



UNIVERSIDAD DE CHILE  
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS  
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA

# FILTRADO BAYESIANO PARA ESTIMACIÓN DE PARÁMETROS ORBITALES EN ESTRELLAS BINARIAS VISUALES

MEMORIA PARA OPTAR AL TÍTULO DE  
INGENIERO CIVIL ELÉCTRICO

RUBÉN MATÍAS CLAVERÍA VEGA

PROFESOR GUÍA:  
RENÉ MÉNDEZ BUSSARD

PROFESOR CO-GUÍA:  
MARCOS ORCHARD CONCHA

MIEMBROS DE LA COMISIÓN:  
MARCOS ORCHARD CONCHA  
JORGE SILVA SÁNCHEZ

SANTIAGO DE CHILE  
2015



RESUMEN DE LA MEMORIA  
PARA OPTAR AL TÍTULO DE  
INGENIERO CIVIL ELÉCTRICO  
POR: RUBÉN MATÍAS CLAVERÍA VEGA  
FECHA: 2015  
PROF. GUÍA: RENÉ MÉNDEZ BUSSARD

## FILTRADO BAYESIANO PARA ESTIMACIÓN DE PARÁMETROS ORBITALES EN ESTRELLAS BINARIAS VISUALES

La masa es la propiedad más crítica de una estrella, ya que determina en gran medida su estructura y evolución. Se estima que un porcentaje mayoritario de las estrellas observadas en el cielo corresponde no a estrellas individuales sino a sistemas estelares múltiples. El estudio del movimiento relativo entre los componentes de un sistema estelar es la principal herramienta para calcular masas estelares; de hecho, la mayor parte de las estrellas fuera del Sistema Solar cuya masa ha podido ser determinada de manera directa corresponde a *estrellas binarias* –sistemas compuestos por dos estrellas ligadas gravitacionalmente–. De esta manera, los sistemas estelares múltiples constituyen la base observacional de la teoría de estructura y evolución estelar.

Dadas las características y el volumen de las observaciones astronómicas, es necesario que los métodos computacionales utilizados en el estudio de sistemas estelares múltiples sean *automáticos* y *robustos* frente a la existencia de *ruido de observación*, *brechas temporales* y *pérdida de datos*.

Con esta motivación, el presente Trabajo de Título propone y evalúa un método de estimación automática de elementos orbitales (y con ello, de la masa total) para *estrellas binarias visuales*. La herramienta desarrollada utiliza un enfoque Bayesiano, planteando un esquema de estimación basado en Filtro de Partículas. En este esquema, los parámetros orbitales son formulados como un vector de estado que evoluciona a través del tiempo –concepto conocido como Evolución Artificial de Parámetros– y la función de verosimilitud corresponde a la caracterización estadística del error cuadrático medio del conjunto de observaciones de posición relativa en el plano del cielo. A fin de reducir la dimensionalidad del problema, se utiliza la representación de Thiele-Innes para la descripción de las órbitas.

El método propuesto es probado sobre datos artificiales basados en el sistema estelar Sirius, con diferentes niveles de cobertura de órbita y error de observación característico. En la mayor parte de los casos estudiados, el método logra estimar los elementos orbitales con gran exactitud y precisión –mejor en la medida que el error observacional característico disminuye y la cobertura de la órbita aumenta–; sin embargo, se lograron identificar casos límite en que, sin importar la calidad de los datos, los parámetros orbitales estimados presentan un error significativo (a pesar de generar órbitas concordantes con las observaciones). Finalmente, se muestra que el algoritmo desarrollado permite no sólo obtener un estimador –un valor puntual dentro del espacio paramétrico–, sino también caracterizar la distribución *a posteriori* de los parámetros orbitales.

# Tabla de contenido

<b>1. Introducción</b>	<b>1</b>
<b>2. Marco teórico</b>	<b>4</b>
2.1. Binarias visuales . . . . .	4
2.1.1. Mecánica celeste . . . . .	6
2.2. Plano de observación y elementos geométricos . . . . .	10
2.2.1. Representación de Thiele-Innes . . . . .	12
2.3. Inferencia Bayesiana . . . . .	13
2.3.1. Estimación de estado . . . . .	15
2.3.2. Métodos secuenciales de Monte Carlo: Filtro de partículas . . . . .	18
2.3.2.1. Muestreo de importancia . . . . .	20
2.3.2.2. Muestreo de importancia secuencial y remuestreo . . . . .	22
2.3.2.3. Filtro de partículas para estimación de parámetros . . . . .	24
<b>3. Metodología e implementación</b>	<b>26</b>
3.1. Delimitación del problema . . . . .	26
3.2. Simulación de órbitas y observaciones . . . . .	29
3.2.1. Modelación de la adquisición de datos . . . . .	32
3.3. Método de estimación . . . . .	34
3.3.1. Propiedades básicas de la estrategia de filtrado . . . . .	34
3.3.2. Definición de vector de estado . . . . .	35

3.3.3.	Distribución inicial . . . . .	36
3.3.4.	Ecuación de evolución . . . . .	38
3.3.5.	Definición de <i>meta-observación</i> y función de verosimilitud . . . . .	39
3.3.6.	Resumen . . . . .	41
<b>4.</b>	<b>Resultados y discusión</b>	<b>42</b>
4.1.	Variación de la calidad de los datos . . . . .	42
4.1.1.	Configuración experimental . . . . .	43
4.1.2.	Caracterización estadística del desempeño . . . . .	46
4.2.	Variación de parámetros orbitales . . . . .	51
4.2.1.	Variaciones de $e$ e $i$ . . . . .	52
4.2.2.	Variación de $\mathcal{T}$ . . . . .	56
4.3.	Uso y alcance del método . . . . .	59
4.3.1.	Etapa 1: Estimación . . . . .	62
4.3.2.	Etapa 2: Exploración . . . . .	67
<b>5.</b>	<b>Conclusiones</b>	<b>71</b>
	<b>Bibliografía</b>	<b>74</b>
<b>A.</b>	<b>Ecuación de movimiento orbital</b>	<b>78</b>
A.1.	Ecuación de Kepler para órbitas ligadas . . . . .	81
A.2.	Leyes de Kepler . . . . .	84
<b>B.</b>	<b>Constantes de Thiele-Innes</b>	<b>87</b>
B.1.	Estimador de mínimos cuadrados . . . . .	87
B.2.	Conversión de Thiele-Innes a Campbell . . . . .	88

# Índice de figuras

2.1. Elipse de la solución del problema de los dos cuerpos . . . . .	8
2.2. Trayectorias elípticas . . . . .	8
2.3. Elementos orbitales de orientación espacial . . . . .	11
3.1. Observaciones de trayectoria en el <i>plano de cielo</i> . . . . .	27
3.2. Plano de observación en relación a órbita real . . . . .	28
4.1. Promedio y desviación estándar del valor estimado de $m_1 + m_2$ . . . . .	47
4.2. Estimación de masa con diferentes valores de $\mathcal{T}$ . . . . .	57
4.3. Observaciones y órbitas con cobertura parcial del período . . . . .	58
4.4. Error característico en cada instante de observación . . . . .	62
4.5. Observaciones y órbitas (real y estimada) . . . . .	63
4.6. Valores observados, reales y estimados de $X, Y$ y $\rho, \theta$ . . . . .	65
4.7. O–C para $X, Y$ y $\rho, \theta$ . . . . .	66
4.8. Distribuciones marginales <i>a posteriori</i> de parámetros <i>semi-fijos</i> . . . . .	68
4.9. Distribuciones marginales <i>a posteriori</i> de parámetros <i>explorados</i> . . . . .	69

# Índice de tablas

4.1. Parámetros estimados para $\sigma_{obs} = 0.015$ [arcsec] . . . . .	48
4.2. Parámetros estimados para $\sigma_{obs} = 0.075$ [arcsec] . . . . .	48
4.3. Parámetros estimados para $\sigma_{obs} = 0.150$ [arcsec] . . . . .	48
4.4. Recordatorio de elementos orbitales de Sirius (1) . . . . .	48
4.5. Recordatorio de elementos orbitales de Sirius (2) . . . . .	53
4.6. Estimación de parámetros para variaciones de $e$ . . . . .	53
4.7. Estimación de parámetros para variaciones de $i$ . . . . .	53
4.8. Ajuste orbital erróneo (Experimento 2) . . . . .	58
4.9. Ajuste orbital erróneo (Experimento 3) . . . . .	63
4.10. Valores reales y estimados (Etapa de exploración) . . . . .	67

# Capítulo 1

## Introducción

Se estima que un porcentaje mayoritario de las estrellas observadas en el cielo corresponde no a estrellas individuales sino a sistemas estelares múltiples. Dentro de estos, los más comunes corresponden a *estrellas binarias*: sistemas compuestos por dos estrellas ligadas gravitacionalmente. Bajo ciertas condiciones<sup>1</sup>, esto significa que ambas estrellas orbitan un centro de masa común, definiendo trayectorias elípticas derivadas de las leyes de movimiento de la Mecánica Clásica, sin interacción gravitacional *individual*<sup>2</sup> significativa con otros cuerpos celestes.

La masa estelar es la propiedad más importante de una estrella, pues es el principal factor que determina su estructura y evolución. Por ejemplo, las estrellas de masa muy pequeña (menores a 0.5 veces la masa solar) evolucionan directamente a Enanas Blancas en la mayoría de los casos; las estrellas de masa pequeña (bajo el rango de [1.8, 2.2] veces la masa solar) suelen desarrollar un núcleo de helio –aunque esto también depende de su composición–; las estrellas de masa intermedia (bajo el rango de [5, 6] veces la masa del Sol) a menudo desarrollan núcleos de carbono-oxígeno; mientras que las estrellas masivas (con un mínimo de 7 a 10 masas solares) suelen colapsar en supernovas al final de sus vidas. Con estas consideraciones, es común que los astrónomos agrupen las estrellas en función de sus masas.

Cuando existe ligazón gravitacional entre dos cuerpos celestes, la relación matemática co-

---

<sup>1</sup>Inexistencia de intercambio de masa entre ambos cuerpos: las estrellas del sistema pueden ser consideradas como objetos puntuales.

<sup>2</sup>Se enfatiza la palabra *individual* porque a nivel de sistema sí puede existir interacción gravitacional con otros cuerpos más lejanos –que perciben un efecto aproximadamente equivalente al que tendría la presencia de un solo astro con la misma masa total del sistema–, sin que dicha interacción afecte mayormente la dinámica interna (i.e., trayectorias orbitales periódicas).

nocida como Tercera Ley de Kepler impone restricciones sobre las órbitas definidas por éstos: dado un valor particular de la masa total  $M$  –suma de las masas individuales–, el período ( $P$ ) y semi-eje mayor ( $a$ ) de la elipse deben cumplir  $a^3/P^2 \propto M$ . Visto de otra forma, una trayectoria orbital corresponde a un solo valor de masa total. Por esta razón, el estudio de las posiciones de los astros en el tiempo es una herramienta clave para el estudio de su masa. Al tratarse cuerpos gravitacionalmente ligados, las estrellas binarias constituyen la base observacional del estudio de la masa estelar, pues las estrellas individuales, al no definir trayectorias orbitales, sólo pueden ser estudiadas en términos de sus características espectroscópicas. De hecho, la mayor parte de las estrellas fuera del Sistema Solar cuya masa ha sido determinada directamente corresponde a estrellas binarias. El estudio de estrellas binarias ha contribuido, entre otras cosas, a caracterizar lo que se conoce como estrellas de Secuencia Principal en el diagrama de Hertzsprung-Russell.

Este trabajo se enfoca en las llamadas *estrellas binarias visuales* –sistemas binarios cuyos cuerpos constituyentes pueden ser identificados individualmente en imágenes astronómicas–. Dadas las características y el volumen de las observaciones astronómicas, es necesario que los métodos computacionales utilizados en su estudio sean *automáticos* y *robustos* frente a la existencia de *ruido de observación*, *brechas temporales* y *pérdida de datos*. En el caso de las estrellas binarias, se suma a la lista el problema de la *incompletitud de órbita*: dada la extensión de los períodos orbitales posibles para un sistema binario –los cuales pueden extenderse hasta varios cientos de años–, frecuentemente las observaciones disponibles abarcan sólo una fracción de la órbita. Por lo demás, es común que los conjuntos de datos contengan observaciones de *calidad variable*, ya que éstas pueden provenir de diferentes fuentes de información: el mismo objeto puede ser observado –dependiendo del instante de observación– desde un observatorio terrestre o de uno satelital, con tecnología óptica o por medio de radioastronomía, y en condiciones ambientales variadas. Cabe mencionar que, a pesar de que el problema de las binarias visuales es conocido desde los inicios de la astronomía moderna (Galileo fue el primero en descubrir una estrella que pareció ser dos, y terminó siendo 6), no se trata en ningún caso de un problema obsoleto. Por el contrario, se espera que la evolución de la tecnología de observación astronómica conduzca –a través proyectos como Gaia [Lindgren *et al.*, 2007]– al descubrimiento de numerosos sistemas múltiples mediante astrometría, incluyendo binarias visuales y exoplanetas (que podrían ser analizados con las mismas herramientas aplicadas a sistemas estelares)

Como una manera de abordar las dificultades expuestas, el presente Trabajo de Título se plantea como **objetivo general** el desarrollar y evaluar un método de estimación automática

de elementos orbitales de *estrellas binarias visuales*, a partir de observaciones en el plano del cielo utilizando Filtrado Bayesiano. Específicamente, se plantea utilizar la técnica conocida como Filtro de Partículas. Las razones que respaldan esta decisión son principalmente dos: (i) estas técnicas permiten combinar conocimiento fenomenológico con conocimiento empírico (i.e., observaciones), siendo aptas para problemas cuya física es conocida; (ii) su enfoque estadístico permite modelar matemáticamente la incertidumbre asociada al problema, y utilizar ese conocimiento en el algoritmo. Como **objetivos específicos** se cuentan:

1. Desarrollar un simulador de trayectorias para estrellas binarias. El simulador debe aceptar elementos orbitales arbitrarios (i.e., cualquier forma, tamaño y orientación de elipse) y permitir la incorporación de ruido artificial de observación.
2. Formular un método basado en Filtrado Bayesiano para estimar los elementos orbitales y su incertidumbre respectiva.
3. Plantear casos de estudio para evaluar el desempeño del método, ejecutar y reportar los experimentos pertinentes.

Este documento se estructura como sigue:

- El Capítulo 2 entrega los fundamentos teóricos y observacionales para la comprensión de este trabajo.
- El Capítulo 3 detalla la solución propuesta.
- El Capítulo 4 plantea, reporta y analiza experimentos numéricos para evaluar el desempeño del método propuesto.
- El Capítulo 5 presenta las conclusiones generales del trabajo realizado.

# Capítulo 2

## Marco teórico

En el presente capítulo se presentan los conceptos fundamentales para la comprensión del trabajo realizado. Para ello, este capítulo se organiza en tres secciones, donde se explican los tópicos de astronomía y estadística pertinentes al tema de memoria: la Sección 2.1 introduce el concepto de estrellas binarias visuales y presenta el modelo dinámico del fenómeno; la Sección 2.2 aborda los problemas asociados a la observación; finalmente, la Sección 2.3 presenta los fundamentos de Filtrado Bayesiano y revisa en detalle la técnica de Filtro de Partículas, utilizada en este trabajo.

### 2.1. Binarias visuales

En astronomía, se denomina *sistema estelar múltiple* o simplemente *estrella múltiple* a aquellos sistemas constituidos por un número reducido de estrellas que interactúan de manera gravitacionalmente significativa, de manera que orbitan en torno a un centro de masa común. De esta manera, un sistema como  $\alpha$ -Cen (*Alpha Centauri*) es un sistema múltiple (específicamente, un sistema binario), mientras que sistemas con un gran número de estrellas y variados niveles de interacción gravitacional, como la *Vía Láctea*, a pesar de que en rigor también son un conjunto de estrellas interactuando gravitacionalmente, corresponden, dependiendo de su naturaleza y extensión, a otras categorías: *galaxias* (en el caso de la Vía Láctea), *cúmulos estelares*, entre otros.

La principal característica de una estrella es su masa, ya que ésta define en gran medida su estructura y evolución [Massey y Meyer, 2001]. La masa estelar puede ser determinada directamente mediante el estudio de su interacción gravitacional con otros objetos celestes, como se detalla en la Subsección 2.1.1. Se cree que al menos la mitad de las estrellas cercanas se encuentran en sistemas múltiples, por lo que las oportunidades de observar el movimiento de estrellas múltiples (y en particular, binarias) a fin de estimar sus masas son numerosas.

En la literatura los sistemas binarios se clasifican, más que por su física interna (masa, temperatura o tipo de órbita), por la técnica utilizada en su detección. Se conoce como *binarias visuales* a aquellos sistemas que pueden ser captados por un telescopio común, de manera que es posible resolver sus cuerpos constituyentes en una imagen. En general, se trata de estrellas cercanas al Sistema Solar con una separación significativa entre sí. A pesar de su relativa facilidad de observación, su identificación suele ser complicada, ya que su período orbital puede ser del orden de varios cientos de años, resultando difícil determinar si un par de estrellas corresponde realmente a un sistema binario, a estrellas pertenecientes a diferentes sistemas que casualmente cruzaron trayectorias, o simplemente a un par de estrellas cercanas en el plano de observación pero a una distancia real considerable entre sí y sin interacción gravitacional significativa (este último fenómeno se conoce como *binarias ópticas*).

Como se muestra en la Sección 2.1.1, la masa total del sistema binario impone condiciones sobre sus propiedades orbitales. Por esa razón, la observación de las trayectorias estelares es clave no sólo para discriminar entre binarias visuales y binarias ópticas, sino también para caracterizar a las primeras en términos de su masa total.

Otros tipos de sistemas binarios son:

- **Binarias eclipsantes:** Son sistemas en que la luminosidad observada presenta variaciones periódicas, producto de la superposición entre ambas estrellas a lo largo de la línea de vista. El estudio de sus propiedades se realiza mediante su curva de luz. Esta técnica se asemeja a la técnica de tránsitos para detección de exoplanetas.
- **Binarias espectroscópicas:** Se trata de sistemas en que sólo una de las estrellas es visible, de modo que la presencia de la estrella secundaria se detecta indirectamente mediante los efectos sobre el objeto observado. En este caso, se estudian las variaciones temporales del

espectro de la estrella observada –corrimiento al rojo o al azul, por efecto Doppler–, las cuales son ocasionadas por el movimiento orbital inducido por la interacción gravitacional con la estrella secundaria. Este método es equivalente a la técnica de velocidad radial para detección de exoplanetas.

### 2.1.1. Mecánica celeste

Interpretar las observaciones astronómicas, y particularmente transformar dichas observaciones en información sobre las propiedades físicas del sistema observado, requiere de un modelo que explique el fenómeno subyacente. A pesar de que el estudio de varios fenómenos astronómicos exige la consideración de enfoques más avanzados –tales como la teoría de la relatividad o la física cuántica–, el estudio de las binarias visuales se basa, al menos en una primera aproximación, en la formulación Newtoniana de la mecánica clásica: Leyes de Movimiento y Ley de Gravitación Universal.

Para efectos de análisis de datos, se considera que la estrella binaria: (i) es un sistema gravitacionalmente aislado, es decir, se desprecia la influencia gravitacional de otros cuerpos; (ii) es *desconectado* o *separado*<sup>1</sup>, es decir, sus estrellas constituyentes no intercambian material y por lo tanto poseen masas fijas; (iii) se constituye de masas puntuales: se desprecian efectos de marea, rotación y campos magnéticos. De esta manera, el sistema binario queda reducido al problema de los dos cuerpos, cuya solución es conocida y se revisa a continuación<sup>2</sup>:

En un sistema estelar binario, se denomina *estrella primaria* a la estrella de mayor luminosidad (en alguna banda de frecuencia) y *estrella secundaria* a la más tenue (en la misma banda). Sean  $\vec{r}_1$  y  $\vec{r}_2$  las posiciones de la primaria y la secundaria, respectivamente, en un sistema de referencia inercial con origen en el centro de masa ( $\frac{m_1\vec{r}_1+m_2\vec{r}_2}{m_1+m_2} = \vec{0}$ ). A partir de  $\vec{r}_1$  y  $\vec{r}_2$  se define el vector de posición relativa (de la estrella secundaria con respecto a la primaria):  $\vec{r} = \vec{r}_2 - \vec{r}_1$ , cuya magnitud es  $r = |\vec{r}|$ . La Segunda Ley de Newton establece que la fuerza total sobre un cuerpo (i.e. la suma vectorial de fuerzas individuales) es igual al producto de la masa por la aceleración. A la vez, el supuesto de que el sistema binario está gravitacionalmente aislado permite afirmar que la única fuerza presente es aquella que las estrellas ejercen mutuamente por efecto de la Ley de

---

<sup>1</sup>Preferentemente se utiliza su nombre en inglés: *detached*

<sup>2</sup>Una derivación detallada de la solución puede encontrarse en el Capítulo 6 de [Green, 1985]

Gravitación Universal. De esta manera, las ecuaciones de movimiento de los cuerpos individuales son:

$$\begin{aligned} m_1 \cdot \ddot{\vec{r}}_1 &= -\vec{F}, \\ m_2 \cdot \ddot{\vec{r}}_2 &= +\vec{F}, \end{aligned} \tag{2.1}$$

donde  $\vec{F} = -G\frac{m_1 \cdot m_2}{r^3}\vec{r}$  es la fuerza ejercida sobre el cuerpo  $m_2$ . A partir de las ecuaciones de fuerza de cada cuerpo es posible plantear una ecuación diferencial para  $\vec{r}$ :

$$\ddot{\vec{r}} = -\frac{G(m_1 + m_2)}{r^3}\vec{r}, \tag{2.2}$$

cuya solución, derivada detalladamente en el Apéndice A, es una elipse cuando la propiedad conocida como *energía orbital específica* ( $\epsilon$ , definida posteriormente) es negativa:

$$r(\nu) = \frac{a(1 - e^2)}{1 + e \cos \nu}. \tag{2.3}$$

El valor  $\nu$ , conocido como *anomalía verdadera* (ver Figura 2.1), corresponde al ángulo formado por el periastro, el foco  $F$  (ocupado por la primaria) y la secundaria. El largo del semi-eje mayor,  $a$ , y la excentricidad,  $e$ , dependen de la energía mecánica total, el momentum angular y la masa del sistema. Específicamente:

$$\begin{aligned} a &= -\frac{\mu}{2\epsilon} \quad (\epsilon < 0), \\ e &= \sqrt{1 + \frac{2\epsilon h^2}{\mu^2}} \quad (0 \leq e \leq 1), \end{aligned}$$

donde  $h = \|\vec{r} \times \vec{v}\|$  es el *momento angular específico*,  $\mu = G(m_1 + m_2)$  el *parámetro orbital estándar* y  $\epsilon$  la *energía orbital específica*, que se define como la suma de la energía cinética y la energía potencial del sistema, dividida por la masa reducida ( $\frac{m_1 \cdot m_2}{m_1 + m_2}$ ).

Si bien el problema de dos cuerpos fue resuelto en términos de la posición relativa,  $\vec{r}$ , las posiciones individuales  $\vec{r}_1$  y  $\vec{r}_2$  se pueden calcular a partir de  $\vec{r}$  como:

$$\begin{aligned} r_1 &= -\frac{m_2}{m_1 + m_2}r, \\ r_2 &= \frac{m_1}{m_1 + m_2}r, \end{aligned} \tag{2.4}$$

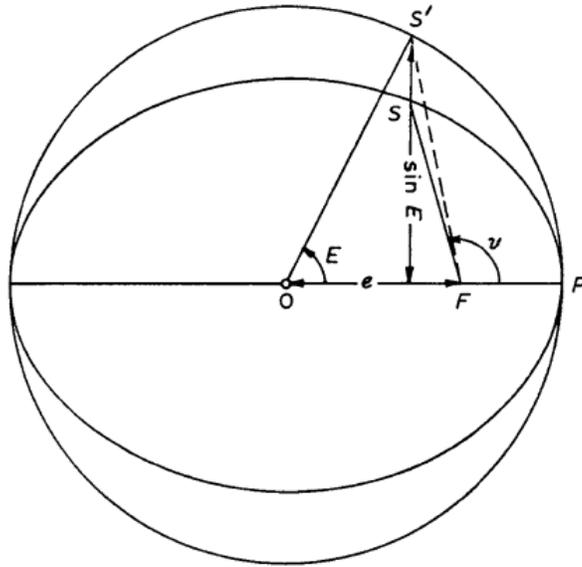


Figura 2.1: Elipse de la solución del problema de los dos cuerpos

de manera que las posiciones de ambas estrellas describen, también, trayectorias elípticas, como se ilustra en la Figura 2.2. Notar que el segmento que une ambas estrellas debe atravesar el centro de masa en todo momento.

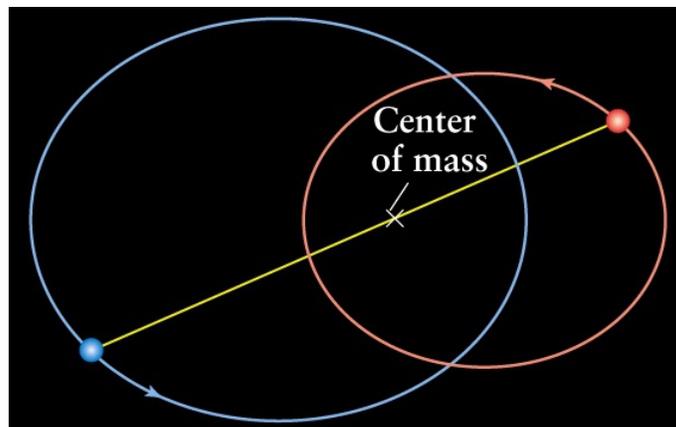


Figura 2.2: Trayectorias elípticas

Hasta este punto se ha descrito la elipse en función del ángulo  $\nu$ , sin embargo, no se ha explicitado la evolución temporal del sistema. Dado que ya se definió una ecuación que describe la separación  $r$  en función de  $\nu$ , basta conocer  $\nu$  en función del tiempo para describir el estado del sistema para cualquier instante  $\tau$ . A continuación, se introducen tres definiciones necesarias para la comprensión de  $\nu(\tau)$ :

- Se denomina  $P$  al período orbital del sistema.
- Se conoce como  $T$ , o *tiempo de paso por el periastro* al instante en que la estrella secundaria se encuentra en el periastro de la órbita (notar que, dada la periodicidad del sistema,  $T$  no es único: el sistema alcanza el punto de menor separación, o periastro, cada  $P$  años).
- Dada una *circunferencia auxiliar*, entendida como una circunferencia de radio igual al semi-eje mayor,  $a$ , y centrada en el mismo punto que la elipse, se define el valor  $E$  o *anomalía excéntrica* como el ángulo (medido desde el centro de la elipse) que forma la proyección perpendicular de la estrella secundaria sobre la circunferencia auxiliar (punto  $S'$  en la Figura 2.1), y el semi-eje mayor de la elipse (que coincide con el periastro). Ver Figura 2.1 para mayor claridad.

Como se explica en el Apéndice A.1, el ángulo  $E$  se relaciona con el tiempo  $\tau$  mediante la llamada Ecuación de Kepler:

$$\pi(\tau - T)/P = M = E - e \sin E. \quad (2.5)$$

A pesar de que esta expresión no posee una solución analítica, el valor de  $E$  puede ser aproximado mediante métodos numéricos (en particular, soluciones basadas en el método Newton-Raphson son frecuentemente utilizadas). Una vez determinado  $E(\tau)$ , es posible determinar  $\nu(\tau)$  por medio de la siguiente relación:

$$\tan \frac{\nu}{2} = \sqrt{\frac{1+e}{1-e}} \tan \frac{E}{2}. \quad (2.6)$$

Finalmente, cabe recalcar que la importancia de caracterizar de manera precisa la órbita de un sistema radica en que los parámetros orbitales están directamente relacionados con la masa total del sistema mediante la Tercera Ley de Kepler:

$$\frac{a^3}{P^2} = \frac{G}{4\pi^2}(m_1 + m_2). \quad (2.7)$$

Cuando  $a$  es expresado en *unidades astronómicas* ( $AU$ ),  $P$  en *año juliano* (365.25 días de 86400 segundos cada uno de acuerdo al Sistema Internacional de Unidades y denominado *yr* en este documento) y las masas en *masas solares* ( $M_\odot$ ), la ecuación anterior se simplifica a:

$$\frac{a^3}{P^2} = (m_1 + m_2). \quad (2.8)$$

## 2.2. Plano de observación y elementos geométricos

El *plano del cielo* o *plano de observación* es el plano perpendicular a la línea de vista entre el observador y el objeto observado. Las observaciones astrométricas corresponden a la proyección de la posición real de los cuerpos en el espacio sobre dicho plano. En general, a menos que se disponga de información adicional sobre los cuerpos observados (por ejemplo, datos de velocidad radial), no es posible inferir la posición real de éstos. En el caso de las binarias visuales, el conocimiento de la dinámica subyacente –el problema de dos cuerpos– permite caracterizar la *órbita real* a partir de la *órbita aparente*, como se conoce a la proyección de la órbita real sobre el plano del cielo. En esta sección se explica la geometría asociada a la observación y la relación entre ambas órbitas.

De acuerdo a lo visto en la Sección 2.1.1, el problema de dos cuerpos puede representarse tanto en términos de las posiciones individuales ( $\vec{r}_1, \vec{r}_2$ ) como en términos del vector de posición relativa  $\vec{r}$ . Las observaciones de estrellas binarias visuales/astrométricas utilizan una representación similar a la segunda, pero con algunas diferencias: (i) el origen del sistema está en la estrella primaria en lugar de en el centro de masa; (ii) describe la posición relativa proyectada en el plano del cielo en vez de la posición relativa real entre los astros. Esta representación comprende tres variables:

- $\tau$  o instante de observación (*época*).
- $\rho(\tau)$ , correspondiente a la magnitud de la separación angular entre ambas estrellas en el instante  $\tau$ .
- $\theta(\tau)$ , correspondiente al ángulo formado entre un vector de referencia –por lo general, la dirección del polo celeste norte– y la estrella secundaria.

