de los datos.

El análisis previo asume que $m_1 = M_{Jup}$, $a_1 = 1$ UA y $M_* = M_\odot$, por lo cual $\sigma^* = \sigma$. Otros valores sin embargo cambiarían el panorama. Sobre la población de planetas extrasolares conocidos, quizás el caso más extremo sea el sistema GJ876 ($m_2/m_1 \sim 3.1$), donde $m_1 = 0.6M_{Jup}$, $a_1 = 0.13$ UA y $M_* = 0.32M_\odot$ (Butler et al. 2006). Aplicando la ecuación (5.35) a los resultados mostrados en la Figura 5.13, encontramos que para $N = 100$ el sistema resonante puede ser identificado incluso

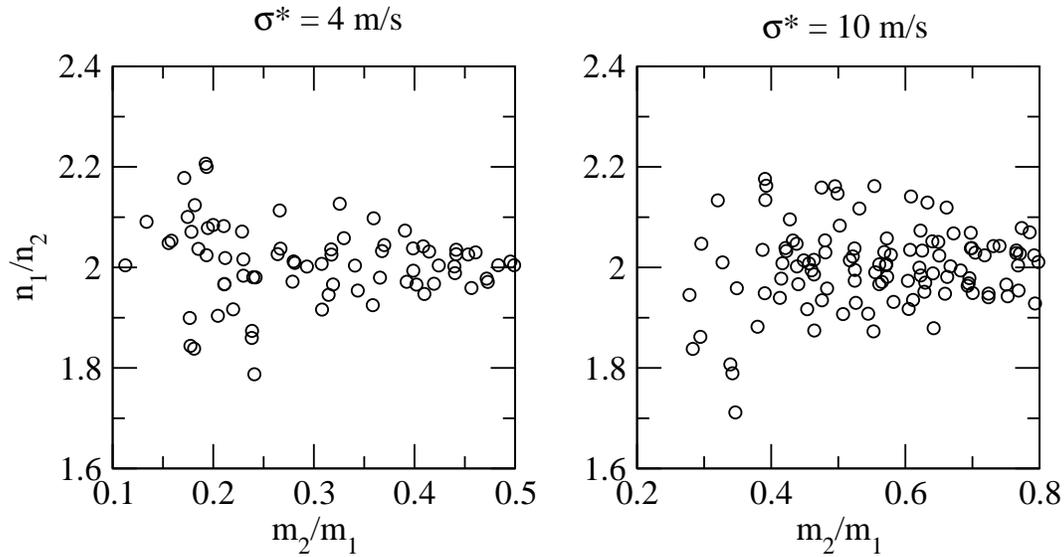


Figura 5.15: Distribución de la razón de movimientos medios n_1/n_2 como función de la razón de masas planetarias m_2/m_1 , de las simulaciones MonteCarlo usadas en la Fig. 5.12. Sólo los sistemas con $\mathcal{F} < 10^{-4}$ fueron dibujados.

para desviaciones estándar en los datos RV del orden de $\sigma \sim 40$ m/s (correspondientes a $\sigma^* \simeq 8$ m/s).

La Figura 5.14 muestra como el límite de detectabilidad para m_2/m_1 varía con la magnitud de m_1 , ahora asumiendo un conjunto fijo de datos $N = 100$. Tanto a_1 como M_* se mantienen en sus valores originales. Una vez más el sistema resonante es detectado siempre que la razón de masas m_2/m_1 se encuentre dentro de ambas curvas de igual color. Podemos ver que la detectabilidad se incrementa significativamente con m_1 , permitiendo un rango más amplio de razones de masa que los previamente mostrados. Sin embargo la posibilidad de un sistema resonante muy masivo debe ser considerada con sumo cuidado, pues las perturbaciones mutuas pueden comprometer la estabilidad dinámica del sistema.

Una de las predicciones del modelo es justificar la dificultad de detectar sistemas resonantes con masas significativamente diferentes, estando este resultado de acuerdo con los sistemas actualmente conocidos. Recordemos que la Figura 5.9 muestra la distribución de razones de masa en función de la razón de períodos para todos los pares planetarios consecutivos. Todos los sistemas planetarios cercanos a la resonancia MMR 2/1 tienen razones de masa m_2/m_1 agrupadas cerca del valor de la unidad, de acuerdo a nuestras predicciones con el presente modelo.

5.2.4. Estimación de errores en los parámetros ajustados

Incluso si el sistema resonante de 2 planetas es detectado, no hay garantía que las masas y los parámetros orbitales se puedan estimar con alguna precisión. Como ejemplo en la Figura 5.15 se muestra la distribución de razones de movimientos medios para aquellos conjuntos de datos sintéticos dados en la Fig. 5.12 que satisfacen el criterio de detectabilidad. Se aprecia una significativa dispersión en los movimientos medios alrededor del valor nominal $n_1/n_2 = 2$, con un porcentaje apreciable de sistemas que serían clasificados como casi-resonantes, pero no atrapados en la MMR. Este efecto presenta una dificultad adicional en la detección de planetas para razones de masa m_2/m_1 muy diferentes

a la unidad.

Estos resultados fueron obtenidos con conjuntos de datos sintéticos que cubren cuatro períodos del planeta exterior. A pesar de que el criterio de detectabilidad es independiente del intervalo de tiempo observacional (siempre que podamos garantizar una buena cobertura de las fases), la precisión de los elementos orbitales debería mejorar con intervalos de tiempo más largos. No obstante, incluso si se identifican correctamente como dentro de la resonancia, los elementos orbitales pueden tener errores significantes y afectar la dinámica deducida de los sistemas ajustados. Para cuantificar este efecto analizamos en los tres casos de la Tabla 5.1: uno con razón de masa mayor que la unidad, igual a la unidad y menor que la unidad.

5.2.4.1. Planeta exterior más masivo

Empezamos analizando el caso donde el planeta exterior es más masivo que su compañero interior (i.e. $m_2/m_1 > 1$). El conjunto de condiciones iniciales se eligió de la Tabla 5.1, caracterizado por razón de masas $m_2/m_1 = 3/1$ y elementos orbitales que producen una pequeña amplitud de oscilación ($\sim 5^\circ$) alrededor de una ACR tipo $(0, 0)$. Se generaron 1000 conjuntos de VR cada uno con $N = 200$ datos distribuidos aleatoriamente sobre cuatro períodos del planeta exterior (i.e. cuatro años de observaciones). Los valores de cada velocidad radial individual fueron elegidos aleatoriamente siguiendo una distribución Gaussiana alrededor del valor exacto y varianza dada por la desviación estándar constante para $\sigma = 7$ m/s (i.e. $\sigma^* = 5.5$ m/s). Estos valores garantizan la detectabilidad de dos planetas y ubican la razón de movimientos medios dentro de la región de conmensurabilidad.

