



UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

"INTEGRADORES SIMPLECTICOS BASADOS EN FUNCIONES GENERATRICES"

T E S I S

QUE PARA OBTENER EL TITULO DE:

A C T U A R I O

P R E S E N T A :

ABRAHAM DE LA ROSA IBARRA



FACULTAD DE CIENCIAS UNAM

DIRECTOR DE TESIS: DR. ARTURO OLVERA CHAVEZ

2005



FACULTAD DE CIENCIAS SECCION ESCOLAR

m. 345099



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

ACT. MAURICIO AGUILAR GONZÁLEZ
Jefe de la División de Estudios Profesionales de la
Facultad de Ciencias
Presente

Comunicamos a usted que hemos revisado el trabajo escrito:

"Integradores simplécticos basados en funciones generatrices"

realizado por Abraham De la Rosa Ibarra

con número de cuenta 09621345-6 , quien cubrió los créditos de la carrera de:

Actuaría

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

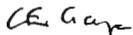
Director de Tesis
Propietario


Dr. Arturo Olvera Chávez

Propietario

Dr. Panagiotis Georgios Panagiotaros

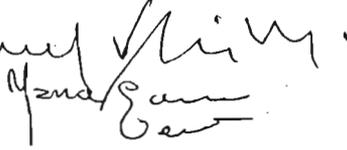
Propietario

Dra. Clara Garza Hume 

Suplente

Dr. Manuel Jesús Falconi Magaña 

Suplente

M. en C. María de Lourdes Velasco Arregui 

Consejo Departamental de Matemáticas


Act. Juan José Torres Alamillo
DE
MATEMÁTICAS

Integradores simplécticos basados en funciones generatrices

Abraham De la Rosa Ibarra

Agradecimientos

Agradezco a mis padres por su apoyo durante toda mi carrera y sin el cual no habría podido completar mis estudios de licenciatura con la realización de esta tesis.

Agradezco al Dr. Arturo Olvera por su dirección, dedicación y paciencia para explicarme durante todo el tiempo que llevó hacer la tesis.

Agradezco a los miembros de mi honorable jurado: Dra. Clara Garza, Dr. Panagiotis Georgios Panagiotaros, M. en C. Ma. de Lourdes Velasco y el Dr. Manuel Falconi por sus correcciones e imprescindibles observaciones.

Agradezco al proyecto de CONACyT G25427-E “Matemáticas no lineales en la física y en la ingeniería” por la beca que me fue otorgada para terminar mi trabajo de tesis.

Abraham De la Rosa Ibarra.

Contenido

Agradecimientos	I
Introducción	IV
1. Métodos Numéricos Clásicos	1
1.1. Métodos de un paso	2
1.1.1. Método de Euler	4
1.1.2. Métodos Runge-Kutta	6
1.2. Métodos multipaso	12
1.2.1. Métodos lineales multipaso	18
1.2.2. Consideraciones sobre los métodos multipaso	28
1.3. Sistemas de ecuaciones	29
2. Sistemas Hamiltonianos	31
2.1. Nociones de cálculo de variaciones	31
2.2. Ecuación de Euler-Lagrange	35
2.3. Coordenadas canónicas	40
2.3.1. Transformada de Legendre	42
2.4. Transformaciones canónicas	43
2.4.1. Ejemplos de transformaciones canónicas	46
2.5. Notación clásica	48
3. Integración Simpléctica	53
3.1. Simplecticidad	53
3.1.1. Transformaciones canónicas y simplecticidad	59
3.2. Integrales primeras	61
3.3. Integradores simplécticos	64
3.4. Integradores basados en función generatriz	66
3.4.1. Orden y simplecticidad del Método	80
4. Simulaciones Numéricas	83
4.1. El oscilador armónico	85
4.1.1. Resultados numéricos	85
4.2. El péndulo ideal	92

4.2.1. Resultados numéricos	93
4.3. La cadena de Toda	104
4.3.1. Antecedentes y desarrollo	104
4.3.2. Implementación numérica de los integradores simpléticos	107
4.3.3. Resultados numéricos	109
Conclusiones	121

Introducción

En el desarrollo de la Mecánica Clásica, ha habido una gran necesidad de resolver las ecuaciones de Newton cuando las fuerzas involucradas son del tipo no lineal. Desde la época de Lagrange y Laplace, los científicos han desarrollado métodos asintóticos y numéricos para obtener soluciones aproximadas de problemas mecánicos que han sido de gran interés para físicos, matemáticos, ingenieros, etc. La necesidad de desarrollar tablas de efemérides astronómicas confiables ha sido una prioridad de muchas naciones a partir del siglo XVII hasta nuestro tiempo, la navegación se hizo confiable gracias a la capacidad que tuvieron los capitanes de barcos y aviones de saber su posición geográfica con mucha precisión y esto se ha logrado con la ayuda del cálculo eficiente de efemérides astronómicas.

Los modelos mecánicos del movimiento de los cuerpos celestes son del tipo conservativo, prácticamente no existe disipación de energía a las escalas de los cuerpos celestes. La Mecánica de cuerpos que conservan energía fue desarrollada ampliamente por Lagrange. La formulación de la Mecánica a partir de principios variacionales dio como consecuencia la aparición de la Mecánica Lagrangiana al final del siglo XVIII. En el siglo XIX, Hamilton introdujo una visión geométrica a la Mecánica Lagrangiana; esta forma de estructurar la Mecánica es lo que ahora conocemos como Mecánica Hamiltoniana, la cual combina la geometría simpléctica con la dinámica de cuerpos que conservan un cierto número de integrales de movimiento durante su evolución. Esta forma de entender la dinámica de los sistemas mecánicos conservativos es muy poderosa y ha sido la herramienta básica para el desarrollo de los viajes espaciales, la colocación de todo tipo de satélites artificiales alrededor de nuestra Tierra así como para alcanzar los últimos planetas del Sistema Solar con naves automáticas que navegan prácticamente sin combustible. En todos los casos ha sido la Mecánica Hamiltoniana la herramienta indispensable para el planteamiento de misiones espaciales y en la determinación de efemérides muy precisas del movimiento de los cuerpos en el Sistema Solar.

La Mecánica Hamiltoniana no sólo se ha utilizado en las escalas en las que trabaja la Mecánica Celeste; el estudio de modelos hidrodinámicos donde no se consideran pérdidas por fricción pueden ser representados como sistemas Hamiltonianos, los cuales dan una muy buena aproximación de la dinámica de fluidos donde hay poca vorticidad. A mediados del siglo XX, el desarrollo de los reactores de fusión llevó a la necesidad de estudiar la dinámica de plasmas

y de partículas cargadas que siguen líneas de campo magnético. Este tipo de problemas se puede plantear como problemas Hamiltonianos y tiene un estrecho acercamiento a los problemas de la Mecánica Celeste. Esta área del conocimiento estimuló fuertemente el desarrollo de los sistemas Hamiltonianos junto con los Sistemas dinámicos.

Es claro entonces que desde los primeros desarrollos de la Mecánica Celeste ha habido una gran necesidad de resolver las ecuaciones que plantea la Mecánica Hamiltoniana desde un punto de vista numérico. Desde finales del siglo XVIII se han desarrollado métodos asintóticos y numéricos para tener una buena aproximación de las soluciones de los sistemas Hamiltonianos. Es a mediados del siglo XX con el advenimiento de la computadora cuando se ha tenido un gran desarrollo de los métodos numéricos para obtener soluciones confiables de los modelos mecánicos que preservan sus integrales primeras. La carrera espacial fue especialmente fructífera en el desarrollo de nuevos modelos numéricos para calcular misiones espaciales con un alto grado de precisión. Sin embargo, el cálculo numérico de las soluciones de sistemas Hamiltonianos ha enfrentado el problema de asegurar la preservación de las constantes de movimiento; la estructura simpléctica del sistema debe preservarse de manera homogénea respecto al error generado por los pasos de integración numéricos. En los últimos veinte años han aparecido un gran número de métodos numéricos cuyo propósito es obtener soluciones numéricas que preserven la estructura simpléctica de forma adecuada. El propósito de esta tesis es el estudiar y examinar un método numérico simpléctico que preserve las integrales de movimiento de forma homogénea respecto al paso de integración, este integrador está basado en la construcción asintótica de una función generatriz que permite obtener una transformación canónica que mapea las nuevas coordenadas sobre el flujo Hamiltoniano.

Aun cuando el tema central de esta tesis es la integración numérica simpléctica de sistemas Hamiltonianos, no se comienza con el estudio de éstos de manera inmediata pues el trabajo se ha estructurado de forma tal que todo el material necesario para comprender cómo surgen este tipo de sistemas y las distintas formas de integrarlos esté contenido en la misma tesis. Claro está que conocimientos de cálculo diferencial e integral, ecuaciones diferenciales ordinarias y álgebra lineal son supuestos.

La estructura que se eligió para introducir al lector en el tema de la integración simpléctica es la siguiente, en el capítulo 1 se dan los elementos clásicos de integración numérica de sistemas de ecuaciones ordinarias en general, después en el capítulo 2 se presenta un tipo particular de sistemas de ecuaciones diferenciales, los sistemas Hamiltonianos, su origen y propiedades teóricas que más tarde serán utilizadas para integrarlos numéricamente. En el capítulo 3 se prosigue con el estudio de estos sistemas, haciendo particular énfasis en cierta estructura geométrica de su dinámica, llamada estructura simpléctica. Éste es el punto central por el cual se hace necesaria una integración distinta a la clásica para estos sistemas, pues buscamos preservar la simplecticidad de su estructura a lo largo de la integración, así, en la parte final de este capítulo se desarrolla el integrador simpléctico que es objeto de esta tesis. En el capítulo 4 se hacen simulaciones numéricas con distintos métodos de integración para algunos sistemas Hamil-

tonianos clásicos con la finalidad de comparar qué tanto respetan la estructura simpléctica de la dinámica del sistema. Los integradores que se implementaron son los integradores clásicos vistos en el capítulo 1 y los desarrollados a la luz de la estructura simpléctica de estos sistemas estudiada en el capítulo 3.

En el desarrollo de la teoría a lo largo de toda la tesis se ha tomado gran parte de los teoremas, lemas y resultados en general de la bibliografía a la que hacemos referencia al final de la tesis, sin embargo es necesario decir que parte de las pruebas, explicaciones e interpretaciones de los resultados presentados no se encuentran o bien exactamente como en la bibliografía o simplemente no se encuentran.

El capítulo 1 está basado principalmente en los libros de Golub [6] y Eldén [3]. Este capítulo trata los métodos numéricos clásicos para resolver sistemas de ecuaciones diferenciales ordinarias con valor inicial. Se da la definición de método de un paso y se comienza con el estudio del método de Euler y del método de Runge-Kutta de orden general, y se ejemplifica el procedimiento mediante el cual se obtiene este último para cualquier orden. Después se da la teoría básica de interpolación de datos para desarrollar los métodos multipaso de Adams-Bashforth y de Adams-Moulton también ejemplificando como obtenerlos para orden arbitrario. A lo largo de este capítulo se dan las definiciones básicas para el cálculo del error, como son, los errores locales de discretización y el orden del método, haciendo uso de éstas para demostrar el orden de éstos métodos. Más tarde estas definiciones serán utilizadas en el análisis del error de los métodos numéricos desarrollados en el capítulo 3.

Los resultados del capítulo 2 están tomados esencialmente de los libros de Ize [8], Gelfand [4] y Goldstein [5]. Los dos primeros sirven como base para el desarrollo de la teoría básica del cálculo de variaciones. Por ejemplo las definiciones y resultados relacionados con la derivada de Gâteaux y Fréchet son tomadas del libro de Ize así como la manera de obtener las ecuaciones de Euler-Lagrange de un problema variacional; sin embargo algunas de las pruebas son un poco más parecidas a las dadas en el libro de Gelfand [4]. La parte teórica del capítulo finaliza con el estudio de transformaciones de coordenadas que hagan que las ecuaciones de Euler-Lagrange en forma canónica se preserven. Se demuestra que dichas transformaciones, llamadas canónicas, pueden ser obtenidas mediante una función, llamada función generatriz. Para la sección que trata de función generatriz, los libros de Gelfand[4] y Goldstein[5] son mayormente utilizados, siendo el de Goldstein[5] el más importante en cuanto a la presentación de los distintos tipos de funciones generatrices así como de su obtención basada en la transformada de Legendre. Esta sección es una de las más importantes, ya que de la teoría de funciones generatrices se desarrollará el integrador simpléctico estudiado en el capítulo 3 de manera teórica y prácticamente en el capítulo 4. El capítulo 2 termina con algunas interpretaciones físicas de los sistemas Hamiltonianos introduciendo de esta manera la notación clásica para esta teoría, notación que se mantendrá hasta el final de la tesis.