Las ecuaciones presentadas en la Sección 2.1.1 describen la solución teórica del problema de los dos cuerpos, pero ignoran las dificultades asociadas a la orientación espacial, pues para la derivación de la Ecuación 2.2 se elige el centro de masa como origen del sistema, y el plano de la elipse como plano de referencia. Sin embargo, al momento de observar una estrella binaria real, su órbita puede tener una orientación arbitraria en el espacio, que generalmente no coincide con el plano del cielo. Es por esto que, junto a los parámetros orbitales  $P$ ,  $T$ ,  $a$  y  $e$ , definidos en la

Sección 2.1.1, se definen los parámetros geométricos  $i$ ,  $\Omega$ ,  $\omega$ , que describen la orientación de la órbita en el espacio:

1. *Inclinación de la órbita* ( $i$ ): Ángulo formado entre el plano orbital –i.e. donde habita la órbita real– y el plano de observación (donde habita la órbita aparente).
2. *Longitud del nodo ascendente* ( $\Omega$ ): Ángulo, en el plano de observación, del *nodo ascendente* respecto al norte astronómico. El *nodo ascendente* es el punto donde la estrella secundaria cruza el plano del cielo moviéndose desde el hemisferio sur al hemisferio norte celeste, mientras que el nodo descendente es el punto donde dicha estrella cruza el plano en sentido contrario. El segmento que une ambos puntos recibe el nombre de *línea de nodos* y corresponde a la intersección de la órbita real con su proyección en el plano del cielo.
3. *Argumento del periastro* ( $\omega$ ): Ángulo que va desde el nodo ascendente hasta el periastro de la órbita real. Indica la orientación de la órbita real dentro de su propio plano.

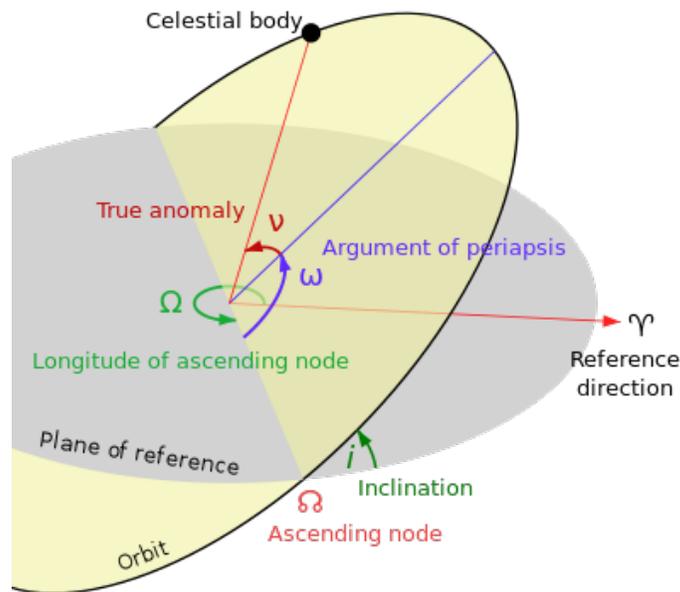


Figura 2.3: Elementos orbitales de orientación espacial

La Figura 2.3<sup>3</sup> ofrece una descripción gráfica de los parámetros  $i$ ,  $\omega$  y  $\Omega$ . El ángulo  $\nu$  es el mismo utilizado en la Sección 2.1.1 (Figura 2.1), es decir, aquel que describe la posición de la estrella secundaria respecto a la primaria (foco de la elipse) respecto al periastro. El vector  $\Upsilon$  define la dirección de referencia (polo celeste norte).

<sup>3</sup>Imagen por Lasumcty, CC-BY-SA-3.0

### 2.2.1. Representación de Thiele-Innes

El conjunto de parámetros  $\{T, P, e, a, \omega, \Omega, i\}$  se conoce como *Elementos de Campbell*. A continuación se presenta una representación equivalente: la representación de Thiele-Innes.

Sean  $x_r$  e  $y_r$  las coordenadas de posición relativa de la estrella secundaria en el plano de órbita real. Se definen las *coordenadas normalizadas* de posición relativa como  $x = x_r/a$ ,  $y = y_r/a$ . Dado un valor de anomalía excéntrica  $E$  –que depende, a su vez, de  $T$ ,  $P$ ,  $e$  y el instante  $\tau$ –, el valor de las *coordenadas normalizadas* se calcula como:

$$\begin{aligned} x(E) &= \cos E - e \\ y(E) &= \sin E \sqrt{1 - e^2}. \end{aligned} \tag{2.9}$$

Dada una configuración orbital definida por valores específicos de  $\{T, P, e, a, \omega, \Omega, i\}$ , la órbita aparente se puede calcular fácilmente a partir de los valores de  $x$  e  $y$  [Van de Kamp y Stearns, 1967]. Se definen las constantes de Thiele-Innes como:

$$\begin{aligned} A &= a(\cos \omega \cos \Omega - \sin \omega \sin \Omega \cos i) \\ B &= a(\cos \omega \sin \Omega + \sin \omega \cos \Omega \cos i) \\ F &= a(-\sin \omega \cos \Omega - \cos \omega \sin \Omega \cos i) \\ G &= a(-\sin \omega \sin \Omega + \cos \omega \cos \Omega \cos i). \end{aligned} \tag{2.10}$$

Entonces las posiciones  $X$ ,  $Y$  en la órbita aparente (es decir, la proyección de las posiciones reales en el plano del cielo) pueden calcularse a partir de  $x$ ,  $y$  y las constantes de Thiele-Innes a través de las siguientes ecuaciones:

$$\begin{aligned} X &= Bx + Gy \\ Y &= Ax + Fy \end{aligned} \tag{2.11}$$

Estas ecuaciones son el resultado de calcular la posición real de la secundaria en función de los parámetros orbitales y agrupar, mediante manipulación algebraica, ciertos valores en las constantes definidas en las Ecuaciones 2.10, para luego rescatar la componente tangencial al plano del cielo. La derivación completa de este resultado puede encontrarse en [Binnendijk, 1960].

Además de proveer una forma sencilla de calcular la órbita aparente a partir de una configuración orbital, las constantes de Thiele-Innes entregan una representación alternativa de los parámetros  $a$ ,  $\Omega$ ,  $\omega$ ,  $i$ , ya que estos valores están matemáticamente relacionados con  $A$ ,  $B$ ,  $F$ ,  $G$  mediante las siguientes expresiones:

$$\begin{aligned} \tan(\omega + \Omega) &= \frac{B - F}{A + G} \\ \tan(\omega - \Omega) &= \frac{-B - F}{A - G} \\ a^2(1 + \cos^2 i) &= A^2 + B^2 + F^2 + G^2 \\ a^2 \cos^2 i &= AG - BF \end{aligned} \tag{2.12}$$

Si bien estas relaciones no son completamente unívocas, puesto que existen ambigüedades de ángulo y cuadrante al calcular  $(a, \Omega, \omega, i)$  a partir de  $(A, B, F, G)$ , se considera que ambas representaciones son equivalentes, pues dichas ambigüedades se resuelven acotando el rango de la solución. En el caso del parámetro  $\Omega$ , por ejemplo, se escoge por defecto la solución entre  $0$  y  $\pi$  y se utilizan datos de velocidad radial para desambiguar su valor real [Van de Kamp y Stearns, 1967], que puede ser la solución seleccionada inicialmente, o su solución en el rango  $[\pi, 2\pi]$ . De esta manera, los elementos de Campbell,  $\{T, P, e, a, \omega, \Omega, i\}$ , pueden ser reemplazados por la representación de Thiele-Innes,  $\{T, P, e, A, B, F, G\}$ . Algunos métodos trabajan directamente con dicha representación [Pourbaix, 1994].

### 2.3. Inferencia Bayesiana

Las teorías de la física nos permiten realizar predicciones: dada una descripción completa del sistema, es posible conocer la evolución de su estado. El problema de predecir la evolución del estado de un sistema es conocido como *modelación*, *simulación* o *problema directo*. Por el contrario, se conoce como *problema inverso* a la tarea de convertir mediciones reales en información acerca del sistema. Estos problemas surgen dentro de numerosas áreas de la ciencia e ingeniería, pudiéndose citar como ejemplo la oceanografía [Wunsch, 1996], la sismología [Lailly, 1983] o la medicina, donde las técnicas de EEG/ECG<sup>4</sup> [Pascual-Marqui, 1999, Rudy, 1987] son

---

<sup>4</sup>Electroencefalografía y electrocardiografía, respectivamente

casos relevantes.

A pesar de que los problema directos suelen tener solución única –al menos en la física determinista–, es frecuente que los problemas inversos no estén *bien planteados* en el sentido Hadamard<sup>5</sup> y tengan, por tanto, múltiples soluciones [Tarantola, 2005]. Por esta razón, al momento de buscar solución a un problema inverso es necesario incluir en el análisis tanto conocimiento previo como sea posible, así como tomar en consideración la incertidumbre asociada a los modelos utilizados y las mediciones realizadas.

Es acá donde adquiere importancia el enfoque *Bayesiano*, pues proporciona una forma de incluir la información previa –conocida en este contexto como información *a priori*– en el análisis. El término *Bayesiano* designa a una serie de conceptos y métodos dentro de las Probabilidades y la Estadística basados en la aplicación del Teorema de Bayes (Ecuación 2.13), el cual entrega una forma de reescribir la probabilidad condicional entre dos afirmaciones  $X$  e  $Y$ . En particular, se conoce como *Inferencia Bayesiana* al tipo de inferencia estadística en la cual las observaciones se emplean para actualizar o inferir la probabilidad de una hipótesis.

$$p(X|Y) = \frac{p(Y|X) p(X)}{p(Y)} \quad (2.13)$$

La notación utilizada en la Ecuación 2.13 no es casual, pues a pesar de que el teorema es válido para dos proposiciones  $X$  e  $Y$  cualesquiera, resulta de particular utilidad cuando el objeto de interés es una variable  $X$  estudiada a través de una observación  $Y$ . Dicha utilidad radica en que  $p(X|Y)$  –llamada probabilidad *a posteriori*– podría no ser calculable directamente, sin embargo, el teorema permite expresarla en función de las expresiones  $p(Y|X)$  –conocida como *verosimilitud*– y la probabilidad *a priori*  $p(X)$ , cuyos valores dependen de la información disponible previamente. Por lo general, la expresión  $p(Y)$  –llamada *probabilidad marginal* de  $Y$ – no tiene mayor importancia en el análisis, ya que su valor corresponde simplemente al factor de normalización necesario para que la probabilidad *a posteriori* cumpla el segundo axioma de probabilidad: la integral de  $p(X|Y)$  entre  $-\infty$  y  $\infty$  debe ser igual a 1.

En lo que sigue del capítulo se plantea el problema de estimar el estado de un sistema variante en el tiempo desde un enfoque Bayesiano y se introduce el concepto de Métodos Secuenciales de

---

<sup>5</sup>Una definición detallada de este concepto puede encontrarse en [Kabanikhin, 2008]

Monte Carlo, detallando la formulación de Filtro de Partículas utilizada en el presente trabajo.

### 2.3.1. Estimación de estado

Fenómenos de diversa índole son caracterizables matemáticamente como un *sistema dinámico*, vale decir, como un sistema cuya evolución temporal puede ser descrita mediante una función de su estado actual  $x_t$  y un conjunto de parámetros  $\theta$ . En ocasiones también se considera una entrada  $u_t$  externa al sistema (la Teoría de Control Automático [Lee y Markus, 1967] es un ejemplo de ello). A continuación se muestra una forma general de la ecuación de evolución, en sus versiones continua (Ecuación 2.14) y discreta (Ecuación 2.15). Este trabajo estará centrado en sistemas de tiempo discreto, pues las mediciones y el procesamiento se realizarán de manera digital.

$$\dot{x}_t = f_t(x_t, u_t | \theta) \quad (2.14)$$

$$x_{t+1} = f_t(x_t, u_t | \theta) \quad (2.15)$$

Cabe mencionar que en ambos casos es necesario conocer la condición inicial del sistema,  $x_0$ , para que la caracterización sea completa. Notar que  $f$  también puede ser dependiente del tiempo.

Las Ecuaciones 2.14 y 2.15 corresponden a un planteamiento determinístico, sin embargo, dada la incertidumbre asociada al estudio de fenómenos naturales –ya sea porque las observaciones son ruidosas o porque el fenómeno en sí mismo incluye un grado de aleatoriedad–, es necesario plantear una variante estocástica de los sistemas dinámicos. En ese planteamiento, el proceso se caracteriza como una cadena de Markov de primer orden con probabilidad de transición  $p_{x_{t+1}|x_t}(x_{t+1}|x_t)$  y cierta condición inicial  $p_{x_0}(x)$ . Los estados  $x_t$  no se caracterizan por un valor puntual sino por una distribución de probabilidad, como se muestra a continuación<sup>6</sup>:

$$x_0 \sim p_{x_0}(x_0) \quad (2.16)$$

$$x_{t+1}|x_t \sim p_{x_{t+1}|x_t}(x_{t+1}|x_t). \quad (2.17)$$

---

<sup>6</sup>Se omite el término  $u_t$  en las ecuaciones, dado que el problema estudiado en este trabajo no incluye ninguna entrada externa.

Frecuentemente, los estados  $x_t$  del sistema de interés no se observan directamente sino que se estudian a través de observaciones  $y_t$ , que dependen del estado  $x_t$  y poseen, igualmente, un grado de incertidumbre:

$$y_t|x_t \sim p_{y_t|x_t}(y_t|x_t). \quad (2.18)$$

Un caso particular de la estructura presentada en las ecuaciones 2.17 y 2.18 son los sistemas caracterizables por ecuaciones de recurrencia no determinísticas. Similares en forma a la Ecuación 2.15, estos modelos relacionan la probabilidad del estado  $x_{t+1}$  con el estado anterior  $x_t$  mediante una función  $f$ , sin embargo agregan componentes no determinísticos a la ecuación. Las Ecuaciones 2.19 y 2.20 se denominan *ecuación de transición* y *ecuación de observación*, respectivamente. El componente no determinístico viene dado por las variables aleatorias i.i.d.  $\omega_t$  (conocida como *ruido de proceso*) y  $v_t$  (*ruido de observación*).

$$x_{t+1} = f_t(x_t, \omega_t) \quad (2.19)$$

$$y_t = h_t(x_t, v_t). \quad (2.20)$$

Notar que, dado que no hay restricción sobre  $f$  y  $h$ , éstas pueden tomar cualquier forma, incluyendo funciones no-lineales. De igual manera, los ruidos  $\omega_t$  y  $v_t$  pueden tener una función de densidad de probabilidad arbitraria.

En *Filtrado Bayesiano*, definido como el proceso de usar observaciones ruidosas para estimar la densidad de probabilidad posterior del estado de un sistema dinámico no necesariamente lineal ni Gaussiano [Haug, 2005], es de interés calcular el valor  $p(x_{0:t}|y_{1:t})$  (que equivale a la distribución *a posteriori* del proceso, dadas las observaciones). Para ello, considérese que la ecuación de evolución de estado define la densidad *a priori*, mientras que la ecuación de observación define la *verosimilitud*, es decir:

$$p(x_{1:t}) = p(x_1) \prod_{j=2}^t p(x_j|x_{j-1}) \quad (2.21)$$

$$p(y_{1:t}|x_{1:t}) = \prod_{j=1}^t p(y_j|x_j) \quad (2.22)$$

En este contexto, la inferencia sobre de los estados  $\{x_t\}_{t \geq 1}$  dado un conjunto de observaciones  $\{y_t\}_{t \geq 1}$  se expresa como la distribución *a posteriori*:

$$p(x_{1:t}|y_{1:t}) = \frac{p(y_{1:t}|x_{1:t})p(x_{1:t})}{p(y_{1:t})}, \quad (2.23)$$

donde la distribución marginal de  $y_t$  es:

$$p(y_{1:t}) = \int_{x \in \mathcal{X}} p(x_{1:t}, y_{1:t}) \cdot dx_{1:t} \quad (2.24)$$

En aplicaciones donde la adquisición de observaciones es secuencial, es conveniente contar con una relación recursiva para la actualización de información. Tomando en consideración que en numerosos casos importa estimar el estado actual del proceso<sup>7</sup>,  $x_t$ , y no todo el camino recorrido por éste, la distribución de interés es  $p(x_t|y_{1:t})$  en lugar de  $p(x_{1:t}|y_{1:t})$ . Es posible derivar una relación recursiva para  $p(x_t|y_{1:t})$  en dos etapas: predicción y actualización [Arulampalam *et al.*, 2002].

El objetivo de la etapa de predicción es calcular  $p(x_t|y_{1:t-1})$ . Para ello se asume que la distribución *a posteriori* del instante anterior,  $p(x_{t-1}|y_{1:t-1})$ , es conocida. Bajo ese supuesto, es posible obtener una expresión para  $p(x_t|y_{1:t-1})$  (Ecuación 2.25) mediante la aplicación de la ley de Chapman-Kolmogorov y haciendo uso de que  $p(x_t|x_{t-1}, y_{1:t-1}) = p(x_t|x_{t-1})$ , lo cual se cumple por ser  $\{x_t\}_{t \geq 0}$  un proceso de Markov de primer orden.

$$p(x_t|y_{1:t-1}) = \int_{x \in \mathcal{X}} p(x_t|x_{t-1})p(x_{t-1}|y_{1:t-1})dx_{t-1} \quad (2.25)$$

Por otro lado, la etapa de actualización busca incorporar la información recibida en el tiempo  $t$  (la observación  $y_t$ ) para mejorar la estimación de  $x_t$  realizada en la etapa de predicción. Para ello, se deriva la Ecuación 2.26 mediante el teorema de Bayes.

$$p(x_t|y_{1:t}) = \frac{p(y_t|x_t)p(x_t|y_{t-1})}{p(y_t|y_{1:t-1})} \quad (2.26)$$

---

<sup>7</sup>En ciertas aplicaciones también se busca mejorar la estimación de valores pasados, proceso conocido como *smoothing*, así como estimar valores futuros, lo que se conoce como *predicción*.

Notar que  $p(x_t|y_{t-1})$  es la distribución obtenida en la etapa de predicción y que  $p(y_t|y_{1:t-1})$  es una constante de normalización, mientras que  $p(y_t|x_t)$  se obtiene de la medición  $y_t$ .

Las Ecuaciones 2.25 y 2.26 presentan una solución óptima para el problema de estimación Bayesiana [Arulampalam *et al.*, 2002]. Sin embargo, ésta no pasa de ser una solución teórica, pues en la mayoría de los casos no es posible encontrar una expresión analítica para las densidades de probabilidad mencionadas. En casos restringidos, básicamente cuando las Ecuaciones 2.19 y 2.20 definen un modelo de espacio-estado lineal y los ruidos  $\omega_t$  y  $v_t$  poseen distribuciones Gaussianas, existe una solución explícita: el conocido filtro de Kalman [Kalman, 1960]. Cuando, por el contrario, la función que describe la evolución del sistema es no-lineal y los ruidos siguen distribuciones de probabilidad no Gaussianas, se hace necesario implementar soluciones *aproximadas* o *sub-óptimas*, tales como el filtro extendido de Kalman, los métodos basados en grilla o métodos secuenciales de Monte Carlo.

### 2.3.2. Métodos secuenciales de Monte Carlo: Filtro de partículas

El término *Métodos de Monte Carlo*, acuñado por Nicholas Metropolis en 1949 [Metropolis y Ulam, 1949], define una clase de algoritmos computacionales basados la selección aleatoria reiterada de muestras de un conjunto a fin de obtener resultados numéricos. Surgidos en el marco del desarrollo de armas atómicas [Candy, 2011], fueron inicialmente concebidos como una alternativa a los métodos tradicionales de integración numérica basados en grilla, y actualmente son también utilizados en problemas de optimización –dando origen a un área conocida como *optimización estocástica*– y en la caracterización de distribuciones de probabilidad.

Métodos convencionales como *Markov Chain Monte Carlo* (MCMC) han sido exitosamente utilizados para aproximar integrales y densidades de probabilidad. Sin embargo, su uso en el problema de inferencia Bayesiana no es directo, ya que éste involucra una secuencia de densidades de probabilidad variantes en el tiempo, mientras que MCMC asume que la densidad objetivo es invariante. Por esta razón, es necesario contar con una formulación secuencial de los métodos de Monte Carlo que permita incorporar las observaciones para mejorar las estimaciones realizadas en forma recursiva.

Uno de los algoritmos que abordan este problema es el *Filtro de Partículas*. Esta herramienta

es definida en [Del Moral *et al.*, 2006] como una técnica numérica para aproximar integrales multidimensionales con respecto a una secuencia de medidas de probabilidad. Dada una sucesión de distribuciones  $\{\pi_t(x_{0:t})\}_{t \geq 1}$ , este método busca estimar dicha secuencia mediante la generación de  $N \gg 1$  muestras aleatorias ponderadas  $\{\omega_t^{(i)}, x_{0:t}^{(i)}\}_{i \in \{1, \dots, N\}}$  –las *partículas*–, donde  $\omega_t^{(i)}$  representa el peso de la muestra  $i$  en el tiempo  $t$ . La idea es aproximar  $\pi_t(x_{0:t})$  mediante una distribución empírica, como se muestra en la Ecuación 2.27:

$$\pi_t^N(x_{0:t}) = \sum_{i=1}^N \omega_t^{(i)} \delta_{x_{0:t}^{(i)}}(x_{0:t}) \quad (2.27)$$

donde  $\delta_\alpha(\cdot)$  es el delta de Kronecker centrado en  $\alpha$ .

En el contexto de Inferencia Bayesiana, la distribución objetivo  $\pi_t(x_{0:t})$  es la posterior  $p(x_{0:t}|y_{1:t})$ , que con este método puede ser estimada mediante la distribución empírica  $\pi_t^N(x_{0:t})$ . De esta manera, de acuerdo a [Andrieu *et al.*, 2001] las esperanzas pueden ser aproximadas mediante:

$$\begin{aligned} \int_{\mathcal{X}^{t+1}} \phi_t(x_{0:t}) \pi_t(x_{0:t}) dx_{0:t} &\simeq \int_{\mathcal{X}^{t+1}} \phi_t(x_{0:t}) \pi_t^N(x_{0:t}) dx_{0:t} \\ &= \int_{\mathcal{X}^{t+1}} \phi_t(x_{0:t}) \sum_{i=1}^N \omega_t^{(i)} \delta_{x_{0:t}^{(i)}}(x_{0:t}) dx_{0:t} \\ &= \sum_{i=1}^N \omega_t^{(i)} \phi_t(x_{0:t}^{(i)}), \end{aligned} \quad (2.28)$$

para alguna función de interés  $\phi_t(\cdot)$ . Así, el problema de inferencia Bayesiana se reduce a la selección de muestras y sus respectivos pesos de manera que la sumatoria de la Ecuación 2.28 converja a  $\mathbb{E}_{\pi(x_{0:t})}(\phi(x_{0:t}))$ .

En las subsecciones siguientes se discutirán los problemas asociados al *muestreo de importancia* tradicional para luego plantear *muestreo de importancia secuencial con remuestreo*. Finalmente, se describe una forma de utilizar el Filtro de Partículas para estimar parámetros: la Evolución Artificial de Parámetros.

### 2.3.2.1. Muestreo de importancia

La estimación de  $\mathbb{E}_{\pi_t(x_{0:t})}(\phi(x_{0:t}))$  puede ser realizada mediante una aproximación de Monte Carlo, esto es, generar  $N$  caminos  $x_{0:t}^{(i)}$  y ajustar sus pesos de manera que su promedio aproxime aritméticamente la integral cuando  $N \rightarrow \infty$ . Una forma de resolver el problema es extraer las  $N$  muestras  $x_{0:t}^{(i)}$  según la distribución de probabilidad  $\pi_t(x_{0:t}) = p(x_{0:t}|y_{1:t})$  y calcular un promedio de la integral con pesos uniformes:

$$\hat{\mathbb{E}}_{\pi_t(x_{0:t})}(\phi(x_{0:t})) = \frac{1}{N} \sum_{i=1}^N \phi(x_{0:t}^{(i)}). \quad (2.29)$$

De esta manera, es la propia distribución  $p(x_{0:t}|y_{1:t})$  la que dará mayor peso a los valores de mayor impacto en la estimación, pues el muestreo favorecerá los valores de mayor probabilidad. Sin embargo, los casos de interés son aquellos donde esta distribución no es conocida o bien no se pueden extraer muestras de ella.

Una forma de evitar la extracción directa de muestras de  $p(x_{0:t}|y_{1:t})$  es suponer la existencia de una densidad  $q(x_{0:t}|y_{1:t})$ , llamada función de importancia, de la cual es posible muestrear y evaluar puntualmente salvo por una constante de normalización. Considerando esta densidad, las esperanzas pueden reescribirse de la siguiente forma [Geweke, 1989]:

$$\begin{aligned} \mathbb{E}_{\pi_t(x_{0:t})}(\phi(x_{0:t})) &= \int_{\mathcal{X}^{t+1}} \phi_t(x_{0:t}) p(x_{0:t}|y_{1:t}) dx_{0:t} \\ &= \frac{\int_{\mathcal{X}^{t+1}} \phi_t(x_{0:t}) \omega(x_{0:t}) q(x_{0:t}|y_{1:t}) dx_{0:t}}{\int_{\mathcal{X}^{t+1}} \omega(x_{0:t}) q(x_{0:t}|y_{1:t}) dx_{0:t}} \end{aligned} \quad (2.30)$$

donde  $\omega(x_{0:t})$  son pesos de importancia dados por un cociente entre la distribución objetivo y la función de importancia (derivada de Radon-Nikodyn):

$$\omega(x_{0:t}) = \frac{p(x_{0:t}|y_{1:t})}{q(x_{0:t}|y_{1:t})}. \quad (2.31)$$

Para que estos pesos estén bien definidos es necesario que el soporte de  $q(\cdot|y_{1:t})$  incluya el de  $p(\cdot|y_{1:t})$ . Notar que el peso en el denominador de la Ecuación 2.30 permite que  $q(\cdot|y_{1:t})$  no sea necesariamente una función de distribución de probabilidad.

De esta forma, es posible construir una aproximación de Monte Carlo de  $\mathbb{E}_{\pi_t(x_{0:t})}(\cdot)$  mediante la extracción de  $N$  muestras según  $q(x_{0:t}|y_{1:t})$ , como se muestra a continuación:

$$\begin{aligned}\hat{\mathbb{E}}_{\pi_t(x_{0:t})}(\phi(x_{0:t})) &= \frac{\sum_{i=1}^N \phi(x_{0:t}^{(i)}) \omega(x_{0:t}^{(i)})}{\sum_{i=1}^N \omega(x_{0:t}^{(i)})} \\ &= \sum_{i=1}^N \phi(x_{0:t}^{(i)}) \bar{\omega}(x_{0:t}^{(i)}),\end{aligned}\tag{2.32}$$

donde  $\bar{\omega}(\cdot)$  son los pesos normalizados. Este enfoque se conoce como *muestreo de importancia* o IS<sup>8</sup>, por sus siglas en inglés.

Notar que para  $N$  finito el estimador  $\hat{\mathbb{E}}_{\pi_t(x_{0:t})}$  es sesgado, pues es el cociente entre dos estimadores. Sin embargo, bajo ciertos supuestos la ley de los grandes números asegura que el estimador converge casi seguramente a la esperanza:

$$\hat{\mathbb{E}}_{\pi}(\phi(x_{0:t})) \xrightarrow{\text{c.s.}} \mathbb{E}_{\pi}(\phi(x_{0:t})).\tag{2.33}$$

Adicionalmente, el teorema central del límite garantiza la convergencia independiente de la dimensión del integrando [Geweke, 1989], lo cual presenta una ventaja del enfoque de Monte Carlo con IS por sobre la integración numérica basada en grilla, cuyo error crece conforme aumenta la dimensionalidad del problema [Candy, 2011].

El muestreo de importancia es ampliamente usado en los métodos de Monte Carlo para reducir la varianza de la estimación. Este método explota el hecho de que ciertos valores tienen una mayor importancia en la estimación que otros (de ahí su nombre). Sin embargo, el uso de funciones de importancia  $q$  no representativas del proceso puede llevar a estimaciones sesgadas, a menos que se ponderen correctamente las estimaciones obtenidas según sus pesos  $\omega$ .

A pesar de que IS entrega un enfoque simple para aproximar integrales definidas invariantes, no es adecuado dentro del contexto de estimación recursiva, ya que cada vez que se desee estimar  $p(x_{0:t}|y_{1:t})$  se necesita todo el conjunto de valores  $y_{1:t}$ , haciendo el procedimiento computacionalmente costoso debido a la alta dimensionalidad de la densidad objetivo. Por lo demás, de acuerdo a [Doucet *et al.*, 2001] el uso de muestreo de importancia en el contexto de estimación recursiva

---

<sup>8</sup>*Importance sampling*

ha dejado de manifiesto que la varianza de los pesos de las estimaciones diverge si no se incluyen rutinas de remuestreo. En resumen, cada vez que una observación  $y_t$  es recibida, interesa simplemente recalcular los pesos y las muestras representativas de la densidad de filtrado de forma secuencial y sin degeneración<sup>9</sup> de los pesos.

### 2.3.2.2. Muestreo de importancia secuencial y remuestreo

La motivación de formular una versión secuencial del muestreo de importancia es disminuir el alto costo computacional asociado a la utilización de toda la secuencia  $y_{1:t}$  cada vez que se quiera estimar  $p(x_{0:t}|y_{1:t})$ . El *muestreo de importancia secuencial* (SIS<sup>10</sup>, por su acrónimo en inglés) se basa en estimar la densidad *a posteriori* de manera recursiva, conservando en el tiempo  $t$  las muestras del instante anterior  $\{x_{0:t-1}^{(i)}; i = 1, \dots, N\}$  y extrayendo solamente las muestras  $\{x_t^{(i)}; i = 1, \dots, N\}$  en base a la observación  $y_t$  y a  $p(x_{0:t-1}|y_{1:t-1})$ . Adicionalmente, para solucionar el problema de la degeneración de pesos es posible introducir una etapa de *remuestreo* de la población de partículas, el que asegura una varianza mínima de los pesos, permitiendo que la densidad empírica  $\pi_t^N(x_{0:t}^{(i)})$  sea representativa de  $p(x_{0:t}|y_{1:t})$ .