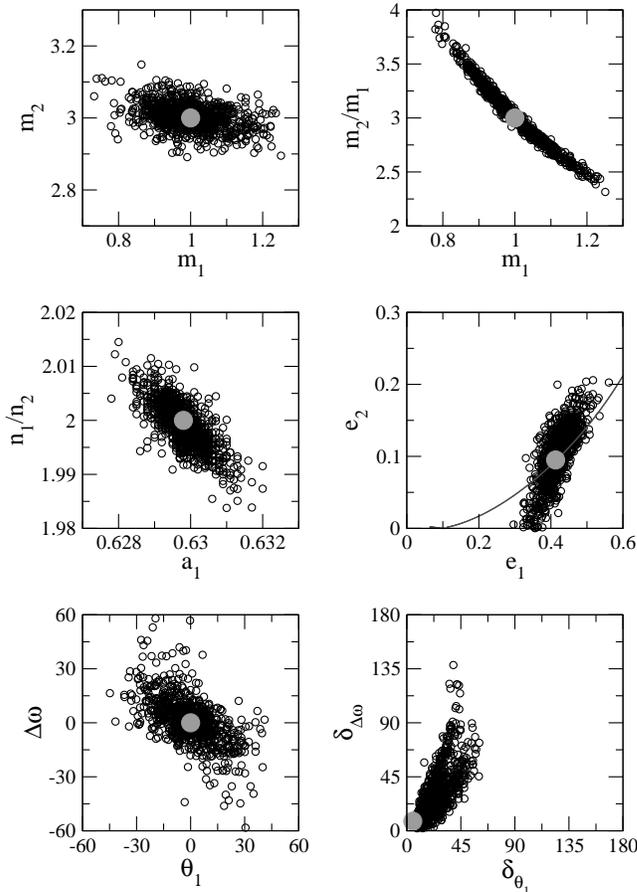


Figura 5.16: Resultados de los mejores ajustes orbitales de 1000 datos sintéticos, correspondientes a la solución nominal dada por el primer conjunto de datos en la Tabla 5.1, con $N = 200$ observaciones y desviación estándar constante de $\sigma = 7$ m/s. Los círculos rellenos grises indican la posición de la solución nominal. En el plano de excentricidades la curva gris muestra la familia ACR de amplitud cero para razón de masa $m_2/m_1 = 3/1$.

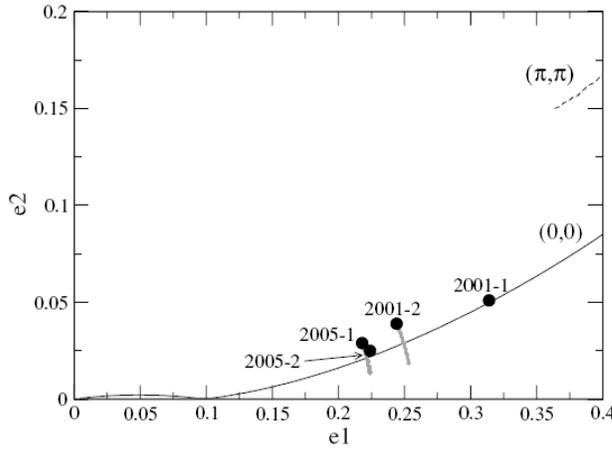


Figura 5.17: Familias de ACRs para la razón de masa $m_2/m_1 = 3.2$. La línea continua gruesa muestra las familias $(0, \pi)$ y $(0, 0)$, mientras que la curva a trazos superior muestra las soluciones ACR del tipo (π, π) . Los círculos rellenos muestran distintos ajustes orbitales para el sistema GJ 876, mientras que las curvas grises muestran la evolución orbital de las condiciones iniciales en este plano. Reproducido de Beaugé et al 2006.

Para cada conjunto de datos se calculó el mejor ajuste orbital para 1 planeta utilizando la metodología especificada en nuestra Sección 2.4.2 utilizando algoritmos genéticos sin ninguna suposición acerca de la solución inicial. Si el ajuste resultante tiene una probabilidad de falsa alarma (FAP) menor que 10^{-4} es aceptado y el ajuste para dos planetas es realizado. La solución de un sólo planeta fue utilizada como primera conjetura y el código identificó la solución con 2 planetas con el mínimo residuo. Una vez más la solución se aceptó sólo si el FAP del ajuste de 2 planetas con respecto a de la solución de 1 planeta era menor que 10^{-4} . Si es aceptado, entonces se integró numéricamente la configuración sobre un período de tiempo de 10^4 períodos orbitales y se estimó la amplitud de oscilación de los ángulos resonantes $\theta_i = 2\lambda_2 - \lambda_1 - \varpi_i$ y la diferencia en longitudes de pericentro $\Delta\varpi = \varpi_2 - \varpi_1$. El ángulo crítico θ_1 corresponde a una resonancia interior (ver Michtchenko et al. 2008a), como se halla en ACR simétricas, mientras que θ_2 es el ángulo librante para resonancias exteriores (ver Michtchenko et al. 2008b) y así aplicable a ACR asimétricas.

Los resultados se muestran en la Fig. 5.16, donde se presentan las distribuciones en varios planos. En cada marco el círculo gris relleno indica la solución nominal y los círculos negros más pequeños indican el mejor ajuste orbital de cada conjunto de datos. La curva gris en el plano de excentricidades muestra la familia ACR tipo $(0, 0)$ para la razón de masa $m_2/m_1=3/1$. Como se esperaba de la elección de N y σ , todos los sistemas ficticios de dos planetas se detectaron con proceso de ajuste y se identificaron dentro de la resonancia 2/1 MMR, y ambos θ_1 y $\Delta\varpi$ oscilan alrededor de la solución de equilibrio. Sin embargo, en el 2% de los casos $\Delta\varpi$ circula, tomando todos los valores entre 0 y 2π , mientras θ_1 libra. Este tipo de movimiento es usualmente referido como un θ_1 -librador y es considerado distinto que un ACR. Sin embargo, tal como mostrara Michtchenko et al. (2008a), incluso dentro de una ACR, la libración de $\Delta\varpi$ es puramente cinemática y no es asociada a ninguna separatriz, siempre que las excentricidades no sean demasiado grandes.

Los dos gráficos superiores en la Fig. 5.16 muestran la distribución de las masas planetarias detectadas. En el gráfico superior izquierdo vemos que la masa mayor (i.e. mayor K) tiene dispersión bastante menor, mientras que m_1 no está bien determinada, consecuentemente produciendo una dispersión mayor en la razón de masas m_2/m_1 (gráfico superior derecho) entre 2.5 y 3.5. Aún así, el semeje mayor y la razón de movimientos medios son bien determinados con pequeña variación alrededor del valor nominal (gráfico central izquierdo).