El capítulo 3 esta basado principalmente en el libro de Harier [7] y en los artículos de P. J. Chanell & C. Scovel [2] y T. J. Stuchi [11]. Este capítulo comienza definiendo el concepto de simplecticidad para una aplicación lineal

y luego lo extiende para funciones diferenciales en general, demuestra algunas propiedades e implicaciones de la simplecticidad de una función como la preservación del área simpléctica. Uno de los resultados más importantes de este capítulo es que para un sistema Hamiltoniano que no depende explícitamente del tiempo, se tiene que una transformación canónica de coordenadas es justo una transformación simpléctica y viceversa, así, una transformación canónica es simpléctica y por tanto tiene toda esta estructura geométrica de preservación de área. Se dedica una sección a las integrales primeras de un sistema Hamiltoniano demostrando el resultado clásico de Frobenius para que una cantidad sea integral primera basada en el corchete de Poisson. Se demuestra el teorema de Poincaré que establece que el flujo de un sistema Hamiltoniano es simpléctico, resultado que permite obtener como integrales primeras las bien conocidas integrales invariantes de Poincaré. Este resultado de Poincaré también permite obtener el integrador simpléctico que buscamos, pues al ser el flujo una transformación simpléctica, estamos en búsqueda de una función generatriz que avance en el tiempo la solución del sistema es decir que transforme las coordenadas en un tiempo dado a las coordenadas en un tiempo posterior. De esta manera habremos integrado exactamente el sistema. Estas ideas son desarrolladas por P. J. Channell & C. Scovel [2] sin embargo en la tesis se presenta una forma alternativa para obtener la ecuación llave que permite obtener en serie de Taylor la función generatriz que buscamos. A su vez se da un desarrollo explícito de la manera de deducir dicho integrador para cualquier orden.

En el capítulo 4 se estudian los sistemas Hamiltonianos del péndulo ideal, el oscilador armónico y la cadena de Toda finita de tres partículas. La manera de estudiar estos sistemas es implementando los integradores clásicos vistos en el capítulo 1 y los simplécticos vistos en el capítulo 3 y comparar la conservación de energía mecánica para los distintos integradores. En el caso de la cadena de Toda se analiza el comportamiento del error relativo de la energía mecánica así como el del momento angular generalizado. Los sistemas del péndulo y el oscilador se eligieron por ser sistemas conocidos, razón por la cual resulta fácil juzgar los resultados numéricos. La cadena de Toda con 3 partículas se eligió por ser un problema, en principio, con 3 grados de libertad (puede ser reducido a uno con 2 grados de libertad) que además involucra un Hamiltoniano exponencial lo que lo hace un sistema inestable desde el punto de vista numérico. Es necesario decir que el objetivo es comparar la preservación de la energía mecánica de los distintos integradores para tiempos de integración grandes.

Capítulo 1

Métodos numéricos clásicos para la solución de ecuaciones diferenciales ordinarias

Consideremos el problema de resolver numericamente la ecuación diferencial

$$\frac{dy}{dx} = f(x, y) \quad a \leq x \quad \text{con condición inicial} \quad y(a) = \hat{y}. \quad (1.1)$$

Así pues, deseamos aproximar la solución y en un conjunto discreto de puntos $x_k \geq a$ con $k = 0, 1, 2, \dots$. En la mayoría de los casos, aproximar la solución de (1.1) en un número infinito de puntos resulta muy difícil, es por esta razón que nos avocaremos a encontrar aproximaciones de y en un conjunto finito de valores mayores o iguales que a . De esta manera, encontraremos una solución numérica¹ de (1.1) en un intervalo de la forma $[a, b]$. Luego, si se desea observar el comportamiento de la solución y cuando $x \rightarrow \infty$ entonces basta con aproximar la solución para b suficientemente grande.

El primer paso para encontrar una solución numérica es introducir una malla de puntos en el intervalo $[a, b]$. Sea esta malla $a = x_0 < x_1 < \dots < x_N$. Supondremos por simplicidad que los puntos de dicha malla están separados uniformemente, es decir que $x_k = a + kh$ con $k = 0, \dots, N$ y donde h es el espacio entre cada elemento de la malla. Denotaremos por $y(x_k)$ el valor de la solución exacta en x_k y por y_k la aproximación a $y(x_k)$ generada por el método numérico. De manera natural, una vez introducida esta notación, surge la primera definición de error; ésta no depende del método numérico que se elija.

¹Una aproximación a la solución de (1.1) en un conjunto finito de valores mayores o iguales que a .

Definición 1.1 (Error global de discretización).

Definimos el error global de discretización en $[a, b]$ por:

$$E(h; [a, b]) = \max_{1 \leq i \leq N} |y_i - y(x_i)| \quad \text{con } x_i = a + ih.$$

También nos referiremos a él indistintamente por $E(h)$.

Intuitivamente esperamos que $E(h) \rightarrow 0$ cuando $h \rightarrow 0$. Pasaremos ahora a la construcción de los métodos llamados de un paso.

1.1. Métodos de un paso

Los métodos que son llamados de un paso son aquellos para los cuales, la construcción de las aproximaciones y_k a la solución exacta está dada por una regla como la siguiente:

$$y_{k+1} = y_k + h\Phi(x_k, y_k) \tag{1.2}$$

donde Φ es una función de \mathbb{R}^2 en \mathbb{R} .

Es importante observar que si la construcción de las aproximaciones sigue la regla anterior, entonces y_{k+1} depende sólo del paso anterior y_k .

Cálculo del error

Esencialmente, los cálculos requeridos para evaluar un método provienen de calcular $f(x_k, y_k)$. Por tanto, el error que se comete al aproximar proviene de dos fuentes: la primera, al discretizar la ecuación diferencial y la segunda, el error de redondeo.

El error de discretización puede clasificarse en dos; el que se produce al dar una aproximación discreta a la solución exacta (llamado error global de discretización) y el que se produce (dependiendo del método que se elija) al discretizar la ecuación diferencial. Por supuesto, estos dos están íntimamente relacionados, como veremos más adelante. En lo sucesivo supondremos que no existe error de redondeo. Daremos ahora las definiciones de error que nos faltan, éstas, a diferencia de la primera, dependen del método numérico que se elija.

Definición 1.2 (Error de discretización en un punto).

Definimos para cualquier método de un paso, el error de discretización en un punto $x \in [a, b]$ por

$$L(x, h) = \frac{1}{h} [y(x+h) - y(x)] - \Phi(x, y(x))$$

y donde Φ es la función que define el método en (1.2).

Definición 1.3 (Error local de discretización).

Definimos para cualquier método de un paso, el error local de discretización en $[a, b]$ por

$$L(h) = \max_{a \leq x \leq b-h} |L(x, h)|$$

Definición 1.4 (Orden del método).

Diremos que el orden del método es el entero positivo p si $L(h) = O(h^p)$.

Con estas definiciones, podemos hacer inmediatamente la siguiente observación:

Observación 1.1.

Si partimos de la solución exacta, entonces, el error producido por el método en un solo paso es h veces el error local de discretización.

Demostración. Por hipótesis, tenemos que $y_k = y(x_k)$, como en la condición inicial, entonces

$$\begin{aligned} y(x_{k+1}) - y_{k+1} &= y(x_{k+1}) - y_k - h\Phi(x_k, y_k) \\ &= y(x_{k+1}) - y(x_k) - h\Phi(x_k, y(x_k)) \\ &= y(x_k + h) - y(x_k) - h\Phi(x_k, y(x_k)) \\ &= hL(x_k, h), \end{aligned}$$

con lo cual queda demostrada la observación. □

Los dos tipos de errores que hemos definido arriba, se relacionan a través de su orden de convergencia, como lo establece el siguiente teorema.

Teorema 1.2.

Si el método numérico definido por la ecuación (1.2) es tal que $L(h) = O(h^p)$ y Φ tiene su derivada parcial respecto a su segundo argumento acotada, entonces $E(h) = O(h^p)$.

Demostración. Denotaremos por e_k a la diferencia $y(x_k) - y_k$, así, podemos escribir $E(h) = \max_{1 \leq n \leq N} |e_n|$.

De la definición de $L(x_k, h)$ se sigue que $y(x_{k+1}) = hL(x_k, h) + h\Phi(x_k, y(x_k)) + y(x_k)$, además, $y_{k+1} = y_k + h\Phi(x_k, y_k)$, por lo tanto, si sustituimos estas expresiones en la definición de e_k obtenemos que:

$$\begin{aligned} e_{k+1} &= y(x_{k+1}) - y_{k+1} \\ &= e_k + h \left[\Phi(x_k, y(x_k)) - \Phi(x_k, y_k) \right] + hL(x_k, h). \end{aligned}$$

Por otro lado, por hipótesis sabemos que $|\partial\Phi/\partial y(x, y)| \leq M_1$ con $(x, y) \in [a, b] \times \mathbb{R}$ y por el teorema del valor medio tenemos que para alguna $0 < \theta < 1$,

$$\begin{aligned} |\Phi(x_k, y(x_k)) - \Phi(x_k, y_k)| &= \left| \frac{\partial\Phi}{\partial y}(x_k, \theta y(x_k) + (1-\theta)y_k) (y(x_k) - y_k) \right| \\ &\leq M_1 |e_k| \end{aligned}$$

de este modo, sustituyendo en la última expresión encontrada para e_{k+1} , tenemos

$$\begin{aligned} |e_{k+1}| &\leq |e_k| + hM_1|e_k| + h|L(x_k, h)| \\ &\leq (1 + hM_1)|e_k| + h|L(h)| \end{aligned}$$

es decir, $\bar{e}_{k+1} \leq c\bar{e}_k + d$, donde $\bar{e}_k = |e_k|$, $c = 1 + hM_1$, $d = hL(h)$ y $\bar{e}_0 = 0$ (\bar{e}_0 es la condición inicial) así,

$$\begin{aligned} \bar{e}_n &\leq c(c\bar{e}_{n-2} + d) + d \leq \dots \leq (1 + c + \dots + c^{n-1})d \\ &= \left(\frac{c^n - 1}{c - 1} \right) d = \frac{[(1 + hM_1)^n - 1]}{hM_1} hL(h) \leq \frac{(1 + hM_1)^n}{M_1} L(h) \\ &\leq \frac{e^{hnM_1}}{M_1} L(h) \quad \text{pues } 1 + x \leq e^x \text{ con } x \geq 0. \end{aligned}$$

por otro lado, la observación (1.1), implica que $\bar{e}_n = h|L(x_n, h)| \leq hL(h) \rightarrow 0$ cuando $h \rightarrow 0$ pues por hipótesis $L(h) = O(h^p)$ y como $0 \leq \bar{e}_n$ entonces $\bar{e}_n \rightarrow 0$ si $h \rightarrow 0$. Teniendo en cuenta la última desigualdad obtenida para \bar{e}_n , podemos afirmar que \bar{e}_n/h^p está acotado por una constante cuando $h \rightarrow 0$ (de nuevo porque $L(h) = O(h^p)$), por lo tanto

$$E(h) = \max_{1 \leq n \leq N} \bar{e}_n = O(h^p).$$

□

En las hipótesis del teorema anterior, cabe remarcar que la condición $L(h) = O(h^p)$ implica condiciones sobre Φ y por tanto sobre f .

1.1.1. Método de Euler

Para obtener este método, es necesario expandir y en su serie de Taylor alrededor de x_k para lo cual supondremos que y es dos veces derivable. Así,

$$y(x) = y(x_k) + y'(x_k)(x - x_k) + \frac{1}{2}y''(\xi_k)(x - x_k)^2 \quad \text{con } \xi_k \in (x_k, x_{k+1}),$$

entonces

$$\begin{aligned} y(x_{k+1}) &= y(x_k) + y'(x_k)(x_{k+1} - x_k) + \frac{1}{2}y''(\xi_k)(x_{k+1} - x_k)^2 \\ &= y(x_k) + y'(x_k)(a + (k+1)h - a - kh) + \frac{1}{2}y''(\xi_k)(a + (k+1)h - a - kh)^2 \\ &= y(x_k) + y'(x_k)(a + kh + h - a - kh) + \frac{1}{2}y''(\xi_k)(a + kh + h - a - kh)^2 \\ &= y(x_k) + hy'(x_k) + \frac{h^2}{2}y''(\xi_k). \end{aligned}$$

Entonces, si y'' es acotada y h es pequeña, se puede ignorar el último término, obteniendo así

$$y(x_{k+1}) \doteq y(x_k) + hf(x_k, y(x_k)).$$

Finalmente hemos obtenido el método de Euler, que se escribe como

$$\begin{aligned} y_{k+1} &= y_k + hf(x_k, y_k) \quad k = 0, \dots, N-1 \\ y_0 &= \hat{y} \end{aligned}$$

Notemos que para este método Φ coincide con f y de acuerdo con nuestra definición de $L(h)$, ésta depende de h , de f y de $[a, b]$, sin embargo, la única dependencia explícita es sobre h .

Por otro lado, geoméricamente, el método de Euler consiste en aproximar la solución en x_{k+1} siguiendo la tangente a la curva solución en x_k .

Podemos enunciar ahora el siguiente teorema:

Teorema 1.3.

Si f tiene su derivada parcial respecto de su segundo argumento acotada y si la solución de (1.1) tiene segunda derivada acotada, entonces las aproximaciones de Euler convergen a la solución exacta cuando $h \rightarrow 0$ y $E(h) = O(h)$.