Para asegurar que las muestras  $\{x_{0:t-1}^{(i)}; i = 1, \dots, N\}$  puedan ser conservadas en el tiempo  $t$ , se debe disponer de una función de importancia  $q(x_{0:t}|y_{1:t})$  que admita a la función de importancia en el instante anterior,  $q(x_{0:t-1}|y_{1:t-1})$ , como distribución marginal:

$$q(x_{0:t}|y_{1:t}) = q(x_{0:t-1}|y_{1:t-1})q(x_t|x_{0:t-1}, y_{1:t}). \quad (2.34)$$

Esto quiere decir que el muestreo de  $x_{0:t}$  desde  $q(x_{0:t-1}|y_{1:t})$  no depende de  $y_t$ . Con este supuesto, es posible obtener muestras  $x_{0:t}^{(i)}$  sin necesidad de alterar  $x_{0:t-1}^{(i)}$ :

$$x_{0:t}^{(i)} = \left[ x_{0:t-1}^{(i)}, x_t^{(i)} \right], \quad (2.35)$$

donde  $x_t^{(i)}$  se extrae según  $q(x_t|x_{0:t-1}^{(i)})$ . Además, es necesario contar con una relación que permita

---

<sup>9</sup>Que la población de partículas degenere quiere decir que su capacidad para aproximar la distribución buscada decae progresivamente. Sin una estrategia que se encargue de hacer frente a este problema, el fenómeno se traduce en el colapso de todos los pesos en una sola partícula.

<sup>10</sup>*Sequential Importance Sampling*

actualizar los pesos de importancia de manera recursiva, la cual está dada por:

$$\bar{\omega}(x_{0:t}^{(i)}) \propto \bar{\omega}(x_{0:t-1}^{(i)}) \frac{p(y_t|x_{0:t}^{(i)})p(x_t^{(i)}|x_{t-1}^{(i)})}{q(x_t^{(i)}|x_{0:t-1}^{(i)}, y_{1:t})}. \quad (2.36)$$

Un caso importante es cuando se considera la densidad *a priori* del estado como función de importancia, es decir,  $q(x_{0:t}|y_{1:t}) = p(x_{0:t})$ . Con esta relación, la recursión para los pesos de importancia queda expresada como:

$$\bar{\omega}(x_{0:t}^{(i)}) = \bar{\omega}(x_{0:t-1}^{(i)})p(y_t|x_{0:t}^{(i)}). \quad (2.37)$$

De esta forma, los pesos pueden ser actualizados en forma recursiva directamente desde las nuevas observaciones evaluadas en la función de verosimilitud.

A pesar de que este enfoque resuelve el problema de manera más eficiente que IS, tiene el inconveniente de que la distribución de los pesos de importancia se vuelve cada vez más disímil, al punto de que sólo una partícula la tiene peso no nulo [Doucet *et al.*, 2001]. Es por ello que surge el enfoque de *remuestreo secuencial de importancia* (SIR por sus siglas en inglés), que aprovecha la estructura de SIS para generar muestras y calcular los pesos según su verosimilitud, pero incorpora una etapa de remuestreo en el que se remueven elementos de bajo peso –y por tanto, de baja verosimilitud– para generar partículas donde la densidad objetivo es alta y así evitar el esfuerzo computacional perdido en la actualización de partículas de poco impacto en la estimación.

Un criterio para decidir cuándo realizar remuestreo es considerar el *tamaño efectivo de muestras* (en inglés *effective samples size*), concepto introducido por [Liu, 1996] y definido en [Kong *et al.*, 1994] como:

$$N_{eff} = \frac{N}{1 + \text{Var}_{(\cdot|y_{1:t})}[\omega(x_{0:t})]}. \quad (2.38)$$

A pesar de que este valor no puede ser calculado, el siguiente estimador es propuesto en [Doucet *et al.*, 2000]:

$$\hat{N}_{eff} = \frac{\left(\sum_{i=1}^N \omega(x_{0:t}^{(i)})\right)^2}{\sum_{i=1}^N \omega(x_{0:t}^{(i)})^2}. \quad (2.39)$$

Básicamente, cada vez que  $\hat{N}_{eff}$  cae bajo cierto umbral, el algoritmo procede a ejecutar una rutina

de remuestreo. Existen varias formas de realizar el remuestreo, siendo el enfoque multinomial [Del Moral, 2004] uno de los más comunes. Este enfoque consiste en extraer muestras (con reemplazo) del conjunto  $\{x_{0:t}^{(i)}\}_{i=1,\dots,N}$  con probabilidad proporcional a su peso. Una forma de extraer estas muestras es utilizar:

$$\hat{x}_{0:t}^{(i)} = \text{Inv}(u_i), i = 1, \dots, N. \quad (2.40)$$

Donde  $\hat{x}_{0:t}^{(i)}$  son las nuevas muestras,  $u_i$  son valores aleatorios generados según  $(U)(0, 1)$  e  $\text{Inv}(\cdot)$  es la inversa de la distribución empírica acumulada de  $x_t$ :

$$\text{Inv}(u_i) = x_{0:t}^{(i)} \Leftrightarrow \left[ \sum_{j=1}^{i-1} \omega(x_{0:t}^{(j)}), \sum_{j=1}^i \omega(x_{0:t}^{(j)}) \right]. \quad (2.41)$$

Cabe mencionar que después del remuestreo los nuevos pesos  $\omega$  son uniformes, ya que las muestras fueron extraídas según la estimación de la distribución posterior del estado.

### 2.3.2.3. Filtro de partículas para estimación de parámetros

A pesar de que el planteamiento inicial de este capítulo es la estimación de estados, en variadas situaciones es de interés estimar los parámetros del modelo utilizado, ya que estos pueden ser desconocidos o bien variantes en el tiempo para horizontes prolongados de observación. Para entender el problema de estimación de parámetros fuera del contexto Bayesiano, se define el vector de parámetros  $\theta$ . El estimador de máxima verosimilitud se obtiene a partir de la maximización de la función de log-verosimilitud:

$$l(\theta) = \log[L(\theta)] = \sum_{t=1}^T \log[p(y_t|y_{1:t-1}, \theta)], \quad (2.42)$$

donde

$$p(y_t|y_{1:t-1}, \theta) = \int p(y_t|x_t, \theta)p(x_t|y_{1:t-1}, \theta) \quad (2.43)$$

debe ser aproximada por Monte Carlo.

La maximización de la Ecuación 2.42 como criterio para estimar  $\theta$  no es siempre directa, y

las aproximaciones de la Ecuación 2.43 hacen ese método altamente poco práctico, debido al alto costo computacional involucrado cuando la estimación de  $\theta$  se desea realizar en cada iteración de tiempo real. En este contexto, se propone como alternativa un enfoque Bayesiano que incorpore al vector de estado  $x_t$  los parámetros  $\theta$  que se desea estimar [Kitagawa y Sato, 2001]:

$$x_t = \begin{bmatrix} x_t \\ \theta \end{bmatrix}. \quad (2.44)$$

Notar que en este modelo  $\theta$  se considera independiente del tiempo. Sin embargo, para poder utilizar este enfoque en conjunto con Filtro de Partículas es necesario que  $\theta$  esté afectado por una fuente de incertidumbre, de manera que tenga su propia función de densidad de probabilidad (f.d.p.) no trivial.

Al considerar el estado variante en el tiempo, [Gordon *et al.*, 2002] sugiere un enfoque para reducir la degeneración de las muestras mediante la incorporación de pequeñas perturbaciones a las partículas (llamadas *roughening penalties*), además de los ruidos inherentes del modelo. Esta idea ha sido extrapolada al problema de identificación de parámetros invariantes, donde –con la inclusión de perturbaciones aleatorias en el proceso de evolución de los parámetros– éstos son vistos como si fueran variantes en el tiempo. Este enfoque es conocido una *evolución artificial de parámetros*. De acuerdo a [Liu y West, 2001], la implementación de esta idea considera la siguiente ecuación de estados:

$$\theta_t = \theta_{t-1} + \epsilon_t, \quad (2.45)$$

donde  $\epsilon_t \sim \mathcal{N}(0, W_t)$ , siendo  $W_t$  una determinada matriz de covarianza.  $\theta_t$  es condicionalmente independiente a las observaciones hasta el instante  $t - 1$ .

La extensión del vector de estados y la incorporación de ruido artificial permiten que el Filtro de Partículas pueda ser usado para estimar conjuntamente el estado no observable y el vector de parámetros. Sin embargo, este enfoque introduce algunas desventajas: al incluir fuentes de incertidumbre externas, las densidades resultantes para el estado original son más difusas que las obtenidas teóricamente en el caso del modelo con parámetros fijos. Por otra parte, el hecho de contar con un vector de estados más extenso introduce un problema de observabilidad en sí mismo, especialmente si no es compensado con una extensión del vector de observación. En otras palabras, se están estimando más valores con las mismas fuentes de información.

# Capítulo 3

## Metodología e implementación

El presente capítulo presenta una descripción detallada de los aspectos metodológicos correspondientes a este trabajo. Para ello, este capítulo se ha estructurado en 3 secciones: la Sección 3.1 entrega una definición formal del problema y enuncia los supuestos simplificadorios adoptados; la Sección 3.2 explica la relevancia de la simulación de órbitas e indica el procedimiento para llevarla a cabo; finalmente, la Sección 3.3 presenta un método de estimación de elementos orbitales basado en Filtro de Partículas, exponiendo los argumentos teóricos y prácticos que respaldan el diseño del algoritmo.

### 3.1. Delimitación del problema

En el caso de las binarias visuales, las observaciones consisten en datos de posición en el plano del cielo en segundos de arco. Dado un conjunto de posiciones observadas en el plano del cielo, como el de la Figura 3.1, se busca estimar los parámetros orbitales definidos en la Sección 2.2. Este problema, conocido como *orbital fitting* en la literatura, ha sido abordado anteriormente desde diversos enfoques: desde los métodos geométricos de fines del siglo XIX y principios del siglo XX (Thiele-Innes-Van Den Bos) hasta algoritmos de optimización heurística como *simulated annealing* [Pourbaix, 1994] o algoritmos genéticos, además de enfoques bayesianos presentados recientemente [Lucy, 2014]. Formalmente, las entradas del problema son:

- $\mathcal{E} = \{\tau_i\}_{i=1,\dots,N}$ : Conjunto de las *épocas* o instantes de observación.

- $\mathcal{P} = \{(X_{obs}(i), Y_{obs}(i))\}_{i=1, \dots, N}$ : Conjunto de valores de posición relativa observados en el plano del cielo, medidos en segundos de arco. El valor  $i$  es un identificador del instante de observación.
- $\mathcal{W} = \{W(i)\}_{i=1, \dots, N}$ : Peso asignado a cada una de las  $N$  observaciones. Se utiliza la letra  $W$  (en mayúscula) para los pesos de las observaciones a fin de diferenciarlos de los pesos del algoritmo de Filtro de Partículas explicado en la Sección 3.3.

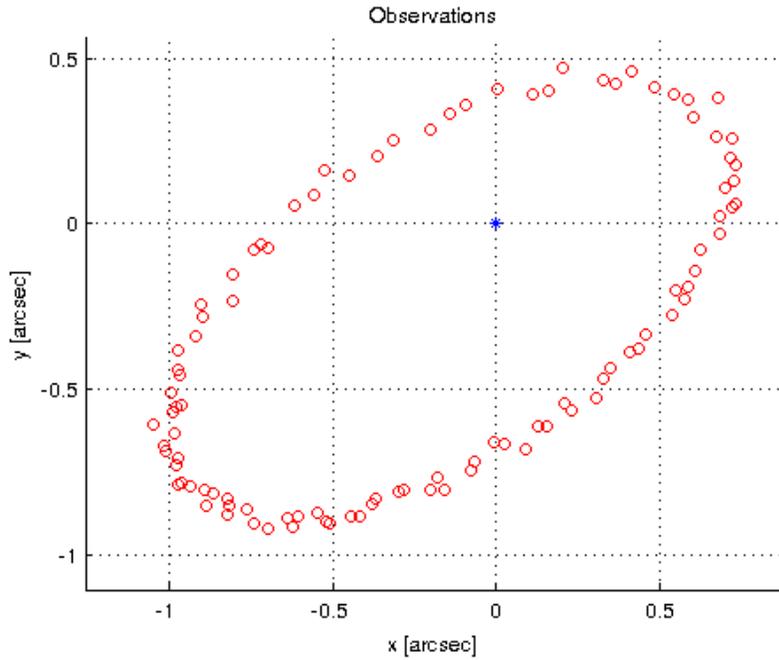


Figura 3.1: Observaciones de trayectoria en el *plano de cielo*

Como una manera de visualizar el problema, se incluye la Figura 3.2, que enfoca las órbitas desde un ángulo arbitrario, contrastando el plano del cielo (representado únicamente por observaciones ruidosas) con la órbita real en el espacio tridimensional. El eje  $Z$  del sistema (normalizado por la distancia de observación) de coordenadas apunta hacia el punto desde donde se realiza observación, de manera que el plano del cielo coincide con el plano  $XY$ .

Uno de los principales objetivos de estimar los parámetros orbitales de un sistema binario es conocer su masa total, que se relaciona con dichos parámetros a través de la Tercera Ley de Kepler. Sin embargo, el hecho de que la información de posición esté dada en segundos de arco en lugar de unidades astronómicas impone que el semi-eje mayor,  $a$ , también deba ser expresado en segundos de arco (este valor se denota  $a''$ ), de modo que no es posible aplicar la Tercera Ley

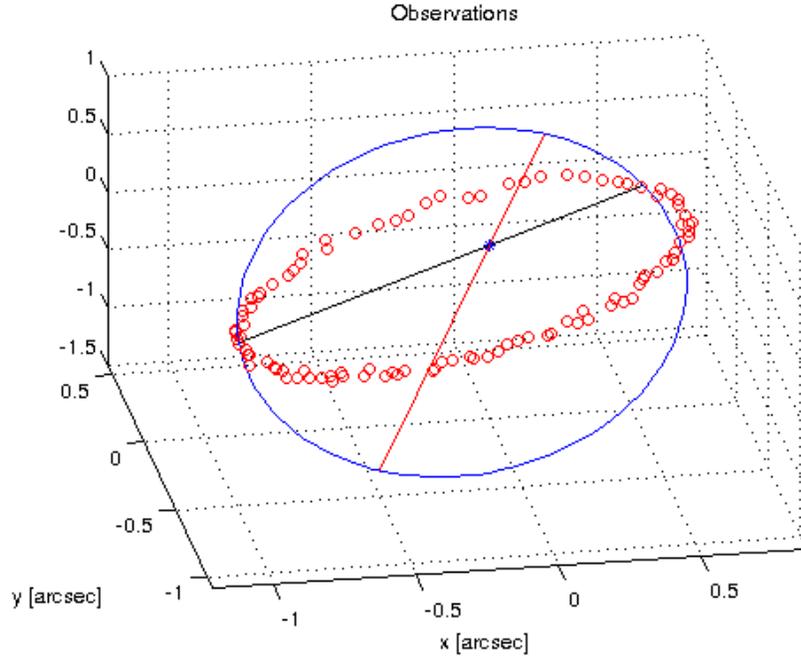


Figura 3.2: Plano de observación en relación a órbita real

de Kepler directamente. No obstante, cuando la paralaje  $\bar{\omega}$  del sistema es conocida, es posible rescatar el valor del semi-eje mayor en unidades de longitud mediante:

$$\left(\frac{a''}{\bar{\omega}}\right)^3 = \frac{G}{4\pi^2}(m_1 + m_2)P^2. \quad (3.1)$$

Los párrafos restantes de esta sección sintetizan la finalidad de este trabajo y los supuestos adoptados.

**Objetivo del trabajo:** Proponer y evaluar un método para estimar, en base a datos astrométricos de estrellas binarias visuales, los 7 parámetros orbitales orbitales y sus incertidumbres respectivas.

**Supuestos simplificadorios:**

1. Como se señala en la Subsección 2.1.1, el sistema observado es: (i) gravitacionalmente aislado y (ii) separado o *detached*. Debe recalcar que el sistema se considera estrictamente binario, por lo que su análisis no incluye la posibilidad de cuerpos adicionales tales como estrellas de menor tamaño (como ocurre en sistemas estelares jerárquicos) o planetas. De esta manera, la dinámica interna del sistema se rige únicamente por las ecuaciones presentadas

en la subsección mencionada.

2. Por la misma razón, el único ruido presente en las observaciones es aquel introducido en el proceso de medición, el cual se supone Gaussiano, bivariado –para introducir perturbación en ambos ejes del plano– y de valor medio nulo –para que sea insesgado–. Si bien este ruido depende mayormente del instrumento de observación utilizado, también podrían considerarse otros factores, tales como las condiciones atmosféricas. Sin embargo, el origen del ruido no es relevante, en la medida que dicha perturbación cumpla el supuesto de Gaussianidad y sea posible cuantificar su varianza característica.
3. Tanto la simulación como el análisis de datos omiten: (i) efectos relativistas, (ii) efectos asociados al movimiento propio, (iii) precesión de los equinoccios y (iv) efectos de marea o torque entre ambos cuerpos. Debido a los efectos (ii) y (iii), los datos de posición ( $\rho$ ,  $\theta$  o equivalentemente  $X$  y  $Y$ ) deben ser corregidos antes del cálculo de órbita en situaciones reales. En este trabajo, los datos de posiciones artificialmente generados están “corregidos” de antemano.
4. Se considera que los parámetros orbitales del sistema estudiado son constantes durante el período de observación (esto no necesariamente es válido para todos los cuerpos celestes; a modo de ejemplo, la excentricidad del planeta Tierra ha ido cambiando con el tiempo).
5. La paralaje del sistema es conocida.
6. Los pesos de las observaciones no pueden ser arbitrariamente asignados: son inversamente proporcionales a la varianza característica del proceso de observación que las generó.

## 3.2. Simulación de órbitas y observaciones

La evaluación del método de estimación propuesto en este trabajo requiere conjuntos de datos sobre los cuales aplicar el algoritmo. En lugar de utilizar datos reales, se opta por generar datos artificiales que simulen observaciones astrométricas. Lo que se busca es generar un conjunto de observaciones ruidosas de posición relativa en el plano del cielo, usando la representación presentada en la Sección 2.2. Para ello, es clave plantear un modelo del sistema observado y un

método para simularlo. Sin embargo, la importancia de la simulación no se limita a la generación de inputs artificiales, sino que además juega un rol fundamental en el método de estimación propuesto en la Sección 3.3 (y más generalmente, en cualquier algoritmo basado en Filtro de Partículas). En esta sección se explica el procedimiento para generar datos artificiales, detallando el *input*, el *output* deseado y los pasos del algoritmo de simulación implementado.

**Entrada del método de simulación:**

- Tiempos de observación:  $\mathcal{E}$  o *epochs*. Consiste en un arreglo con los instantes en que se realizan las observaciones. No hay restricciones sobre el patrón de muestreo: éste puede ser no uniforme, tener largos períodos sin observaciones o incluso pérdida de datos durante períodos activos de observación.  $\mathcal{E} = [\tau_1, \dots, \tau_N]$
- Elementos orbitales:  $[P, T, e, a, \omega, \Omega, i]$ . Son los parámetros “reales” de un sistema orbital hipotético.
- Caracterización del ruido de observación:  $[\mu_v, \Sigma_v]$ . Estos valores corresponden al valor medio y la matriz de covarianza del error de observación, que se asume Gaussiano de media nula, no correlacionado y dependiente únicamente del instrumento de medición. Específicamente, la observaciones en el plano del cielo se modelan como:

$$\begin{bmatrix} x_{obs}(\tau) \\ y_{obs}(\tau) \end{bmatrix} = \begin{bmatrix} x_{proj}(\tau) \\ y_{proj}(\tau) \end{bmatrix} + \begin{bmatrix} v_x(\tau) \\ v_y(\tau) \end{bmatrix},$$

donde  $x_{proj}$  y  $y_{proj}$  son la proyección en el plano del cielo de los valores reales de posición y el ruido  $[v_x(\tau), v_y(\tau)]$  sigue una distribución Gaussiana multivariada  $\mathcal{N}(\mu_v, \Sigma_v)$ , siendo  $\mu_v$  un vector bidimensional de valor nulo y:

$$\Sigma_v = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix}.$$

**Salida del método de simulación:**  $\{(X_{obs}(\tau_i), Y_{obs}(\tau_i))\}_{i=1, \dots, n}$ , conjunto de observaciones ruidosas de posición relativa en el plano del cielo, como el de la Figura 3.1, en las épocas  $\mathcal{E}$  señaladas en el *input*.

**Algoritmo de simulación:** Conceptualmente, el proceso de generación de observaciones artificiales sigue las etapas indicadas a continuación:

1. Simulación de órbita real en 3 dimensiones.
2. Proyección de la órbita real sobre plano del cielo.
3. Adición de ruido de observación.

El procedimiento específico se muestra en el pseudo-código de abajo. Por simplicidad, se utilizan las constantes de Thiele-Innes y la Ecuación 2.11 para calcular la proyección de las posiciones.

---

**Algoritmo 1** Algoritmo de generación de observaciones orbitales

---

```

procedure GENERATEOBS( $\mathcal{E}, T, P, e, a, \Omega, \omega, i, \sigma_x, \sigma_y$ )
   $A \leftarrow a(\cos \omega \cos \Omega - \sin \omega \sin \Omega \cos i)$ 
   $B \leftarrow a(\cos \omega \sin \Omega + \sin \omega \cos \Omega \cos i)$ 
   $F \leftarrow a(-\sin \omega \cos \Omega - \cos \omega \sin \Omega \cos i)$ 
   $G \leftarrow a(-\sin \omega \sin \Omega + \cos \omega \cos \Omega \cos i)$ 
   $X_{obs} \leftarrow \text{array}(N)$ 
   $Y_{obs} \leftarrow \text{array}(N)$ 
  for  $i = 1 : N$  do
     $\tau_i \leftarrow \mathcal{E}(i)$ 
     $M \leftarrow \frac{2\pi}{P}(\tau_i - T)$ 
     $E \leftarrow \text{solveKeplerEq}(M, e)$ 
     $X_i \leftarrow \cos E - e$ 
     $Y_i \leftarrow (1 - e^2)^{\frac{1}{2}} \sin E$ 
     $X_{proj}^i \leftarrow A \cdot X_i + F \cdot Y_i$ 
     $Y_{proj}^i \leftarrow B \cdot X_i + G \cdot Y_i$ 
     $X_{obs}(i) \leftarrow X_{proj}^i + v_x, v_x \sim \mathcal{N}(0, \sigma_x^2)$ 
     $Y_{obs}(i) \leftarrow Y_{proj}^i + v_y, v_y \sim \mathcal{N}(0, \sigma_y^2)$ 
  end for
  return  $X_{obs}, Y_{obs}$ 
end procedure

```

---

Para la simulación de órbitas hipotéticas, es decir, órbitas (proyectadas) no afectadas por ruido, basta ejecutar el Algoritmo 1 con  $\sigma_x = \sigma_y = 0$ . El método *solveKeplerEq* entrega el valor de la *anomalía excéntrica*,  $E$ , dado un valor específico de *anomalía media*,  $M$ , y excentricidad,  $e$ . La implementación utilizada en este trabajo corresponde al método de Newton-Raphson.

---

**Algoritmo 2** Método iterativo para resolver la Ecuación de Kepler

---

```
procedure SOLVEKEPLEREQ( $M, e$ )  
   $\epsilon \leftarrow 0.0001$   
   $E_n \leftarrow M$ ;  
   $E_{sol} \leftarrow E_n - (E_n - e \cdot \sin E_n - M)/(1 - e \cdot \cos E_n)$   
  while  $E_{sol} - E_n > \epsilon$  do  
     $E_n \leftarrow E_{sol}$   
     $E_{sol} \leftarrow E_n - (E_n - e \cdot \sin E_n - M)/(1 - e \cdot \cos E_n)$   
  end while  
  return  $E_{sol}$   
end procedure
```

---

### 3.2.1. Modelación de la adquisición de datos

El muestreo de datos está modelado por el conjunto  $\mathcal{E}$  definido anteriormente, pues son los instantes de adquisición,  $\tau_i$ , los que definen si éste es uniforme, con o sin pérdida de datos, o si la observación es continua o interrumpida en el tiempo. A fin de emular los instantes de adquisición de una campaña astronómica real, la generación de  $\mathcal{E}$  debe tomar en cuenta las situaciones que pueden ocurrir durante dicho proceso:

- **Existencia de períodos sin observaciones:** Es común que la observación de un objeto astronómico sea interrumpida por largos períodos. Esto puede deberse a motivos naturales –ocultamiento del objeto de interés por parte de otro cuerpo celeste– o a limitaciones sobre los tiempos de observación asignados en los observatorios. Asimismo, debe tenerse en consideración que . En este trabajo, los períodos sin observaciones se identifican como *brechas temporales*, mientras que los períodos en que se adquieren datos con cierta regularidad se denominan *períodos de observación activa*.
- **Pérdida de datos:** Operaciones regulares de mantenimiento, fallas técnicas ocasionales o condiciones atmosféricas no favorables a la observación podrían afectar la regularidad de la adquisición de datos, incluso durante períodos de observación continua. Por ejemplo, una campaña destinada a realizar observaciones astrométricas cada 24 horas podría sufrir una interrupción temporal a causa de una falla en un telescopio, imposibilitando la adquisición de datos durante un breve período. De manera similar, un cielo inusualmente nublado podría forzar el descarte de los datos obtenidos bajo esas condiciones.

- **Variabilidad en la calidad de los datos:** En general, las fuentes de información para el estudio de un objeto astronómico son múltiples. En el caso de estrellas binaria visuales, especialmente aquellas de período largo (cientos de años), es común que la información disponible provenga de observaciones: (i) dispersas en el tiempo; (ii) obtenidas con instrumentos de diferente naturaleza: desde observaciones visuales con telescopio en el Siglo XIX hasta las obtenidas con modernos equipos radioastronómicos en la segunda mitad del Siglo XX. Por esta razón, las observaciones reciben diferentes pesos  $W_i$  en función de alguna medida de calidad.

El procedimiento utilizado en este trabajo para generar los instantes o *épocas* de observación se detalla en el Algoritmo 3. El método se basa en dividir un período en intervalos regulares, diferenciar los *períodos de observación activa* de los *gaps temporales* y finalmente descartar de manera aleatoria algunas *épocas*. Los argumentos de dicho algoritmo son: (i) *TimeObs*, cantidad de años desde la primera hasta la última observación; (ii) *nInt*, número de intervalos iguales en que es dividido el período *TimeObs*; (iii) *Indices*, conjunto de índices identificados como períodos de observación activa; (iv) *LossRate*, probabilidad de que una época se descarte. Este método no genera los *gaps* temporales de acuerdo a un modelo subyacente, sino que el usuario los introduce manualmente a través del parámetro *Indices*. El pseudo-código presentado hace uso libre de expresiones típicas de MATLAB.

---

**Algoritmo 3** Procedimiento para generar  $\mathcal{E}$

---

```

procedure GENERATEEPOCHS(TimeObs, nInt, Indices, LossRate)
   $\mathcal{E} \leftarrow \text{linspace}(0, \textit{TimeObs}, \textit{nInt} + 1)$            ▷ Período dividido en intervalos regulares
   $\mathcal{E} \leftarrow \mathcal{E}(\textit{Indices})$                                ▷ Sólo permanecen épocas de observación activa
   $n \leftarrow \text{length}(\mathcal{E})$ 
  while  $i \leq n$  do                                         ▷ Ciclo descarta aleatoriamente algunos instantes
     $\textit{loss} \leftarrow \text{Bernoulli}(\textit{LossRate})$            ▷ Realización de v.a. de Bernoulli con  $p = \textit{LossRate}$ 
    if  $\textit{loss} = 1$  then
       $\mathcal{E}(i) \leftarrow \{\}$                                  ▷ Se descarta el instante  $\mathcal{E}(i)$ 
    end if
     $i \leftarrow i + 1$ 
  end while
  return  $\mathcal{E}$ 
end procedure

```

---

### 3.3. Método de estimación

En este trabajo se aborda el problema de *orbital fitting* desde un enfoque Bayesiano. Específicamente, se utiliza un Método Secuencial de Monte Carlo (o Filtro de Partículas) para realizar una estimación secuencial de los parámetros orbitales. A pesar de que el Filtro de Partículas fue concebido como un método para estimar el estado de un sistema, en este trabajo se formula de manera que las variables estimadas sean los parámetros  $\{P, T, e, a, \omega, \Omega, i\}$ , utilizando el concepto de Evolución Artificial de Parámetros. En esta sección se detalla el método propuesto y los argumentos teóricos y prácticos que respaldan su diseño.

#### 3.3.1. Propiedades básicas de la estrategia de filtrado

Los argumentos en favor de utilizar Filtro de Partículas en lugar de Filtro de Kalman en el problema estimación de parámetros orbitales son: (i) la no-linealidad del modelo del sistema –las ecuaciones orbitales– y no-linealidad de la ecuación de observación –en este caso, el error cuadrático medio definido en la Subsección 3.3.5–; (ii) flexibilidad para la caracterización estadística de los ruidos, que no necesariamente son Gaussianos.

El método propuesto se basa en la formulación clásica del Filtro de Partículas, es decir, aquella en que:

1. Se toma la densidad a priori como densidad de importancia:  $q(x_j|x_{j-1}^k, y_j) = p(x_j^k|x_{j-1}^k)$  (con  $k$  y  $j$  siendo los identificadores de la partícula y de la iteración, respectivamente).
2. Los pesos se actualizan, consecuentemente, según  $w^k(j) \propto w^k(j-1)p(y(j)|x^k(j))$ .
3. La decisión de hacer *resampling* en una iteración dada se realiza dependiendo de si la estimación del número efectivo de partículas,  $\hat{N}_{eff}$ , cae bajo cierto umbral.

La decisión de utilizar esta versión del Filtro se basa en motivos principalmente prácticos: la facilidad de implementación y abundancia de referencias en la literatura. Sin embargo, dada la naturaleza del problema, tampoco existen razones teóricas para considerar otras variantes del éste, como podría ocurrir, por mencionar un ejemplo, en problemas financieros, donde la

alta variabilidad del objeto estudiado –series de tiempo financieras– sugiere utilizar un enfoque sensible al riesgo [Mundnich *et al.*, 2013].

### 3.3.2. Definición de vector de estado

A diferencia de las aplicaciones tradicionales del Filtrado Bayesiano, como los problemas de *tracking* de objetos [Khan *et al.*, 2004], en este trabajo las variables de interés no son los valores del estado de un sistema en un instante dado, sino sus parámetros. El enfoque de Evolución Artificial de Parámetros presentado en la Sección 2.3 se basa en extender el espacio de estados de un modelo con los parámetros de interés, a fin de realizar estimación simultánea de estados y parámetros. Puntualmente, las variables de interés de este trabajo son los elementos orbitales (parámetros) más que las posiciones sucesivas del sistema estelar en el plano del cielo (estados), de modo que el tradicional *vector de estado*<sup>1</sup> del Filtro de Partículas,  $\vec{X}$ , puede plantearse directamente como  $\vec{X} = (P, T, e, a, \omega, \Omega, i)$  (notar que está compuesto íntegramente por parámetros).