La distribución de las soluciones en el plano de excentricidades muestra una interesante correlación con la familia ACR de amplitud cero (mostrada como línea gris). Este resultado no era esperado pues no hay relación evidente entre el proceso de ajuste y la topología de la MMR. Sin embargo

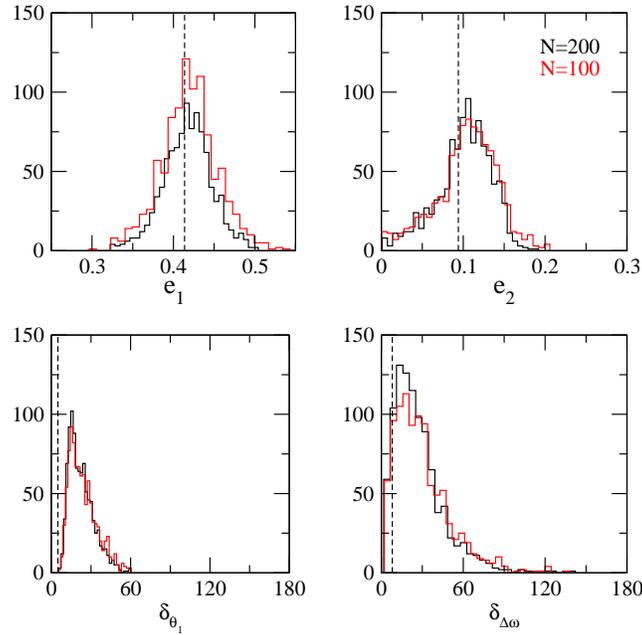


Figura 5.18: Distribución de excentricidades y amplitudes de oscilación del ángulo resonante θ_1 y $\Delta\varpi$ para el caso $m_2/m_1 = 3/1$. Los valores nominales son indicados por líneas verticales a trazos. Los códigos de color son utilizados para identificar distintos tamaños de conjuntos de datos N .

está indicando que, incluso con la determinación imprecisa de las excentricidades, no necesariamente nos alejaremos demasiado de la familia ACR y consecuentemente de las soluciones estables. Recordemos que determinaciones sucesivas de los planetas del sistema resonante GJ876 (relación de masa cercana a $m_2/m_1 = 3/1$, ver Figura 5.17) muestra un comportamiento similar. A pesar que en los ajustes orbitales las excentricidades variaron significativamente, todas las soluciones orbitales se encontraron cercanas a la ACR estable.

Finalmente los dos gráficos inferiores muestran la variación en las variable angulares. El gráfico de la izquierda da la distribución de los valores calculados de θ_1 y $\Delta\varpi$. Más interesante, en el marco de la derecha se muestra la distribución de amplitudes de oscilación alrededor de la ACR. Se puede apreciar una dispersión importante en el comportamiento dinámico de los ajustes calculados. Tal como en el caso de la excentricidad, la amplitud tiene un límite inferior en cero que causa un efecto sistemático que conduce al hecho de estimar la amplitud de libración sea sobreestimada respecto al valor verdadero.

Esta figura muestra que incluso bajo circunstancias favorables dadas por un conjunto de datos grandes e incertezas observacionales razonables, los errores en los ajustes orbitales son todavía significativos, especialmente en las excentricidades y variables angulares. Incluso si las ACR son identificadas como tales, las amplitudes de oscilación son incrementadas bastante más allá de sus valores nominales.

La Figura 5.18 muestra la distribución de las excentricidades ajustadas y las amplitudes de oscilación de las variables angulares para $N = 200$ (histogramas en negro) y $N = 100$ (histogramas en rojo). Los valores nominales correspondientes al sistema original se marcaron como líneas verticales a trazos. En todos los casos las V_r tienen tiempos de observación aleatoriamente distribuidos a lo largo de 4 períodos del planeta exterior y $\sigma = 7$ m/s.

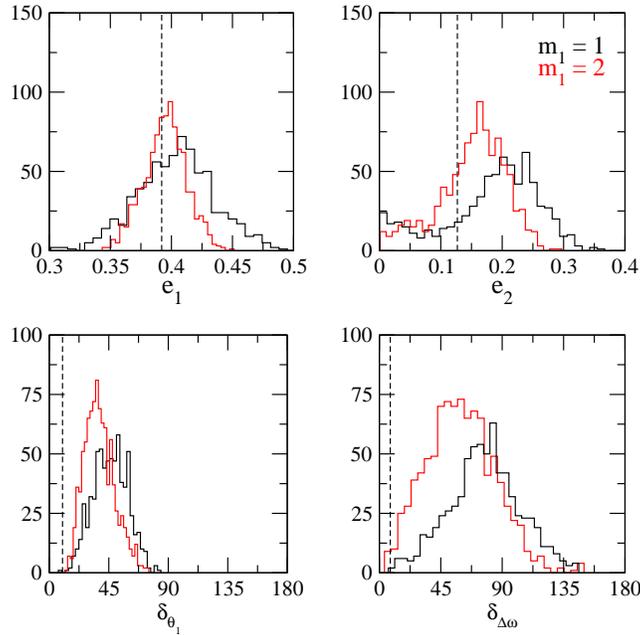


Figura 5.19: Histogramas de soluciones para razones de masa $m_2/m_1 = 1/1$. En este caso se utilizaron conjuntos de datos de $N = 100$ y los histogramas negros son los obtenidos para $m_1 = 1M_{\text{Jup}}$ mientras que los rojos son los calculados con $m_1 = 2M_{\text{Jup}}$.

Los 2 marcos superiores corresponden a excentricidades; e_1 a la izquierda y e_2 a la derecha. La distribución de las excentricidades del planeta interior es medianamente simétrica con respecto al valor real (~ 0.42), y es evocativo de los resultados mostrados para un solo planeta (Shen y Turner 2008 y Sección 5.1.2). A pesar que la dispersión es un poco mayor para un número menor de observaciones N , parece haber un pequeño error sistemático y el pico del histograma es cercano a la línea vertical. En el caso de la excentricidad del planeta exterior, la distribución aparenta ser menos simétrica y existe un sesgo hacia valores más grandes de e_2 .

Los dos dibujos inferiores muestran la distribución de las amplitudes de oscilación del ángulo resonante θ_1 (izquierda) y $\Delta\varpi$ (derecha). Una vez más los valores nominales se muestran como líneas verticales y corresponden a pequeñas amplitudes de oscilación. Como se muestra en la Figura 5.16, existe un importante incremento en las amplitudes conduciendo a una tendencia sistemática hacia soluciones con amplitudes de oscilación más grandes. Sorprendentemente el efecto es el mismo para conjuntos de datos más pequeños. En el caso de θ_1 prácticamente no hay soluciones con amplitudes más grandes que 60° . Lo mismo sin embargo no ocurre para $\Delta\varpi$. Incluso para grandes conjuntos de datos, algunas soluciones tienen amplitudes de 180 grados, que corresponden a circulaciones de la diferencia de longitudes de pericentro.

5.2.5. Planetas de masas iguales

El análisis previo puede ser repetido para razones de masa de la unidad ($m_2/m_1 = 1$) y eligiendo condiciones iniciales cercanas a una solución ACR estable cuyos elementos orbitales se encuentran en la Tabla 5.1. Esta ACR en particular fue elegida de tal forma que el momento angular total del sistema sea aproximadamente el mismo que en el sistema anterior; esto implica principalmente que la suma

del cuadrado de las excentricidades sea del mismo orden.

La Figura 5.19 muestra el histograma con la distribución de las excentricidades (arriba), amplitud de libración de θ_1 (abajo izquierda) y amplitud de oscilación de $\Delta\varpi$ (abajo derecha). Contrariamente a la Figura 5.18, el número de datos fue mantenido constante en $N = 100$ y el código de color se usa para distintas masas del planeta interno. Los histogramas en negro se construyeron con los datos generados por $m_1 = M_{\text{Jup}}$, mientras que en rojo se generaron con los datos tomando $m_1 = 2M_{\text{Jup}}$. Este último valor genera una amplitud de señal RV que es similar al caso discutido en la sección anterior, esto permite una comparación más sencilla de la dispersión de las soluciones estables.