Demostración. Expandamos en serie de Taylor $y(x+h)$ alrededor de x :

$$\begin{aligned} y(x+h) &= y(x) + y'(x)(x+h-x) + \frac{1}{2}y''(\xi)(x+h-x)^2 \\ &= y(x) + hy'(x) + \frac{h^2}{2}y''(\xi) \quad \text{con} \quad \xi \in (x, x+h). \end{aligned}$$

Entonces

$$\begin{aligned} y(x+h) - y(x) &= hy'(x) + \frac{h^2}{2}y''(\xi) \\ &= hf(x, y(x)) + \frac{h^2}{2}y''(\xi), \end{aligned}$$

luego, de acuerdo a nuestra definición de $L(h)$ obtenemos

$$\begin{aligned} L(h) &= \max_{a \leq x \leq b-h} \left| \frac{1}{h} [hf(x, y(x)) + \frac{h^2}{2}y''(\xi)] - f(x, y(x)) \right| \\ &= \max_{a \leq x \leq b-h} \left| \frac{h}{2}y''(\xi) \right|. \end{aligned}$$

Ahora, como $\xi \in (x, x+h)$ y $x \in [a, b-h]$, entonces $a \leq x < \xi < x+h \leq b$, es decir $\xi \in [a, b]$, así

$$L(h) = \frac{h}{2} \max_{a \leq \xi \leq b} |y''(\xi)| \leq \frac{h}{2} M,$$

donde M es la constante que por hipótesis sabemos que existe y que acota a y'' en $[a, b]$. Por tanto como $\frac{h}{2}M = O(h)$ entonces $L(h) = O(h)$. Como por

hipótesis tenemos que f tiene su derivada parcial respecto a su segundo argumento acotada y $f = \Phi$, podemos aplicar el teorema (1.2) para garantizar que $E(h) = O(h)$. Así, retomando la observación (1.1), podemos escribir

$$|y(x_{k+1}) - y_{k+1}| \leq hL(x_k, h) \leq hL(h) = hO(h) \xrightarrow{h \rightarrow 0} 0.$$

Por tanto las aproximaciones de Euler convergen a la solución exacta cuando h tiende a cero. □

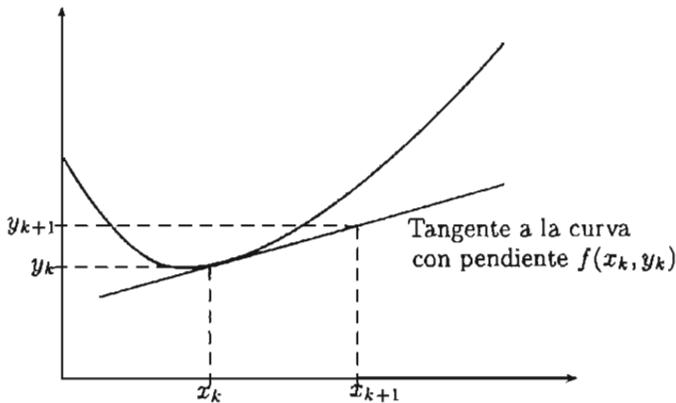


Figura 1.1: Interpretación geométrica del método de Euler

1.1.2. Métodos Runge-Kutta

En esta sección estudiaremos los llamados métodos de Runge-Kutta. El método de Runge-Kutta de orden R tiene la siguiente regla de construcción:

$$y_{n+1} = y_n + h \sum_{r=1}^R a_r k_r \quad (1.3)$$

donde,

$$k_1 = f(x_n, y_n) \quad y$$

$$k_r = f\left(x_n + hp_r, y_n + h \sum_{s=1}^{r-1} q_{r,s} k_s\right)$$

para $r \in \{2, \dots, R\}$.

En la ecuación para k_r aparecen los coeficientes p , q y a , los cuales serán determinados al encontrar k_r . Para este método de un paso, de acuerdo a nuestra definición, $\Phi = \sum_{r=1}^R a_r k_r$.

A continuación calcularemos los métodos de Runge-Kutta de orden 2 y 3 a modo de ejemplos particulares del método general expuesto en las líneas precedentes. Para obtener el primero de ellos es necesario fijar $R = 2$, por tanto, este método queda descrito, teóricamente, de acuerdo con (1.3) como

$$y_{n+1} = y_n + h(a_1 k_1 + a_2 k_2)$$

en donde,

$$\begin{aligned} k_1 &= f(x_n, y_n) \\ k_2 &= f(x_n + hp_2, y_n + hq_{2,1}k_1). \end{aligned}$$

En lo sucesivo, si no se manifiesta explícitamente, supondremos que todas las funciones están evaluadas en x_n y y_n . Ahora, para obtener los valores de los coeficientes a , p y q , debemos aproximar a $y(x_n + h)$ a través de su serie de Taylor truncada hasta el término de orden 2, esto es:

$$y(x_n + h) = y(x_n) + \dot{y}(x_n)h + \frac{1}{2}\ddot{y}(x_n)h^2,$$

es decir, tenemos que:

$$y_{n+1} = y_n + hf + \frac{h^2}{2}(f_y f + f_x),$$

ya que $\dot{y} = f$ y $\ddot{y} = f_y f + f_x$ de acuerdo a la ecuación diferencial (1.1). Por otro lado, la definición de nuestro método nos dice que

$$\begin{aligned} y_{n+1} &= y_n + h\{a_1 f + a_2 f(x_n + hp_2, y_n + hq_{2,1}k_1)\} \\ &= y_n + ha_1 f + ha_2(f + hp_2 f_x + hq_{2,1} f f_y) \\ &= y_n + hf(a_1 + a_2) + h^2 a_2(p_2 f_x + q_{2,1} f f_y), \end{aligned}$$

donde la segunda igualdad se obtiene al expandir en serie de Taylor² a la función $f(x_n + hp_2, y_n + hq_{2,1}k_1)$ al rededor del punto (x_n, y_n) y sustituir el valor de k_1 , y la última se obtiene simplemente al reagrupar los términos de acuerdo al exponente de h . Si comparamos las dos últimas expresiones que hemos obtenido para y_{n+1} obtenemos el siguiente sistema de ecuaciones:

$$\begin{aligned} a_1 + a_2 &= 1 \\ a_2 p_2 &= \frac{1}{2} \\ a_2 q_{2,1} &= \frac{1}{2}. \end{aligned}$$

Como se puede apreciar, este sistema tiene una infinidad de soluciones, ya que las dos últimas ecuaciones dependen sólo de a_2 , dejando libre el valor de a_1 . La

²Ver S. Lange [9].

elección clásica es $a_1 = a_2 = \frac{1}{2}$ y $p_2 = q_{2,1} = 1$. Éste es el llamado método de Heun, que queda escrito de la siguiente manera

$$y_{k+1} = y_k + \frac{h}{2}[f(x_k, y_k) + f(x_{k+1}, y_k + hf(x_k, y_k))]. \quad (1.4)$$

Una observación interesante es que hemos reemplazado $f(x_k, y_k)$ del método de Euler por un promedio de f evaluado en lugares diferentes, es decir que Φ en vez de ser simplemente f , queda descrita mediante la siguiente ecuación:

$$\Phi(x, y) = \frac{1}{2}[f(x, y) + f(x + h, y + hf(x, y))].$$

Éste método, por supuesto es de segundo orden, es decir $L(h) = O(h^2)$.

Obtendremos ahora el método de Runge-Kutta de orden 3 con la finalidad de esbozar un poco más el procedimiento general para obtener este método de cualquier orden.

Del mismo modo que hicimos antes, a través de la serie de Taylor, pero esta vez expandiendo hasta el término de orden 3, podemos obtener aproximaciones a la solución con la siguiente forma:

$$y_{n+1} = y_n + hf + \frac{h^2}{2}(f_y f + f_x) + \frac{h^3}{6}(f^2 f_{yy} + 2ff_{xy} + ff_y^2 + f_x f_y + f_{xx}),$$

esto ya que:

$$\begin{aligned} \ddot{y} &= \frac{d}{dx}(f_y f + f_x) = f\left(\frac{df_y}{dy} \frac{dy}{dx} + \frac{df_y}{dx}\right) + f_y\left(\frac{df}{dy} \frac{dy}{dx} + \frac{df}{dx}\right) + \frac{df_x}{dy} \frac{dy}{dx} + \frac{df_x}{dx} \\ &= f^2 f_{yy} + 2ff_{xy} + ff_y^2 + f_x f_y + f_{xx}, \end{aligned}$$

donde, claro está, hemos supuesto que f es dos veces continuamente diferenciable.

Ahora, escribiremos el método de Runge-Kutta de acuerdo con nuestra definición pero fijando $R = 3$, entonces, queda:

$$\begin{aligned} y_{n+1} &= y_n + h(a_1 k_1 + a_2 k_2 + a_3 k_3) \\ &= y_n + hf(a_1 + a_2) + h^2 a_2 (p_2 f_x + q_{2,1} f f_y) + ha_3 k_3, \end{aligned}$$

con k_1 y k_2 como en el caso anterior y $k_3 = f(x_n + hp_3, y_n + h(q_{3,1} k_1 + q_{3,2} k_2))$ de acuerdo a la relación que define a las constantes k . La segunda igualdad se obtiene haciendo uso de los cálculos realizados para el caso $R = 2$. Debemos entonces encontrar una expresión para $ha_3 k_3$ en términos de f y sus derivadas, para ello, utilizaremos la expansión en serie de Taylor de f truncada hasta el

término necesario para obtener h^3 ; desarrollemos pues ha_3k_3 :

$$\begin{aligned}
 ha_3k_3 &= ha_3f(x_n + hp_3, y_n + h(q_{3,1}k_1 + q_{3,2}k_2)) \\
 &= ha_3\left\{f + hp_3f_x + h(q_{3,1}k_1 + q_{3,2}k_2)f_y + \frac{1}{2}[h^2p_3^2f_{xx} \right. \\
 &\quad \left. + 2h^2p_3(q_{3,1}k_1 + q_{3,2}k_2)f_{xy} + h^2(q_{3,1}k_1 + q_{3,2}k_2)^2f_{yy}]\right\} \\
 &= ha_3\left\{f + hp_3f_x + h[q_{3,1}f + q_{3,2}(f + hp_2f_x + hq_{2,1}ff_y)]f_y + \frac{h^2}{2}[p_3^2f_{xx} \right. \\
 &\quad \left. + 2p_3f(q_{3,1} + q_{3,2})f_{xy} + f^2(q_{3,1} + q_{3,2})^2f_{yy}]\right\} \\
 &= hf a_3 + h^2 a_3[p_3f_x + (q_{3,1} + q_{3,2})ff_y] + h^3 \frac{a_3}{2}[2q_{3,2}f_y(p_2f_x + q_{2,1}ff_y) \\
 &\quad + p_3^2f_{xx} + 2p_3f(q_{3,1} + q_{3,2})f_{xy} + f^2(q_{3,1} + q_{3,2})^2f_{yy}].
 \end{aligned}$$

Así, si sustituimos esta expresión para ha_3k_3 en la ecuación que define nuestro método y agrupamos de acuerdo al exponente de h , tenemos que:

$$\begin{aligned}
 y_{n+1} &= y_n + hf(a_1 + a_2 + a_3) + h^2[(a_2p_2 + a_3p_3)f_x + (a_2q_{2,1} + a_3q_{3,1} \\
 &\quad + a_3q_{3,2})ff_y] + h^3 \frac{a_3}{2}[2q_{3,2}p_2f_xf_y + 2q_{3,2}q_{2,1}ff_y^2 \\
 &\quad + p_3^2f_{xx} + 2p_3(q_{3,1} + q_{3,2})f_{xy} + (q_{3,1} + q_{3,2})^2f_{yy}^2].
 \end{aligned}$$

De este modo si comparamos término a término esta última expresión de y_{n+1} con la obtenida a través de la expansión en serie de Taylor, obtenemos el siguiente sistema de ecuaciones,

$$\begin{aligned}
 a_1 + a_2 + a_3 &= 1 \\
 a_2p_2 + a_3p_3 &= \frac{1}{2} \\
 a_2q_{2,1} + a_3q_{3,1} + a_3q_{3,2} &= \frac{1}{2} \\
 a_3q_{3,2}p_2 &= \frac{1}{6} \\
 a_3q_{3,2}q_{2,1} &= \frac{1}{6} \\
 a_3p_3^2 &= \frac{1}{3} \\
 a_3p_3(q_{3,1} + q_{3,2}) &= \frac{1}{3} \\
 a_3(q_{3,1} + q_{3,2})^2 &= \frac{1}{3}.
 \end{aligned}$$

Al resolver este sistema, encontramos las constantes que definen completamente el método Runge-Kutta de orden 3. Se puede verificar fácilmente que estas constantes son: $a_1 = a_2 = a_3 = \frac{1}{3}$, $p_2 = \frac{1}{2}$, $p_3 = 1$, $q_{2,1} = \frac{1}{2}$, $q_{3,1} = 0$ y $q_{3,2} = 1$. De esta manera, podemos escribir este método como sigue,

$$y_{k+1} = y_k + \frac{h}{3}(k_1 + k_2 + k_3)$$

donde,

$$\begin{aligned}k_1 &= f(x_k, y_k) \\k_2 &= f\left(x_k + \frac{h}{2}, y_k + \frac{h}{2}f(x_k, y_k)\right) \\k_3 &= f\left(x_k + h, y_k + hf\left(x_k + \frac{h}{2}, y_k + \frac{h}{2}f(x_k, y_k)\right)\right)\end{aligned}$$

De manera totalmente análoga a como hemos hecho se obtiene el más conocido de los métodos de Runge-Kutta, este es de orden 4 y se escribe como sigue

$$y_{k+1} = y_k + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4)$$

donde,

$$\begin{aligned}k_1 &= f(x_k, y_k) \\k_2 &= f\left(x_k + \frac{h}{2}, y_k + \frac{h}{2}k_1\right) \\k_3 &= f\left(x_k + \frac{h}{2}, y_k + \frac{h}{2}k_2\right) \\k_4 &= f(x_k + h, y_k + hk_3)\end{aligned}$$

A continuación demostraremos un teorema que ejemplifica la manera de calcular el orden de un método de un paso, el cual nos servirá, para probar que el método de Heun es de segundo orden.