Bajo ese enfoque, las partículas corresponderían a muestras puntuales del espacio paramétrico. Formalmente, puede definirse la partícula  $k$  en la iteración  $j$ -ésima como  $p_k(j) = [T_k(j), P_k(j), e_k(j), a_k(j), \omega_k(j), \Omega_k(j), i_k(j)]$ . Sin embargo, con la finalidad de reducir la dimensionalidad del espacio de estado, el presente trabajo utiliza una representación alternativa: la representación de Thiele-Innes presentada en la Subsección 2.2.1. Si bien esta representación consta, igualmente, de 7 elementos ( $T, P, e, B, A, G$  y  $F$ ), tiene la particularidad de que 4 de sus parámetros –las constantes de Thiele-Innes:  $B, A, G$  y  $F$ – pueden ser estimados en función de los elementos restantes ( $T, P$  y  $e$ ) y las observaciones disponibles ( $X, Y$ ) a través de un estimador de mínimos cuadrados **definido analíticamente**. La derivación de dicho estimador, explicada en el Apéndice B.1, aprovecha la dependencia lineal de los valores teóricos de  $X$  e  $Y$  en plano del cielo respecto a las constantes de Thiele-Innes. Las fórmulas involucradas en el cálculo del estimador también pueden consultarse en el apéndice mencionado. En esta sección el estimador se tratará, por simplicidad de notación, como una función computacional:

$$\left[ \hat{B}, \hat{A}, \hat{G}, \hat{F} \right] = \text{leastSquaresEstimation}(\mathcal{E}, X, Y, T, P, e, W) \quad (3.2)$$

---

<sup>1</sup>Se denomina *vector de estado* para mantener consistencia con la literatura, a pesar de que corresponde, en rigor, a un vector de parámetros.

Donde *leastSquaresEstimation* realiza los cálculos introducidos en el Apéndice B.1 en función de los argumentos entregados. La variable  $W$  corresponde a los pesos asignados a cada observación.

Este estimador entrega los mejores valores estimados de  $B$ ,  $A$ ,  $G$  y  $F$  –en el sentido de mínimos cuadrados– dado un valor de  $T$ ,  $P$  y  $e$ . En particular, si los valores de  $T$ ,  $P$  y  $e$  se conocen con exactitud, la estimación de  $B$ ,  $A$ ,  $G$  y  $F$  tenderá a acercarse al valor real. Si, por el contrario, los valores estimados de  $T$ ,  $P$  y  $e$  distan de los valores reales, los valores de  $B$ ,  $A$ ,  $G$  y  $F$  serán también erróneos, a pesar de minimizar el error cuadrático medio para un conjunto de observaciones dado.

La acción de dejar las constantes de Thiele-Innes en función del resto de los parámetros y de las observaciones permite reducir el espacio de búsqueda del Filtro de Partículas a sólo 3 dimensiones: las correspondientes a  $T$ ,  $P$ ,  $e$ . De esta manera, una partícula individual queda definida simplemente por  $p_k(j) = [T_k(j), P_k(j), e_k(j)]$ , aunque no debe perderse de vista que dicha partícula define implícitamente valores de  $\hat{B}_k(j)$ ,  $\hat{A}_k(j)$ ,  $\hat{G}_k(j)$ ,  $\hat{F}_k(j)$ : basta reemplazar  $T$ ,  $P$ ,  $e$  por  $P_k(j)$ ,  $T_k(j)$ ,  $e_k(j)$  en la Ecuación 3.2.

Esta reducción permite aumentar la eficiencia del método, puesto que al disminuir la dimensión del espacio de estados el Filtro puede funcionar con un número menor de partículas y aún así obtener resultados satisfactorios bajo alguna métrica (error porcentual o error cuadrático medio, por ejemplo). Sin embargo, tiene la desventaja de reducir el espacio de soluciones: dado un conjunto de observaciones **ruidosas** y valores puntuales de  $T$ ,  $P$  y  $e$ , el mínimo global del error medio cuadrático (que es lo que calcula la Ecuación 3.2) no necesariamente coincide con los parámetros de la órbita verdadera, ya que se descartan soluciones factibles –en general, mínimos locales– que podrían coincidir con el valor real.

### 3.3.3. Distribución inicial

La distribución inicial de las partículas depende del grado de conocimiento acerca de la órbita. Un desconocimiento absoluto de un parámetro orbital determinado se reflejaría en una distribución inicial uniforme de amplio rango, mientras que el conocimiento de una aproximación inicial permitiría plantear distribuciones más *localizadas* dentro del espacio de parámetros, posiblemente descartando algunas soluciones factibles (valores de parámetros capaces de generar un conjunto

de observaciones similar).

La reducción del vector de estado presentada en la subsección anterior permite reducir, de igual forma, la dimensión de la distribución inicial: sólo es necesario tener una distribución a priori para  $T$ ,  $P$  y  $e$ . Se define el parámetro alternativo  $\mathcal{T}$  como la fracción del período donde ocurre el paso por el periastro:

$$\mathcal{T} = \frac{T}{P}. \quad (3.3)$$

De esta manera, las partículas toman la forma  $p_k(j) = [\mathcal{T}_k(j), P_k(j), e_k(j)]$ . Con ello, tanto la excentricidad  $e$  como el tiempo de paso por el periastro, redefinido como  $\mathcal{T}$  están acotados en el rango  $[0, 1)$  independiente de las condiciones particulares del sistema observado. Sin embargo, en sistemas binarios reales el período orbital  $P$  puede tomar valores que van desde unos pocos minutos (5.4 [min] en el caso de la estrella binaria HM Cancri [Roelofs *et al.*, 2010]) hasta varios cientos de años, por lo que realizar una estimación previa de este parámetro resulta crítico. La caracterización de la distribución inicial es un problema diferente al proceso de filtrado, el cual simplemente recibe la distribución inicial como una más de sus entradas. En consistencia con eso, el software desarrollado para este trabajo ofrece dos opciones:

1. Ejecutar una rutina de *estimación amplia* para determinar un rango inicial del vector de parámetros  $[\mathcal{T}, P, e]$ , bajo el supuesto de una distribución inicial uniforme. Esta rutina se resume en los siguientes pasos: (i) estimar  $e$  como  $\hat{e} = (\rho_{max} - \rho_{min}) / (\rho_{max} + \rho_{min})$ , donde  $\rho$  es la magnitud de la posición relativa en la órbita aparente, y  $\mathcal{T}$  como  $\hat{\mathcal{T}} = \text{argmin}_i \rho(\tau_i)$ ; (ii) generar, con  $e = \hat{e}$ ,  $\mathcal{T} = \hat{\mathcal{T}}$  y  $a = 1$ , trayectorias para los períodos candidatos  $P_i^c \in \{1, 2, \dots, 999, 1000\}$ ; (iii) seleccionar los 25 períodos candidatos con menor error cuadrático medio respecto a los datos normalizados (i.e., valores de  $\rho$  divididos por  $\hat{a} = (\rho_{min} + \rho_{max})/2$ ) y agruparlos en el subconjunto  $\mathcal{P}^{25}$ ; (iv) finalmente, entregar el rango [mín  $\mathcal{P}^{25}$ , máx  $\mathcal{P}^{25}$ ] como distribución inicial para  $P$ . A pesar de que los rangos iniciales de  $\mathcal{T}$  y  $e$  podrían determinarse a partir de  $\hat{e}$  y  $\hat{\mathcal{T}}$ , estos no generan –de acuerdo a la evaluación empírica realizada en el desarrollo de este trabajo– un impacto tan crítico en el proceso de filtrado como la estimación inicial de  $P$ ; en consecuencia, dichos rangos se fijan en  $[0, 1)$  por simplicidad (y para no descartar soluciones factibles).
2. Ingresar manualmente una distribución inicial a través de un archivo de configuración. Esta opción es útil cuando se trabaja con datos astronómicos reales –posiblemente estudiados en investigaciones previas– sobre los cuales ya existe un grado de conocimiento, o bien cuando

se desee determinar una distribución inicial mediante un algoritmo alternativo (i.e., no el descrito en el punto anterior).

### 3.3.4. Ecuación de evolución

Bajo la idea de Evolución Artificial de Parámetros, se define una ecuación de evolución para que las partículas *explore*n aleatoriamente diferentes valores del espacio paramétrico:

$$p_k(j) = p_k(j - 1) + \epsilon_k(j), \quad (3.4)$$

donde  $\epsilon_k(j)$  es una variable aleatoria Gaussiana multivariada de dimensión 3 y componentes no correlacionadas correspondientes a  $\mathcal{T}$ ,  $P$  y  $e$ . Las varianzas del ruido en cada dimensión –la diagonal de la matriz de covarianza de  $\epsilon$ – es un *input* del algoritmo, que puede elegirse de acuerdo al conocimiento previo acerca de los datos, de manera que, por ejemplo, la desviación estándar en la dimensión correspondiente a  $P$  sea al menos un orden de magnitud menor que los valores comprendidos en la distribución inicial. En este trabajo, la distribución de  $\epsilon_k(j)$  permanece fija, a pesar de que en la formulación original [Liu y West, 2001] se presente una ecuación de actualización (es decir, puede ser dependiente de la iteración  $j$ ).

Dado que tanto  $\mathcal{T}$  como  $e$  deben estar acotados en el rango  $[0, 1)$ , y que el período  $P$  debe cumplir  $P > 0$ , la aplicación de la Ecuación 3.4 puede generar, dada la aleatoriedad del ruido  $\epsilon_k$ , valores no válidos, como excentricidades fuera del rango  $[0, 1)$  o períodos negativos. Como medida correctiva, la rutina de filtrado aplica la operación *mod* 1 sobre los valores resultantes de la actualización de partículas en las dimensiones correspondientes a  $\mathcal{T}$  y  $e$ . En el caso del período  $P$ , se define un período  $P_{max}$  y se aplica la operación *mod*  $P_{max}$  para evitar valores negativos. Cabe agregar que después de este paso se calculan las constantes de Thiele-Innes correspondientes a los valores de  $\mathcal{T}$ ,  $P$  y  $e$  de la partícula, tal como se explicó en la Subsección 3.3.2.

### 3.3.5. Definición de *meta-observación* y función de verosimilitud

Dado un conjunto de observaciones  $\{\tau_i, X_i, Y_i, W_i\}_{i=1, \dots, N}$  y una partícula  $p_k(j)$ , se define la *meta-observación* –observación sobre las observaciones–  $\mathcal{Y}_k(j)$  como el error cuadrático medio de los valores observados respecto a la trayectoria definida por los parámetros de la partícula:

$$\mathcal{Y}_k(j) = \sum_{i=1}^N W_i^x \cdot [X_i - X_{comp}^{i,k,j}]^2 + \sum_{i=1}^N W_i^y \cdot [Y_i - Y_{comp}^{i,k,j}]^2, \quad (3.5)$$

donde  $X_{comp}^{i,k,j}$ ,  $Y_{comp}^{i,k,j}$  son los valores teóricos de la órbita aparente, dado el set de parámetros definido por la partícula  $p_k(j)$ , y el instante de observación  $\tau_i$  (estos valores se calculan con las ecuaciones orbitales presentadas en la Subsección 2.1.1 e implmentadas en el Algoritmo 1, aunque considerando ruidos nulos). De acuerdo a la restricción establecida en la Sección 3.1, los pesos asignados a las observaciones,  $W_i$ , son inversamente proporcionales a la varianza. Específicamente,  $W_i^x = \frac{1}{N} \cdot \frac{1}{\sigma_x^2(i)}$ , siendo  $\sigma_x(i)$  la desviación estándar del error de observación característico en el instante  $\tau_i$  en el eje  $x$  del plano del cielo (la aseveración análoga se cumple para  $W_i^y$ ). De esta manera, la *meta-observación* queda definida como:

$$\mathcal{Y}_k(j) = \frac{1}{N} \sum_{i=1}^N \left( \frac{1}{\sigma_x^2(i)} [X_i - X_{comp}^{i,k,j}]^2 + \frac{1}{\sigma_y^2(i)} [Y_i - Y_{comp}^{i,k,j}]^2 \right) \quad (3.6)$$

Lo que cambia en cada iteración del filtro no son las observaciones de posición relativa –siempre se utilizan el conjunto completo–, sino el valor de la partícula, lo cual genera diferentes valores de  $\mathcal{Y}_k(j)$ .

La restricción sobre los pesos  $W_i$  de las observaciones tiene como finalidad facilitar la caracterización estadística del error cuadrático medio, cuya derivación se revisa a continuación:

1. Considerar el supuesto de que el ruido de observación es Gaussiano, es decir, la diferencia entre el valor observado y el valor real sigue una distribución normal de valor medio nulo y cierta varianza característica. Con ello, los términos elevados al cuadrado dentro de la sumatoria de la Ecuación 3.6 (es decir,  $1/\sigma_x(i) \cdot [X_i - X_{comp}^{i,k,j}]$  y  $1/\sigma_y(i) \cdot [Y_i - Y_{comp}^{i,k,j}]$ ) corresponderían, cada uno por separado, a una variable aleatoria Gaussiana de valor medio nulo y varianza unitaria, pues las diferencias entre la observación y el valor computado están divididas por la desviación estándar de la observación (o *normalizadas*). De esta

manera, al tomar el cuadrado se obtiene una variable aleatoria  $\chi_1^2$ .

2. La suma de dos variables aleatorias  $\chi_1^2$  (en este caso,  $1/\sigma_x^2(i) \cdot [X_i - X_{comp}^{i,k,j}]^2$  y  $1/\sigma_y^2(i) \cdot [Y_i - Y_{comp}^{i,k,j}]^2$ ) sigue una distribución  $\chi_2^2$ .
3. Por lo tanto, la sumatoria de la Ecuación 3.6 corresponde a la media aritmética de  $N$  variables aleatorias  $\chi_2^2$ :

$$\mathcal{Y}_k(j) = \bar{X} = \frac{1}{N} \sum_{i=1}^N X_i, \text{ con } X_i \sim \chi_{deg}^2, \text{ } deg = 2, \quad (3.7)$$

correspondiendo  $X_i$  a las variables aleatorias  $\chi_2^2$  descritas en el punto anterior. Las expresiones con la forma de la Ecuación 3.7 siguen una distribución *Gamma*( $\alpha = N \cdot deg/2, \theta = 2/N$ ), lo que permite concluir que  $\mathcal{Y}_k(j) \sim \text{Gamma}(\alpha = N, \theta = 2/N)$  cuando la partícula coincide con los valores reales de los elementos orbitales.

En vista de la conclusión enunciada en el último punto, se propone la función de densidad de probabilidad de la distribución *Gamma* como una medida de cuán *creíble* o *verosímil* es el conjunto de parámetros definido por una partícula individual: si dicho conjunto se asemeja a los valores reales de los elementos orbitales, el error cuadrático medio  $\mathcal{Y}_k(j)$  debería adquirir valores altamente probables según la función de densidad de probabilidad (f.d.p.) mencionada; si, por el contrario, la partícula posee valores lejanos a los parámetros orbitales reales, generará un valor de error cuadrático medio menos probable. Bajo este argumento, se propone realizar la actualización de pesos de las partículas según:

$$w_k(j) = w_k(j-1) \cdot \mathcal{L}_{\mathcal{Y}}(p_k(j)|\mathcal{Y}_k(j)), \quad (3.8)$$

donde  $\mathcal{L}_{\mathcal{Y}}$  corresponde a la verosimilitud, definida como la f.d.p. de la distribución *Gamma*:

$$\mathcal{L}_{\mathcal{Y}}(\Theta|x) = P_{\mathcal{Y}}(x|\Theta) = \frac{1}{\Gamma(\alpha)\theta^\alpha} x^{\alpha-1} e^{-\frac{x}{\theta}}, \quad (3.9)$$

siendo  $\Theta$  el conjunto de parámetros definido por la partícula<sup>2</sup> y  $x$  un valor arbitrario adquirido por meta-observación  $\mathcal{Y}$ .  $\alpha$  y  $\theta$  son los parámetros de la distribución *Gamma* definidos anteriormente. El valor  $e$  de la ecuación corresponde al número de Euler y no a la excentricidad.

---

<sup>2</sup>En la notación tradicional de Estadística y Probabilidad, se utiliza  $\theta$  (en minúscula) para denotar vectores de parámetros. En este caso se utiliza  $\Theta$  (en mayúscula) para diferenciarlo del parámetro de escala de la distribución *Gamma*.

### 3.3.6. Resumen

La idea general del algoritmo es explorar el espacio de parámetros mediante Evolución Artificial de Parámetros y rescatar las partículas más *verosímiles* (en el sentido de tener valores comparativamente altos de  $\mathcal{L}_Y$ ). Esto permite realizar una estimación de la distribución *a posteriori* (i.e., *dadas* las observaciones) de los elementos orbitales del sistema estudiado a través de un conjunto acotado de muestras del espacio paramétrico (las partículas con sus respectivos pesos). A continuación se muestra, en forma de pseudo-código, un resumen del método propuesto:

---

**Algoritmo 4** Algoritmo de estimación de parámetros orbitales

---

```

procedure ESTIMATEORBITALELEMENTS( $\mathcal{E}, X_{obs}, Y_{obs}, W_{obs}, initDist$ )
   $N_{part} \leftarrow N_{part}^*$  ▷ Se fija el número de partículas en algún valor
   $N_{it} \leftarrow N_{it}^*$  ▷ Se fija el número de iteraciones en algún valor
   $R_{th} \leftarrow R_{th}^*$  ▷ Se fija el umbral de remuestreo en algún valor
  for  $j = 1 : N_{it}$  do
    for  $k = 1 : N_{part}$  do
      if  $j = 1$  then
         $p_k(j) \sim initDist$  ▷ Inicialización de partícula según distribución inicial
      else
         $W_k(j) \sim \mathcal{N}(0, \Sigma_\epsilon)$  ▷ Generación ruido de evolución artificial
         $p_k(j) \leftarrow p_k(j-1) + W_k(j)$  ▷ Aplicación de ecuación de evolución
         $[B, A, G, F] \leftarrow leastSquaresEstimate(\mathcal{E}, X_{obs}, Y_{obs}, \mathcal{T}_k(j), P_k(j), e_k(j), W_{obs})$ 
▷ Cálculo de elementos de Thiele-Innes
         $\mathcal{Y}_k(j) \leftarrow MSE(\mathcal{E}, X_{obs}, Y_{obs}, W_{obs}, p_k(j), B, A, G, F, W_{obs})$ 
▷ Cálculo de meta-observación
         $w_k(j) \leftarrow w_k(j-1) \cdot \mathcal{L}_Y(\mathcal{Y}_k(j))$  ▷ Actualización de pesos
      end if
    end for
     $w_{(1:N_{part})}(j) \leftarrow w_{(1:N_{part})}(j) / \sum_{k=1}^{N_{part}} w_k(j)$  ▷ Normalización de pesos
    if  $1 / \sum_{k=1}^{N_{part}} w_k^2(j) < R_{th}$  then
       $[p_{(1:N_{part})}(j), w_{(1:N_{part})}(j)] \leftarrow resampling(p_{(1:N_{part})}(j), w_{(1:N_{part})}(j))$ 
    end if
  end for
  return  $p_{(1:N_{part})}(N_{it}), w_{(1:N_{part})}(N_{it})$  ▷ Rutina retorna partículas y pesos
end procedure

```

---

Finalmente, se estiman los elementos orbitales del sistema estudiado mediante la suma de las partículas ponderadas por sus respectivos pesos (promedio ponderado). La conversión desde la representación de Thiele-Innes a la de Campbell se realiza a través del procedimiento explicado en el Apéndice B.2.

# Capítulo 4

## Resultados y discusión

Esta sección tiene como objetivo evaluar el desempeño del método de estimación propuesto en el Capítulo 3. Para ello, se plantean tres experimentos destinados a probar el algoritmo en diferentes escenarios: la Sección 4.1 estudia la precisión y exactitud del método en función de la calidad de los datos de entrada; la Sección 4.2 estudia cómo cambian los resultados cuando se plantean configuraciones orbitales diferentes a las de la Sección 4.1; finalmente, la Sección 4.3 describe la realización de un caso particular, entregando mayor detalle sobre la información que es posible inferir a través de la técnica desarrollada en este trabajo. Por cada experimento, se reportan sus resultados y se adjunta un análisis sobre éstos.

### 4.1. Variación de la calidad de los datos

Este experimento busca explorar el desempeño del método en función de la *calidad de los datos*. Para ello, se considera una simulación basada en un sistema estelar real (ver 4.1.1) y la generación de observaciones artificiales de diversa calidad. La *calidad de los datos* está definida a través de dos parámetros:

- Calidad de observación: se define la variable  $\sigma_{obs}$  [*arcsec*] como la desviación estándar característica del ruido de la observación. Esto supone que el ruido es aditivo, Gaussiano, de media nula y de igual varianza en cada coordenada del plano del cielo (ejes  $X$  e  $Y$ ).
- Tiempo total de observación: se define  $\mathbf{f}$  como la *fracción del período orbital* durante la cual se reciben observaciones de posición relativa. Sea  $P_r$  [*yr*] el período orbital real del

sistema estelar estudiado y  $t_0$  [yr] el momento en que comienza una campaña astronómica hipotética; entonces, el *tiempo total de observación* abarca desde el instante  $t_0$  hasta el instante  $t_0 + \mathbf{f} \cdot P_r$ , con  $\mathbf{f} \in (0, 1]$ . Notar que la fracción  $\mathbf{f}$  puede ser conocida con exactitud sólo en el contexto de una simulación computacional, pues depende del período orbital  $P_r$ , que es precisamente uno de los parámetros que se desean estimar.

Este experimento busca realizar una caracterización *estadística* del desempeño del método en términos de la **precisión** y la **exactitud** alcanzada. Para ello, por cada combinación de  $\sigma_{obs}$  y  $\mathbf{f}$  (i.e. por cada tipo de *calidad de datos*) se generan  $N_{rep}$  diferentes conjuntos de observaciones, para luego ejecutar el método de estimación sobre cada uno de ellos. Sobre las  $N_{rep}$  repeticiones del proceso de estimación realizadas para un valor de  $(\sigma_{obs}, \mathbf{f})$ , se calcula:

- El **promedio** (aritmético) del valor estimado de la masa total y de cada uno de los elementos orbitales.
- La **desviación estándar** del valor estimado de la masa total y de cada uno de los elementos orbitales.

Lo que resta de esta sección se conforma de dos subsecciones: la Subsección 4.1.1 entrega los aspectos importantes sobre la generación de datos artificiales (el sistema estelar tomado como referencia, los supuestos sobre el proceso de observación) y la configuración del Filtro de Partículas utilizado, mientras que la Subsección 4.1.2 desarrolla un experimento para explorar el desempeño del algoritmo a través de diferentes combinaciones de  $\sigma_{obs}$  y  $\mathbf{f}$ .

#### 4.1.1. Configuración experimental

Con la finalidad de evaluar el método de estimación en un escenario realista, se escogieron los parámetros del sistema estelar Sirius –sistema largamente estudiado por la comunidad astronómica– como referencia para realizar las simulaciones. A continuación se muestran los valores de este sistema con su respectivo margen de error (la unidad [yr] se refiere al año juliano definido en la Subsección 2.1.1):

- Paralaje:  $\pi = 379.21 \pm 1.58$  [mas] ( $\pi = 0.37921$  [arcsec]).
- Distancia:  $8.60 \pm 0.04$  [ly].

- Período orbital:  $P = 50.090 \pm 0.055$  [yr]
- Longitud del semi-eje mayor:  $a = 7.50 \pm 0.04$  [arcsec].
- Excentricidad:  $e = 0.5923 \pm 0.0019$ .
- Inclinação:  $i = 136.53 \pm 0.43$  [°]
- Longitud del nodo ascendente:  $\Omega = 44.57 \pm 0.44$  [°]
- Tiempo de paso por el periastro:  $T = 1894.130 \pm 0.015$  [yr]
- Argumento del periastro:  $\omega = 147.27 \pm 0.54$  [°]

A pesar de que los valores presentados en esta lista son, en última instancia, una estimación (aunque una estimación respaldada por la comunidad astronómica, cabe mencionar), para efectos de simulación se les considera valores reales. El margen de error señalado junto a cada valor es meramente informativo y no cumple ninguna función relevante en este estudio.

Este trabajo contempla el supuesto de que, en el problema de estimación de parámetros orbitales, la calidad de un conjunto de datos está determinada en mayor medida por la *idoneidad de los instantes de observación* que por la cantidad de muestras. En este contexto, la *idoneidad de los instantes de observación* se mide por: (i) cobertura de la órbita (lo ideal es tener observaciones durante todo el período o, al menos, durante una fracción considerable); (ii) separación temporal adecuada entre las muestras: en consideración de la Segunda Ley de Kepler, ésta debe ser menor cuando la estrella secundaria se encuentra cerca del periastro y mayor cuando se aproxima al apoastro. En la misma línea, y de acuerdo a antecedentes empíricos, un tamaño muestral suficiente para caracterizar una órbita bordearía las 12 muestras, pero éstas deben cubrir la totalidad de la órbita. Por las razones expuestas, en este experimento el número de observaciones,  $N_{samp}$ , se mantiene fijo en 15 independiente del valor de  $\mathbf{f}$ .

La ejecución del método de estimación requiere que los instantes o épocas de observación, así como el tiempo de paso por el periastro,  $T$ , estén acotados en el rango  $[0, P_{obs}]$ , siendo  $P_{obs}$  el largo del período de observación. Para ello, basta sustraer el valor del primer instante de observación,  $t_0$ , de las épocas  $\mathcal{E}$  y del parámetro  $T$ . En esta simulación, se elige arbitrariamente que el inicio del período de observación es el año  $t_0 = 1930$ , con lo que el tiempo de paso por el periastro más próximo se encuentra en el instante  $1894.130 + 50.090 = 1944.220$ . Sustrayendo el valor  $t_0$  definido anteriormente,  $T$  queda reformulado como  $T = 14.220$ . En la representación de  $T$  como

fracción del período  $P$  (ver explicación en Subsección 3.3.4), el tiempo de paso por el periastro queda definido por  $\mathcal{T} = 14.220/50.090 = 0.2839$ .

El Filtro de Partículas utilizado en esta sección utiliza  $N_p = 500$  partículas y remuestreo multinomial. El umbral de remuestreo es  $N_{th} = 0.5$ . El número de iteraciones del filtro es fijado en 40.

La distribución *a priori* de los parámetros iniciales<sup>1</sup> es calculada de manera independiente para cada conjunto de observaciones a través de la rutina descrita en la Subsección 3.3.3. Es decir, a pesar de que este experimento es una simulación y los parámetros reales son conocidos –lo que permite asignar manualmente un rango acotado que contenga al valor real  $P$ –, cada realización del proceso de estimación se ejecuta como si lo único conocido de antemano fueran las posiciones observadas ( $X$  e  $Y$ ) y los instantes de observación (conjunto  $\mathcal{E}$ ), haciendo necesario ejecutar una rutina auxiliar para obtener un rango inicial de los parámetros a estimar.

---

<sup>1</sup>En este caso sólo son de interés  $T$ ,  $e$  y especialmente  $P$ , ya que el resto es calculado mediante un estimador de mínimos cuadrados.

### 4.1.2. Caracterización estadística del desempeño

El objetivo de este apartado es, conocida la configuración experimental utilizada, plantear (detalladamente), reportar (sus resultados) y discutir la prueba realizada para caracterizar estadísticamente el desempeño del método. A continuación, se presentan las características específicas del experimento:

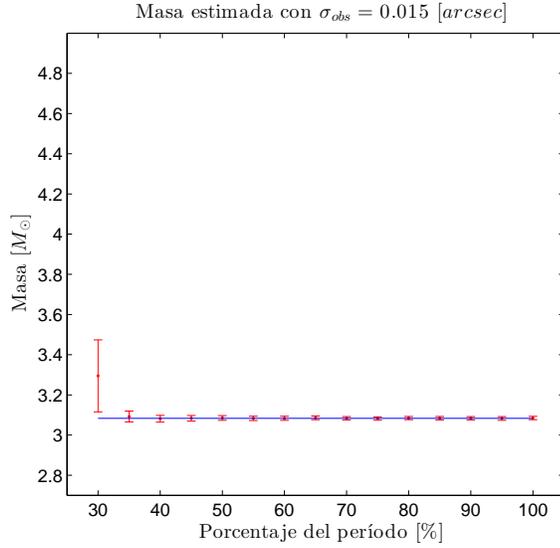
- Se definen tres valores para la desviación estándar del error de observación:
  - $\sigma_{obs} = 15 \text{ mas}$  (0.015 [arcsec]) (error de observación bajo).
  - $\sigma_{obs} = 75 \text{ mas}$  (0.075 [arcsec]) (error de observación medio).
  - $\sigma_{obs} = 150 \text{ mas}$  (0.150 [arcsec]) (error de observación alto).

Se asume un error Gaussiano de media nula e igual varianza en cada coordenada del plano del cielo, es decir,  $\sigma_x = \sigma_y = \sigma_{obs}$  (con correlación nula).

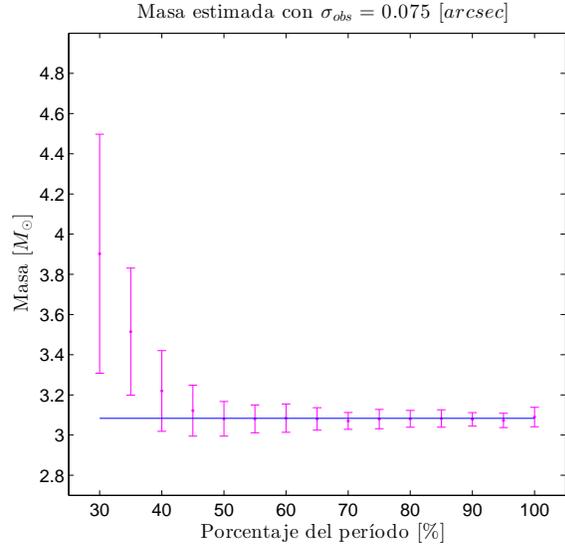
- Se define un arreglo con diferentes períodos de observación, (parametrizados por  $\mathbf{f}$ ). En este caso,  $\mathbf{f} = 30\%, 35\%, \dots, 95\%, 100\%$ .
- Cada simulación considera un conjunto de observaciones de tamaño  $N_{samp} = 15$ . Las muestras son equiespaciadas en el intervalo temporal  $[0, \mathbf{f} \cdot P]$ .
- Se realizan  $N_{rep} = 80$  repeticiones para cada combinación de  $(\sigma_{obs}, \mathbf{f})$ . Cada realización considera un nuevo conjunto de valores de posición relativa observada, aunque generados en el mismo intervalo de tiempo (definido por  $\mathbf{f}$ ) y por la misma distribución (definida por  $\sigma_{obs}$ ).
- Sobre cada conjunto de repeticiones se calcula el valor medio y la desviación estándar de los valores estimados de parámetros orbitales y la masa total.