Como antes, la excentricidad del planeta interno es aproximadamente bien establecido. Para ambos valores de m_1 la distribución es centrada cerca del valor nominal, a pesar del hecho que notamos una menor dispersión en los resultados para masas más grandes (i.e. amplitudes mayores de la señal RV). Sin embargo, la excentricidad del planeta exterior, muestra una distribución compleja, con un corrimiento sistemático notorio del valor medio. También como antes todos los ajustes orbitales llevan a grandes amplitudes de libración, especialmente la diferencia de longitudes de pericentro. El valor medio de este ángulo se encuentra cercano a 90° . A pesar que este sesgo decrece significativamente con valores más grandes de m_1 , estos son todavía más grandes que los presentados en la Figura 5.18.

5.2.5.1. Planeta exterior menos masivo

Finalmente analizamos el caso donde $m_2/m_1 < 1$. Las condiciones iniciales dadas por la Tabla 5.1 corresponden a amplitud de oscilación pequeña alrededor de una ACR asimétrica con $m_2/m_1 = 1/3$. Una vez más los histogramas en rojo son obtenidos para la masa interior planetaria incrementada ($m_1 = 3M_{\text{Jup}}$) que produce en la señal de RV una magnitud similar a aquella amplitud de $m_2/m_1 = 3/1$.

A pesar que ambos planetas se detectan en todos los conjuntos de datos, para $m_1 = M_{\text{Jup}}$ sólo el 26 % de las órbitas resultantes reproducen la ACR. En el 68 % de los casos la configuración resultante fue no-resonante y ni θ_2 o $\Delta\varpi$ libraban alrededor del valor de equilibrio dado por la ACR, sino que en su lugar circulaban tomando todos los valores entre cero y 2π . Dada la alta excentricidad del planeta exterior, estas órbitas inevitablemente producen movimiento inestable con desintegración del sistema en tiempos de escalas pequeños. En el remanente 6 % de los casos los mejores ajustes orbitales conducen a θ_2 -libradores pero con la diferencia de pericentros ($\Delta\varpi$) circulando.

Estos números mejoran significativamente para $m_1 = 3M_{\text{Jup}}$, donde las desintegraciones del sistema ocurren sólo en el 23 % de los casos y las mayoría del resto corresponden a ACR estables. Incluso, esto implica que para desviaciones estándar tan bajas como $\sigma^* = 2$ m/s, un cuarto de los datos sintéticos fallan en identificar el movimiento resonante y producen ajustes orbitales dinámicamente inestables.

La distribución de los elementos de soluciones estables ACR se muestran en la Figura 5.20. Debido a que eliminamos los casos no-resonantes, los histogramas aparecen similares a los de otras razones de masa, con una estimación gruesa de la excentricidad del planeta más interno mucho más dispersa que el externo en e_2 . Para $m_1 = 1M_{\text{Jup}}$ también notamos indicaciones de una distribución bimodal para la excentricidad del planeta exterior, con un pico centrado aproximadamente alrededor de $e_2 = 0.45$ y el otro pico cercano a 0.6. La bi-modalidad está relacionada con una bifurcación de la familia ACR para esta razón de masas y altas excentricidades (ver Michtchenko et al. 2008b). Sin embargo, el histograma en rojo muestra una distribución de e_2 más localizada.

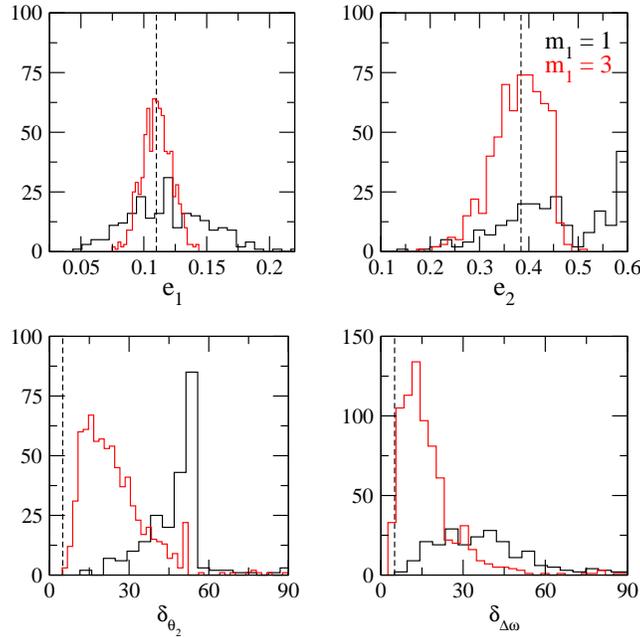


Figura 5.20: Histogramas similares a los de la Figura 5.19, pero para razón de masas $m_2/m_1 = 1/3$. Los histogramas en negro fueron obtenidos con $m_1 = 1M_{\text{Jup}}$, mientras que aquellos en rojo con $m_1 = 3M_{\text{Jup}}$.

Como con otras razones de masas, la distribución de amplitudes de oscilación/libración también muestra un sesgo marcado hacia valores mayores. A pesar que este decrece para valores más grandes de m_1 , incluso para $m_1 = 3M_{\text{Jup}}$ prácticamente ninguno de los ajustes orbitales reproduce la oscilación de pequeña amplitud alrededor de una ACR asimétrica. Finalmente, el histograma en negro muestra la distribución de la amplitud de libración de θ_2 , con un pequeño pico alrededor de 60 grados, que a grandes rasgos corresponde al máximo de amplitud permitido por una ACR asimétrica. Las amplitudes mayores ubican al sistema fuera del dominio resonante, que puede explicar la gran proporción de resultados no-resonantes en los ajustes orbitales.

5.2.6. Aplicación a exosistemas reales

La Tabla 5.2 resume los resultados dinámicos posibles de los ajustes orbitales para tres relaciones de masa. A pesar que todos los conjuntos de datos se construyeron para la misma desviación estándar $\sigma = 7$ m/s, hay una marcada diferencia en los resultados. Para $m_2/m_1 = 3/1$ la mayoría de los ajustes orbitales identifican correctamente la ACR y consecuentemente producen movimiento estable para ambos planetas en tiempos de escala de 10^4 períodos orbitales. El caso inverso, $m_2/m_1 = 1/3$ es significativamente menos preciso, probablemente debido a la existencia de varios dominios de movimientos estables, cada uno asociado a una familia ACR distinta (Michtchenko et al. 2008b). Cada dominio es separado por una región caótica, y son menos robustos que aquellos encontrados para razón de masas mayores que la unidad.