Teorema 1.4.

Si en el método (1.2) consideramos

$$\Phi(x, y) = c_2 f(x, y) + c_3 f(x + c_1 h, y + c_1 h f(x, y)),$$

entonces dicho método es de segundo orden si se satisface que

$$c_2 + c_3 = 1 \quad y \quad c_1 c_2 = \frac{1}{2}.$$

Demostración. Primero, expandamos en serie de Taylor a $f(x + c_1 h, y + c_1 h f(x, y))$ alrededor del punto (x, y) y sustituyamos esta expresión en Φ , entonces Φ queda escrita como

$$\begin{aligned}\Phi &= c_2 f + c_3 [f + c_1 h f_y + c_1 h f_x + O(h^2)] \\&= (c_2 + c_3) f + c_1 c_3 h (f_y + f_x) + O(h^2).\end{aligned}$$

Por otro lado, si expandimos en serie de Taylor a la solución de (1.1) evaluada en $x + h$ alrededor de x , obtenemos que

$$\begin{aligned} \frac{1}{h}[y(x+h) - y(x)] &= \dot{y}(x) + \frac{h}{2}\ddot{y}(x) + O(h^2) \\ &= f + \frac{h}{2}\frac{df}{dx} + O(h^2) \\ &= f + \frac{h}{2}(ff_y + f_x) + O(h^2). \end{aligned}$$

De este modo, de acuerdo con nuestra definición de $L(h)$, sustituyendo las igualdades encontradas para Φ y $1/h[y(x+h) - y(x)]$, tenemos que

$$L(h) = (1 - c_2 - c_3)f + h\left(\frac{1}{2} - c_1c_3\right)(ff_y + f_x) + O(h^2),$$

de donde, al sustituir las condiciones sobre las constantes c_1 , c_2 y c_3 , concluimos que el método es de segundo orden para cualquier f con derivadas acotadas. \square

Así, es fácil ver que para el método de Heun, Φ tiene la forma necesaria para aplicar el teorema anterior. Eligiendo las constantes como $c_2 = c_3 = \frac{1}{2}$ y $c_1 = 1$ tenemos que se satisfacen las condiciones del teorema. Por tanto, la ecuación (1.4) no define un solo método de segundo orden, sino toda una familia de métodos dependiendo de la elección de las constantes. Mediante un procedimiento análogo se puede demostrar que efectivamente el método de Runge-Kutta que se obtiene al fijar $R = 4$ es de cuarto orden.

A continuación, presentamos una gráfica que muestra la diferencia entre aproximar con el método de Euler y el de Heun.

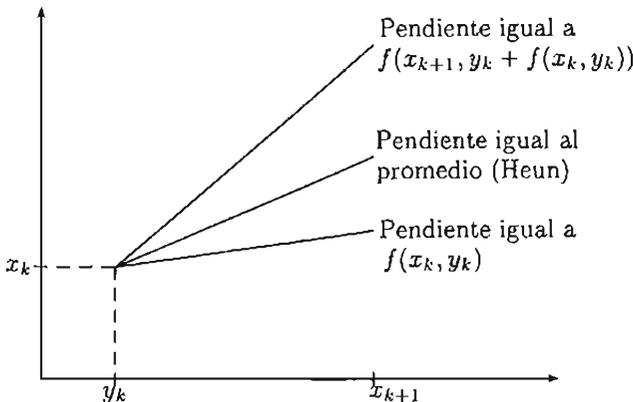


Figura 1.2: Interpretación geométrica del método de Heun.

1.2. Métodos multipaso

En esta sección construiremos métodos numéricos para aproximar la solución de la ecuación (1.1), que a diferencia de los que hemos construido en la sección anterior, el valor y_{k+1} no dependerá sólo de la información en el punto anterior x_k sino en la información de varios puntos anteriores, $x_k, x_{k-1}, \dots, x_{k-m}$. Intuitivamente, debemos ganar precisión de esta manera, sin embargo, para construir estos métodos, es necesaria cierta teoría previa.

Interpolación polinomial

El problema que en esta sección trataremos, es el de encontrar un polinomio $p(x)$ que interpole el conjunto de datos $\{(x_0, y_0), \dots, (x_n, y_n)\} \subset \mathbb{R}^2$, es decir que

$$p(x_i) = y_i \quad (i = 0, \dots, n).$$

Al conjunto $\{x_0, x_1, \dots, x_n\}$ se le llama **conjunto de nodos** y al conjunto $\{y_0, y_1, \dots, y_n\}$ **conjunto de valores asociados** al conjunto de nodos.

Empezaremos con la definición de interpolación

Definición 1.5 (Función interpoladora).

Dado un conjunto de datos $\{(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)\} \subset \mathbb{R}^2$ decimos que una función g de \mathbb{R} en \mathbb{R} interpola los datos si,

$$g(x_i) = y_i, \quad (i = 0, 1, \dots, n).$$

Para los fines que perseguimos, estamos interesados en las funciones g que además de interpolar un conjunto de datos, también son polinomios, a estas funciones les llamaremos **polinomios interpoladores**.

Así, para un conjunto dado de datos y su polinomio interpolador, es necesario observar que el grado de este polinomio dependerá de la cantidad de datos que se quiera interpolar; por ejemplo, si se tuvieran los datos en solo dos nodos, bastaría un polinomio de grado 1 para interpolarlos, no siendo así si se tuvieran en tres nodos, ya que en este caso, si no estuvieran en una línea recta, se requeriría de un polinomio de grado al menos dos para interpolarlos. Por supuesto, habría una infinidad de polinomios de grado 3 o mayor que los interpolarían.

El siguiente teorema, nos ofrece un resultado acerca de la existencia y grado del polinomio interpolador de un conjunto de datos.

Teorema 1.5 (Existencia y unicidad del polinomio interpolador).

Dado un conjunto de datos $\{(x_0, y_0), \dots, (x_n, y_n)\} \subset \mathbb{R}^2$, tal que todos los nodos x_i ($i = 0, \dots, n$) son distintos, entonces existe un único polinomio interpolador de dichos datos con grado menor o igual a n .

Demostración. Probaremos primero la existencia por inducción sobre n . Para $n = 1$, podemos tomar

$$p_1(x) = y_0 + \frac{y_1 - y_0}{x_1 - x_0}(x - x_0)$$

ya que este polinomio es de grado a lo más 1 y además $p_1(x_0) = y_0$ y $p_1(x_1) = y_1$. Ahora haremos nuestra hipótesis de inducción, supondremos cierto el teorema para $n = k - 1$, es decir, que dado un conjunto de k datos existe un polinomio de grado menor o igual a $k - 1$ que los interpola. Probaremos ahora que el teorema sigue siendo válido cuando $n = k$.

Supongamos que $p_{k-1}(x)$ es un polinomio de grado menor o igual a $k - 1$ tal que, $p_{k-1}(x_i) = y_i$, con $i = 0, 1, \dots, k - 1$. Consideremos ahora el siguiente polinomio,

$$p_k(x) = p_{k-1}(x) + c(x - x_0)(x - x_1) \cdots (x - x_{k-1}).$$

Claramente, este polinomio es de grado menor o igual a k , además, $p_k(x_i) = y_i$ con $i = 0, 1, \dots, k - 1$ ya que el segundo sumando se anula en x_i cuando $i = 0, 1, \dots, k - 1$ y el primero, p_{k-1} , vale y_i . Por otro lado, como los nodos son distintos, podemos determinar c de tal manera que $p_k(x_k) = y_k$. Esto se logra evaluando p_k en x_k e igualando a y_k . Así se obtiene que,

$$c = \frac{y_k - p_{k-1}(x_k)}{(x_k - x_0)(x_k - x_1) \cdots (x_k - x_{k-1})}.$$

Por lo tanto p_k es un polinomio de grado menor o igual a k que interpola al conjunto de datos $\{(x_0, y_0), \dots, (x_k, y_k)\}$. De este modo queda probada la existencia.

Para probar la unicidad, supongamos que existen p y q , polinomios distintos de grado menor o igual a n , que interpolan los $n + 1$ datos. Consideremos ahora $r = p - q$, es claro que r es un polinomio de grado menor o igual a n que tiene al menos $n + 1$ ceros, pues,

$$r(x_i) = p(x_i) - q(x_i) = y_i - y_i = 0 - 0 \quad (i = 0, \dots, n),$$

ya que p y q son polinomios interpoladores. Sin embargo, de acuerdo al teorema fundamental del álgebra r solo puede tener a lo más n raíces, por lo cual $r = 0$, concluyendo que $p = q$. □

Polinomio interpolador de Newton

En esta sección veremos cómo calcular explícitamente el polinomio interpolador que el teorema (1.5) garantiza que existe para un conjunto de datos $\{(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)\} \subset \mathbb{R}^2$. Esto lo lograremos obteniendo los coeficiente del polinomio p , que usamos en la prueba del teorema (1.5), y que a continuación escribimos,

$$p(x) = c_0 + c_1(x - x_0) + c_2(x - x_0)(x - x_1) + \cdots + c_n(x - x_0)(x - x_1) \cdots (x - x_{n-1}).$$

Así, los coeficientes c_0, c_1, \dots, c_n , son determinados a través de las ecuaciones:

$$p(x_i) = y_i \quad (i = 0, \dots, n).$$

Esto se logra evaluando p en x_i e igualando con y_i , es decir, obtenemos el siguiente sistema de ecuaciones:

$$\begin{aligned} y_0 &= c_0 \\ y_1 &= c_0 + c_1(x_1 - x_0) \\ y_2 &= c_0 + c_1(x_2 - x_0) + c_2(x_2 - x_0)(x_2 - x_1) \\ &\vdots \\ y_n &= c_0 + c_1(x_n - x_0) + \cdots + c_n(x_n - x_0) \cdots (x_n - x_{n-1}) \end{aligned}$$

Si escribimos este sistema en términos de una matriz, es fácil ver que la matriz asociada a este sistema es triangular inferior, y como $x_i \neq x_j$ si $i \neq j$, entonces es posible calcular los coeficientes a través de sustitución hacia adelante. De esta manera, se puede observar en general que c_k depende solo del conjunto de nodos $\{x_0, \dots, x_k\}$ y de sus valores asociados $\{y_0, \dots, y_k\}$.

Este hecho junto con la unicidad del polinomio interpolador dan sentido a la siguiente definición.

Definición 1.6.

Dado un conjunto de datos $\{(x_0, y_0), \dots, (x_n, y_n)\}$, definimos, a $y[x_0, x_1, \dots, x_n]$ como el coeficiente c_n del polinomio interpolador de los datos, escrito de la siguiente forma:

$$p(x) = c_0 + c_1(x - x_0) + c_2(x - x_0)(x - x_1) + \cdots + c_n(x - x_0)(x - x_1) \cdots (x - x_{n-1}).$$

Nuestro objetivo ahora, es encontrar una fórmula recursiva que nos permita encontrar los coeficientes c_k . Una primera observación es que p puede ser visto como p_n en la formulación recursiva que escribimos a continuación.

$$\begin{aligned} p_0(x) &= c_0, \\ p_k(x) &= p_{k-1}(x) + c_k(x - x_0)(x - x_1) \cdots (x - x_{k-1}), \quad (k = 1, \dots, n) \end{aligned} \quad (1.5)$$

donde p_{k-1} es un polinomio de grado menor o igual a $k - 1$ que interpola los datos $\{(x_0, y_0), \dots, (x_{k-1}, y_{k-1})\}$. Con esta formulación recursiva de p , es posible observar que $y[x_0, \dots, x_k]$ con $0 \leq k < n$, es igual al coeficiente c_k del polinomio que interpola los datos $\{(x_0, y_0), \dots, (x_n, y_n)\}$ y no sólo al coeficiente c_k del polinomio que interpola los $k + 1$ primeros puntos.

Para calcular c_k recursivamente, demostraremos que podemos escribir p_k alternativamente, como lo afirma el siguiente lema.

Observación 1.6.