Los resultados de la estimación de la masa total  $M = m_1 + m_2$  son presentados en la Figura 4.1. Dicha figura muestra las estimaciones (promedio) de masa total con su respectiva desviación estándar para los tres diferentes regímenes de observación (error bajo, medio y alto), además de un gráfico que resume el error porcentual de estimación en los tres casos.

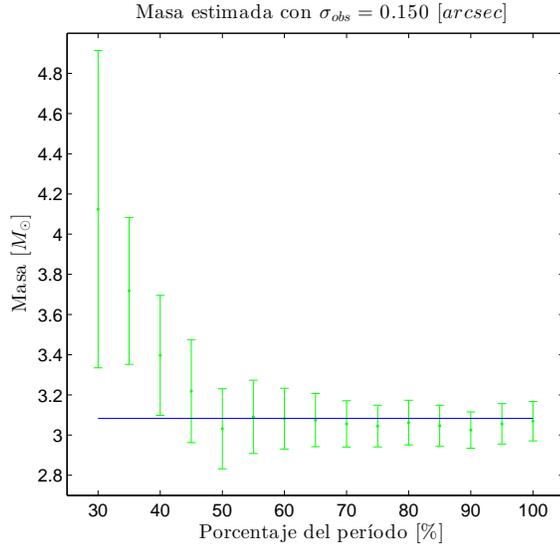
A fin de facilitar la comparación visual de la estimación de masa, las Subfiguras 4.1a, 4.1b y 4.1c son presentadas en la misma escala. Por razones de espacio, los resultados para la estimación



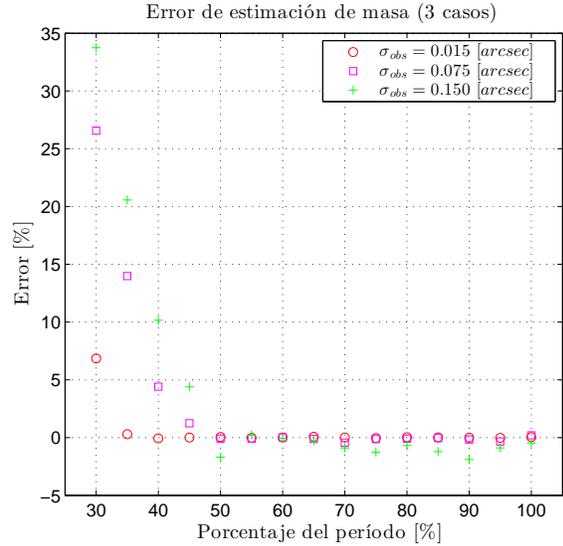
(a) Barra de error para  $\sigma_{obs} = 0.015$  [arcsec]



(b) Barra de error para  $\sigma_{obs} = 0.075$  [arcsec]



(c) Barra de error para  $\sigma_{obs} = 0.150$  [arcsec]



(d) Gráfico de error porcentual en tres escenarios

Figura 4.1: Promedio y desviación estándar del valor estimado de  $m_1 + m_2$ .

de los elementos orbitales es presentada en formato tabla. Las Tablas 4.1, 4.2 y 4.3 resumen los resultados para los casos con error de observación  $\sigma_{obs} = 0.015$  [arcsec],  $\sigma_{obs} = 0.075$  [arcsec] y  $\sigma_{obs} = 0.150$  [arcsec], respectivamente (la Tabla 4.4 se añade para facilitar la comparación numérica entre valores reales y valores estimados). El contenido de las casillas corresponde al promedio (antes del símbolo  $\pm$ ) seguido de la desviación estándar.

Tabla 4.1: Parámetros estimados para  $\sigma_{obs} = 0.015$  [arcsec]

<b>f</b>	$T$ [yr]	$P$ [yr]	$e$	$a$ [arcsec]	$\omega$ [°]	$\Omega$ [°]	$i$ [°]
30 [%]	0.312 ± 0.021	46.225 ± 2.640	0.579 ± 0.009	7.258 ± 0.159	153.425 ± 4.666	46.586 ± 1.427	134.992 ± 1.216
35 [%]	0.285 ± 0.003	49.914 ± 0.502	0.591 ± 0.003	7.489 ± 0.030	147.367 ± 0.448	44.567 ± 0.230	136.443 ± 0.253
40 [%]	0.284 ± 0.002	50.135 ± 0.289	0.592 ± 0.002	7.503 ± 0.017	147.325 ± 0.231	44.623 ± 0.199	136.545 ± 0.160
45 [%]	0.284 ± 0.001	50.106 ± 0.217	0.592 ± 0.001	7.502 ± 0.013	147.298 ± 0.186	44.590 ± 0.184	136.523 ± 0.116
50 [%]	0.284 ± 0.001	50.073 ± 0.154	0.592 ± 0.001	7.500 ± 0.009	147.275 ± 0.215	44.565 ± 0.193	136.528 ± 0.093
55 [%]	0.284 ± 0.001	50.104 ± 0.140	0.592 ± 0.001	7.501 ± 0.011	147.278 ± 0.205	44.576 ± 0.181	136.533 ± 0.097
60 [%]	0.284 ± 0.001	50.073 ± 0.130	0.592 ± 0.001	7.498 ± 0.010	147.245 ± 0.160	44.557 ± 0.151	136.539 ± 0.081
65 [%]	0.284 ± 0.001	50.082 ± 0.097	0.592 ± 0.001	7.501 ± 0.010	147.233 ± 0.157	44.548 ± 0.123	136.508 ± 0.105
70 [%]	0.284 ± 0.001	50.097 ± 0.094	0.592 ± 0.001	7.501 ± 0.008	147.288 ± 0.183	44.585 ± 0.152	136.522 ± 0.097
75 [%]	0.284 ± 0.000	50.099 ± 0.066	0.592 ± 0.001	7.500 ± 0.007	147.294 ± 0.155	44.589 ± 0.126	136.522 ± 0.097
80 [%]	0.284 ± 0.000	50.094 ± 0.067	0.592 ± 0.001	7.501 ± 0.007	147.284 ± 0.141	44.584 ± 0.114	136.513 ± 0.092
85 [%]	0.284 ± 0.000	50.087 ± 0.060	0.592 ± 0.001	7.500 ± 0.007	147.264 ± 0.167	44.568 ± 0.135	136.523 ± 0.093
90 [%]	0.284 ± 0.000	50.082 ± 0.051	0.592 ± 0.001	7.499 ± 0.006	147.291 ± 0.198	44.592 ± 0.162	136.535 ± 0.089
95 [%]	0.284 ± 0.000	50.090 ± 0.041	0.592 ± 0.001	7.499 ± 0.006	147.264 ± 0.197	44.559 ± 0.164	136.537 ± 0.099
100 [%]	0.284 ± 0.000	50.091 ± 0.034	0.592 ± 0.001	7.501 ± 0.006	147.301 ± 0.163	44.600 ± 0.132	136.516 ± 0.085

Tabla 4.2: Parámetros estimados para  $\sigma_{obs} = 0.075$  [arcsec]

<b>f</b>	$T$ [yr]	$P$ [yr]	$e$	$a$ [arcsec]	$\omega$ [°]	$\Omega$ [°]	$i$ [°]
30 [%]	0.373 ± 0.051	39.829 ± 4.536	0.561 ± 0.013	6.912 ± 0.232	166.515 ± 11.045	50.156 ± 3.090	131.388 ± 2.991
35 [%]	0.328 ± 0.030	44.132 ± 3.465	0.555 ± 0.024	7.178 ± 0.181	151.742 ± 3.226	44.709 ± 0.787	133.325 ± 2.086
40 [%]	0.297 ± 0.019	48.140 ± 2.720	0.579 ± 0.019	7.399 ± 0.143	147.530 ± 1.237	44.026 ± 1.178	135.712 ± 1.162
45 [%]	0.287 ± 0.013	49.592 ± 2.151	0.588 ± 0.014	7.475 ± 0.122	147.179 ± 1.248	44.365 ± 1.331	136.384 ± 0.665
50 [%]	0.284 ± 0.009	50.159 ± 1.573	0.593 ± 0.011	7.502 ± 0.098	147.183 ± 1.221	44.456 ± 1.221	136.567 ± 0.422
55 [%]	0.284 ± 0.007	50.146 ± 1.113	0.593 ± 0.008	7.502 ± 0.078	147.281 ± 0.983	44.596 ± 0.924	136.535 ± 0.513
60 [%]	0.283 ± 0.006	50.227 ± 1.043	0.594 ± 0.008	7.513 ± 0.067	147.272 ± 0.904	44.573 ± 0.779	136.432 ± 0.539
65 [%]	0.284 ± 0.004	50.105 ± 0.669	0.593 ± 0.005	7.498 ± 0.046	147.356 ± 0.827	44.657 ± 0.694	136.544 ± 0.475
70 [%]	0.283 ± 0.003	50.231 ± 0.530	0.593 ± 0.005	7.503 ± 0.052	147.313 ± 0.758	44.613 ± 0.641	136.533 ± 0.531
75 [%]	0.283 ± 0.003	50.263 ± 0.444	0.594 ± 0.005	7.514 ± 0.042	147.129 ± 0.773	44.421 ± 0.695	136.408 ± 0.535
80 [%]	0.283 ± 0.003	50.196 ± 0.442	0.593 ± 0.005	7.508 ± 0.042	147.330 ± 0.800	44.587 ± 0.693	136.427 ± 0.619
85 [%]	0.283 ± 0.002	50.183 ± 0.352	0.594 ± 0.005	7.508 ± 0.030	147.028 ± 0.903	44.376 ± 0.797	136.339 ± 0.484
90 [%]	0.283 ± 0.002	50.132 ± 0.262	0.594 ± 0.004	7.500 ± 0.023	147.025 ± 0.885	44.344 ± 0.788	136.553 ± 0.451
95 [%]	0.283 ± 0.001	50.178 ± 0.179	0.594 ± 0.004	7.500 ± 0.026	147.004 ± 0.827	44.320 ± 0.683	136.528 ± 0.399
100 [%]	0.284 ± 0.001	50.108 ± 0.161	0.593 ± 0.004	7.506 ± 0.033	147.139 ± 1.015	44.484 ± 0.881	136.429 ± 0.487

Tabla 4.3: Parámetros estimados para  $\sigma_{obs} = 0.150$  [arcsec]

<b>f</b>	$T$ [yr]	$P$ [yr]	$e$	$a$ [arcsec]	$\omega$ [°]	$\Omega$ [°]	$i$ [°]
30 [%]	0.390 ± 0.060	38.468 ± 5.007	0.560 ± 0.030	6.857 ± 0.244	170.512 ± 13.475	51.195 ± 4.730	130.160 ± 4.423
35 [%]	0.346 ± 0.032	42.158 ± 3.576	0.541 ± 0.027	7.090 ± 0.184	153.614 ± 4.169	44.762 ± 1.631	131.753 ± 2.397
40 [%]	0.315 ± 0.026	45.739 ± 3.692	0.561 ± 0.026	7.268 ± 0.195	148.220 ± 2.007	43.560 ± 2.100	134.849 ± 1.970
45 [%]	0.297 ± 0.022	48.222 ± 3.572	0.578 ± 0.025	7.399 ± 0.184	146.649 ± 2.557	43.490 ± 2.528	136.152 ± 1.274
50 [%]	0.280 ± 0.020	51.106 ± 3.459	0.598 ± 0.024	7.544 ± 0.208	147.545 ± 2.295	44.962 ± 2.353	136.634 ± 0.976
55 [%]	0.285 ± 0.017	50.093 ± 3.032	0.591 ± 0.022	7.495 ± 0.180	147.246 ± 2.377	44.406 ± 2.253	136.655 ± 0.871
60 [%]	0.282 ± 0.015	50.546 ± 2.655	0.594 ± 0.020	7.536 ± 0.173	147.373 ± 1.875	44.673 ± 1.574	136.436 ± 1.207
65 [%]	0.283 ± 0.011	50.292 ± 1.944	0.594 ± 0.016	7.508 ± 0.130	147.095 ± 1.926	44.397 ± 1.679	136.531 ± 1.120
70 [%]	0.282 ± 0.010	50.528 ± 1.780	0.595 ± 0.016	7.517 ± 0.130	147.474 ± 1.316	44.683 ± 1.218	136.451 ± 1.233
75 [%]	0.281 ± 0.008	50.662 ± 1.234	0.599 ± 0.012	7.523 ± 0.100	147.017 ± 1.365	44.166 ± 1.277	136.362 ± 1.106
80 [%]	0.281 ± 0.008	50.564 ± 1.329	0.597 ± 0.016	7.527 ± 0.088	146.984 ± 1.809	44.235 ± 1.676	136.332 ± 1.080
85 [%]	0.280 ± 0.006	50.731 ± 1.129	0.601 ± 0.014	7.531 ± 0.085	146.192 ± 2.055	43.414 ± 1.930	136.327 ± 1.102
90 [%]	0.280 ± 0.005	50.699 ± 0.709	0.601 ± 0.011	7.511 ± 0.064	146.094 ± 2.030	43.253 ± 1.870	136.439 ± 1.099
95 [%]	0.282 ± 0.004	50.372 ± 0.523	0.598 ± 0.011	7.504 ± 0.062	146.414 ± 2.458	43.752 ± 2.348	136.411 ± 1.037
100 [%]	0.283 ± 0.003	50.196 ± 0.390	0.597 ± 0.010	7.498 ± 0.064	146.426 ± 2.541	43.762 ± 2.252	136.594 ± 0.875

Tabla 4.4: Recordatorio de elementos orbitales de Sirius (1)

	$T$ [yr]	$P$ [yr]	$e$	$a$ [arcsec]	$\omega$ [°]	$\Omega$ [°]	$i$ [°]	$M$ [ $M_{\odot}$ ]
Valor real	0.2839	50.090	0.592	7.500	147.27	44.57	136.53	3.0835

Las Subfiguras 4.1a, 4.1b y 4.1c indican, concretamente, que las 80 repeticiones<sup>2</sup> del algoritmo de estimación realizadas para una combinación de  $(\sigma_{obs}, \mathbf{f})$  se acercan, en promedio, al valor indicado por el punto y que, a la vez, la mayoría de los valores estimados se encuentran dentro del rango definido por la barra (desviación estándar). En estas gráficas se puede apreciar que la estimación de  $M = m_1 + m_2$  experimenta una mejora sustancial a medida que la calidad de observaciones –medida en función de la cobertura del período,  $\mathbf{f}$ , y el error de observación,  $\sigma_{obs}$ – aumenta. Específicamente, a medida que  $\sigma_{obs}$  disminuye la estimación es más exacta –más cercana al valor real– y más precisa –los valores calculados por el algoritmo fluctúan en un rango menor–. Por otra parte, a medida que aumenta la cobertura del período orbital, la precisión y exactitud de la masa estimada también mejora. Esta mejora es especialmente notoria cuando  $\mathbf{f}$  sobrepasa cierto *valor crítico* (el primero en que la estimación alcanza un valor cercano a la masa real): antes de dicho umbral, la masa tiende a ser sobreestimada; después de él, la estimación no mejora sustancialmente y se mantiene relativamente constante. Dicho valor crítico se puede identificar, de manera aproximada, en 35 %, 45 % y 50 % para las subfiguras 4.1a, 4.1b y 4.1c, respectivamente.

La sobreestimación de la masa total cuando la cobertura del período está bajo el valor crítico se explica principalmente por la subestimación del período  $P$  (ver tablas 4.1, 4.2 y 4.3 y recordar la relación  $a^3/P^2 \propto m_1 + m_2$ ). Si bien el valor de  $a$  también es subestimado por el algoritmo cuando la calidad de la información es menor<sup>3</sup>, lo que en teoría disminuiría el valor de masa estimada, la magnitud de la diferencia con el valor real de  $a$  no alcanza a compensar el error para el parámetro  $P$ , produciéndose una estimación de masa hasta 35% mayor que el valor real (ver Figura 4.1d).

Cuando el valor  $\mathbf{f}$  de un conjunto de observaciones es bajo, es común que existan diversas configuraciones orbitales capaces de calzar con las posiciones observadas. Una ilustración de este caso puede verse en la Figura 4.2, si bien el escenario no es exactamente el mismo que en esta sección. Para los casos estudiados en este experimento, la tendencia del algoritmo a converger a valores de  $P$  inferiores al real se debe principalmente a la distribución *a priori* utilizada en estos casos, calculada a través de la rutina descrita en la Subsección 3.3.3. Frecuentemente, el rango de  $P$  entregado por dicha rutina entrega rangos que contienen más de un período factible, pudiendo incluso dejar fuera el período real (que, sin embargo, puede ser alcanzado a

---

<sup>2</sup>Recordar que las repeticiones implican la generación de un conjunto **nuevo** de observaciones, aunque estadísticamente equivalente al anterior (definido por valores puntuales de  $\mathbf{f}$ ,  $\sigma_{obs}$ ).

<sup>3</sup>El peor caso se da cuando  $\sigma_{obs} = 0.150$  [arcsec] y  $\mathbf{f} = 30\%$ , ver Tabla 4.3.

través de las variaciones generadas por Evolución Artificial de Parámetros si no se encuentra a una distancia significativa del rango inicial). De esta manera, existe la probabilidad de que las partículas del Filtro encuentren primero un período diferente al real –pero perteneciente a una configuración factible– y se acumulen en torno a este valor. Esto es una desventaja asociada a la aleatoriedad del método, que podría ser abordada a través de otras variedades del Filtro de Partículas (por ejemplo, una versión que fuerce el muestreo de puntos que no necesariamente tengan alta verosimilitud dadas las observaciones).

Respecto a la estimación de los elementos orbitales no mencionados hasta el momento, se puede acotar, en base a la información provista por las Tablas 4.1, 4.2 y 4.3, que: (i) para la configuración orbital estudiada, la estimación es insesgada (en la Sección 4.2 se verá que hay configuraciones para las que el método entrega como *output* conjuntos de parámetros factibles, pero erróneos); (ii) sigue el mismo patrón que la estimación de masa total, es decir, la precisión y exactitud mejoran a medida que aumenta  $\mathbf{f}$  y disminuye  $\sigma_{obs}$ .

Finalmente, cabe mencionar que en todos los casos estudiados hasta este punto se han obtenido estimadores razonables de la masa total, en el sentido de encontrarse en un rango de magnitud cercano al valor real de ésta. En la situación más desfavorable ( $\mathbf{f} = 30\%$ ,  $\sigma_{obs} = 0.150$  [*arcsec*]) el estimador es apenas 35% mayor a la masa real. Por un lado, esto es congruente con el descubrimiento empírico realizado en [Eggen, 1967]: *el valor de la masa total puede estimarse razonablemente incluso si  $a$  y  $P$  no están bien estimados individualmente*. Esto se explica, en líneas generales, gracias a que tan pronto como la estrella secundaria muestra un patrón de movimiento significativamente distinto al movimiento rectilíneo, la curvatura (y con ello, la aceleración) de la órbita se comienza a hacer patente. Por otro lado, debe observarse que el tiempo de paso por el periastro está cubierto por el período de observación en todos los casos, incluso cuando  $\mathbf{f}$  es bajo. Además, el hecho de que la excentricidad posea un valor grande (dentro de su dominio) facilita la estimación, pues los valores de  $T$  y  $\omega$  degeneran a medida que  $e$  se acerca a 0. Estas condiciones son altamente favorables y hacen que los resultados obtenidos no sean completamente generalizables, pues los instantes cercanos al periastro son los que mejor permiten caracterizar la curvatura de la órbita.

## 4.2. Variación de parámetros orbitales

El experimento presentado en la Sección 4.1 contempla únicamente variaciones en la calidad de los datos, sin embargo, es natural preguntarse si la capacidad de estimación del algoritmo se ve afectada por los valores reales de los parámetros o, visto de otra forma, si existen configuraciones orbitales más *difíciles* de estimar que otras. Considerar, a modo de ejemplo, el caso límite en que la inclinación del plano orbital es  $i = 90^\circ$ ; en tal caso, las posiciones observadas en el plano del cielo quedan reducidas, salvo errores observacionales, a una línea en el plano  $XY$ , por lo que es de esperar que la calidad de estimación empeore. Este experimento busca explorar cómo se ve afectada la estimación si algunos parámetros orbitales del sistema toman valores distintos a los utilizados en la Sección 4.1. Dada la dimensionalidad del problema, el proceso de probar todas las combinaciones de valores resultaría exhaustivo (incluso si se considera un grillado no tan fino en cada componente del vector de parámetros). Por esa razón, esta prueba sólo incluye los parámetros que teóricamente podrían tener mayor impacto en la estimación, los cuales se indican a continuación (con la justificación pertinente):

- **Excentricidad  $e$ :** Cuando el valor de este parámetro se acerca a 0, los valores de  $T$ ,  $\omega$  y  $\Omega$  se hacen más difíciles de estimar, pues la órbita se asemeja a una circunferencia. Si  $e$  alcanza el valor nulo, los parámetros mencionados derechamente pierden significado.
- **Inclinación  $i$ :** A medida que el valor de  $i$  se acerca a  $90^\circ$  el conjunto de observaciones de la órbita tiende a colapsar en una línea, perdiendo su carácter bidimensional.
- **Tiempo de paso por el periastro:** Dada la representación alternativa utilizada en este trabajo ( $\mathcal{T} \in [0, 1]$ ), no es necesario cambiar el tiempo de paso por el periastro como tal (que en el caso de Sirius ocurre en los años 1894.130, 1944.220, etc.) sino el año en que se inicia la observación. En la Sección 4.1 se fijó el año  $t_0 = 1930$  como inicio de la campaña, resultando  $\mathcal{T} = 0.2839$ , de manera que el instante 1944.00 está cubierto para todo valor de  $\mathbf{f}$ . De manera natural, surge la pregunta de qué sucede cuando el tiempo de paso por el periastro no se encuentra dentro de la fracción de órbita cubierta por los datos.

### 4.2.1. Variaciones de $e$ e $i$

Con las consideraciones manifestadas previamente sobre los elementos  $e$  e  $i$ , se plantea un experimento con las siguientes características:

- La desviación estándar del error de observación es fijada en  $\sigma_{obs} = 0.075$  [*arcsec*] (75 *mas*). Al igual que en las pruebas presentadas en la Sección 4.1, el error posee una distribución Gaussiana de media nula e igual varianza en cada coordenada del plano del cielo, es decir,  $\sigma_x = \sigma_y = \sigma_{obs}$  (con correlación nula).
- Se utiliza un período de observación fijo ( $\mathbf{f} = 90\%$ ).
- Cada simulación considera un conjunto de observaciones de tamaño  $N_{samp} = 15$ . Las muestras son equiespaciadas en el intervalo temporal  $[0, \mathbf{f} \cdot P]$ .
- Se utiliza como referencia el sistema descrito en la Sección 4.1, pero con los valores  $i = 0^\circ, 18^\circ, 36^\circ, 54^\circ, 72^\circ, 90^\circ$  y  $e = 0, 0.19, 0.38, 0.57, 0.76, 0.95$ .
- Se realizan  $N_{rep} = 80$  repeticiones para cada valor de  $i$  y para cada valor de  $e$  (por separado, es decir, variando sólo un parámetro a la vez). Cada realización considera un nuevo conjunto de valores de posición relativa observada, aunque generados en el mismo intervalo de tiempo (definido por  $\mathbf{f}$ ) y por la misma distribución (definida por  $\sigma_{obs}$ ).
- Sobre cada conjunto de repeticiones se calcula el valor medio y la desviación estándar de los valores estimados de parámetros orbitales y la masa total.

Al igual que en la sección anterior, se utiliza el Filtro de Partículas con la siguiente configuración:  $N_p = 500$  partículas, remuestreo multinomial, umbral de remuestreo  $N_{th} = 0.5$  y número de iteraciones igual a 40. La determinación del rango inicial de  $P$  se realiza a través de la rutina explicada en la Subsección 3.3.3.

Los resultados de este experimento se muestran en las tablas 4.6 y 4.7. Se incluye la Tabla 4.5 para facilitar la comparación de los valores estimados con los de la órbita de referencia.

Tabla 4.5: Recordatorio de elementos orbitales de Sirius (2)

	$T$ [yr]	$P$ [yr]	$e$	$a$ [arcsec]	$\omega$ [°]	$\Omega$ [°]	$i$ [°]	$M$ [ $M_{\odot}$ ]
Valor real	0.2839	50.090	0.592	7.500	147.27	44.57	136.53	3.0835

Tabla 4.6: Estimación de parámetros para variaciones de  $e$

Parámetro	$e = 0$	$e = 0.19$	$e = 0.38$	$e = 0.57$	$e = 0.76$	$e = 0.95$
$\mathcal{T}$	$0.440 \pm 0.211$	$0.284 \pm 0.002$	$0.283 \pm 0.002$	$0.283 \pm 0.002$	$0.283 \pm 0.002$	$0.284 \pm 0.003$
$P$ [yr]	$50.224 \pm 0.191$	$50.151 \pm 0.179$	$50.186 \pm 0.217$	$50.157 \pm 0.249$	$50.150 \pm 0.316$	$50.189 \pm 0.467$
$e$	$0.005 \pm 0.002$	$0.191 \pm 0.003$	$0.382 \pm 0.004$	$0.572 \pm 0.004$	$0.761 \pm 0.005$	$0.951 \pm 0.005$
$a$ [arcsec]	$7.507 \pm 0.032$	$7.500 \pm 0.033$	$7.499 \pm 0.031$	$7.502 \pm 0.035$	$7.513 \pm 0.035$	$7.534 \pm 0.163$
$\omega$ [°]	$171.700 \pm 73.369$	$147.150 \pm 0.903$	$147.097 \pm 0.914$	$147.066 \pm 1.014$	$146.850 \pm 1.309$	$146.003 \pm 5.735$
$\Omega$ [°]	$44.256 \pm 0.630$	$44.346 \pm 0.639$	$44.369 \pm 0.738$	$44.410 \pm 0.923$	$44.266 \pm 1.141$	$43.179 \pm 5.274$
$i$ [°]	$136.318 \pm 0.444$	$136.454 \pm 0.485$	$136.474 \pm 0.449$	$136.459 \pm 0.529$	$136.429 \pm 0.552$	$137.037 \pm 1.857$
$M$ [ $M_{\odot}$ ]	$3.076 \pm 0.034$	$3.076 \pm 0.037$	$3.071 \pm 0.043$	$3.078 \pm 0.046$	$3.092 \pm 0.039$	$3.117 \pm 0.176$

Tabla 4.7: Estimación de parámetros para variaciones de  $i$

Parámetro	$i = 0$ [°]	$i = 18$ [°]	$i = 36$ [°]	$i = 54$ [°]	$i = 72$ [°]	$i = 90$ [°]
$\mathcal{T}$	$0.284 \pm 0.001$	$0.284 \pm 0.001$	$0.284 \pm 0.002$	$0.282 \pm 0.002$	$0.259 \pm 0.013$	$0.228 \pm 0.024$
$P$ [yr]	$50.119 \pm 0.203$	$50.103 \pm 0.210$	$50.168 \pm 0.268$	$50.310 \pm 0.299$	$54.206 \pm 2.575$	$62.880 \pm 7.383$
$e$	$0.593 \pm 0.004$	$0.593 \pm 0.003$	$0.593 \pm 0.004$	$0.596 \pm 0.004$	$0.634 \pm 0.019$	$0.667 \pm 0.029$
$a$ [arcsec]	$7.532 \pm 0.028$	$7.502 \pm 0.034$	$7.500 \pm 0.033$	$7.506 \pm 0.032$	$7.718 \pm 0.144$	$8.152 \pm 0.379$
$\omega$ [°]	$94.854 \pm 51.307$	$146.601 \pm 4.507$	$147.300 \pm 1.264$	$146.880 \pm 0.690$	$145.597 \pm 0.919$	$148.390 \pm 1.769$
$\Omega$ [°]	$97.042 \pm 51.400$	$45.249 \pm 4.490$	$44.611 \pm 1.226$	$44.953 \pm 0.512$	$45.781 \pm 0.519$	$44.560 \pm 0.197$
$i$ [°]	$7.099 \pm 2.201$	$18.027 \pm 1.328$	$35.958 \pm 0.663$	$54.007 \pm 0.366$	$72.717 \pm 0.489$	$90.044 \pm 0.210$
$M$ [ $M_{\odot}$ ]	$3.120 \pm 0.043$	$3.085 \pm 0.039$	$3.074 \pm 0.045$	$3.065 \pm 0.038$	$2.877 \pm 0.120$	$2.549 \pm 0.238$

En la Tabla 4.6 se observa que la estimación de la masa total  $M$  no se ve significativamente afectada por la excentricidad de la órbita, manteniéndose relativamente constante. El caso  $e = 0.95$  escapa ligeramente del patrón mencionado, pues el promedio de los valores estimados está más alejado del valor real y la estimación es un levemente más dispersa (i.e., con mayor desviación estándar) que en el resto de los casos. Sin embargo, se puede concluir que, en general, la estimación de masa es de gran precisión y exactitud. La calidad de observación fijada para este experimento explica parcialmente el resultado, pues en vista de los resultados de la Sección 4.1 se puede afirmar que con un 90% de cobertura del período real –segmento que en este caso incluye el tiempo de paso por el periastro– y error de observación de magnitud media, la ambigüedad respecto a los parámetros  $a$  y  $P$  disminuye considerablemente.