Así es aparente que los planetas resonantes con razones de masa menores que la unidad no son solo más difíciles de detectar, sino incluso cuando son detectados los ajustes orbitales usualmente producen ajustes orbitales con razones de periodos mal determinadas y/o soluciones dinámicamente inestables. A pesar que estos resultados son válidos para la resonancia 2/1, y los planetas 55Cnc-c y 55Cnc-d se localizan en cercanías de la conmensurabilidad 3/1 (con razón de masa $m_2/m_1 \simeq 1/5$),

Tabla 5.2: Porcentaje de resultados dinámicos de los mejores ajustes orbitales para distintas razones de masa m_2/m_1 . Para ACR simétricos el ángulo resonante es tomado como θ_1 , mientras que θ_2 es adoptado para soluciones asimétricas (ver Michtchenko et al. 2008ab). Los valores de m_1 son dados en unidades de M_{Jup} .

m_2/m_1		no-resonantes	θ_i -librador	ACR
3/1	$N = 200$	0 %	2 %	98 %
	$N = 100$	0 %	12 %	88 %
1/1	$m_1 = 1$	9 %	21 %	70 %
	$m_1 = 2$	2 %	2 %	96 %
1/3	$m_1 = 1$	68 %	6 %	26 %
	$m_1 = 3$	23 %	1 %	76 %

parece posible esperar resultados similares, particularmente debido a que las ACR asimétricas están involucradas en ambos casos. Esto puede ser una posible explicación de por que los sucesivos ajustes orbitales de los planetas de *55Cnc* alternativamente producen movimiento resonante y no-resonante, mientras que la misma incerteza no está presente para exosistemas con otras relaciones de masa. Otro ejemplo es dado por el sistema HD37124 con dos planetas en la vecindad de la MMR 2/1. A pesar que el mejor ajuste orbital indica $m_2/m_1 \simeq 1.2$ y una ACR asimétrica (Baluev 2008), existen otras soluciones que ponen los cuerpos fuera de la conmensurabilidad (Vogt et al. 2005).

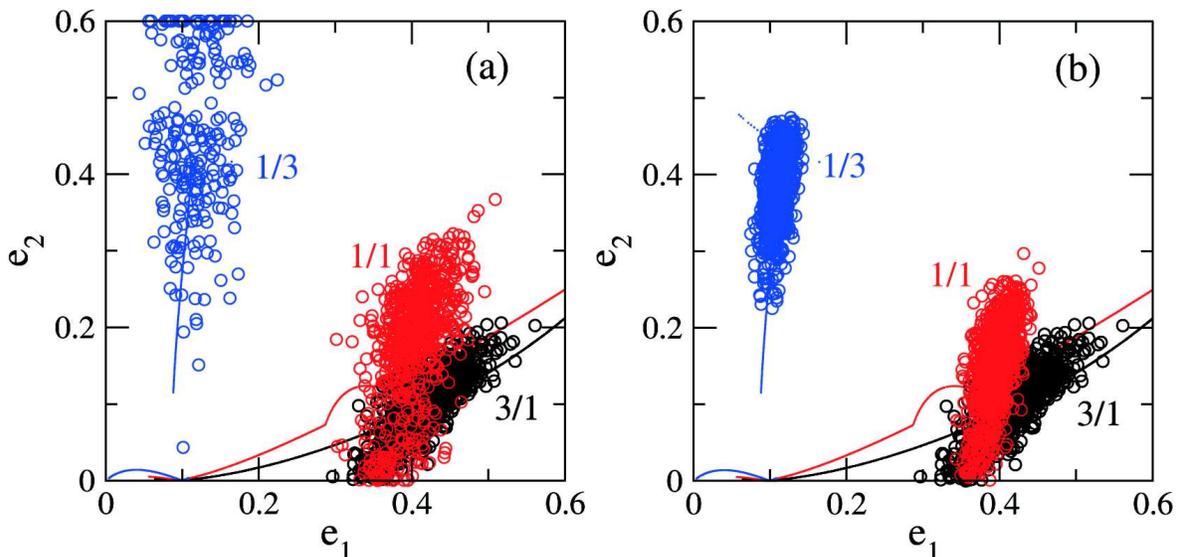


Figura 5.21: Distribución de excentricidades planetarias obtenidas de los mejores ajustes orbitales de conjuntos orbitales sintéticos, comparados con las familias de ACR. Los códigos de color identifican distintas razones de masas. Negro : $m_2/m_1 = 3/1$, rojo : $m_2/m_1 = 1/1$ y azul : $m_2/m_1 = 1/3$. **(a)**: Todas las masas planetarias internas son iguales a $1M_{\text{Jup}}$. **(b)**: masas planetarias internas aumentadas: $m_1 = 2M_{\text{Jup}}$ para $m_2/m_1 = 1/1$ y $m_1 = 3M_{\text{Jup}}$ para $m_2/m_1 = 1/3$.

La Figura 5.21 compara la distribución de las excentricidades encontradas para $m_2/m_1 = 3/1$ (círculos negros adentro), $m_2/m_1 = 1/1$ (rojo) y $m_2/m_1 = 1/3$ (azul). Solo aquellos ajustes orbitales que producen ACR estables fueron incluidos. Las curvas de color corresponden a soluciones ACR de amplitud cero para cada razón de masas. El gráfico de la izquierda fue dibujado utilizando $m_1 = M_{\text{Jup}}$ en los tres casos. Para el de la derecha usamos los valores incrementados de masa m_1 dados en la leyenda de la figura.

Para $m_2/m_1 = 1/3$ notamos que la mayoría de las soluciones se encuentran relativamente cerca

a la familia de ACR estables, sin embargo es importante recordar que estos corresponden a menos de la mitad de los datos ajustes totales. Una tendencia similar es observada para $m_2/m_1 = 3/1$, excepto que ahora todos los ajustes orbitales fueron dibujados. Sin embargo, en el caso de masas iguales $m_2/m_1 = 1/1$, no hay correlación aparente entre la distribución de excentricidades y las familias de ACR. Así, aparentemente para razones de masas mayores, hay una probabilidad más grande que incluso ajustes orbitales no tan buenos se encuentren próximos a configuraciones ACR estables. Para razones de masas cercanas a la unidad, incluso para pequeñas desviaciones estándar pueden producir grandes amplitudes de libración o movimiento inestable. Incluso con masa planetaria incrementadas, es decir mejores ajustes orbitales, la misma correlación es observada. De hecho, la gran razón S/N de los datos de la velocidad radial mantiene una dispersión similar en e_2 , sin embargo la estimación del valor de e_1 es más precisa.

Capítulo 6

Conclusiones

A lo largo de esta tesis hemos desarrollado una estrategia general de búsqueda de soluciones, en problemas de minimización altamente no lineales, en una amplia variedad de situaciones relacionadas con la detección de planetas a partir de datos de RV. La combinación de métodos no-determinísticos y evolutivos ha sido muy satisfactoria. Se determinó que una rutina que contenga *Algoritmos Genéticos* + *Simplex* + *Annealing* (en ese orden) es una herramienta óptima y efectiva para calcular residuos y ubicar los mínimos en las soluciones para detectar exoplanetas con Velocidad Radial.

Alternativamente, se emplearon varias herramientas para obtener errores y correlaciones alrededor de los mejores ajustes. Tanto las *Cadenas Markov MonteCarlo* como la generación de muestras sintéticas con *Resampling*, son útiles para conocer el comportamiento de la función residuo y buscar los valores dentro de la *región de confianza* que sean dinámicamente estables y compatibles con las observaciones disponibles. El método jackknife brinda información acerca de la estabilidad estadística y permite estimar como un único dato puede cambiar los valores en los parámetros del mejor ajuste. Además sabremos cuanto podemos esperar que varíen los elementos planetarios con nuevas observaciones. También se trabajó con la incompletitud del modelado en las observaciones, tratando de mejorarlo adicionando un nuevo cuerpo al modelo del sistema. Los análisis sobre la forma de la función residuo en distintos planos, desarrollados en *región de confianza*, mostraron ser útiles a la hora de identificar la existencia de varios mínimos con casi el mismo valor de residuo.