En la formulación (1.5) para p , es posible escribir

$$p_k(x) = p_{k-1}^{(1)}(x) + \frac{x - x_0}{x_k - x_0} (p_{k-1}^{(2)}(x) - p_{k-1}^{(1)}(x)),$$

con $p_{k-1}^{(1)}$ y $p_{k-1}^{(2)}$ polinomios de grado menor o igual a $k - 1$ y tales que interpolan a los conjuntos de datos $\{(x_0, y_0), \dots, (x_{k-1}, y_{k-1})\}$ y $\{(x_1, y_1), \dots, (x_k, y_k)\}$ respectivamente.

Demostración. Los polinomios $p_{k-1}^{(1)}(x)$ y $p_{k-1}^{(2)}(x)$ existen por el teorema (1.5). Veremos ahora que el polinomio $q(x) = p_{k-1}^{(1)}(x) + (x - x_0/x_k - x_0)(p_{k-1}^{(2)}(x) - p_{k-1}^{(1)}(x))$ interpola los datos $\{(x_0, y_0), \dots, (x_k, y_k)\}$ y que su grado es menor o igual a k . Lo segundo es claro. Procederemos ahora a demostrar que q interpola dichos datos. Es claro que

$$q(x_0) = p_{k-1}^{(1)}(x_0) = y_0$$

ya que el segundo término de q se anula. Ahora para $i = 1, \dots, k-1$, tenemos que

$$\begin{aligned} q(x_i) &= p_{k-1}^{(1)}(x_i) + \frac{x_i - x_0}{x_k - x_0} (p_{k-1}^{(2)}(x_i) - p_{k-1}^{(1)}(x_i)) \\ &= p_{k-1}^{(1)}(x_i) + \frac{x_i - x_0}{x_k - x_0} (y_i - y_i) \\ &= p_{k-1}^{(1)}(x_i) \\ &= y_i \end{aligned}$$

y finalmente, para x_k , obtenemos que:

$$\begin{aligned} q(x_k) &= p_{k-1}^{(1)}(x_k) + \frac{x_k - x_0}{x_k - x_0} (p_{k-1}^{(2)}(x_k) - p_{k-1}^{(1)}(x_k)) \\ &= p_{k-1}^{(2)}(x_k) \\ &= y_k. \end{aligned}$$

Por lo tanto, q interpola datos $\{(x_0, y_0), \dots, (x_k, y_k)\}$ y por la unicidad del polinomio interpolador y de acuerdo al planteamiento (1.5), concluimos que $p_k = q$. □

La observación anterior nos permitirá demostrar el siguiente resultado, que finalmente nos proporciona una manera recursiva de calcular los coeficientes c_k , ya que estos son iguales a $y[x_0, \dots, x_k]$.

Lema 1.7.

Dado un conjunto de de datos $\{(x_0, y_0), \dots, (x_n, y_n)\}$, entonces, se satisfacen las siguientes igualdades:

$$\begin{aligned} y[x_i] &= y_i \quad (i = 0, \dots, n) \\ y[x_0, x_1, \dots, x_k] &= \frac{y[x_1, x_2, \dots, x_k] - y[x_0, x_1, \dots, x_{k-1}]}{x_k - x_0} \quad (k = 1, \dots, n). \end{aligned}$$

Demostración. Para probar que $y[x_i] = y_i$ basta con remitirse a la definición (1.6) y observar de acuerdo con el planteamiento (1.5) que $p_0(x_i) = y_i = c_0 = y[x_i]$. Por otro lado, también del planteamiento (1.5), se puede ver que, de acuerdo con la definición (1.6), $y[x_0, \dots, x_k]$ es el coeficiente de x^k en p_k . Por

otro lado, de acuerdo con la última observación el coeficiente de x^k es $(1/(x_k - x_0))(c_{k-1}^{(2)} - c_{k-1}^{(1)})$ donde $c_{k-1}^{(2)}$ y $c_{k-1}^{(1)}$ son los coeficientes de x^{k-1} de los polinomios $p_{k-1}^{(2)}$ y $p_{k-1}^{(1)}$ respectivamente. Sin embargo, de la definición (1.6), tenemos que $c_{k-1}^{(2)} = y[x_1, \dots, x_k]$ y $c_{k-1}^{(1)} = y[x_0, \dots, x_{k-1}]$. Por lo tanto si igualamos los coeficientes de x^k que obtenemos, por un lado de la última observación y por el otro, del planteamiento (1.5), concluimos que:

$$y[x_0, x_1, \dots, x_k] = \frac{y[x_1, x_2, \dots, x_k] - y[x_0, x_1, \dots, x_{k-1}]}{x_k - x_0}.$$

□

Gracias a este lema, es posible calcular los coeficientes c_k del polinomio interpolador de Newton, que finalmente queda escrito de la siguiente manera:

$$p(x) = y[x_0] + y[x_0, x_1](x - x_0) + y[x_0, x_1, x_2](x - x_0)(x - x_1) + \dots + y[x_0, x_1, \dots, x_n](x - x_0)(x - x_1) \dots (x - x_{n-1}) \quad (1.6)$$

Ahora bien, para los fines que perseguimos, el caso de interés es cuando los **nodos son equidistantes**, es decir cuando

$$x_i = x_0 + ih, \quad (i = 0, \dots, n).$$

Veremos que cuando esto pasa, el polinomio interpolador de Newton toma una forma particular. Para ello, es necesario introducir la siguiente definición.

Definición 1.7.

Dado un conjunto de números $\{y_0, \dots, y_n\}$, definimos el símbolo $\Delta^k y_i$ con $i \in \{0, \dots, n-1\}$, de manera recursiva como

$$\begin{aligned} \Delta y_i &= y_{i+1} - y_i \\ \Delta^k y_i &= \Delta^{k-1} y_{i+1} - \Delta^{k-1} y_i \quad (k = 2, \dots, n-i). \end{aligned}$$

Por convención, consideraremos que $\Delta^0 y_i = y_i$. Estamos ahora en condiciones de enunciar y probar el siguiente lema.

Lema 1.8.

Dado un conjunto de de datos $\{(x_0, y_0), \dots, (x_n, y_n)\} \subset \mathbb{R}^2$, con nodos equidistantes se tiene que

$$y[x_0, x_1, \dots, x_k] = \frac{\Delta^k y_0}{h^k \cdot k!}, \quad (k = 0, \dots, n).$$

Demostración. Haremos inducción sobre k . El caso $k = 0$ es trivial, pues, el lema (1.7) implica que $y[x_0] = y_0$ y de acuerdo con la convención hecha para $\Delta^0 y_i$, se tiene que $\Delta^0 y_0 / h^0 \cdot 0! = y_0$.

Empecemos la inducción con $k = 1$. El lema (1.7) nos dice que

$$y[x_0, x_1] = \frac{y[x_1] - y[x_0]}{x_1 - x_0} = \frac{y_1 - y_0}{x_1 - x_0} = \frac{\Delta y_0}{h},$$

donde la última igualdad se da gracias a que los nodos son equidistantes y a la definición (1.7). Hagamos ahora el paso inductivo, es decir, supongamos que para $1 < k \leq n$, se tiene que:

$$y[x_0, x_1, \dots, x_{k-1}] = \frac{\Delta^{k-1} y_0}{h^{k-1} \cdot (k-1)!}$$

por otro lado, por el lema (1.7), podemos escribir,

$$y[x_0, x_1, \dots, x_k] = \frac{y[x_1, x_2, \dots, x_k] - y[x_0, x_1, \dots, x_{k-1}]}{x_k - x_0} \quad (1.7)$$

sin embargo, por la hipótesis que hemos hecho, se tiene que:

$$y[x_1, x_2, \dots, x_k] = \frac{\Delta^{k-1} y_1}{h^{k-1} \cdot (k-1)!} \quad \text{y} \quad y[x_0, x_2, \dots, x_{k-1}] = \frac{\Delta^{k-1} y_0}{h^{k-1} \cdot (k-1)!}$$

entonces la ecuación (1.7), queda escrita de la siguiente manera:

$$\begin{aligned} y[x_0, x_1, \dots, x_k] &= \frac{1}{x_k - x_0} \left(\frac{\Delta^{k-1} y_1}{h^{k-1} \cdot (k-1)!} - \frac{\Delta^{k-1} y_0}{h^{k-1} \cdot (k-1)!} \right) \\ &= \frac{1}{x_k - x_0} \left(\frac{\Delta^{k-1} y_1 - \Delta^{k-1} y_0}{h^{k-1} \cdot (k-1)!} \right) \\ &= \frac{1}{hk} \left(\frac{\Delta^{k-1} y_1 - \Delta^{k-1} y_0}{h^{k-1} \cdot (k-1)!} \right) \\ &= \frac{\Delta^{k-1} y_1 - \Delta^{k-1} y_0}{h^k \cdot k!} \\ &= \frac{\Delta^k y_0}{h^k \cdot k!} \end{aligned}$$

esto, ya que $x_k = x_0 + kh$ y por la definición (1.7). Esto concluye la inducción y por lo tanto el lema es cierto. \square

Con este último lema, podemos escribir el polinomio interpolador de Newton, dado por la ecuación (1.6), en el caso de nodos equidistantes como sigue:

$$p(x) = y_0 + \frac{\Delta y_0}{h}(x - x_0) + \frac{\Delta^2 y_0}{2h^2}(x - x_0)(x - x_1) + \dots + \frac{\Delta^n y_0}{n!h^n}(x - x_0)(x - x_1) \cdots (x - x_{n-1}). \quad (1.8)$$

En la siguiente sección veremos, que para construir los métodos multipaso, resulta más útil conocer el polinomio interpolador de Newton no precisamente en la forma expresada por la ecuación (1.8), sino en una muy similar que ahora construiremos. En primer lugar, definiremos un símbolo muy parecido a Δ .

Definición 1.8.

Dado un conjunto de números $\{y_0, \dots, y_n\}$, definimos el símbolo $\bar{\Delta}^k y_i$ ($i = 0, \dots, n-1$), de manera recursiva como

$$\begin{aligned}\bar{\Delta} y_i &= y_i - y_{i+1} \\ \bar{\Delta}^k y_i &= \bar{\Delta}^{k-1} y_i - \bar{\Delta}^{k-1} y_{i+1} \quad (k = 2, \dots, n-i).\end{aligned}$$

La expresión del polinomio interpolador de Newton, que estamos buscando, está en términos de este símbolo que acabamos de definir.

Debemos ahora encontrar una relación entre Δ y $\bar{\Delta}$, para ello demostraremos el siguiente lema.

Lema 1.9.

Dado un conjunto de números $\{y_0, \dots, y_n\}$, entonces, para $i = 0, \dots, n-1$ se satisface

$$\Delta^k y_i = (-1)^k \bar{\Delta}^k y_i \quad (k = 1, \dots, n-i).$$

Demostración. Usaremos un razonamiento inductivo sobre k . Empecemos con $k = 1$,

$$\Delta y_i = y_{i+1} - y_i = -\bar{\Delta} y_i$$

supongamos ahora que para $1 \leq k < n-i$, se tiene que $\Delta^k y_i = (-1)^k \bar{\Delta}^k y_i$, entonces,

$$\begin{aligned}\Delta^{k+1} y_i &= \Delta^k y_{i+1} - \Delta^k y_i = (-1)^k \bar{\Delta}^k y_{i+1} - (-1)^k \bar{\Delta}^k y_i \\ &= (-1)^k (\bar{\Delta}^k y_{i+1} - \bar{\Delta}^k y_i) = (-1)(-1)^k (\bar{\Delta}^k y_i - \bar{\Delta}^k y_{i+1}) \\ &= (-1)^{k+1} \bar{\Delta}^{k+1} y_i,\end{aligned}$$

utilizando simplemente la hipótesis de inducción y las definiciones (1.7) y (1.8). Con esto se completa la inducción y por tanto el lema es cierto. \square

Así, gracias a este lema, obtenemos la expresión, del polinomio (1.8), que estábamos buscando. Esta es:

$$\begin{aligned}p(x) &= y_0 - \frac{\bar{\Delta} y_0}{h}(x-x_0) + \frac{\bar{\Delta}^2 y_0}{2h^2}(x-x_0)(x-x_1) - \frac{\bar{\Delta}^3 y_0}{6h^3}(x-x_0)(x-x_1)(x-x_2) \\ &\quad + \dots + (-1)^n \frac{\bar{\Delta}^n y_0}{n!h^n}(x-x_0)(x-x_1) \cdots (x-x_{n-1}). \quad (1.9)\end{aligned}$$

1.2.1. Métodos lineales multipaso

Estos métodos, como cualquier otro método multipaso, utilizan, para calcular y_{k+1} , la información en m puntos anteriores, pero con una forma particular, esta es:

$$y_{k+1} = \sum_{i=1}^m \alpha_i y_{k+1-i} + h \sum_{i=0}^m \beta_i f_{k+1-i} \quad (1.10)$$

donde, $f_j = f(x_j, y_j)$ y $m \in \mathbb{N}$. Este método es llamado lineal, ya que, y_{k+1} es una combinación lineal de y_i y f_i con $k - (m - 1) \leq i \leq k$. Daremos ahora una definición.

Definición 1.9.

Diremos que el método (1.10), es *explícito* si $\beta_0 = 0$ y que es *implícito* si $\beta_0 \neq 0$.