Respecto a la estimación de elementos orbitales, ésta tiende a coincidir, salvo pequeñas diferencias, con el valor real (es decir, no converge en soluciones factibles diferentes a los valores reales). Esto se cumple, en particular, con los valores modificados de  $e$ . A pesar de estos buenos resultados generales, hay ciertos aspectos que vale la pena analizar en detalle:

- Elementos orbitales cuando  $e = 0$ : Los parámetros  $\mathcal{T}$  y  $\omega$  presentan un valores lejanos a los de la órbita original y con una variabilidad notoria en comparación con los casos donde  $e > 0$ . Esto se explica porque, al tratarse de una órbita circular, dichos parámetros pierden sentido. Recordar que  $\mathcal{T}$  y  $\omega$  son el *tiempo de paso por el periastro* y el *argumento del periastro*; a ser todos los puntos equidistantes al foco (que coincide con el centro), el periastro simplemente no existe, o bien todos los puntos pueden ser tratados como tal y, por lo tanto, cualquier valor de  $\mathcal{T}$  y  $\omega$  (en los respectivos dominios) satisface las observaciones. La razón por la que estos parámetros no toman valores en todo su dominio es que el algoritmo no estima una órbita estrictamente circular, sino que cae en valores de excentricidad cercanos a  $e = 0.005$ .
- Elementos orbitales cuando  $e = 0.95$ : Se puede apreciar que la variabilidad, representada por la desviación estándar, aumenta para todos los parámetros, con la excepción de  $\mathcal{T}$  y  $e$ . Una explicación de este hecho puede encontrarse en la Segunda Ley de Kepler: la excesiva velocidad angular con la que pasa la estrella secundaria por los puntos cercanos al periastro hace que el muestreo temporalmente esquiapaciado de las posiciones observadas no permita caracterizar detalladamente ese fragmento de órbita, de manera que un rango de valores ligeramente más amplio puede encajar con las posiciones observadas. Cuando  $e$  es menor, la diferencia de la separación y las velocidades angulares alcanzadas en torno al periastro no son *tan* diferentes a las del resto de la órbita.

En la Tabla 4.7 se aprecia una clara degradación en la calidad de la estimación de la masa a partir de (aproximadamente) los  $70^\circ$  de inclinación, ocasionada a su vez por la degradación de la estimación de  $a$  y  $P$ . En los casos  $i = 72^\circ$  y  $i = 90^\circ$  el algoritmo, en promedio, converge a soluciones cercanas, pero incorrectas, de los parámetros  $\mathcal{T}$ ,  $P$ ,  $e$  y  $a$  (esto se deriva del hecho de que los valores de la órbita real ni siquiera están dentro de los rangos de error indicados en la tabla). Una posible explicación podría encontrarse en: (i) la rutina (Subsección 3.3.3) para calcular el rango inicial del período, que ocasionalmente deja fuera al período real; (ii) la acumulación de partículas del Filtro en torno al primer valor factible encontrado, que en este problema no necesariamente coincide con el valor real (la existencia de esta ambigüedad, i.e. más de un valor factible en el rango estudiado, posiblemente no existe para valores menores de  $i$ ). Notar que la rutina para determinar la distribución *a priori* del período ignora factores de

orientación; por esta razón, es de esperar que su desempeño frente a escenarios más desafiantes (i.e. más cercanos a casos límites) en ese sentido empeore. Otros aspectos importantes de este resultado son explicados a continuación:

- Caso  $i = 0^\circ$  (órbita en posición frontal): Los elementos  $\omega$  y  $\Omega$  presentan valores alejados a los de la órbita de referencia y amplios rangos de error. La raíz de esto es una ambigüedad de ángulo: de la Ecuación 2.10 se desprende que cuando  $i = 0^\circ$ ,  $A = G = a \cos \omega + \Omega$  y  $B = -F = a \sin \omega + \Omega$ , de manera que  $\omega - \Omega$  queda indeterminado (ver Ecuación 2.12).
- Caso  $i = 90^\circ$ : Llama la atención la precisión y exactitud alcanzada para el parámetro  $\Omega$ , siendo mejor que para el resto de los casos a pesar de que la estimación de masa es notoriamente peor. Esto se debe a que cuando la órbita se encuentra inclinada en ángulo recto respecto al plano del cielo, colapsa en una línea inclinada exactamente en un ángulo  $\Omega$  respecto a la dirección de referencia (Norte astronómico), independientemente del resto de los parámetros.

### 4.2.2. Variación de $\mathcal{T}$

Con las consideraciones manifestadas previamente acerca del parámetro  $\mathcal{T}$ , se plantea un experimento similar al de la Sección 4.1 (en el sentido de incluir variación en  $\mathbf{f}$  y  $\sigma_{obs}$ ), pero contemplando además variaciones en  $\mathcal{T}$ :

- Se definen dos valores distintos para el instante en que comienza la observación (recordar que el tiempo de paso por el periastro ocurre en 1944.220):
  - Año 1930, resultando  $\mathcal{T} = 14.220/50.090 = 0.2839$  (igual que en la Sección 4.1).
  - Año 1900, resultando  $\mathcal{T} = 44.220/50.090 = 0.8828$  (el tiempo de paso por el periastro sólo es cubierto cuando la observación abarca más que el 88.28 % del período orbital).
- Se definen tres valores para la desviación estándar del error de observación:
  - $\sigma_{obs} = 15 \text{ mas}$  (0.015 [arcsec])
  - $\sigma_{obs} = 75 \text{ mas}$  (0.075 [arcsec])
  - $\sigma_{obs} = 150 \text{ mas}$  (0.150 [arcsec])

Se asume un error Gaussiano de media nula e igual varianza en cada coordenada del plano del cielo, es decir,  $\sigma_x = \sigma_y = \sigma_{obs}$  (con correlación nula).

- Se define un arreglo con diferentes períodos de observación, (parametrizados por  $\mathbf{f}$ ). En este caso,  $\mathbf{f} = 30\%$ ,  $32\%$ ,  $\dots$ ,  $98\%$ ,  $100\%$ .
- Cada simulación considera un conjunto de observaciones de tamaño  $N_{samp} = 15$ . Las muestras son equiespaciadas en el intervalo temporal  $[0, \mathbf{f} \cdot P]$ .
- Se realizan  $N_{rep} = 10$  repeticiones para cada combinación de  $(\sigma_{obs}, \mathbf{f})$ . Cada realización considera un nuevo conjunto de valores de posición relativa observada, aunque generados en el mismo intervalo de tiempo (definido por  $\mathbf{f}$ ) y por la misma distribución (definida por  $\sigma_{obs}$ ).
- Sobre cada conjunto de repeticiones se calcula el valor medio de la masa total.

Los resultados de este experimento son mostrados a continuación:

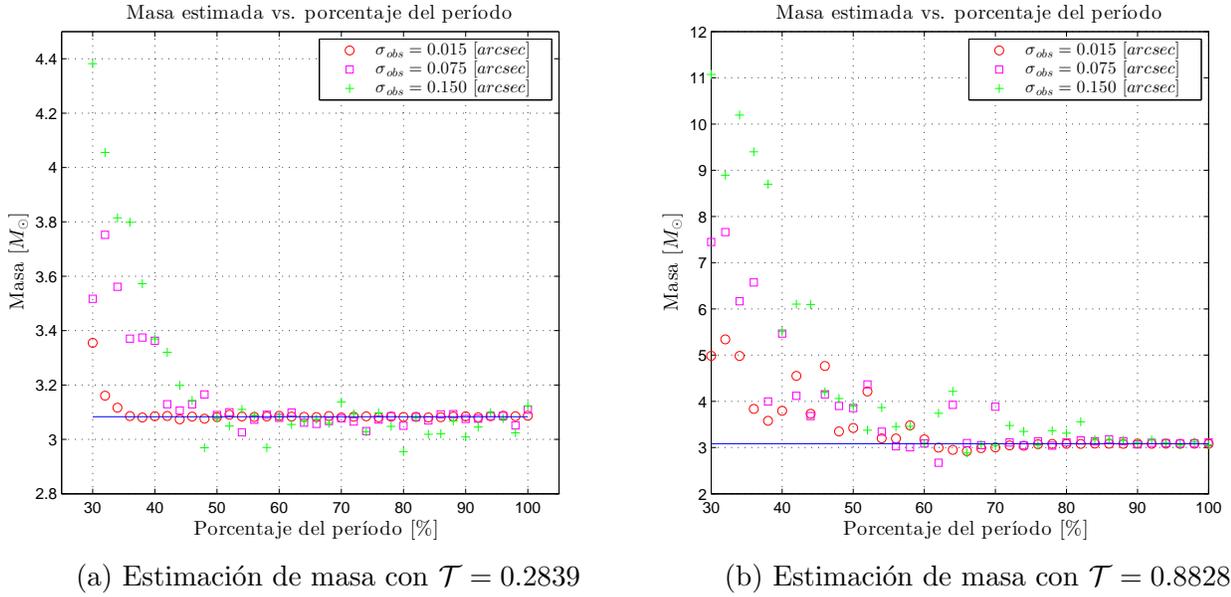


Figura 4.2: Estimación de masa con diferentes valores de  $\mathcal{T}$

El contraste entre las Figuras 4.2a y 4.2b (notar que no están en la misma escala) da cuenta de la importancia de disponer de observaciones en torno al periastro. Mientras la Subfigura 4.2a muestra concordancia absoluta con los resultados obtenidos en la Sección 4.1 –es un experimento muy similar–, la Figura 4.2b exhibe una sobreestimación de magnitud comparativamente mayor (hasta  $11 [M_{\odot}]$  en el caso de  $\mathbf{f} = 30\%$  y  $\sigma_{obs} = 0.150 \text{ [arcsec]}$ ). La sobreestimación de  $M$  se explica por la convergencia del algoritmo en valores de órbita que, si bien son capaces de calzar con las posiciones observadas (es decir, presentan un error cuadrático medio similar al de una órbita bien estimada), no corresponden a la órbita real. Una de las fuentes de ambigüedad es la cobertura limitada de órbita, tal como se discutió en la Sección 4.1; sin embargo, en el caso de la Figura 4.2b, con  $\mathcal{T} = 0.8828$ , la situación se agrava por la ausencia del periastro (o puntos cercanos) entre las observaciones, pues el segmento de órbita capturado no logra caracterizar la curvatura de la órbita. Naturalmente, a medida que la cobertura de órbita  $\mathbf{f}$  se acerca al  $100\%$  la ambigüedad desaparece. En la Figura 4.3 se muestra un caso particular de ambigüedad de órbita (obtenido para  $\sigma_{obs} = 0.075 \text{ [arcsec]}$  y  $\mathbf{f} = 34\%$ ). La Tabla 4.8 compara los valores estimados por el algoritmo en el caso particular ilustrado en la Figura 4.3 con los valores reales. Observar cómo una órbita con menor período ( $P = 41.7772 \text{ [yr]}$ ), mayor excentricidad ( $e = 0.7124$ ), mayor largo del semi-eje principal ( $a = 8.0190$ ) y distinta orientación respecto a la órbita real puede “explicar” las observaciones de manera satisfactoria.

Se mantiene, no obstante, el patrón observado en los resultados de la Sección 4.1: a medida que  $\sigma_{obs}$  disminuye y la cobertura de órbita aumenta, la estimación mejora. Se puede postular, igualmente, la existencia de un *umbral* o *valor crítico* de  $\mathbf{f}$  a partir del cual la estimación de masa se concentra en torno al valor real; sin embargo, éste es más difuso –existen algunas variaciones inesperadas entre valores sucesivos de  $\mathbf{f}$ , lo que da cuenta de una mayor varianza (no medida explícitamente)– y se encuentra definitivamente desplazado hacia la derecha (el umbral podría situarse en  $\mathbf{f} = 60\%$ , aproximadamente).

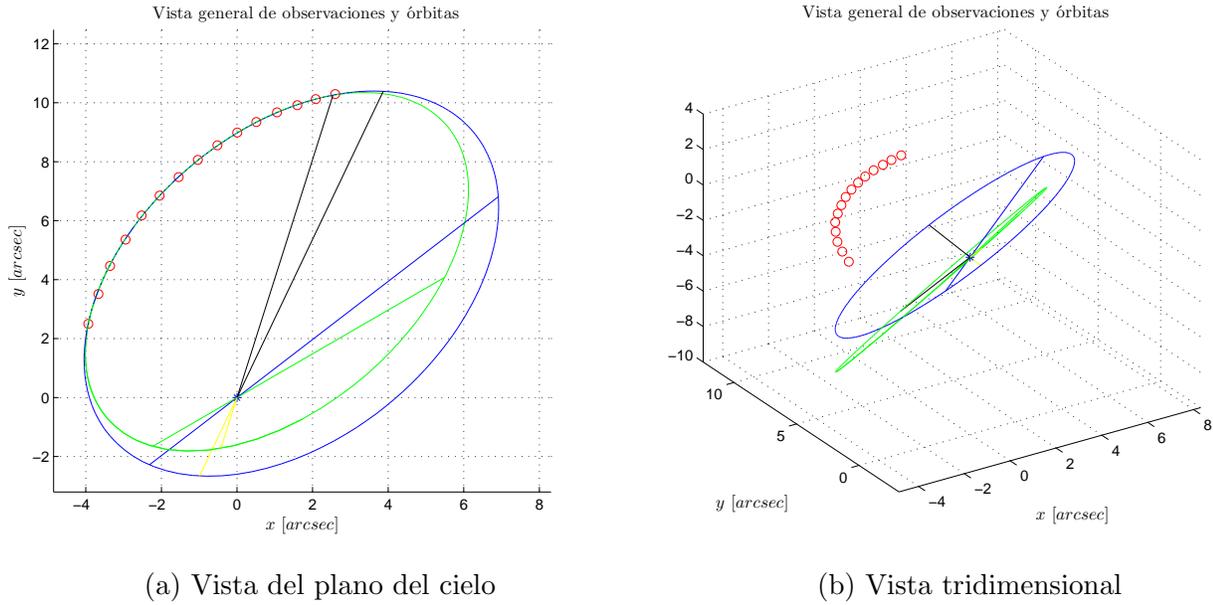


Figura 4.3: Observaciones y órbitas con cobertura parcial del período

Tabla 4.8: Ajuste orbital erróneo (Experimento 2)

Caso	$\mathcal{T}$	$P$ [yr]	$e$	$a$ [arcsec]	$\omega$ [°]	$\Omega$ [°]	$i$ [°]	$M$ [ $M_{\odot}$ ]
Real	0.8828	50.0900	0.5923	7.5000	147.2700	44.5700	136.5300	3.0835
Estimado	0.8814	41.7772	0.7124	8.0190	126.4047	36.5954	127.4378	5.4181

**Nota:** Por razones de espacio, la simbología de la Figura 4.3 se omitió de la imagen y es explicada textualmente en este párrafo: (i) la elipse de color azul representa la órbita real; (ii) la elipse de color verde representa la órbita estimada; (iii) las circunferencias representan posiciones observadas (a mayor error de observación, más oscuras); (iv) las rectas color azul y verde representan las líneas de nodos (intersección del plano orbital con el plano del cielo) para las órbitas real y estimada, respectivamente; (v) las rectas amarillas corresponden al segmento que une la estrella primaria (en el foco de la elipse) con el periastro (en ambas órbitas); (vi) las rectas negras corresponden al segmento que une la estrella primaria (en el foco de la elipse) con el apoastro (en ambas órbitas). También debe notarse que las posiciones observadas son expresadas en segundos de arco para emular las observaciones reales en el plano del cielo, y esta unidad de medida se mantiene para el caso de la vista tridimensional. El eje  $Z$  no está rotulado pues los segundos de arco no tienen significado físico en dicha dimensión.

### 4.3. Uso y alcance del método

A diferencia de los experimentos anteriores, el caso presentado en esta sección no busca validar estadísticamente el método de estimación propuesto, sino mostrar en detalle una **realización particular** de éste. La motivación más obvia para un reporte de esta naturaleza es la de servir al lector como **ejemplo** o **manual de uso**, no sólo indicando los pasos a seguir sino también –y aún más importante– dando a conocer, a través de gráficos representativos, el *output* del algoritmo en su totalidad, ya que el enfoque estadístico de las secciones anteriores obliga a omitir parte de la información. Adicionalmente, esta experiencia sirve para atisbar el **alcance** o la **utilidad** de la herramienta desarrollada más allá de la función presentada en las Secciones 4.1 y 4.2 (i.e. la estimación de valores puntuales de los elementos orbitales). Los siguientes párrafos contienen el planteamiento de este experimento.

**Sistema simulado:** En consideración de los resultados de la Sección 4.2, donde se aprecia que el comportamiento del método de estimación tiende a empeorar cuando la inclinación  $i$  se acerca a  $0^\circ$ , en esta sección se plantea un sistema estelar hipotético destinado a *desafiar* al algoritmo, en el sentido de que las observaciones sugieren no una, sino varias combinaciones de parámetros capaces de calzar con la órbita aparente.

- Paralaje:  $\pi = 50.00$  [mas] ( $\pi = 0.050$  [arcsec]).
- Período orbital:  $P = 50.090$  [yr].
- Longitud del semi-eje mayor:  $a = 1$  [arcsec].
- Excentricidad:  $e = 0.34$ .
- Inclinación:  $i = 6.00^\circ$
- Longitud del nodo ascendente:  $\Omega = 6.00^\circ$
- Tiempo de paso por el periastro:  $\mathcal{T} = 0.3637$  [yr].
- Argumento del periastro:  $\omega = 6.00^\circ$

Los valores del semi-eje mayor y de la paralaje fueron escogidos de manera que el sistema estelar simulado posea una masa total similar a la de Sirius, pero a la vez se encuentre a una

mayor distancia respecto al observador. Cabe mencionar que el tiempo de paso por el periastro no requiere un punto específico en el tiempo (como el año 1894.130 en el caso de Sirius), pues al tratarse de un sistema hipotético, el valor mencionado se puede representar directamente mediante el parámetro alternativo  $\mathcal{T}$ , definido en la Subsección 3.3.4.

**Generación de observaciones:** A diferencia de los experimentos anteriores, el muestreo no es uniforme en el tiempo, sino generado mediante la rutina presentada en la Subsección 3.2.1 (con los parámetros de entrada siguientes:  $TimeObs = 50.090$  [yr],  $nInt = 89$ ,  $Indices = [2 : 17, 25 : 40, 70 : 78]$ ,  $LossRate = 0.04$ ). Además, la calidad de los datos es variable: las muestras tomadas los primeros 12 años del período de observación tienen un error de observación de  $\sigma_{obs} = 0.08$  [arcsec], valor que disminuye a  $\sigma_{obs} = 0.04$  [arcsec] en las observaciones posteriores. La idea detrás de esta decisión es modelar (gruesamente) el hecho de que el paso de los años trae consigo avances tecnológicos, que en este caso se traducen en observaciones astrométricas de mayor calidad.

**Procedimiento:** Esta prueba se divide en dos etapas:

1. **Estimación:** La primera utiliza el Filtro de Partículas tal como se planteó en la Sección 3.3, a fin de obtener una primera estimación de la f.d.p. de los elementos orbitales (representada por las partículas del filtro).
2. **Exploración:** La segunda etapa hace uso de una modificación del filtro de la Sección 3.3, con las siguientes características: (i) se utiliza la representación clásica de la órbita:  $\mathcal{T}$ ,  $P$ ,  $e$ ,  $a$ ,  $\omega$ ,  $\Omega$ ,  $i$ ; (ii) se realiza Evolución Artificial de Parámetros en todas las dimensiones, es decir, no se aplica la reducción de dimensionalidad descrita en la Subsección 3.3.2; (iii) para los parámetros  $T$ ,  $P$ ,  $e$ ,  $a$ , se utiliza como distribución *a priori* la obtenida en la primera etapa; (iv) para los parámetros de orientación ( $\omega$ ,  $\Omega$ ,  $i$ ), se utiliza una distribución *a priori* uniforme en el rango de validez de cada parámetro ( $[0, 2\pi]$   $[0, \pi]$   $[0, \pi]$ , respectivamente); (v) el número de iteraciones se reduce a 15 y el número de partículas aumenta a  $N_p = 15000$ .

Antes de la presentación de resultados –realizada en las Subsecciones 4.3.1 y 4.3.2–, es necesario hacer ciertas acotaciones sobre la segunda etapa:

- La motivación de este procedimiento es la **exploración** del espacio de soluciones sin la

restricción de la reducción de dimensionalidad, pues el estimador de mínimos cuadrados usado en la etapa 1 (y en los experimentos anteriores) deja fuera un universo de soluciones factibles que ocasionalmente coinciden con los valores reales del sistema y, sin embargo, quedan descartadas. De esa manera, el estimador de mínimos cuadrados podría imponer un sesgo sobre los parámetros  $a$ ,  $\omega$ ,  $\Omega$ ,  $i$ .

- Sobre los parámetros estimados por mínimos cuadrados en la etapa 1: como se puede apreciar en la Sección 4.2, los parámetros orbitales que presentan mayor error cuando  $i$  se acerca a  $0^\circ$  de estimación son los relacionados con la orientación. El largo del semi-eje mayor, por el contrario, es estimado con gran exactitud y precisión en la mayoría de los casos, al igual que  $\mathcal{T}$ ,  $P$ ,  $e$ . Esa es la razón por la cual su distribución no es uniforme, sino basada en la distribución de partículas de la etapa 1.
- Puesto que después de la primera etapa de estimación los valores de  $T$ ,  $P$ ,  $e$  y  $a$  suelen quedar localizadas en un rango acotado dentro el espacio de soluciones (con mayor o menor exactitud, pero, por regla general, con considerable precisión), durante la etapa 2 pueden considerarse como valores conocidos o *semi-fijos*, mientras que los parámetros  $\omega$ ,  $\Omega$ ,  $i$  son forzados a tomar valores en todo su rango factible.
- El número de partículas  $N_p = 15000$  está justificado por el punto anterior. Este valor es varias veces superior al  $N_p = 500$  utilizado en la primera etapa, aunque aún computacionalmente tratable<sup>4</sup>. El mayor conocimiento de los elementos orbitales –existencia de parámetros *semi-fijos*– permite acotar el rango de búsqueda en algunas dimensiones, priviligando la variabilidad en los componentes donde se busca explorar soluciones factibles ( $i$ ,  $\omega$ ,  $\Omega$ ). Notar que si no existiese conocimiento previo de la órbita y se explorara un amplio rango de valores para  $T$ ,  $P$ ,  $e$  y  $a$ , es probable que ninguna partícula alcance un valor factible<sup>5</sup>.

---

<sup>4</sup>Sin embargo, es un número insuficiente para permitir el descubrimiento de soluciones mediante fuerza bruta: tener en cuenta que la raíz séptima de 15000 es levemente menor a 4, por lo que un grillado con esa cantidad de puntos permitiría no más de 4 valores distintos por cada componente del vector de parámetros.

<sup>5</sup>Nota: Cuando las observaciones se comparan con las posiciones teóricas definidas por una partícula de configuración orbital excesivamente diferente a una configuración factible (en el sentido de generar error medio cuadrático extremadamente alto), el valor de la verosimilitud (distribución teórica del error medio cuadrático) es computacionalmente igual a 0. Si esto ocurre para todas las partículas, no es posible realizar estimación alguna.

### 4.3.1. Etapa 1: Estimación

Esta subsección detalla la primera etapa de estimación. En primer lugar, se muestra la Figura 4.4, la cual permite observar: (i) en qué instantes se obtienen datos (cuando no se recibe observación, la coordenada  $Y$  aparece en blanco); (ii) la desviación estándar  $\sigma_{obs}$  del error de observación asociada a cada instante o época. Notar que en el primer período de observación activa (aquel con  $\sigma_{obs} = 0.08$  [arcsec]) existe pérdida de información poco antes del año 5. El total de observaciones es 40.

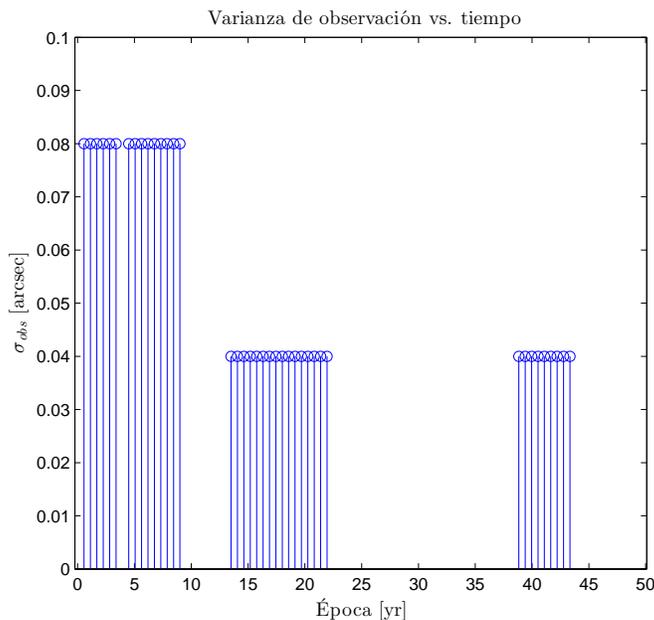
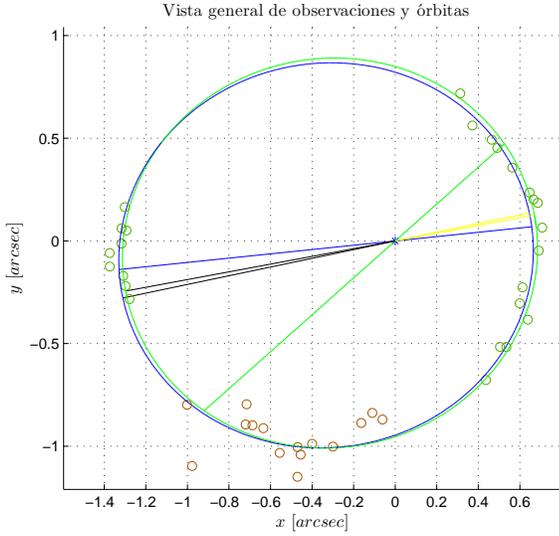
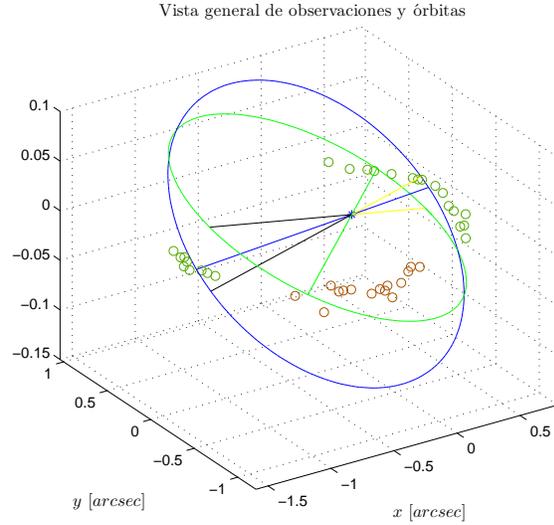


Figura 4.4: Error característico en cada instante de observación

Al momento de ejecutar el método de estimación, la única información disponible es la de los datos de posición y la de los instantes de observación. No existe una distribución *a priori* de los parámetros a estimar, por lo que se utiliza la rutina descrita en la Subsección 3.3.3 a fin de obtener un rango inicial de valores para el período (la primera generación de partículas distribuye uniforme en ese rango). Para las observaciones artificiales generadas en este experimento, el rango inicial calculado por dicha rutina es [48.00 yr, 72.00 yr], el cual contiene el período orbital real  $P = 50.090$  [yr]. Con ese input, se procede a realizar la primera etapa de estimación, cuyos resultados numéricos se encuentran resumidos en la Tabla 4.9. La Figura 4.5, por otra parte, entrega dos vistas de las posiciones observadas junto a la órbita real y la órbita estimada.



(a) Vista del plano del cielo



(b) Vista tridimensional

Figura 4.5: Observaciones y órbitas (real y estimada)

Tabla 4.9: Ajuste orbital erróneo (Experimento 3)

Caso	$\mathcal{T}$	$P$ [yr]	$e$	$a$ [arcsec]	$\omega$ [°]	$\Omega$ [°]	$i$ [°]	$M$ [ $M_{\odot}$ ]
Real	0.3637	50.0900	0.3400	1.0000	6.0000	6.0000	6.0000	3.1885
Estimado	0.3655	49.7161	0.3190	1.0015	328.6447	41.9729	3.9677	3.2517

**Nota:** Por razones de espacio, la simbología de la Figura 4.5 se omitió de la imagen y es explicada textualmente en este párrafo: (i) la elipse de color azul representa la órbita real; (ii) la elipse de color verde representa la órbita estimada; (iii) las circunferencias representan posiciones observadas (a mayor error de observación, más oscuras); (iv) las rectas color azul y verde representan las líneas de nodos (intersección del plano orbital con el plano del cielo) para las órbitas real y estimada, respectivamente; (v) las rectas amarillas corresponden al segmento que une la estrella primaria (en el foco de la elipse) con el periastro (en ambas órbitas); (vi) las rectas negras corresponden al segmento que une la estrella primaria (en el foco de la elipse) con el apoastro (en ambas órbitas). También debe notarse que las posiciones observadas son expresadas en segundos de arco para emular las observaciones reales en el plano del cielo, y esta unidad de medida se mantiene para el caso de la vista tridimensional. El eje  $Z$  no está rotulado pues los segundos de arco no tienen significado físico en dicha dimensión.

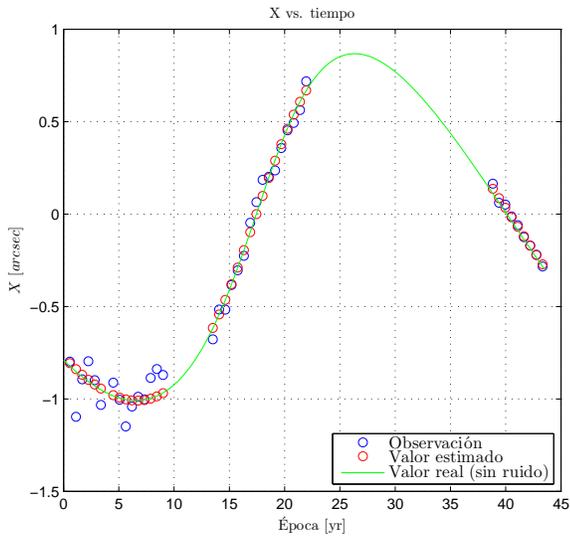
El caso presentado por la Figura 4.5 y la Tabla 4.9 muestra una ambigüedad de naturaleza distinta a la explorada en las secciones anteriores. En esta ocasión, la cobertura de órbita es relativamente buena (contiene periastro, apoastro y un segmento entre ambas), de manera que el algoritmo converge a una órbita cuya proyección coincide (aproximadamente) con la de la órbita real en todos sus tramos. Contraste con la Figura 4.3, donde la solución es capaz de explicar las observaciones disponibles pero es notoriamente discordante con la órbita real en el segmento orbital no cubierto. La ambigüedad, en este caso, se encuentra en la orientación espacial, pues

en la Tabla 4.9 se observan diferencias significativas entre los valores reales y estimados de  $\omega$ ,  $\Omega$  y, en menor medida,  $i$ . La diferencia entre ambas órbitas puede apreciarse visualmente en la Figura 4.5b.