En particular en el caso de HD82943, la configuración con mínimo *wrms* produce inestabilidad dinámica. Sin embargo las soluciones estables pueden encontrarse en la vecindad, todas compatibles con co-rotaciones apsidales (ACR) en las cuales los cuerpos tienen ambos ángulos, el resonante y la diferencia de longitud de pericentros, librando alrededor de cero. La importancia relativa de cada mínimo varía con el número de observaciones N , favoreciendo configuraciones estables o inestables. Sobre el final de esta exposición demostramos la dificultad de establecer la configuración resonante basados en los procesos de detección de RV.

Existe una evidencia intrigante en la posibilidad de existencia de un tercer planeta para el sistema HD82943, con aproximadamente la mitad de la masa de Júpiter y período orbital de ~ 900 . Las consideraciones de estabilidad, compatibles con migración planetaria, apuntan a que los cuerpos se pueden encontrar en una resonancia de Laplace, donde los tres planetas se encuentran en sucesivas MMR 2/1 y ACRs dobles asimétricas. Sin embargo, en este sistema, las observaciones cubren solo dos veces el período orbital del tercer planeta exterior hipotético; consecuentemente es posible que nuevas observaciones confirmen este resultado.

Por otro lado, analizamos el límite de detectabilidad de dos planetas en la vecindad de la resonancia de movimientos medios 2/1, como función del número de datos N y desviación estándar σ de los valores de V_r . Para valores de σ y N comparables a los exosistemas reales, encontramos una tendencia importante a detectar sistemas con $m_2/m_1 \sim 1$. Los planetas con razones de masa mucho mayores o menores que la unidad son más difíciles de separar a partir de los datos Doppler. Esto no solamente ayuda a explicar la falta de sistemas con razones de masa distintas a la unidad, sino que también indica que el número de sistemas resonantes en existencia puede ser actualmente mayor que los catalogados. Se mejoró la teoría de detección de uno y dos planetas y se lograron establecer límites de confianza para las detecciones.

También se analizó la dispersión en las masas y elementos orbitales de los exosistemas detectados. A pesar que para razones de masa grandes la mayoría de los datos sintéticos efectivamente son detectados en sistemas resonantes, para valores bajos de m_2/m_1 una gran proporción de ajustes orbitales no son detectados dentro de la resonancia y erróneamente son indicados como configuraciones no resonantes. Inclusive dentro del subgrupo de órbitas resonantes, notamos una gran dispersión en los valores estimados de la excentricidad del planeta exterior y la media de la distribución es siempre mayor que el valor nominal de e_2 . Esto está de acuerdo con varios sistemas reales en la MMR 2/1, particularmente HD82943. Más aún, las soluciones nominales corresponden a pequeñas amplitudes de libración alrededor de las ACR, mientras que existe una tendencia a estimar grandes amplitudes de oscilación que incluso comprometen la estabilidad dinámica en tiempos de escala largos.

Esto presenta una posible explicación a la pregunta de por que, con la excepción de los cuerpos del sistema resonante GJ876, todos los otros sistemas resonantes tienen ajustes orbitales compatibles con grandes amplitudes de oscilación. A pesar que trabajos recientes de Sándor et al. (2007) y Crida et al. (2008) proponen que las grandes amplitudes de libración pueden ser causadas por procesos dinámicos en el pasado (e.g. scattering planetarios o dispersión de disco), los resultados presentes indican que esta característica observacional puede no ser real, sino una simple consecuencia de tratar con pequeños conjuntos de datos con grandes incertezas en los valores de V_r .

Para sistemas multiplanetarios es importante recalcar que la precisión numérica absoluta de los elementos orbitales no es en sí misma el único dato a considerar, tanto como la diversidad en el comportamiento dinámico dentro de la región de incerteza. En otras palabras, si los dos planetas no son resonantes (con bajas excentricidades), incluso cambios significativos en los elementos orbitales no necesariamente producirán diferentes tipos de movimiento. Inversamente si los planetas se encuentran en una resonancia MMR y (preferiblemente) tienen excentricidades altas a moderadas, inclusive pequeños cambios en los elementos orbitales pueden significar la diferencia entre movimiento estable e inestables.

Hemos recalcado varias veces que se debe tener precaución en el uso de las soluciones estables de los sistemas exoplanetarios resonantes como evidencia a favor (o en su contra) de distintos mecanismos de formación. En particular HD82943 puede ser un caso extremo y sirve como ejemplo de cuan dependientes son los ajustes orbitales con los datos observacionales (Sección 3.4). Para este sistema en particular es difícil estipular con un grado de certeza cuales son los verdaderos elementos orbitales o masas mínimas. Más aún, las nuevas observaciones pueden producir inclusive cambios significativos en las configuraciones, pero muy probablemente dentro de las regiones establecidas por métodos MonteCarlo expuestos en esta tesis.

Este trabajo resume la investigación desarrollada durante el doctorado y brinda herramientas numéricas, así como experiencias con métodos de minimización no determinísticos, en espacios multiparamétricos. Sin dudas muchas de las incertezas comentadas en un principio, comienzan a ser acotadas con la combinación de técnicas de detección. A lo largo de estos años de investigación hemos visto

cómo las técnicas avanzan y continúan mejorando, teniendo una evolución incesante. Los sistemas detectados son cada vez más complejos y más variados. En particular existen esfuerzos marcados para utilizar combinación de datos de tránsitos y velocidad radial. Afortunadamente el desarrollo expuesto continúa siendo útil con otras técnicas de detección (como *tránsitos*) y además abre posibilidades para un sin fin de aplicaciones futuras.