Cálculo del error

Para los métodos descritos por la ecuación (1.10), definiremos de manera natural, el error de discretización en un punto como sigue,

Definición 1.10.

Para los métodos lineales multipaso (1.10), definimos el error de discretización en un punto por:

$$L(x, h) = \frac{1}{h} \left[y(x+h) - \sum_{i=1}^m \alpha_i y(x - (i-1)h) \right] - \sum_{i=0}^m \beta_i y'(x - (i-1)h)$$

donde y es la función que satisface la ecuación (1.1).

Las definiciones de error local de discretización y orden para estos métodos están dadas por las definiciones (1.3) y (1.4) respectivamente. De acuerdo con la definición (1.10), para cualquier elección de $m \in \mathbb{N}$ y constantes α_i y β_i , es posible calcular el error de discretización en un punto expandiendo y y y' en serie de Taylor alrededor de x . Esto lo haremos, en las siguientes secciones, para algunos casos particulares e incluso, para estos mismos casos, probaremos algunos resultados acerca de su orden.

Métodos de Adams

Para construir este tipo de métodos, es necesario integrar la ecuación (1.1) sobre el intervalo $[x_k, x_{k+1}]$, de este modo, se obtienen las siguientes igualdades:

$$y(x_{k+1}) - y(x_k) = \int_{x_k}^{x_{k+1}} y'(x) dx = \int_{x_k}^{x_{k+1}} f(x, y(x)) dx. \quad (1.11)$$

Sin embargo, la integral de f puede ser muy difícil de calcular, motivo por el cual, sería útil sustituir f por algún polinomio p . Para encontrar este polinomio, supongamos que $y_{k+m}, \dots, y_k, y_{k-1}, \dots, y_{k-N}$, son aproximaciones a la solución en $x_{k+m}, \dots, x_k, x_{k-1}, \dots, x_{k-N}$ y donde la malla determinada por las x_i es equiespaciada ($x_i = x_0 + ih$) como hemos venido suponiendo. Entonces, si denotamos por f_i a $f(x_i, y_i)$ con $i = k - N, \dots, k + m$, tenemos que las f_i son aproximaciones a $f(x, y(x))$ en los puntos $x_{k+m}, \dots, x_k, x_{k-1}, \dots, x_{k-N}$. De este modo, tomaremos al polinomio p como el polinomio interpolador, de grado menor que $N + m + 1$, de los datos (x_i, f_i) con $i = k - N, \dots, k + m$, que sabemos que existe gracias al teorema 1.5. Así, podemos sustituir la integral de f , en (1.11), por la integral de

p que en principio podemos hacer de manera explícita para obtener los llamados métodos de Adams que a continuación escribimos,

$$y_{k+1} = y_k + \int_{x_k}^{x_{k+1}} p(x) dx. \quad (1.12)$$

Es importante notar que si $m \neq 0$ el método de Adams definido por la ecuación (1.12) resulta implícito, ya que ambos lados de dicha ecuación involucran los valores y_{k+1}, \dots, y_{k+m} . Así, para determinar estos valores son necesarias más ecuaciones. Si $m = 0$, el método resulta explícito. Ejemplos de ambas cosas son los métodos de Adams-Moulton y Adams-Bashforth respectivamente.

Métodos de Adams-Bashforth

Para este caso particular de método de Adams, se tiene que $m = 0$. Como un primer ejemplo, si $N = 0$, el polinomio p , de grado 0, que interpola a f_k es la misma constante f_k y al integrar esta constante se obtiene el método de Euler, que como sabemos es un método de un paso. Por esta razón, estamos interesados en el caso $N \neq 0$. Como un primer ejemplo, obtendremos el método de Adams-Bashforth cuando $N = 1$. De acuerdo con nuestra definición de métodos de Adams (1.12), en este caso, p , que denotaremos por p_1 , es el polinomio lineal que interpola los datos (x_{k-1}, f_{k-1}) y (x_k, f_k) , que aplicando la igualdad (1.9), para el polinomio interpolador de Newton, p_1 queda escrito como sigue:

$$p_1(x) = f_k - \frac{\bar{\Delta} f_k}{h}(x - x_k).$$

Basta ahora integrar p_1 en el intervalo $[x_k, x_{k+1}]$,

$$\begin{aligned} \int_{x_k}^{x_{k+1}} p_1(x) dx &= \int_{x_k}^{x_{k+1}} f_k - \frac{\bar{\Delta} f_k}{h}(x - x_k) dx \\ &= f_k h - \frac{\bar{\Delta} f_k}{h} \left(\int_{x_k}^{x_{k+1}} x dx - \int_{x_k}^{x_{k+1}} x_k dx \right) \\ &= f_k h - \frac{\bar{\Delta} f_k}{h} \left(\frac{x_{k+1}^2 - x_k^2}{2} - h x_k \right) \\ &= f_k h - \frac{\bar{\Delta} f_k}{h} \left(\frac{(x_k - x_{k+1})(x_k + x_{k+1})}{2} - h x_k \right) \\ &= f_k h - \bar{\Delta} f_k \left(\frac{x_{k+1} + x_k}{2} - x_k \right) \\ &= f_k h - \frac{\bar{\Delta} f_k}{2} h. \end{aligned}$$

Las igualdades anteriores se deben a que los nodos de la malla son equidistantes. De esta manera, de acuerdo con la forma general de los métodos de Adams (1.12), el método de Adams-Bashforth con $N = 1$, queda descrito por la siguiente ecuación,

$$y_{k+1} = y_k + h f_k - \frac{\bar{\Delta} f_k}{2} h = y_k + \frac{h}{2}(3f_k - f_{k-1}). \quad (1.13)$$

Como veremos más adelante éste es un método de segundo orden. Calcularemos ahora, el método de Adams-Bashforth cuando $N = 2$. En este caso buscamos el polinomio interpolador de los datos (x_{k-2}, f_{k-2}) , (x_{k-1}, f_{k-1}) y (x_k, f_k) . De nueva cuenta, utilizando la ecuación (1.9), tenemos que el polinomio que buscamos es:

$$p_2(x) = p_1(x) + \frac{\bar{\Delta}^2 f_k}{2h^2}(x - x_k)(x - x_{k-1}),$$

repetiendo el procedimiento que antes hicimos con p_1 , pero ahora con p_2 , obtendremos la integral de p_2 sobre el intervalo $[x_k, x_{k+1}]$, así,

$$\int_{x_k}^{x_{k+1}} p_2(x) dx = \int_{x_k}^{x_{k+1}} p_1(x) dx + \frac{\bar{\Delta}^2 f_k}{2h^2} \int_{x_k}^{x_{k+1}} (x - x_k)(x - x_{k-1}) dx. \quad (1.14)$$

De la ecuación anterior, como conocemos la integral de p_1 , sólo falta calcular la siguiente integral:

$$\begin{aligned} \int_{x_k}^{x_{k+1}} (x - x_k)(x - x_{k-1}) dx &= \int_{x_k}^{x_{k+1}} x^2 - x(x_k + x_{k-1}) + x_k x_{k-1} dx \\ &= \int_{x_k}^{x_{k+1}} x^2 dx + (x_k + x_{k-1}) \int_{x_k}^{x_{k+1}} x dx + (x_k x_{k-1}) \int_{x_k}^{x_{k+1}} 1 dx \end{aligned} \quad (1.15)$$

Calculemos estas integrales por separado, primero,

$$\begin{aligned} \int_{x_k}^{x_{k+1}} x^2 dx &= \frac{1}{3}(x_{k+1}^3 - x_k^3) = (x_{k+1} - x_k)(x_k^2 + x_{k+1}x_k + x_{k+1}^2) \\ &= \frac{h}{3}(x_k^2 + x_k x_{k+1} + x_{k+1}^2), \end{aligned} \quad (1.16)$$

donde la segunda igualdad se obtiene factorizando $h = x_{k+1} - x_k$. Calculemos ahora el segundo sumando de la expresión (1.15), este es:

$$\begin{aligned} -(x - x_k)(x - x_{k-1}) \int_{x_k}^{x_{k+1}} x dx &= -\frac{1}{2}(x - x_k)(x - x_{k-1})(x_{k+1}^2 - x_k^2) \\ &= -\frac{h}{2}(x_k + x_{k-1})(x_k - x_{k+1}), \end{aligned} \quad (1.17)$$

y por último,

$$(x_k x_{k-1}) \int_{x_k}^{x_{k+1}} 1 dx = (x_k x_{k-1})(x_{k+1} - x_k) = h(x_k x_{k-1}). \quad (1.18)$$

Sumando las ecuaciones (1.16), (1.17) y (1.18), de acuerdo con la expresión (1.15), obtenemos que:

$$\begin{aligned} \int_{x_k}^{x_{k+1}} (x - x_k)(x - x_{k-1}) dx &= h \left[\frac{1}{3}(x_k^2 + x_k x_{k+1} + x_{k+1}^2) - \frac{1}{2}(x_k + x_{k-1})(x_k - x_{k+1}) \right. \\ &\quad \left. + (x_k x_{k-1}) \right], \end{aligned}$$

sin embargo, el término entre corchetes es posible reducirlo, de modo que la ecuación anterior, queda de la siguiente manera:

$$\begin{aligned} \int_{x_k}^{x_{k+1}} (x - x_k)(x - x_{k-1}) dx &= h \left[\frac{1}{6}(x_{k+1} - x_k)(x_k + 2x_{k+1} - 3x_{k-1}) \right] \\ &= h \left[\frac{h}{6}(x_k + 2x_{k+1} - 3x_{k-1}) \right] \\ &= \frac{h^2}{6} [(x_k - x_{k-1}) + 2(x_{k+1} - x_{k-1})] \\ &= \frac{h^2}{6} (h + 2 \cdot 2h) = \frac{5}{6} h^3. \end{aligned}$$

Así, si sustituimos la ecuación anterior en (1.14), obtenemos que:

$$\begin{aligned} \int_{x_k}^{x_{k+1}} p_2(x) dx &= \int_{x_k}^{x_{k+1}} p_1(x) dx + \frac{\bar{\Delta}^2 f_k}{2h^2} \cdot \frac{5}{6} h^3 \\ &= \frac{h}{2} (3f_k - f_{k-1}) + \frac{5}{12} h (f_{k-2} - 2f_{k-1} + f_k), \end{aligned}$$

donde, la última igualdad se da, utilizando los cálculos hechos para p_1 en el caso $N = 1$ y porque $\bar{\Delta}^2 f_k = \bar{\Delta} f_{k-1} - \bar{\Delta} f_k = f_{k-2} - 2f_{k-1} + f_k$. De este modo, si agrupamos términos en la última ecuación, concluimos que:

$$\int_{x_k}^{x_{k+1}} p_2(x) dx = \frac{h}{12} (23f_k - 16f_{k-1} + 5f_{k-2}).$$

Habiendo calculado la integral de p_2 , podemos concluir que el método de Adams-Bashforth, cuando $N = 2$, queda descrito según (1.12) como sigue:

$$y_{k+1} = y_k + \frac{h}{12} (23f_k - 16f_{k-1} + 5f_{k-2}). \quad (1.19)$$

Con un procedimiento totalmente análogo, es posible calcular el método de Adams-Bashforth para N arbitraria, por supuesto, los cálculos se vuelven más complicados ya que el conjunto de datos que se interpolan crece y por tanto el grado del polinomio que los interpola también. Por ejemplo, cuando $N = 3$, el polinomio que se requiere para interpolar los datos, es un polinomio cúbico, así, el método de Adams-Bashforth, en este caso, queda descrito por la siguiente ecuación:

$$y_{k+1} = y_k + \frac{h}{24} (55f_k - 59f_{k-1} + 37f_{k-2} - 9f_{k-3}). \quad (1.20)$$

Podemos observar que todos los métodos de Adams-Bashforth son métodos lineales multipaso de acuerdo con la ecuación (1.10). También, de acuerdo con la definición 1.9, los métodos (1.13), (1.19) y (1.20) son explícitos, ya que si los escribimos en la forma de (1.10), tenemos que $\beta_0 = 0$ en los tres casos. De hecho, el método de Adams-Bashforth con N arbitraria es explícito, ya que el conjunto de datos que se deben interpolar, nunca involucra al valor f_{k+1} y por tanto el polinomio que se debe integrar tampoco.