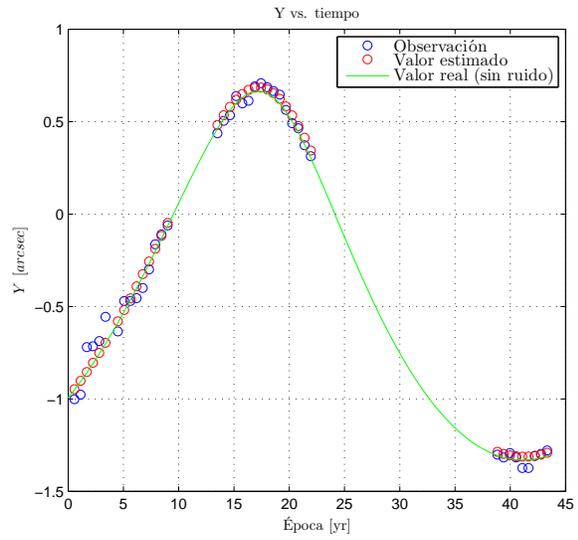
La explicación de esta ambigüedad tiene su origen en la indefinición del ángulo  $\Omega - \omega$  a medida que  $i$  tiende a  $0^\circ$  (como se explicó en la Sección 4.2). En esta sección, el valor real de  $i$  es positivo pero cercano a  $0^\circ$ , de manera que aún existe cierta ambigüedad. Geométricamente, cuando  $i$  es nulo o cercano a cero, se puede conocer relativamente bien el ángulo  $\Omega + \omega$ , pues corresponde, de manera exacta cuando  $i = 0^\circ$  y de manera aproximada cuando sólo está cerca de dicho valor, al ángulo del periastro (en la Figura 4.5a se observa que los segmentos de color amarillo casi coinciden entre sí; de hecho, la suma de los valores estimados  $\Omega = 41.9729^\circ$  y  $\omega = 328.6447^\circ$  es equivalente a  $10.6176^\circ$ , muy cercana al valor real  $12.00^\circ$ ); sin embargo, resulta difícil conocer el ángulo específico en que se encuentra la línea de nodos (o, lo que es lo mismo, conocer  $\Omega - \omega$ ), pues el plano del cielo y el plano orbital se encuentran (casi) acoplados. Respecto a esta última afirmación, notar que: (i) en la Figura 4.5b, las rectas azul y verde (líneas de nodos de las órbitas real y estimada, respectivamente) se encuentran significativamente alejadas entre sí; (ii) la resta entre los valores estimados  $\Omega = 41.9729^\circ$  y  $\omega = 328.6447^\circ$  es  $-286.6718^\circ$ , muy alejada del valor real de  $0^\circ$ .

La idea de extender el algoritmo mediante una etapa de exploración, como la presentada en la Subsección 4.3.2, surge precisamente a raíz de las ambigüedades existentes en casos como el recién presentado.

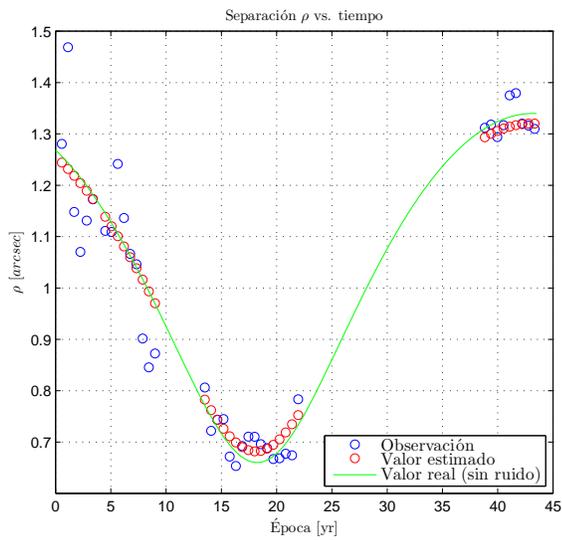
Antes de pasar a la Subsección 4.3.2, se presentan algunos gráficos de uso común en esta área de la astronomía. La Figura 4.6 agrupa las secuencias temporales de las posiciones en el plano del cielo (observada, estimada y real), tanto en términos de  $X$  e  $Y$  como de  $\rho$  y  $\theta$ . La Figura 4.7 agrupa las secuencias temporales de valores de **O–C** para  $X$ ,  $Y$ ,  $\rho$  y  $\theta$ . En astronomía, el término **O–C** hace referencia a la diferencia entre el valor **observado** (observación visual/astrométrica en este caso) con el valor **computado** o **calculado** (el valor de posición que correspondería, de acuerdo a los parámetros orbitales estimados, al cada instante de observación).



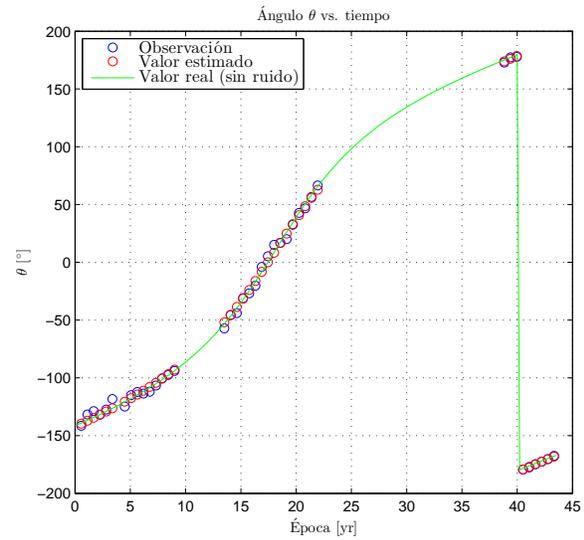
(a) Valores de  $X$  en el tiempo



(b) Valores de  $Y$  en el tiempo

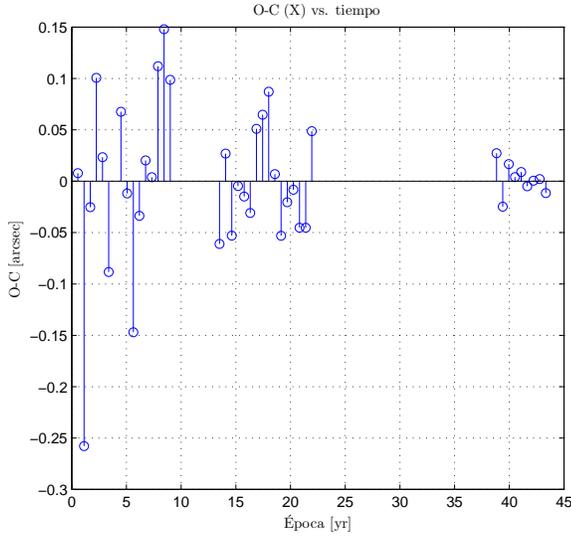


(c) Valores de  $\rho$  en el tiempo

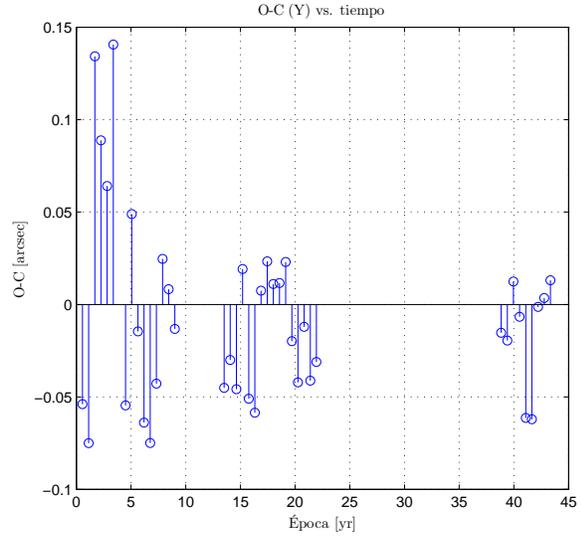


(d) Valores de  $\theta$  en el tiempo

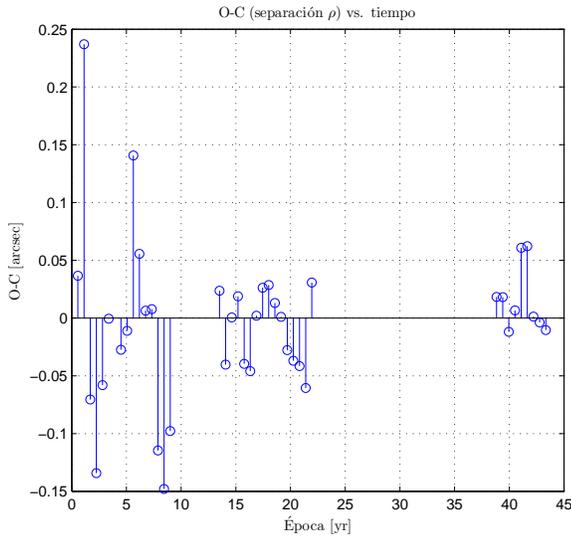
Figura 4.6: Valores observados, reales y estimados de  $X$ ,  $Y$  y  $\rho$ ,  $\theta$



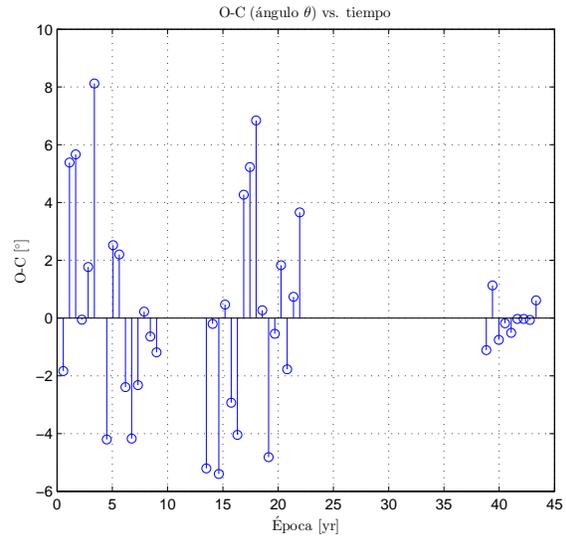
(a) O-C de  $X$  en el tiempo



(b) O-C de  $Y$  en el tiempo



(c) O-C de  $\rho$  en el tiempo



(d) O-C de  $\theta$  en el tiempo

Figura 4.7: O-C para  $X$ ,  $Y$  y  $\rho$ ,  $\theta$

La información presentada en las Figuras 4.6 y 4.7 es, de alguna manera, redundante con la de la Figura 4.5, sólo que los valores numéricos y la dimensión temporal aparecen de manera más explícita. Notar que en el primer período de observación es el que presenta mayor **O-C**, pues la varianza del error observacional es mayor que en los períodos siguientes. Actualmente, los valores típicos de **O-C** para sistemas estelares reales suelen ser menores que los obtenidos en este trabajo; sin embargo, también la calidad de observación suele ser mejor que la utilizada en esta simulación (algunos instrumentos pueden alcanzar una precisión por debajo de los  $0.001 [arcsec]$ ).

### 4.3.2. Etapa 2: Exploración

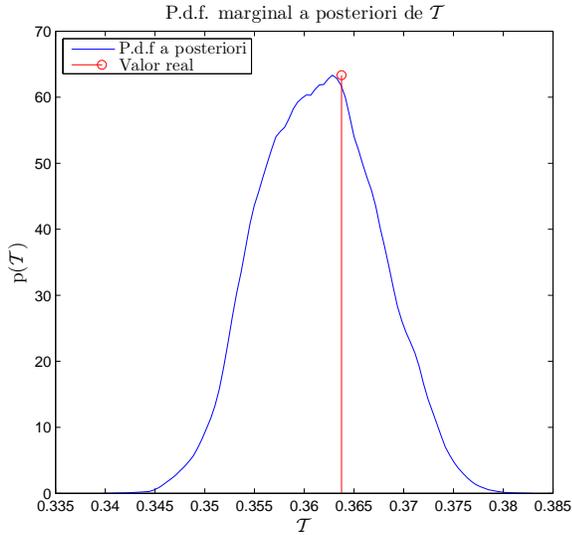
Esta subsección presenta los resultados de la etapa de **exploración**, detallada en el apartado titulado **Procedimiento** al inicio de la Sección 4.3. Debe mencionarse, a modo de recordatorio, que esta etapa del experimento consiste, en términos generales, en usar la estimación llevada a cabo en la primera etapa como información *a priori* de una nueva ejecución del Filtro de Partículas, usando el mismo conjunto de observaciones pero realizando, esta vez, Evolución Artificial de Parámetros en los siete componentes del vector de estado, a fin de identificar soluciones factibles en todo el espacio paramétrico. Los resultados son entregados a través de la Tabla 4.10 y las Figuras 4.8, 4.9.

Tabla 4.10: Valores reales y estimados (Etapa de exploración)

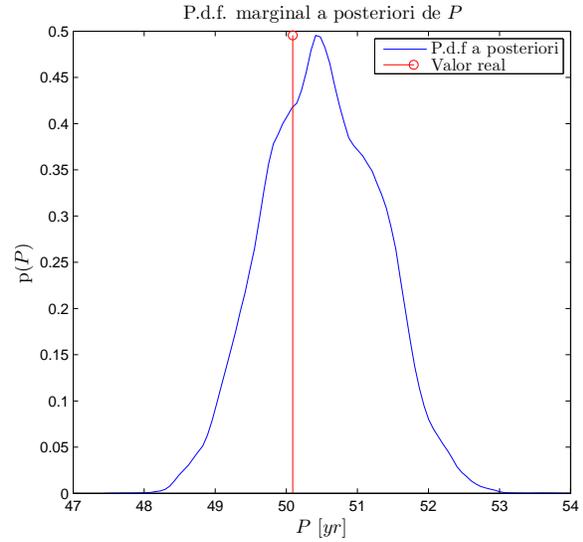
Caso	$\mathcal{T}$	$P$ [yr]	$e$	$a$ [arcsec]	$\omega$ [°]	$\Omega$ [°]	$i$ [°]	$M$ [ $M_{\odot}$ ]
Real	0.3637	50.0900	0.3400	1.0000	6.0000	6.0000	6.0000	3.1885
Estimado (1)	0.3655	49.7161	0.3190	1.0015	328.6447	41.9729	3.9677	3.2517
Estimado (2)	0.3615	50.5009	0.3353	1.0049	279.3857	76.1346	7.6031	3.1832

En la Tabla 4.10 se muestran los originales del sistema estelar (primera fila), los valores estimados en la primera etapa (segunda fila) y los valores estimados en la etapa de exploración (tercera fila). Se observa que, con la excepción de  $e$ , ninguno de los elementos orbitales estimados presenta mejoras respecto a los valores estimados en la primera etapa. Sin embargo, debe recordarse que el objetivo primordial de esta etapa es caracterizar el espacio de soluciones, más que encontrar una solución única. Incluso, la estimación de parámetros por medio del promedio ponderado (aproximación del valor esperado) utilizada en los valores de la tabla pierde el sentido, pues en ciertas componentes los valores factibles tienden más bien a agruparse en torno a *peaks* (ver Figuras 4.9a y 4.9b). De hecho, al visualizar la órbita artificial generada por los valores paramétricos de la tercera fila de la Tabla 4.10, ésta no muestra gran concordancia con la órbita real, a diferencia de lo observado en la Figura 4.5. De esta manera, un enfoque más apropiado para realizar la estimación sería realizar algún tipo de *clustering* para identificar las modas, o bien utilizar el *máximo a posteriori* (que sería equivalente a seleccionar la moda de máxima verosimilitud).

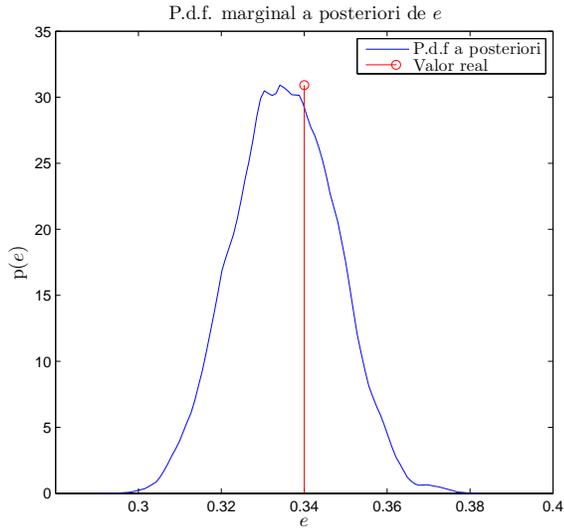
Las Figuras 4.8 y 4.9 constituyen un mero acercamiento a la distribución *a posteriori* calculada por el Filtro de Partículas, puesto que la dimensionalidad del problema impide visualizar p.d.f. de las siete componentes en un único gráfico.



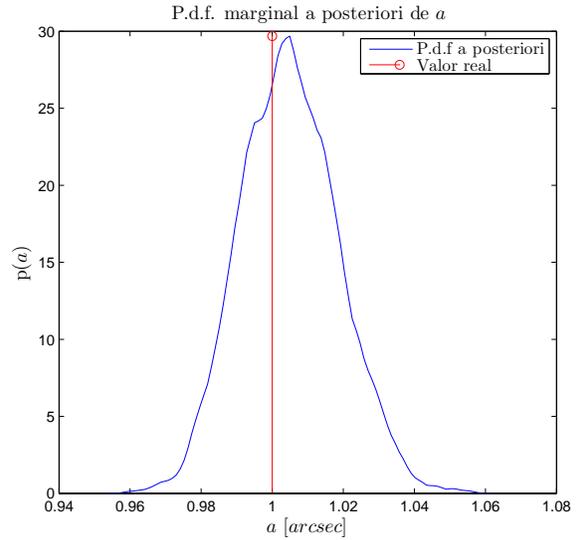
(a) Tiempo de paso por el periastró



(b) Período orbital



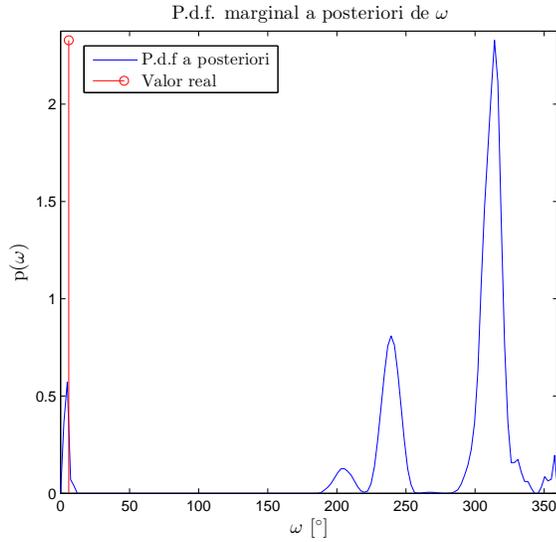
(c) Excentricidad



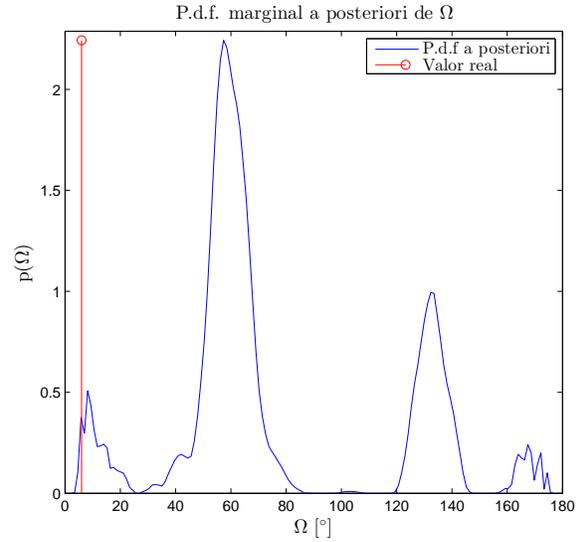
(d) Largo del semi-eje mayor

Figura 4.8: Distribuciones marginales *a posteriori* de parámetros *semi-fijos*

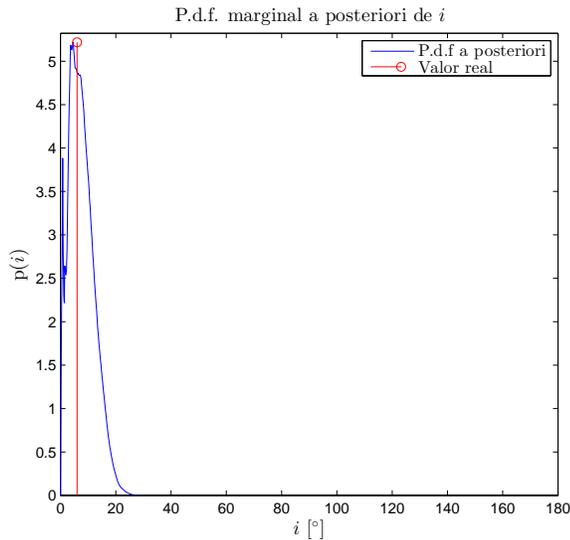
Las Figura 4.8 contiene aproximaciones de la distribución marginal *a posteriori* para los parámetros identificados como *conocidos* o *semi-fijos*:  $\mathcal{T}$ ,  $P$ ,  $e$  y  $a$ . Estas aproximaciones se obtuvieron luego de la ejecución del Filtro de Partículas sin reducción de dimensionalidad y fueron calculadas a través de la aplicación de un *kernel* de tipo Epanechnikov [Epanechnikov, 1969] a las partículas de la iteración final del filtro, con sus respectivos pesos  $w$



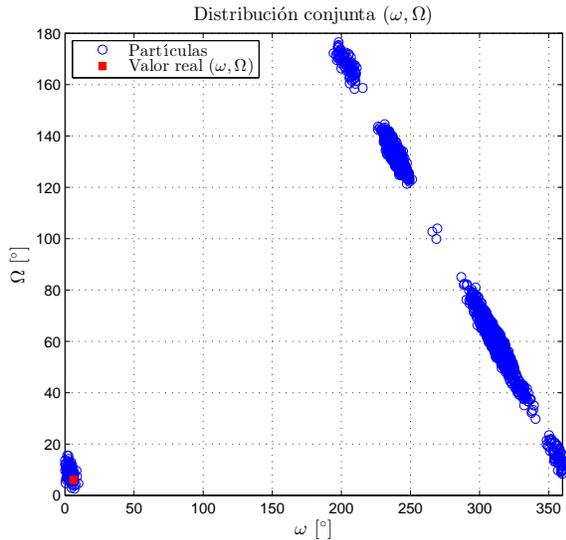
(a) Argumento del periastro



(b) Longitud del nodo ascendente



(c) Inclínación orbital



(d) Distribución conjunta  $(\omega, \Omega)$

Figura 4.9: Distribuciones marginales *a posteriori* de parámetros *explorados*

Las Figura 4.9 contiene aproximaciones de la distribución marginal *a posteriori* de los parámetros identificados *explorados*  $(\omega, \Omega, i)$ , además de una aproximación de la distribución conjunta *a posteriori* de  $\omega$  y  $\Omega$ . Las distribuciones marginales individuales se obtuvieron de la misma manera que las presentadas en la Figura 4.8, mientras que la distribución conjunta de la Subfigura 4.9d fue obtenida como una proyección de las partículas de la última iteración del filtro sobre el espacio bidimensional  $(\omega, \Omega)$  (por simplicidad, ya que la aplicación de un *kernel* en espacios de más de una dimensión es más difícil).

En la Figura 4.8 se observa que la estimación de los parámetros  $\mathcal{T}$ ,  $P$ ,  $e$  y  $a$  alcanza considerable exactitud y precisión. Esto resulta esperable, puesto que su distribución *a priori* acota estos parámetros a una región determinada. Las distribuciones presentadas por Figura 4.9, por el contrario, presentan gran dispersión: en el caso de  $\omega$  y  $\Omega$ , existen múltiples *peaks*; el parámetro  $i$ , en cambio, está acotado a una región continua, pero dicha región es bastante extensa (entre  $0^\circ$  y  $20^\circ$ , aproximadamente). La Subfigura 4.9d permite identificar las combinaciones de  $\omega$  y  $\Omega$  que son capaces de explicar las posiciones observadas (notar que la combinación real está dentro de una región factible). Cabe mencionar que dicha gráfica corresponde a una proyección y deja fuera las correlaciones existentes con resto de los parámetros: es posible que los valores de  $P$  o  $i$  (por ejemplo) asociados a las partículas que se encuentran en el extremo inferior izquierdo de la Subfigura 4.9d tiendan a ser distintos a los de partículas pertenecientes a otras regiones identificadas como factibles.

La identificación de regiones de factibilidad facilita la simulación de diferentes escenarios posibles –cada uno con diferentes supuestos acerca de la órbita– y, a partir de ello, estudiar qué tipo de datos –astrométricos o de velocidad radial– y en qué instantes de observación serían necesarios para descartar soluciones factibles, dada la información disponible, y acercarse a una solución única. Es posible, por ejemplo, que en ciertos escenarios la obtención de observaciones astrométricas durante un período adicional sea suficiente para descartar un conjunto de configuraciones factibles; mientras que en otros casos la ambigüedad sea propia de la órbita y que, independientemente de la calidad de los datos astrométricos, la única manera de encontrar una solución definitiva sea complementar la información con datos de velocidad radial<sup>6</sup>.

---

<sup>6</sup>Tener en cuenta que la ecuación de velocidad radial,  $K(\cos(\omega+\nu)+e\cos\omega)$ , incluye algunos de los parámetros orbitales estudiados en este trabajo, por lo que añade una restricción.

# Capítulo 5

## Conclusiones

En este trabajo se propone y evalúa una metodología para estimar los elementos orbitales de una estrella binaria visual de manera automática. La herramienta propuesta utiliza un enfoque Bayesiano, basándose específicamente en el algoritmo de Filtro de Partículas, en combinación con la idea de Evolución Artificial de Parámetros. Adicionalmente, se aprovechan las propiedades de la representación de Thiele-Innes para reducir la dimensionalidad del problema.

El problema estudiado tiene más de un siglo de antigüedad, y aunque varios métodos heurísticos han sido utilizados para su tratamiento ([Pourbaix, 1994] y [Schaefer *et al.*, 2006] utilizan *simulate annealing* y una heurística de Monte Carlo, respectivamente), el enfoque Bayesiano sólo ha sido propuesto en años recientes [Lucy, 2014]. Este enfoque tiene sólidos fundamentos teóricos, lo que supone una ventaja por sobre los métodos heurísticos.

La aplicación de este método es extensible a cualquier conjunto de cuerpos celestes ligados gravitacionalmente. En particular, se espera que el avance de las tecnologías de observación astronómica facilite el descubrimiento de planetas extrasolares a través de datos astrométricos (hasta ahora la mayoría de los exoplanetas han sido descubiertos a través de la técnicas de velocidad radial o de tránsitos). Ya sea por el descubrimiento de nuevas binarias visuales o de exoplanetas, las herramientas diseñadas para la caracterización de órbitas –como la presentada en este trabajo– son relevantes para la astronomía, pese a la antigüedad del problema.

De acuerdo a los resultados, el algoritmo propuesto en este trabajo puede, bajo algunos supuestos, lidiar cómodamente con varios problemas presentes en la observación astronómica:

muestreo irregular, incompletitud de órbita, errores observacionales y variabilidad en la calidad de la información. Esto queda demostrado con los resultados presentados en las Secciones 4.1, 4.2, donde en la mayoría de los casos los parámetros son estimados de manera exacta y precisa (mejor a medida que aumenta la cobertura de órbita y disminuye el error de observación). Se detectaron, sin embargo, situaciones en que el método falla en la estimación de algunos parámetros: aunque las configuraciones orbitales encontradas por el algoritmo expliquen de manera relativamente exitosa las observaciones disponibles, no coinciden con los valores reales. Al respecto, se identificaron dos principales fuentes de ambigüedad:

- Existencia de órbitas de diversa naturaleza (i.e., con variados valores de  $P$ ,  $e$  y  $a$ ) que se ajustan a las observaciones disponibles. Este caso suele ocurrir cuando la cobertura de la órbita es muy pobre (una fracción muy pequeña del período real) y se agrava cuando las observaciones no cubren el tiempo de paso por el periastro. Un caso particular de este fenómeno es presentado por la Figura 4.3 en la Sección 4.2.
- Ambigüedad de orientación: Casos donde a pesar existir una buena estimación de los parámetros más relevantes de la órbita<sup>1</sup> (excentricidad, período y largo del semi-eje mayor), existen diversas orientaciones (i.e. combinaciones de  $\omega$ ,  $\Omega$  e  $i$ ) que podrían originar la órbita aparente observada. Esto ocurre cuando ciertos parámetros se acercan a un valor límite (en particular, en este trabajo se logró detectar  $i = 0^\circ$  como un caso conflictivo). Este caso es estudiado en la Sección 4.3.

Frente a ello, este trabajo propone una extensión del método principal que levanta la restricción impuesta por la reducción de dimensionalidad y explora, a partir de cierto conocimiento previo acerca de la órbita, el espacio de soluciones, a fin de encontrar conjuntos de configuraciones orbitales factibles (Subsección 4.3.2). Una virtud de esta propuesta es que podría facilitar la decisión de cuándo y qué tipo de datos adquirir para eliminar las ambigüedades.

Finalmente, se enumeran algunas líneas de trabajo que pueden surgir a partir de lo realizado:

- Incorporación de datos de velocidad radial en el proceso de estimación (solución combinada) o, incluso, desarrollar una metodología similar a la propuesta en este trabajo basada

---

<sup>1</sup>El resto de los parámetros son meramente relativos al observador, y no determinan ni la masa ni la energía total del sistema.

exclusivamente en datos de velocidad radial (tener en cuenta que a veces son los únicos datos disponibles).

- Comparación del desempeño del algoritmo propuesto con el de otros métodos propuestos en la literatura, utilizando datos reales.
- Analizar el comportamiento del método en casos donde se rompe alguno de los supuestos simplificadorios. Por ejemplo, aplicación a sistemas estelares jerárquicos (recordar que las órbitas Keplerianas utilizadas en este trabajo son, en ciertos casos, una buena aproximación de la trayectoria, incluso si la cantidad de cuerpos presentes en el sistema no es estrictamente igual a dos).
- Extensión del método a casos donde sólo  $\rho$  o sólo  $\theta$  son conocidos en ciertos intervalos de tiempo.
- Proponer procedimientos para lidiar con la ambigüedad de cuadrante.