Referencias

- [1] Adams, F.C., Laughlin, G., Bloch, A.M. 2008, ApJ, 683, 1117
- [2] Andrae, R. 2010, arXiv:1009.2755v1
- [3] Anglada-Escude, G., López-Morales, M., Chambers, J. E. 2009, ApJ, 709, 168
- [4] Bailes, M., Lyne, A.G., Shemar, S.L. 1991, Nature 352, 311
- [5] Bakos, G. Á., Kovács, G., Torres, G., Fischer, D. A., Latham, D. W., et al. 2007, ApJ, 670, 826
- [6] Barlow, R. 1993, Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences, Wiley VCH, ISBN: 978-0-471-92295-7
- [7] Barnes, J. W. 2007, PASP, 119, 2007
- [8] Baluev, R.V. 2008, CeMDA, 102, 297
- [9] Baranne, A., Queloz, D., Mayor, M., et al. 1996, A&AS, 119, 373
- [10] Bean, J. L., McArthur, B. E., Benedict, G. F., Harrison, T., Bizyaev, D., et al. 2007, AJ, 134, 749
- [11] Bean, J. L., Seifahrt, A., Hartman, H., Nilsson, H., Reiners, A., et al. 2010, ApJ, 711, 19
- [12] Beaugé, C., Ferraz-Mello, S., Michtchenko, T.A. 2003, ApJ, 593, 1124
- [13] Beaugé, C., Michtchenko, T. A., Ferraz-Mello, S. 2006, MNRAS, 365, 1160
- [14] Beaugé, C., Sándor, Zs., Érdi, B., Süli, Á. 2007a, A&A, 463, 359
- [15] Beaugé, C., Ferraz-Mello, S., Michtchenko, T.A. 2007b, Planetary Masses and Orbital Parameters from Radial Velocity Measurements. In: *Extrasolar Planets: formation, Detection and Dynamics*, Ed. Dvorak, R., Wiley-VCH, Weinheim.
- [16] Beaugé, C., Giuppone, C.A., Ferraz-Mello, S., Michtchenko, T.A. 2008, MNRAS, 385, 2151
- [17] Benítez-Llambay, P., Masset, F., Beaugé, C. 2011, A&A, 528, 171
- [18] Bennett, D. P., Anderson, J., Bond, I. A., Udalski, A., Gould, A. 2006, ApJ, 647, 171
- [19] Bennett, D. P. 2010, ApJ, 716, 1408
- [20] Boss, A. P., Basri, G., Kumar, S. S., Liebert, J., Martín, E., et al. 2003, IAUS, 211, 529
- [21] Brown, T. M., Charbonneau, D., Gilliland, R. L., Noyes, R. W., Burrows, A. 2001, ApJ, 552, 699
- [22] Brown, R. A. 2004, ApJ, 610, 1079
- [23] Burrows, A., Marley, M., Hubbard, W.B., Lunine, J.I., Guillot, T., et al. 1997, ApJ, 491, 856
- [24] Campbell, B., Walker, G. A. H., Yang, S. 1988, ApJ, 331, 902
- [25] Cassen, P., Guillot, T., Quirrenbach, A. 2006, *Saas-Fee Advanced Course 31: Extrasolar planets*, Berlin, Alemania
- [26] Charbonneau, P. 1995, ApJS, 101, 309
- [27] Charbonneau, P., Knapp, B. 1995, NCAR Technical Note TN-418+IA
- [28] Charbonneau, P. 2002a, NCAR Technical Note TN-450+IA
- [29] Charbonneau, P. 2002b, NCAR Technical Note TN-451+STR
- [30] Charbonneau, D., Brown, T. M., Noyes, R. W., Gilliland, R. L. 2002, ApJ, 568, 377
- [31] Cincotta, P. M., Méndez, M., Núñez, J. A. 1995, ApJ, 449, 231
- [32] Cincotta, P. M., Helmi, A., Méndez, M., Núñez, J. A., Vucetich, H. 1999, MNRAS, 302, 582
- [33] Chauvin, G., Lagrange, A.-M., Dumas, C., Zuckerman, B., Mouillet, D., et al. 2004, A&A, 425, 29

- [34] Claret, A. 2000, *A&A*, 363, 1081
- [35] Clampin, M., Krist, J. E., Ardila, D. R., Golimowski, D. A., Hartig, G. F., et al. 2003, *AJ*, 126, 385
- [36] Correia, A. C. M., Udry, S., Mayor, M., Laskar, J., Naef, D., et al. 2005, *A&A*, 440, 751
- [37] Correia, A. C. M. 2008, Precision Spectroscopy, In *Astrophysics, Proceedings of the ESO/Lisbon/Aveiro*, Eds Santos N.C., et al., Garching, Alemania, p. 207-210
- [38] Cox, A. N. 2000, In: *Allen's Astrophysical Quantities*, New York, AIP, p. 355-357
- [39] Crida, A., Sándor, Zs., Kley, W. 2008, *A&A*, 483, 325
- [40] Crida, A., Masset, F., Morbidelli, A. 2009, *ApJ*, 705, 148
- [41] Cumming, A., Butler, R. P., Marcy, G. W., Vogt, S. S., Wright, J. T., Fischer, D. A. 2008, *PASP*, 120, 531
- [42] Cumming, A., Marcy, G.W., Butler, R.P. 1999, *ApJ*, 526, 890
- [43] Cumming, A. 2004, *MNRAS*, 354, 1165
- [44] Cumming, A., Butler, R. P., Marcy, G. W., Vogt, S. S., Wright, J. T., Fischer, D. A. 2008, *PASP*, 120, 531
- [45] D'Angelo, G., Lubow, S.H. 2008, *ApJ*, 685, 560
- [46] Eggenberger, A., Udry, S. 2010, *EAS Publications Series*, 41, 27
- [47] Fabrycky, D., Winn, J. 2009, *ApJ*, 696, 1230
- [48] Ferraz-Mello, S. 1981, *AJ*, 86, 619
- [49] Ferraz-Mello, S., Quast, G. 1987, Ed. Kleczek, J., *Excercises in Astronomy*, p. 231-235
- [50] Ferraz-Mello, S., Beaugé, C., Michtchenko, T. A. 2003, *CeMDA*, 87, 99
- [51] Ferraz-Mello, S., Michtchenko, T. A., Beaugé, C. 2005, *AJ*, 621, 473
- [52] Fischer, D., Marcy, G., Butler, R., Vogt, S. S., Laughlin, G., Henry, G. W., et al. 2008, *ApJ*, 675, 790
- [53] Ford, E. B. 2005, *AJ*, 129, 1706
- [54] Foster, G. 1995, *AJ*, 109, 1889
- [55] Gaudi, S. B. 2005, *ApJ*, 628, L73
- [56] Gaudi, S. B., Winn, J. 2007, *ApJ*, 655, 550
- [57] Gaudi, S. B. 2010, arXiv:1002.0332v2
- [58] Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B. 2003, *Bayesian Data Analysis*. New York: Chapman & Hall/CRC
- [59] Giménez, A. 2006, *ApJ*, 650, 408
- [60] Giuppone, C.A., Tadeu dos Santos, M., Beaugé, C., Ferraz-Mello, S., Michtchenko, T.A. 2009, *ApJ*, 699, 1321
- [61] Giuppone, C.A., Beaugé, C., Michtchenko, T.A., Ferraz-Mello, S. 2010, *MNRAS*, 407, 390
- [62] Goldberg, D. E. 1989, Genetic Algorithms, In: *Search, Optimization, & Machine Learning*, Addison-Wesley
- [63] Goździewski, K., Konacki, M. 2006, *ApJ*, 647, 573
- [64] Goździewski, K., Konacki, M. 2005, *ApJ*, 622, 1136
- [65] Gregory, P. C. 2005a, *ApJ*, 631, 1198
- [66] Gregory, P. C. 2005b, *Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica Support*, Ed. Gregory, C., ISBN 0-521-84150-X. Cambridge University Press, UK
- [67] Hatzes, A. P., Cochran, W. D., Endl, M., McArthur, B., Paulson, D. B., et al. 2003, *ApJ*, 599, 1383
- [68] Heavens, A. 2009, arXiv:0906.0664v3
- [69] Herbst, W., Dhital, S., Francis, A., Lin, L., Tresser, N., Williams, E. 2006, *PASP*, 118, 828
- [70] Howard, R. 1984, *Annual Review A&A*, 22, 131
- [71] Howard, A., Johnson, J., Marcy, G., Fischer, D., Wright, J. et al. 2010, arXiv1003.3444H
- [72] Jacob, W.S. 1855, *MNRAS*, 15, 228
- [73] Ji, J., Liu, L., Kinoshita, H., Zhou, J., Nakai, H., Li, G. 2003, *ApJ*, 591, 57
- [74] Jones, B. W. 2008, *International Journal of Astrobiology*, 7, 279
- [75] Kalas, P., Graham, J. R., Clampin, M. 2005, *Nature*, 435, 1067
- [76] Kley, W., Lee, M.H., Murray, N., Peale, S.J. 2005, *A&A*, 437, 727