Los métodos (1.13), (1.19) y (1.20), son métodos de segundo, tercer y cuarto orden. Probaremos que el método (1.13) es de segundo orden, las pruebas de que (1.19) y (1.20) son de tercer y cuarto son totalmente análogas. De acuerdo con la definición 1.10, tenemos que para el método (1.13), el error de discretización en un punto x , está dado por:

$$L(x, h) = \frac{1}{h} [y(x+h) - y(x)] - \frac{1}{2} [3y'(x) - y'(x-h)]. \quad (1.21)$$

Sin embargo, si expandimos en serie de Taylor a y y y' alrededor de $x+h$ y $x-h$ respectivamente, obtenemos que:

$$\begin{aligned} y'(x-h) &= y'(x) - hy''(x) + O(h^2), \\ y(x+h) &= y(x) + hy'(x) + \frac{h^2}{2}y''(x) + O(h^3), \end{aligned}$$

y por lo tanto,

$$\begin{aligned} 3y'(x) - y'(x-h) &= 2y'(x) + hy''(x) + O(h^2), \\ y(x+h) - y(x) &= hy'(x) + \frac{h^2}{2}y''(x) + O(h^3). \end{aligned}$$

Si sustituimos estas dos ecuaciones en (1.21), obtenemos que:

$$\begin{aligned} L(x, h) &= y'(x) + \frac{h}{2}y''(x) + \frac{O(h^3)}{h} - y'(x) - \frac{h}{2}y''(x) - \frac{O(h^2)}{2} \\ &= O(h^2). \end{aligned}$$

De esta manera, tenemos que:

$$L(h) = \max_{a \leq x \leq b-h} |L(x, h)| = O(h^2).$$

Así concluimos que el método (1.13) es de segundo orden. En la sección siguiente, analizaremos un tipo de método muy similar a los métodos de Adams-Bashforth.

Métodos de Adams-Moulton

Como vimos en la formulación de los métodos de Adams, para calcular y_{k+1} , es posible no sólo utilizar la información en x_k y puntos anteriores, x_{k-1}, \dots, x_{k-N} , como en los métodos de Adams-Bashforth, sino también información en puntos posteriores a x_k . El caso más simple, es considerar la información en los puntos $x_{k+1}, x_k, x_{k-1}, \dots, x_{k-N}$. Dicho de otro modo, el caso en el cual, de acuerdo a la formulación general de los métodos de Adams, $m = 1$, es el más simple y da lugar a los llamados métodos de Adams-Moulton.

La construcción de este tipo de métodos, es totalmente análoga a la construcción de los métodos de Adams-Bashforth. Construiremos los métodos de Adams-Moulton para los casos $N = 0$ y $N = 2$. De acuerdo con la formulación

de los métodos de Adams, si $N = 0$, entonces, en (1.12), p es el polinomio de grado uno que interpola los datos (x_k, f_k) y (x_{k+1}, f_{k+1}) y que denotaremos por p_1 . Al igual que hicimos antes, utilizaremos la formulación (1.9) del polinomio interpolador de Newton para encontrar p_1 . Así,

$$p_1(x) = f_{k+1} - \frac{\bar{\Delta} f_{k+1}}{h}(x - x_{k+1})$$

entonces,

$$\begin{aligned} \int_{x_k}^{x_{k+1}} p_1(x) dx &= h f_{k+1} - \frac{\bar{\Delta} f_{k+1}}{h} \int_{x_k}^{x_{k+1}} (x - x_{k+1}) dx \\ &= h f_{k+1} - \frac{\bar{\Delta} f_{k+1}}{h} \left(\int_{x_k}^{x_{k+1}} x^2 dx - x_{k+1} \int_{x_k}^{x_{k+1}} 1 dx \right) \\ &= h f_{k+1} - \frac{\bar{\Delta} f_{k+1}}{h} \left[\frac{1}{2}(x_{k+1}^2 - x_k^2) - x_{k+1}(x_{k+1} - x_k) \right] \\ &= h f_{k+1} - \frac{\bar{\Delta} f_{k+1}}{h} \left[\frac{1}{2}(x_{k+1} - x_k)(x_{k+1} + x_k) - h x_{k+1} \right] \\ &= h f_{k+1} - \frac{\bar{\Delta} f_{k+1}}{h} \left[\frac{h}{2}(x_{k+1} + x_k) - h x_{k+1} \right] \\ &= h f_{k+1} - \frac{\bar{\Delta} f_{k+1}}{2} (x_{k+1} + x_k - 2x_{k+1}) \\ &= h f_{k+1} + \frac{\bar{\Delta} f_{k+1}}{2} (x_{k+1} - x_k) \\ &= \frac{h}{2} (2f_{k+1} + f_k - f_{k+1}) = \frac{h}{2} (f_{k+1} + f_k). \end{aligned} \quad (1.22)$$

De esta manera, podemos concluir que el método de Adams-Moulton cuando $N = 0$ es:

$$y_{k+1} = y_k + \frac{h}{2} (f_{k+1} + f_k). \quad (1.23)$$

Este método es de segundo orden. Calcularemos ahora el método de Adams-Moulton cuando $N = 2$. En este caso, el polinomio que interpola los datos (x_{k+1}, f_{k+1}) , (x_k, f_k) , (x_{k-1}, f_{k-1}) y (x_{k-2}, f_{k-2}) queda descrito por la siguiente ecuación:

$$\begin{aligned} p_3(x) &= f_{k+1} - \frac{\bar{\Delta} f_{k+1}}{h}(x - x_{k+1}) + \frac{\bar{\Delta}^2 f_{k+1}}{2h^2}(x - x_{k+1})(x - x_k) \\ &\quad - \frac{\bar{\Delta}^3 f_{k+1}}{6h^3}(x - x_{k+1})(x - x_k)(x - x_{k-1}). \end{aligned}$$

Al observar esta última expresión, podemos escribir la integral de p_3 como sigue:

$$\int_{x_k}^{x_{k+1}} p_3(x) dx = \int_{x_k}^{x_{k+1}} p_1(x) dx + I - II, \quad (1.24)$$

donde,

$$I = \frac{\bar{\Delta}^2 f_{k+1}}{2h^2} \int_{x_k}^{x_{k+1}} (x - x_{k+1})(x - x_k) dx,$$

$$II = \frac{\bar{\Delta}^3 f_{k+1}}{6h^3} \int_{x_k}^{x_{k+1}} (x - x_{k+1})(x - x_k)(x - x_{k-1}) dx.$$

De la ecuación (1.24) ya hemos calculado la integral de p_1 , sin embargo, falta calcular las integrales I y II . Para I , calcularemos la siguiente integral:

$$\begin{aligned} \int_{x_k}^{x_{k+1}} (x - x_{k+1})(x - x_k) dx &= \int_{x_k}^{x_{k+1}} (x^2 - (x_{k+1} + x_k)x + x_k x_{k+1}) dx \\ &= \frac{1}{3}(x_{k+1}^3 - x_k^3) - \frac{1}{2}(x_{k+1} + x_k)(x_{k+1}^2 - x_k^2) + (x_k x_{k+1})(x_{k+1} - x_k) \\ &= \frac{1}{3}(x_{k+1} - x_k)(x_{k+1}^2 + x_k x_{k+1} + x_k^2) - \frac{1}{2}(x_{k+1} + x_k)^2(x_{k+1} - x_k) \\ &\quad + (x_k x_{k+1})(x_{k+1} - x_k) \\ &= \frac{h}{3}(x_{k+1}^2 + x_k x_{k+1} + x_k^2) - \frac{h}{2}(x_{k+1} + x_k)^2 + h(x_k x_{k+1}) \\ &= \frac{h}{6}(2x_{k+1}^2 + 2x_k^2 + 2x_k x_{k+1} - 3x_{k+1}^2 - 6x_k x_{k+1} - 3x_k^2 + 6x_k x_{k+1}) \\ &= \frac{h}{6}(-x_{k+1}^2 - x_k^2 + 2x_k x_{k+1}) = -\frac{h}{6}(x_{k+1} - x_k)^2 = -\frac{h}{6} \cdot h^2 \\ &= -\frac{h^3}{6} \end{aligned}$$

por lo tanto, gracias a esta última igualdad, podemos escribir,

$$I = -\frac{\bar{\Delta}^2 f_{k+1}}{2h^2} \cdot \frac{h^3}{6} = -\frac{h}{12} \bar{\Delta}^2 f_{k+1}. \quad (1.25)$$

Para calcular II , basta calcular la siguiente integral:

$$\begin{aligned} \int_{x_k}^{x_{k+1}} (x - x_{k+1})(x - x_k)(x - x_{k-1}) dx \\ &= \int_{x_k}^{x_{k+1}} (x^2 - (x_{k+1} + x_k)x + x_k x_{k+1})(x - x_{k-1}) dx \\ &= \int_{x_k}^{x_{k+1}} (x^3 - (x_{k+1} + x_k + x_{k-1})x^2 + ((x_{k-1} + x_k)x_{k+1} + x_{k-1}x_k) \\ &\quad - x_{k-1}x_k x_{k+1}) dx \\ &= \frac{1}{4}(x_{k+1}^4 - x_k^4) - \frac{1}{3}(x_{k+1} + x_k + x_{k-1})(x_{k+1}^3 - x_k^3) \\ &\quad + \frac{1}{2}((x_{k-1} + x_k)x_{k+1} + x_{k-1}x_k)(x_{k+1}^2 - x_k^2) - (x_{k+1}x_k x_{k-1})(x_{k+1} - x_k) \\ &= \frac{1}{4}(x_{k+1} - x_k)(x_{k+1} + x_k)(x_{k+1}^2 + x_k^2) \end{aligned}$$

$$\begin{aligned}
& -\frac{1}{3}(x_{k+1} + x_k + x_{k-1})(x_{k+1} - x_k)(x_{k+1}^2 + x_{k+1}x_k + x_k^2) \\
& + \frac{1}{2}((x_{k-1} + x_k)x_{k+1} + x_{k-1}x_k)(x_{k+1} - x_k)(x_{k+1} + x_k) \\
& - (x_{k+1}x_kx_{k-1})(x_{k+1} - x_k) \\
= & \frac{h}{4}(x_{k+1} + x_k)(x_{k+1}^2 + x_k^2) - \frac{h}{3}(x_{k+1} + x_k + x_{k-1})(x_{k+1}^2 + x_{k+1}x_k + x_k^2) \\
& + \frac{h}{2}((x_{k-1} + x_k)x_{k+1} + x_{k-1}x_k)(x_{k+1} + x_k) - h(x_{k+1}x_kx_{k-1}) \\
= & \frac{h}{24} \left[6(x_{k+1} + x_k)(x_{k+1}^2 + x_k^2) - 8(x_{k+1} + x_k + x_{k-1})(x_{k+1}^2 + x_{k+1}x_k + x_k^2) \right. \\
& \left. + 12((x_{k-1} + x_k)x_{k+1} + x_{k-1}x_k)(x_{k+1} + x_k) - 24(x_{k+1}x_kx_{k-1}) \right] \\
= & \frac{h}{24}(-2)(x_{k+1} - x_k)^2(x_{k+1} - 2x_{k-1} + x_k) \\
= & -\frac{h}{12}(x_{k+1} - x_k)^2(x_{k+1} - x_{k-1} + x_k - x_{k-1}) \\
= & -\frac{h}{12} \cdot h^2(2h + h) = -\frac{h^4}{4}.
\end{aligned}$$

Así, de acuerdo con la última igualdad, tenemos que:

$$II = -\frac{\bar{\Delta}^3 f_{k+1}}{6h^3} \cdot \frac{h^4}{4} = -\frac{h}{24} \bar{\Delta}^3 f_{k+1}. \quad (1.26)$$

Finalmente, si integramos los cálculos hechos para la integral de p_1 en (1.22) y las ecuaciones (1.25) y (1.26) en (1.24), obtenemos que:

$$\int_{x_k}^{x_{k+1}} p_3(x) dx = \frac{h}{2}(f_{k+1} + f_k) - \frac{h}{12} \bar{\Delta}^2 f_{k+1} + \frac{h}{24} \bar{\Delta}^3 f_{k+1}. \quad (1.27)$$

Sin embargo, sabemos que:

$$\bar{\Delta}^2 f_{k+1} = \bar{\Delta} f_k - \bar{\Delta} f_{k+1} = f_{k-1} - 2f_k + f_{k+1} \quad (1.28)$$

y también que,

$$\begin{aligned}
\bar{\Delta}^3 f_{k+1} &= \bar{\Delta}^2 f_k - \bar{\Delta}^2 f_{k+1} = \bar{\Delta} f_{k-1} - \bar{\Delta} f_k - \bar{\Delta} f_k + \bar{\Delta} f_{k+1} \\
&= f_{k-2} - 3f_{k-1} + 3f_k - f_{k+1}.
\end{aligned} \quad (1.29)$$

Si sustituimos la ecuaciones (1.28) y (1.29) en (1.27), obtenemos que:

$$\begin{aligned}
\int_{x_k}^{x_{k+1}} p_3(x) dx &= \frac{h}{24} \left[12(f_{k+1} + f_k) - 2(f_{k-1} - 2f_k + f_{k+1}) + f_{k-2} - 3f_{k-1} \right. \\
&\quad \left. + 3f_k - f_{k+1} \right] \\
&= \frac{h}{24} (9f_{k+1} + 19f_k - 5f_{k-1} + f_{k-2}).
\end{aligned} \quad (1.30)$$

Con lo que finalmente, el método de Adams-Moulton, cuando $N = 2$, queda escrito de la siguiente manera:

$$y_{k+1} = y_k + \frac{h}{24}(9f_{k+1} + 19f_k - 5f_{k-1} + f_{k-2}). \quad (1.31)$$

Este es un método de cuarto orden como demostraremos más adelante.