# Bibliografía

- [Andrieu *et al.*, 2001] Andrieu, C., Doucet, A., y Punskeya, E. (2001). Sequential monte carlo methods for optimal filtering. En *Sequential Monte Carlo Methods in Practice*, pp. 79–95. Springer.
- [Arulampalam *et al.*, 2002] Arulampalam, M. S., Maskell, S., Gordon, N., y Clapp, T. (2002). A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *Signal Processing, IEEE Transactions on*, 50(2):174–188.
- [Binnendijk, 1960] Binnendijk, L. (1960). Properties of double stars; a survey of parallaxes and orbits. *NASA Planetary Data System*, 1.
- [Candy, 2011] Candy, J. V. (2011). *Bayesian signal processing: Classical, modern and particle filtering methods*, volumen 54. John Wiley & Sons.
- [Del Moral, 2004] Del Moral, P. (2004). *Feynman-Kac Formulae: Genealogical and Interacting Particle Systems with Applications*. Springer.
- [Del Moral *et al.*, 2006] Del Moral, P., Doucet, A., y Jasra, A. (2006). Sequential monte carlo samplers. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(3):411–436.
- [Doucet *et al.*, 2001] Doucet, A., De Freitas, N., y Gordon, N. (2001). *Sequential Monte Carlo methods in practice*. Springer.
- [Doucet *et al.*, 2000] Doucet, A., Godsill, S., y Andrieu, C. (2000). On sequential monte carlo sampling methods for bayesian filtering. *Statistics and computing*, 10(3):197–208.

- [Eggen, 1967] Eggen, O. (1967). Masses of visual binary stars. *Annual Review of Astronomy and Astrophysics*, 5(1):105–138.
- [Epanechnikov, 1969] Epanechnikov, V. A. (1969). Non-parametric estimation of a multivariate probability density. *Theory of Probability & Its Applications*, 14(1):153–158.
- [Geweke, 1989] Geweke, J. (1989). Bayesian inference in econometric models using monte carlo integration. *Econometrica: Journal of the Econometric Society*, pp. 1317–1339.
- [Gordon *et al.*, 2002] Gordon, N. J., Salmond, D. J., y Smith, A. F. (2002). Novel approach to nonlinear/non-gaussian bayesian state estimation. En *IEE Proceedings F (Radar and Signal Processing)*, volumen 140, pp. 107–113. IET.
- [Green, 1985] Green, R. M. (1985). *Spherical astronomy*. Cambridge University Press.
- [Haug, 2005] Haug, A. (2005). A tutorial on bayesian estimation and tracking techniques applicable to nonlinear and non-gaussian processes. *MITRE Corporation, McLean*.
- [Kabanikhin, 2008] Kabanikhin, S. (2008). Definitions and examples of inverse and ill-posed problems. *Journal of Inverse and Ill-Posed Problems*, 16(4):317–357.
- [Kalman, 1960] Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Fluids Engineering*, 82(1):35–45.
- [Khan *et al.*, 2004] Khan, Z., Balch, T., y Dellaert, F. (2004). An mcmc-based particle filter for tracking multiple interacting targets. En *Computer Vision-ECCV 2004*, pp. 279–290. Springer.
- [Kitagawa y Sato, 2001] Kitagawa, G. y Sato, S. (2001). Monte carlo smoothing and self-organising state-space model. En *Sequential Monte Carlo methods in practice*, pp. 177–195. Springer.
- [Kong *et al.*, 1994] Kong, A., Liu, J. S., y Wong, W. H. (1994). Sequential imputations and bayesian missing data problems. *Journal of the American statistical association*, 89(425):278–288.

- [Lailly, 1983] Lailly, P. (1983). The seismic inverse problem as a sequence of before stack migrations. En *Conference on inverse scattering: theory and application*, pp. 206–220. Society for Industrial and Applied Mathematics, Philadelphia, PA.
- [Lee y Markus, 1967] Lee, E. B. y Markus, L. (1967). Foundations of optimal control theory. Technical report, DTIC Document.
- [Lindgren *et al.*, 2007] Lindgren, Lennart and Babusiaux, Carine and Bailer-Jones, C and Bastian, Ulrich and Brown, Anthony GA and Cropper, Mark and Høg, Erik and Jordi, Carme and Katz, David and Van Leeuwen, F and others (2007). The gaia mission: science, organization and present status. *Proceedings of the International Astronomical Union*, 3(S248):217–223.
- [Liu y West, 2001] Liu, J. y West, M. (2001). Combined parameter and state estimation in simulation-based filtering. En *Sequential Monte Carlo methods in practice*, pp. 197–223. Springer.
- [Liu, 1996] Liu, J. S. (1996). Metropolized independent sampling with comparisons to rejection sampling and importance sampling. *Statistics and Computing*, 6(2):113–119.
- [Lucy, 2014] Lucy, L. (2014). Mass estimates for visual binaries with incomplete orbits. *Astronomy & Astrophysics*, 563:A126.
- [Massey y Meyer, 2001] Massey, P. y Meyer, M. (2001). Stellar masses. *Encyclopedia of Astronomy and Astrophysics*, 1:1882.
- [Metropolis y Ulam, 1949] Metropolis, N. y Ulam, S. (1949). The monte carlo method. *Journal of the American statistical association*, 44(247):335–341.
- [Mundnich *et al.*, 2013] Mundnich, K., Orchard, M. E., Silva, J. F., y Parada, P. (2013). Volatility estimation of financial returns using risk-sensitive particle filters. *Studies in Informatics and Control*, 22(3):297–306.
- [Pascual-Marqui, 1999] Pascual-Marqui, R. D. (1999). Review of methods for solving the eeg inverse problem. *International journal of bioelectromagnetism*, 1(1):75–86.
- [Pourbaix, 1994] Pourbaix, D. (1994). A trial-and-error approach to the determination of the orbital parameters of visual binaries. *Astronomy and Astrophysics*, 290:682–691.

- [Roelofs *et al.*, 2010] Roelofs, G. H., Rau, A., Marsh, T. R., Steeghs, D., Groot, P. J., y Nelemans, G. (2010). Spectroscopic evidence for a 5.4 minute orbital period in hm cancri. *The Astrophysical Journal Letters*, 711(2):L138.
- [Rudy, 1987] Rudy, Y & Messinger-Rapport, B. (1987). The inverse problem in electrocardiography: solutions in terms of epicardial potentials. *Critical reviews in biomedical engineering*, 16(3):215–268.
- [Schaefer *et al.*, 2006] Schaefer, G., Simon, M., Beck, T., Nelan, E., y Prato, L. (2006). Dynamical mass estimates for incomplete orbits: Young multiple stars in taurus and ophiuchus. *The Astronomical Journal*, 132(6):2618.
- [Tarantola, 2005] Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*. Society for Industrial and Applied Mathematics.
- [Van de Kamp y Stearns, 1967] Van de Kamp, P. y Stearns, C. L. (1967). Principles of astrometry. *American Journal of Physics*, 35(10):974–975.
- [Wunsch, 1996] Wunsch, C. (1996). *The ocean circulation inverse problem*. Cambridge University Press.

# Apéndice A

## Ecuación de movimiento orbital

En este apéndice se explica paso a paso la derivación de la solución del problema de los dos cuerpos, para finalmente deducir las Leyes de Kepler a partir de los resultados analíticos obtenidos. El punto de partida es la ecuación diferencial de posición relativa presentada en la Sección 2.1.1:

$$\ddot{\vec{r}} = -\frac{G(m_1 + m_2)}{r^3}\vec{r} \quad (\text{A.1})$$

Considerar, por brevedad,  $\mu = G(m_1 + m_2)$ , con lo que la ecuación queda expresada como:

$$\ddot{\vec{r}} = -\frac{\mu}{r^3}\vec{r} \quad (\text{A.2})$$

Todas las conclusiones importantes sobre el problema de los cuerpos surgen de esta ecuación diferencial de segundo orden. Dado que la posición relativa y su segunda derivada comparten la misma dirección, el producto cruz entre ambos vectores es nulo:

$$\vec{r} \times \ddot{\vec{r}} = 0 \quad (\text{A.3})$$

Al integrar esta igualdad, se llega a que:

$$\vec{r} \times \dot{\vec{r}} = \vec{h} \quad (\text{A.4})$$

Notar que  $\vec{h}$  es un vector constante proporcional a  $\vec{L}$ , el momento angular total del sistema:

$$\begin{aligned}
 \vec{L} &= \vec{r} \times \vec{p} \\
 &= \vec{r} \times \frac{m_1 \cdot m_2}{m_1 + m_2} \dot{\vec{r}} \\
 &= \frac{m_1 \cdot m_2}{m_1 + m_2} \vec{r} \times \dot{\vec{r}} \\
 &= \frac{m_1 \cdot m_2}{m_1 + m_2} \vec{h}
 \end{aligned}
 \tag{A.5}$$

El valor  $\vec{h}$  recibe el nombre de *momento angular específico*. De las ecuaciones A.4, A.5 se desprende que el movimiento orbital está confinado a un plano, pues el momento angular  $\vec{L}$  es perpendicular al movimiento en todo instante ( $\vec{r} \cdot \vec{L} = 0$ ).

Considerar el producto punto de la ecuación A.2 con la primera derivada de  $\vec{r}$  (el factor 2 se agrega para facilitar un resultado posterior):

$$2 \dot{\vec{r}} \cdot \ddot{\vec{r}} = -2 \frac{\mu}{r^3} \dot{\vec{r}} \cdot \vec{r} \tag{A.6}$$

Como  $d(\dot{r}^2)/dt = 2\dot{\vec{r}} \cdot \ddot{\vec{r}}$ , la primitiva del lado izquierdo de la ecuación es  $\dot{r}^2$ . Por otro lado, dado que  $\vec{r} \cdot \dot{\vec{r}} = r\dot{r}$  (leer Demostración), se tiene la siguiente igualdad para el lado derecho:

$$\begin{aligned}
 -2 \frac{\mu}{r^3} \dot{\vec{r}} \cdot \vec{r} &= -2 \frac{\mu}{r^3} r\dot{r} \\
 &= -2 \frac{\mu}{r^2} \dot{r}
 \end{aligned}
 \tag{A.7}$$

Cuya primitiva es  $2\frac{\mu}{r}$ . Con estos resultados, se concluye que la integral de la ecuación A.6 es:

$$\dot{r}^2 = 2\frac{\mu}{r} + C \tag{A.8}$$

Donde  $C$  es la constante de integración. El significado físico de este valor se revisará posteriormente.

### Demostración $\vec{r} \cdot \dot{\vec{r}} = r\dot{r}$

Considerar  $\vec{r}$  y  $\dot{\vec{r}}$  en coordenadas cartesianas:

$$\begin{aligned}\vec{r} &= (r \cos \theta, r \sin \theta, 0) \\ \dot{\vec{r}} &= (\dot{r} \cos \theta - r \sin \theta \dot{\theta}, \dot{r} \sin \theta + r \cos \theta \dot{\theta}, 0)\end{aligned}\tag{A.9}$$

El hecho de que el movimiento orbital esté relegado a un plano permite tomar dicho plano como plano de referencia (ejes  $x$  e  $y$ ) y el vector perpendicular a éste como eje  $z$  del sistema de coordenadas, siendo los valores de  $\vec{r}$  y  $\dot{\vec{r}}$  en dicho eje nulos. Calculando el producto punto de los valores definidos anteriormente:

$$\begin{aligned}\vec{r} \cdot \dot{\vec{r}} &= r\dot{r} \cos^2 \theta - r^2 \cos \theta \sin \theta \dot{\theta} + \\ &\quad r\dot{r} \sin^2 \theta + r^2 \cos \theta \sin \theta \dot{\theta} \\ &= r\dot{r}\end{aligned}$$

Con lo que se concluye que  $\vec{r} \cdot \dot{\vec{r}} = r\dot{r}$ .

Con la representación adoptada en la ecuación A.9, el cálculo de la ecuación A.4 conduce a:

$$r^2 \dot{\theta} = \|\vec{h}\| = h\tag{A.10}$$

Mientras que la ecuación A.8 toma la forma:

$$\dot{r}^2 + r^2 \dot{\theta}^2 = 2\frac{\mu}{r} + C\tag{A.11}$$

Que junto con el resultado de la ecuación A.10, puede ser escrita de manera equivalente:

$$\dot{r}^2 + \frac{h^2}{r^2} = 2\frac{\mu}{r} + C\tag{A.12}$$

Esta representación elimina la variable angular  $\theta$  de la ecuación, permitiendo expresar la ecuación de movimiento exclusivamente en función de  $r$ .

## A.1. Ecuación de Kepler para órbitas ligadas

La constante  $C$  es proporcional a la energía total del sistema. Valores positivos de  $C$  corresponden a órbitas no ligadas, i.e. órbitas donde la energía total es suficiente para permitir que los dos cuerpos escapen de la influencia gravitacional del otro. En esos casos, la órbita relativa es una hipérbola. Se analizará el caso donde  $C$  es negativo. Sin pérdida de generalidad,  $C$  puede escribirse como:

$$C = -\frac{\mu}{a} \quad (\text{A.13})$$

A pesar de que  $a$  puede verse como un simple valor auxiliar para expresar una constante,  $C$ , en términos de  $\mu$ , dicho valor posee un significado geométrico, el cual se explica posteriormente. Con esta expresión para  $C$ , la ecuación A.12 puede reescribirse como:

$$\dot{r}^2 + h^2/r^2 = \mu(2/r - 1/a) \quad (\text{A.14})$$

Este valor debe ser siempre positivo, por lo que  $r$  debe ser menor a  $2a$ . Esto implica que la órbita está ligada (ningún cuerpo se *escapa*). Incluso en esta forma (aún sin solución para  $r$  en función de  $t$ ), la ecuación A.14 es un resultado importante, pues ofrece una fórmula directa para calcular la velocidad instantánea si se conoce la posición en dicho instante. Además, la dirección de velocidad de la estrella secundaria puede ser obtenida de la ecuación A.10, que entrega la componente transversal de la velocidad como:

$$V_T = r\dot{\theta} = \frac{h}{r} \quad (\text{A.15})$$

Un problema más difícil es la derivación de la posición de la secundaria para todo instante  $t$ . Dicho problema es abordado a continuación.

Sea  $\omega$  el ángulo en que la separación  $r$  alcanza su menor valor y  $T$  el tiempo en que  $r$  alcanza ese valor.  $\omega$  recibe el nombre de *argumento del periastro*, mientras que  $T$  se conoce como *tiempo de paso por el periastro*. Se define  $\nu$ , conocida como *anomalía verdadera*, como  $\nu = \theta - \omega$ . Reescribiendo la ecuación A.14 como:

$$\dot{r}^2 = -\frac{\mu}{ar^2}(r^2 - 2ar + h^2\frac{a}{\mu}) \quad (\text{A.16})$$

Por brevedad, se define  $F(r) = (r^2 - 2ar + h^2 \frac{a}{\mu})$ . Para cumplir con la igualdad, el valor de  $F(r)$  debe menor o igual que 0. Sean  $R_1$  y  $R_2$  las dos raíces del polinomio, con  $R_2 \geq R_1$ . Se puede argumentar que para que el polinomio cumpla la desigualdad establecida, el valor de  $r$  debe cumplir que  $R_1 \leq r \leq R_2$ .

Dada una forma cuadrática  $p(x) = ax^2 + bx + c$ , sus raíces  $x_1, x_2$  tienen la forma  $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ , de modo que se cumple que  $x_1 + x_2 = -b/a$  y  $x_1 \cdot x_2 = c/a$ . En este caso, se llega a que:

$$R_1 + R_2 = 2a \quad (\text{A.17})$$

$$R_1 \cdot R_2 = \frac{h^2 a}{\mu} \quad (\text{A.18})$$

Se puede argumentar, una vez presentada la solución (elíptica) del problema de los dos cuerpos, que  $R_1$  y  $R_2$  corresponden al periastro y apoastro, respectivamente. Se define el parámetro  $e$ , identificado posteriormente como la excentricidad orbital, como:

$$e = \frac{R_2 - R_1}{R_2 + R_1} \quad (\text{A.19})$$

Lo que, haciendo uso de la ecuación A.17, permite escribir las raíces como:

$$\begin{aligned} R_1 &= a(1 - e) \\ R_2 &= a(1 + e) \end{aligned} \quad (\text{A.20})$$

Esto conduce, en combinación con la ecuación A.18, a una importante relación para el momento angular específico:

$$h^2 = \mu a(1 - e^2) \quad (\text{A.21})$$

Asumiendo que  $\vec{r}$  describe una trayectoria elíptica, se define la *circunferencia auxiliar* como una circunferencia de radio igual al parámetro  $a$ , centrada en el mismo punto que la elipse. De esta manera, se propone la siguiente parametrización de  $r$ :

$$r = a(1 - e \cos E) \quad (\text{A.22})$$

Donde la variable  $E$  o *anomalía excéntrica* corresponde al ángulo medido desde el centro de la circunferencia auxiliar, que forma la proyección de la estrella secundaria (en el sentido de la línea perpendicular al eje principal), y el eje principal de la elipse (ver Figura 2.1). El valor de  $E$  es nulo en el periastro e incrementa hasta  $2\pi$  a medida que recorre la órbita. Con este cambio de variables, el polinomio cuadrático de la ecuación A.16 toma la forma:

$$F(r) = -a^2 e^2 \sin^2 E \quad (\text{A.23})$$

De modo que la ecuación A.16 puede ser reescrita de la siguiente manera:

$$\begin{aligned} \dot{r}^2 &= -\frac{\mu}{ar^2} \cdot F(r) / \sqrt{(\quad)} \\ \frac{dr}{dt} &= \sqrt{-\frac{\mu}{ar^2} \cdot F(r)} \\ \frac{d(a(1 - e \cos E))}{dt} &= \sqrt{\frac{\mu}{a}} \cdot \frac{1}{a(1 - e \cos E)} \cdot a e \sin E \\ ae \sin E \frac{dE}{dt} &= \sqrt{\frac{\mu}{a}} \cdot \frac{1}{a(1 - e \cos E)} \cdot a e \sin E \end{aligned}$$

La última igualdad puede ser reordenada como:

$$(1 - e \cos E)dE = \sqrt{\frac{\mu}{a^3}} dt \quad (\text{A.24})$$

Se define la variable  $n$ , llamada *movimiento medio*, de acuerdo a:

$$n^2 a^3 = \mu \quad (\text{A.25})$$

Recordando  $E$  debe ser nulo en el periastro, al integrar a ambos lados la ecuación A.24 resulta:

$$M \equiv n \cdot (t - T) = E - e \sin E \quad (\text{A.26})$$

Donde la variable  $M$  es conocida como *anomalía media* y  $T$  corresponde al *tiempo de paso por el periastro*. Este importante resultado es conocido como Ecuación de Kepler. Notar que esta

ecuación no admite una solución analítica, sin embargo, es posible aproximar  $E$  dado  $t$  a través de métodos numéricos.

Hasta este punto, tanto  $r$  como  $t$  pueden expresarse en términos de  $E$ , y lo mismo es posible para la anomalía verdadera  $\nu$ . A continuación, se deriva una expresión que relaciona  $\nu$  con  $E$ : Dado que  $\nu = \theta - \omega$ , la ecuación A.10 puede ser expresada como  $r^2\dot{\nu} = h$ , que a su vez, mediante las expresiones A.21, A.22, A.26, puede ser escrita como:

$$\frac{d\nu}{dE} = \frac{\sqrt{1-e^2}}{1-e\cos E} \quad (\text{A.27})$$

Aplicando el cambio de variable a  $\tan \frac{1}{2}E$  a la formulación integral de la ecuación A.27, puede verificarse que:

$$\int \frac{\sqrt{1-e^2}}{1-e\cos E} dE = 2 \tan^{-1} \left[ \sqrt{\frac{1+e}{1-e}} \tan \frac{1}{2}E \right] \quad (\text{A.28})$$

Puesto que tanto  $\nu$  como  $E$  son iguales a 0 en el periastro, se concluye que:

$$\tan \frac{1}{2}\nu = \sqrt{\frac{1+e}{1-e}} \tan \frac{1}{2}E \quad (\text{A.29})$$

La importancia de la Ecuación de Kepler radica que permite conocer, en conjunto con la ecuación A.29, la posición relativa  $(r, \nu)$  de la estrella secundaria para cualquier instante  $t$ .

## A.2. Leyes de Kepler

Las leyes de Kepler pueden considerarse una formulación geométrica de los resultados analíticos obtenidos previamente. A continuación se derivan las Tres Leyes de Kepler a partir dichos resultados:

- **Primera Ley:** *Los cuerpos se desplazan alrededor de la estrella principal describiendo órbitas elípticas.*

Considerar el proceso de integración que conduce a la solución paramétrica de  $\nu$  (ecuación

A.29). Dicha procedimiento se puede escribir como:

$$\frac{d\nu}{dE} = \frac{\sqrt{1-e^2}}{1-e\cos E} \Leftrightarrow \tan \frac{1}{2}\nu = \sqrt{\frac{1+e}{1-e}} \tan \frac{1}{2}E \quad (\text{A.30})$$

Invirtiendo los roles de  $\nu$  y  $E$  en el proceso de integración, se llega a:

$$\tan \frac{1}{2}E = \sqrt{\frac{1-e}{1+e}} \tan \frac{1}{2}\nu \Rightarrow \frac{dE}{d\nu} = \frac{\sqrt{1-e^2}}{1+e\cos \nu} \quad (\text{A.31})$$

Usando las ecuaciones A.30 y A.31:

$$\frac{d\nu}{dE} \cdot \frac{dE}{d\nu} = 1 = \frac{1-e^2}{(1-e\cos E)(1+e\cos \nu)} \quad (\text{A.32})$$

Usando la parametrización  $r = r(E)$  presentada en la ecuación A.22, se concluye que:

$$r = \frac{a(1-e^2)}{1+e\cos \nu} \quad (\text{A.33})$$

Que puede ser reconocida como la ecuación estándar de una elipse en coordenadas polares, en concordancia con la Primera Ley de Kepler. Con esto, es posible identificar, finalmente, a  $\nu$  como la anomalía excéntrica, a  $e$  como la excentricidad y  $a$  como el semi-eje mayor.

- **Segunda Ley:** *La estrella secundaria se mueve con velocidad areolar constante.*

Considerar el instante  $t_0$  y un pequeño intervalo de tiempo  $dt$ . El área descrita durante ese intervalo por el vector radial  $\vec{r}$  es aproximadamente el triángulo definido por el centro de la elipse, la posición del cuerpo secundario en  $t_0$  y la posición del mismo en  $t_0 + dt$ . La base de dicho triángulo es  $r(t)$ , y su altura igual a la velocidad transversal multiplicada por el tiempo transcurrido:  $V_t \cdot dt$ . De este modo el área resulta ser  $\frac{1}{2}rV_t dt$ . Al calcular el área descrita por unidad de tiempo, se obtiene la expresión  $\frac{1}{2}rV_t$ , la cual de acuerdo a la ecuación A.15 resulta ser constante e igual a  $\frac{1}{2}h$ . Esto justifica la Segunda Ley de Kepler.

- **Tercera Ley:** *El cociente entre el cubo del semi-eje mayor y el cuadrado del período es proporcional a la masa de la estrella principal.*

La Tercera Ley está contenida en la definición del movimiento medio en la ecuación A.25. Dado que  $n$  es la tasa de incremento constante de la anomalía media  $M$ , es también la tasa

de incremento constante de la anomalía excéntrica y la anomalía verdadera. Como cada uno de estos valores aumenta  $2\pi$  en un período orbital  $P$ , se tiene que:

$$n = \frac{2\pi}{P} \quad (\text{A.34})$$

Como consecuencia, usando la definición de  $\mu$ , la ecuación A.25 puede ser reformulada como:

$$\frac{a^3}{P^2} = \frac{G}{4\pi^2}(m_1 + m_2) \quad (\text{A.35})$$

Este resultado relaciona las masas de ambas estrellas con el período de revolución y el tamaño de su órbita relativa, siendo la base de la determinación de masas en Astronomía. Notar que, a diferencia de la Primera y la Segunda Ley de Kepler, la formulación original de esta ley es sólo aproximadamente verdadera, puesto que sólo toma en cuenta la masa del cuerpo principal (que es una buena aproximación sólo cuando el cuerpo secundario es un planeta o una estrella de masa muy inferior a la primaria). La ecuación A.35, en cambio, entrega una relación válida para todos los casos.

# Apéndice B

## Constantes de Thiele-Innes

### B.1. Estimador de mínimos cuadrados

En este apéndice se explica cómo obtener un estimador de mínimos cuadrados de las constantes de Thiele-Innes a partir de un conjunto de observaciones  $\{X_{obs}(i), Y_{obs}(i)\}_{i=1, \dots, N}$  y valores teóricos de  $T$ ,  $P$  y  $e$  (los definidos por la partícula  $k$  en la iteración  $j$  del Filtro de Partículas). Considerar la meta-observación  $\mathcal{Y}$  definida en la Sección 3.3 como un error cuadrático medio de las coordenadas computadas  $(X_{comp}, Y_{comp})$  respecto a las observaciones reales en el plano del cielo  $(X_{obs}, Y_{obs})$ :

$$\mathcal{Y} = MSE = \sum_{i=1}^N w_i \cdot [X_{obs}(i) - X_{comp}(i)]^2 + \sum_{i=1}^N w_i \cdot [Y_{obs}(i) - Y_{comp}(i)]^2 \quad (\text{B.1})$$

Donde, para evitar el exceso de notación, se omitieron los índices  $j$  y  $k$ . De acuerdo a la Sección 2.2, dados  $T$ ,  $P$  y  $e$  y un instante de tiempo –en este caso, los instantes  $\tau_i$  de observación–, es posible calcular las *coordenadas normalizadas*  $x$  e  $y$  de acuerdo a la ecuación 2.9. Con las coordenadas normalizadas y la relaciones definidas en la ecuación 2.11, las diferencias entre la observación y el valor computado quedan descritas por:

$$\begin{aligned} X_{obs}(i) - X_{comp}(i) &= X_{obs}(i) - [B \cdot x(i) + G \cdot y(i)] \\ Y_{obs}(i) - Y_{comp}(i) &= Y_{obs}(i) - [A \cdot x(i) + F \cdot y(i)] \end{aligned} \quad (\text{B.2})$$

Dada la dependencia lineal de los valores teóricos  $X_{comp}$  e  $Y_{comp}$  respecto a las coordenadas normalizadas, es posible calcular el estimador de mínimos cuadrados para las incógnitas  $B$ ,  $G$ ,  $A$ ,  $F$  de manera no iterativa. Además, el primer término de la ecuación B.1 depende sólo del par  $(B, G)$ , mientras que el segundo de  $(A, F)$ . Por lo tanto, los estimadores de  $(B, G)$  se obtienen al minimizar el primer término y los de  $(A, F)$  al minimizar el segundo, de manera que el problema se reduce a resolver dos pares de ecuaciones lineales independientes, cada una con dos incógnitas. Sin detallar los pasos de la solución del conjunto de ecuaciones –que involucran la derivación de la ecuación B.1 con respecto a las constantes de Thiele-Innes–, a continuación se muestra el conjunto de fórmulas resultante. Se definen, por brevedad, los siguientes valores:

$$\begin{aligned}
\alpha &= \sum_i w_i x(i)^2 & \beta &= \sum_i w_i y(i)^2 & \gamma &= \sum_i w_i x(i) y(i) \\
r_{11} &= \sum_i w_i X_{obs}(i) x(i) & r_{12} &= \sum_i w_i X_{obs}(i) y(i) \\
r_{21} &= \sum_i w_i Y_{obs}(i) x(i) & r_{22} &= \sum_i w_i Y_{obs}(i) y(i)
\end{aligned} \tag{B.3}$$

De modo que estimadores de las constantes de Thiele-Innes quedan definidos por:

$$\begin{aligned}
\hat{B} &= \frac{\beta \cdot r_{11} - \gamma \cdot r_{12}}{\Delta} & \hat{G} &= \frac{\alpha \cdot r_{12} - \gamma \cdot r_{11}}{\Delta} \\
\hat{A} &= \frac{\beta \cdot r_{21} - \gamma \cdot r_{22}}{\Delta} & \hat{F} &= \frac{\alpha \cdot r_{22} - \gamma \cdot r_{21}}{\Delta}
\end{aligned} \tag{B.4}$$

Donde  $\Delta = \alpha \cdot \beta - \gamma^2$ .

## B.2. Conversión de Thiele-Innes a Campbell

Una vez que obtenidos los estimadores de mínimos cuadrados  $(\hat{B}, \hat{G}, \hat{A}, \hat{F})$ , es necesario recuperar los valores de Campbell. Para  $\omega$  y  $\Omega$ , se debe resolver el siguiente par de ecuaciones:

$$\begin{aligned}
\omega + \Omega &= \arctan\left(\frac{B - F}{A + G}\right) \\
\omega - \Omega &= \arctan\left(\frac{-B - F}{A - G}\right)
\end{aligned} \tag{B.5}$$

De manera que  $\sin \omega + \Omega$  tenga el mismo signo que  $B - F$  y  $\sin \omega - \Omega$  el mismo signo que  $-B - F$ . El valor resultante de  $\Omega$  podría violar la convención de que  $\Omega \in (0, \pi)$ . Por esta razón, si  $\Omega < 0$ , se reasignan los valores de  $\omega$  y  $\Omega$  de acuerdo a  $\omega = \pi + \omega$  y  $\Omega = \pi + \Omega$ . Si  $\Omega > \pi$ , se reasigna  $\omega$  y  $\Omega$  de acuerdo a  $\omega = \omega - \pi$  y  $\Omega = \Omega - \pi$ .

Para el semi-eje mayor  $a$  y la inclinación  $i$ , se calculan las siguientes variables auxiliares:

$$\begin{aligned}
 k &= \frac{A^2 + B^2 + F^2 + G^2}{2} \\
 m &= A \cdot G - B \cdot F \\
 j &= \sqrt{k^2 - m^2}
 \end{aligned}
 \tag{B.6}$$

Con lo que  $a$  e  $i$  quedan determinados por:

$$\begin{aligned}
 a &= \sqrt{j + k} \\
 i &= \arccos \left( \frac{m}{a^2} \right)
 \end{aligned}
 \tag{B.7}$$