- [77] Kovalevsky, J. 1995, *Modern Astrometry*, ISBN # 3-540-57023-3
- [78] Konacki, M., Maciejewski, A. J., Wolszczan, A. 2000, *ApJ*, 544, 921
- [79] Konacki, M., Maciejewski, A. J. 1999, *ApJ*, 518, 442
- [80] Kurth, R. 1967, *Introduction to Stellar Statistics*, Pergamon Press, Oxford
- [81] Lawton, A. T., Wright, P. 1989, *British Interplanetary Society, Journal* 42, 335
- [82] Lee, M. H., Peale, S. J. 2002, *ApJ*, 567, 596
- [83] Lee, M. H., Butler, R. P., Fischer, D. A., Marcy, G. W., Vogt, S. S. 2006, *ApJ*, 641, 1178
- [84] Lin, D. N. C., Bodenheimer, P., Richardson, D. C. 1996, *Nature*, 380, 606
- [85] Lomb, N. R. 1976, *Ap&SS*, 39, 447
- [86] Lovis, C., Mayor, M., Pepe, F., Alibert, Y., Benz, W., et al. 2006, *Nature*, 441, 305
- [87] Lovis C., Ségransan D., Mayor M., Udry S., Benz W., et al. 2010, *ESO, Press Release*, 1035
- [88] Lyne, A.G., Bailes, M. 1992, *Nature* 355. 213
- [89] Mandel, K., Agol, E. 2002, *ApJ*, 580, L171
- [90] Marcy, G. W., Butler, R. P. 1992, *PASP*, 104, 270
- [91] Masset, F.S., Snellgrove, M.D. 2001, In: *Planetary Systems in the Universe: Observation, Formation and Evolution*, Eds. Penny, A.J, et al., ASP Conference Series
- [92] Mayor, M., Queloz, D. 1995, *Nature*, 378, 355
- [93] Mao, S., Paczynski, B. 1991, *ApJ*, 374, L37
- [94] Mayor, M., Queloz, D. 1995, *Nature*, 378, 355
- [95] Michtchenko, T.A., Ferraz-Mello, S. 2001, *AJ*, 122, 474
- [96] Michtchenko, T. A., Ferraz-Mello, S., Beaugé, C. 2006a, *Icarus*, 181, 555
- [97] Michtchenko, T. A., Beaugé, C., Ferraz-Mello, S. 2006b, *CeMDA*, 94, 411
- [98] Michtchenko, T.A., Beaugé, C, Ferraz-Mello, S. 2008a, *MNRAS*, 387, 747
- [99] Michtchenko, T.A., Beaugé, C, Ferraz-Mello, S. 2008b, *MNRAS*, 391, 227
- [100] Michtchenko, T.A., Ferraz-Mello, S., Beaugé, C. 2005, *Icarus*, 181, 2, 555
- [101] Morbidelli, A., Crida, A. 2007, *Icarus*, 191, 158
- [102] Moorhead, A.V., Adams, F.C. 2005, *Icarus*, 178, 517
- [103] Murray, C. D., Dermott, S. F. 2000, *Solar System Dynamics*. Cambridge Univ. Press, Cambridge
- [104] Nagasawa, M., Ida, S. Bessho, T. 2008, *ApJ*, 678, 498
- [105] Niedzielski, A., Gozdziewski, K., Wolszczan, A., Konacki, M., Nowak, G., et al. 2009, *ApJ*, 693, 276
- [106] Ohta, Y., Taruya, A., Suto, Y. 2005, *ApJ*, 622, 1118
- [107] Papaloizou, J. C. B., Terquem, C. 2006, *Rep. Prog. Phys.*, 69, 119
- [108] Pepe, F., Correia, A.C.M., Mayor, M. et al. 2007, *A&A*, 462, 769
- [109] Pepe, F., Mayor, M., Galland, F., et al. 2002a, *A&A*, 388, 632
- [110] Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P. 1992, *Numerical Recipes*, Cambridge University Press, UK
- [111] Queloz, D., Henry, G. W., Sivan, J. P., Baliunas, S. L., et al. 2001, *A&A*, 379, 279
- [112] Quirrenbach, A. 2006, *IAUS*, 232, 109
- [113] Rasio, F. A., Ford, E. B. 1996, *Science*, 274, 954
- [114] Rivera, E. J., Lissauer, J. J., Butler, R. P., Marcy, G. W., Vogt, S. S., et al. 2005, *ApJ*, 634, 625
- [115] Rivera, E. J., Laughlin, G., Butler, R. P., Vogt, S. S., Haghhighipour, N., et al. 2010, *ApJ*, 719, 890
- [116] Saar, S. H., Donahue, R. A. 1997, *ApJ*, 485, 319
- [117] Sándor, Zs., Kley, W., Klagyivik, P. 2007, *A&A*, 472, 981
- [118] Scargle, J. D. 1982, *ApJ*, 263, 835
- [119] Seager, S., Mallen-Ornelas, G. 2003, *ApJ*, 585, 1083
- [120] See, T. J. J. 1896, *AJ*, 16, 17
- [121] Shen, Y., Turner, E.L. 2008, *ApJ*, 685, 553

- [122] Sherrill, T. J. 1999, *Journal for the history of astronomy* 30. Retrieved on 2007-08-27
- [123] Spiegel, D. S., Burrows, A., Milsom, J. A. 2010, arXiv: 1008.5150v1
- [124] Stellingwerf, R. F. 1978, *ApJ*, 224, 953
- [125] Sumi, T, Bennett, D. P, Bond, I.A, Udalski, A., Batista, V., et al. 2010, *ApJ*, 710, 1641
- [126] Telesco, C. M., Fisher, R. S., Piña, R. K., Knacke, R. F., Dermott, S. F., et al. 2000, *ApJ*, 530, 329
- [127] Udalski, A., Paczynski, B., Zebun, K., Szymanski, M., Kubiak, M., et al 2002, *Acta Astronomica*, 52, 1
- [128] van de Kamp, P. 1969, *AJ*, 74, 757
- [129] Vogt S., Butler P., Marcy G., Fischer D., Henry G., Laughlin G., Wright J., Johnson A. 2005, *ApJ*, 632, 638
- [130] Walker, G. A. H., Bohlender, D. A., Walker, A. R., Irwin, A. W., Yang, S. L. S., Larson, A. 1992, *AJL*, 396
- [131] Wackerly D., Mendenhall III, W., Scheaffer R. 2003, *Estadística Matemática Con Aplicaciones*. Thomson, Mexico, ISBN: 9706861947
- [132] Whitworth, A. P., Zinnecker, H. 2004, *A&A*, 427, 299
- [133] Wolszczan, A., Frail, D. A. 1992, *Nature*, 355. 145
- [134] Wright, J. T. 2005, *PASP*, 117, 657