Una observación importante es que en las ecuaciones (1.23) y (1.31), f_{k+1} es desconocido, ya que necesitamos y_{k+1} para calcular $f_{k+1} = f(x_{k+1}, y_{k+1})$ y y_{k+1} es desconocido (es el valor que deseamos calcular). Por este motivo, las ecuaciones (1.23) y (1.31) tienen como incógnita al valor y_{k+1} . Es por ello que los métodos de Adams-Moulton son llamados implícitos, es decir que de acuerdo con la definición 1.9, $\beta_0 \neq 0$ para cualquier método de Adams-Moulton (con N arbitraria).

Ahora probaremos que el método (1.31) es un método cuarto orden. Para ello, de acuerdo con la definición (1.10), tenemos que, para este método, la siguiente igualdad se cumple:

$$hL(x, h) = [y(x+h) - y(x)] - \frac{h}{24} [9y'(x+h) + 19y'(x) - 5y'(x-h) + y'(x-2h)]. \quad (1.32)$$

Para acotar la expresión (1.32), desarrollemos primero $y(x+h)$ en su serie de Taylor alrededor de x :

$$y(x+h) = y(x) + hy'(x) + \frac{h^2}{2}y^{(2)}(x) + \frac{h^3}{6}y^{(3)}(x) + \frac{h^4}{24}y^{(4)}(x) + O(h^5)$$

o bien,

$$y(x+h) - y(x) = hy'(x) + \frac{h^2}{2}y^{(2)}(x) + \frac{h^3}{6}y^{(3)}(x) + \frac{h^4}{24}y^{(4)}(x) + O(h^5). \quad (1.33)$$

Ahora bien, si expandemos también en serie de Taylor a $y'(x+h)$, $y'(x-h)$ y $y'(x-2h)$ alrededor de x , tenemos que:

$$\begin{aligned} y'(x+h) &= y'(x) + hy^{(2)}(x) + \frac{h^2}{2}y^{(3)}(x) + \frac{h^3}{6}y^{(4)}(x) + O(h^4) \\ y'(x-h) &= y'(x) - hy^{(2)}(x) + \frac{h^2}{2}y^{(3)}(x) - \frac{h^3}{6}y^{(4)}(x) + O(h^4) \\ y'(x-2h) &= y'(x) - 2hy^{(2)}(x) + 2h^2y^{(3)}(x) - \frac{8}{6}h^3y^{(4)}(x) + O(h^4) \end{aligned} \quad (1.34)$$

de donde,

$$9y'(x+h) = 9y'(x) + 9hy^{(2)}(x) + \frac{9}{2}h^2y^{(3)}(x) + \frac{9}{6}h^3y^{(4)}(x) + O(h^4) \quad (1.35)$$

$$-5y'(x-h) = -5y'(x) + 5hy^{(2)}(x) - \frac{5}{2}h^2y^{(3)}(x) + \frac{5}{6}h^3y^{(4)}(x) + O(h^4), \quad (1.36)$$

así, si utilizamos las ecuaciones (1.34), (1.35) y (1.36) obtenemos que

$$\begin{aligned} 9y'(x+h) + 19y'(x) - 5y'(x-h) + y'(x-2h) \\ = 24y'(x) + 12hy^{(2)} + 4h^2y^{(3)}(x) + h^3y^{(4)}(x) + O(h^4) \end{aligned}$$

por tanto, de acuerdo con la ecuación (1.33),

$$\begin{aligned} \frac{h}{24} [9y'(x+h) + 19y'(x) - 5y'(x-h) + y'(x-2h)] \\ = y(x+h) - y(x) + \frac{h}{24} O(h^4) = y(x+h) - y(x) + O(h^5), \end{aligned}$$

si sustituimos esta ecuación en (1.32), tenemos que:

$$hL(x, h) = O(h^5),$$

que finalmente se transforma en

$$L(x, h) = \frac{O(h^5)}{h} = O(h^4).$$

De esta manera, tenemos que:

$$L(h) = \max_{a \leq x \leq b-h} |L(x, h)| = O(h^4).$$

Con lo que concluimos que el método (1.31) es de cuarto orden.

1.2.2. Consideraciones sobre los métodos multipaso

Una primera observación que se puede hacer, es que para cualquiera de los métodos de Adams-Moulton o Adams-Bashforth, resulta imposible aplicarlos desde el inicio, por ejemplo para $k = 0$, pues se requiere de evaluar f en puntos que no existen. Veamos esto con un ejemplo. Considere el método (1.20), si quisiéramos calcular y_1 , tendríamos que calcular los valores f_0 , f_{-1} , f_{-2} y f_{-3} , es decir, tenemos que calcular f en valores que no existen (exceptuando f_0). De esta manera, el método (1.20), solo se puede utilizar para $k \geq 3$. De manera análoga, los métodos de Adams-Moulton presentan el mismo problema. Una manera de solucionar el problema de obtener los valores iniciales (en nuestro ejemplo y_1 , y_2 y y_3), es utilizar un método de un paso, por ejemplo el de Runge-Kutta de cuarto orden o cualquiera que hayamos estudiado, sin embargo, debemos tener cuidado con la precisión de los métodos que combinamos, por ejemplo, para calcular los valores iniciales en el método (1.20), debemos utilizar un método de tercer orden, pues de lo contrario, se perdería la precisión que habíamos calculado. En general, para calcular los valores iniciales de un método multipaso, debemos usar un método que tenga el mismo orden de precisión que el método multipaso que finalmente utilizaremos.

Para calcular los valores iniciales que se necesitan para arrancar un método multipaso podemos, como hemos dicho, buscar un método de un paso que tenga

el mismo orden de precisión o, podríamos pensar, en usar un método de un paso para el primer valor inicial, uno de dos pasos para el segundo y así sucesivamente hasta encontrar los necesarios. Sin embargo, esta última forma tiene el inconveniente, de que usualmente el orden de precisión en los métodos multipaso está estrechamente relacionado con los puntos necesarios para arrancarlo, lo que provocaría que los órdenes de precisión no coincidieran. Una manera de corregir este inconveniente, es reducir el tamaño del paso, introduciendo así más puntos intermedios logrando aumentar la precisión de los métodos multipaso utilizados en la obtención de los valores iniciales. De nuevo para la obtención de los valores que aproximan a y en dichos puntos intermedios serán necesarios métodos con el mismo orden de precisión.

Los métodos de Adams-Moulton presentan adicionalmente un problema más, son implícitos, es decir que para obtener el valor y_{k+1} es necesario resolver una ecuación no lineal, como por ejemplo la ecuación (1.31), lo que representa un problema que no necesariamente es sencillo de resolver. Para resolver dicho inconveniente, se utiliza un tipo de método llamado predictor-corrector, éste utiliza un valor "prededido" de y_{k+1} , que denotaremos por $y_{k+1}^{(p)}$, para calcular, a través del método implícito, un nuevo valor "corregido" de y_{k+1} . Un ejemplo de método predictor-corrector, es la combinación de los métodos (1.20) y (1.31) que por supuesto, son de cuarto orden. Naturalmente, como el método de Adams-Bashforth es explícito, lo utilizaremos para calcular $y_{k+1}^{(p)}$ y el de Adams-Moulton para obtener el valor corregido de y_{k+1} . Así, pues el método predictor-corrector queda descrito por las siguientes ecuaciones:

$$\begin{aligned} y_{k+1}^{(p)} &= y_k + \frac{h}{24}(55f_k - 59f_{k-1} + 37f_{k-2} - 9f_{k-3}), \\ f_{k+1}^{(p)} &= f(x_{k+1}, y_{k+1}^{(p)}), \\ y_{k+1} &= y_k + \frac{h}{24}(f_{k+1}^{(p)} + 19f_k - 5f_{k-1} + f_{k-2}). \end{aligned}$$

El método anterior, puede utilizarse tantas veces como sea necesario para obtener valores de y_{k+1} tan precisos como se requieran, de hecho, dicho método puede ser utilizado como un procedimiento iterativo para resolver la ecuación no lineal (1.31).

1.3. Sistemas de ecuaciones

En esta sección, haremos los comentarios pertinentes para la aplicación de los métodos numéricos, expuestos previamente, a sistemas de ecuaciones diferenciales ordinarias. La generalización inmediata del problema (1.1), a sistemas de ecuaciones, queda descrita por el siguiente sistema:

$$\frac{dy}{dx} = f(x, y(x)) \quad a \leq x \quad \text{con condición inicial} \quad y(a) = \hat{y} \quad (1.37)$$

donde, por supuesto se tiene que,

$$\mathbf{y}(x) = \begin{pmatrix} y_1(x) \\ y_2(x) \\ \vdots \\ y_n(x) \end{pmatrix} \quad \text{y} \quad \mathbf{f}(x, \mathbf{y}(x)) = \begin{pmatrix} f_1(x, y_1(x), \dots, y_n(x)) \\ f_2(x, y_1(x), \dots, y_n(x)) \\ \vdots \\ f_n(x, y_1(x), \dots, y_n(x)) \end{pmatrix}$$

el vector $\hat{\mathbf{y}}$ denota al vector constante $\mathbf{y}(a)$, es decir los valores iniciales para las componentes de \mathbf{y} .

La generalización de los métodos resulta inmediata si se utiliza la notación vectorial, por ejemplo el método de Heun, para sistemas, queda de la siguiente manera:

$$\mathbf{y}_{k+1} = \mathbf{y}_k + \frac{h}{2} \left[\mathbf{f}(x_k, \mathbf{y}_k) + \mathbf{f}(x_{k+1}, \mathbf{y}_k + h\mathbf{f}(x_k, \mathbf{y}_k)) \right].$$

De forma totalmente análoga pero más compleja, se sigue la generalización de los métodos de orden mayor y de los métodos de Adams.

Capítulo 2

Sistemas Hamiltonianos

2.1. Nociones de cálculo de variaciones

En la teoría del cálculo de variaciones se estudian funciones, en nuestro caso reales, cuyas variables son a su vez funciones, curvas o superficies. Estas funciones de funciones se conocen como funcionales. En general, un funcional J , es una función de un espacio vectorial ¹ X , cuyos elementos por lo general serán funciones en \mathbb{R} . Sin embargo, para poder realizar cálculo en estos espacios es necesario el concepto de norma, por lo que en general se supondrá que los funcionales están definidos sobre espacios normados. La definición de continuidad para funcionales resulta natural.

Definición 2.1 (Continuidad de funcionales).

Si J es un funcional de X en \mathbb{R} , diremos que es continuo en $x_0 \in X$ si para toda $\epsilon > 0$, existe $\delta > 0$ tal que

$$|J(x) - J(x_0)| < \epsilon \quad \text{siempre que} \quad \|x - x_0\| < \delta.$$

La linealidad de un funcional queda definida como sigue:

Definición 2.2 (Funcional lineal).

Si J es un funcional de X en \mathbb{R} , diremos que es lineal si para cualesquiera $x, y \in X$ y $\alpha \in \mathbb{R}$, se tiene que

(1) $J(\alpha x) = \alpha J(x)$ (homogeneidad)

(2) $J(x + y) = J(x) + J(y)$ (aditividad).

A continuación enunciamos un lema que suele ser útil y que involucra los dos conceptos antes definidos.

¹También llamados espacios lineales.

Lema 2.1.

Sea J un funcional lineal definido sobre el espacio lineal normado X , entonces las siguientes propiedades son equivalentes:

- (1) J es continuo.
- (2) J es continuo en cero.
- (3) $\|J\| = \sup_{\|x\|=1} \{|J(x)|\} \leq \infty$.

Demostración. Resulta trivial (2) a partir de (1). Supongamos pues que (2) se cumple; en este caso, dada $\epsilon > 0$ existe $\delta > 0$ tal que si $\|x\| \leq \delta$ entonces $|J(x)| = |J(x) - J(0)| \leq \epsilon$ ya que al ser lineal J , se tiene que $J(0) = 0$.

Sea $x \in X$ tal que $\|x\| = 1$, entonces, como $\|\delta x\| = \delta \leq \delta$, tenemos que

$$|J(x)| = |J(\delta x)|/\delta \leq \epsilon/\delta,$$

de este modo, ϵ/δ es cota superior del conjunto $\{|J(x)| : \|x\| = 1\}$, por tanto $\|J\| \leq \epsilon/\delta$ y concluimos que (2) implica (3).

Para probar que (3) implica (1), consideremos $x_0 \in X$ y $\epsilon > 0$. Sea $x \in X$ distinta de x_0 , entonces,

$$|J(x) - J(x_0)| = |J(x - x_0)| = |x - x_0| |J((x - x_0)/|x - x_0|)| \leq |x - x_0| \|J\|$$

por lo tanto si escogemos $\delta = \epsilon/\|J\|$ tendremos que J es continuo en 0. □

Como nuestro interés principal es el estudio de extremos de funcionales, se hace necesaria la noción de diferencial y derivada direccional, a continuación se enuncia el concepto análogo al de derivada direccional.

Definición 2.3 (Derivada de Gâteaux).

Si J es un funcional de X en \mathbb{R} , diremos que J tiene una derivada de Gâteaux en $x_0 \in X$ en la dirección $h \in X$, si,

- (1) existe una vecindad del cero para la cual $g(t) = J(x_0 + th)$ está definida.
- (2) $g(t)$ es derivable en 0.

En tal caso, la derivada de Gâteaux se denota por $D_h J(x_0)$ y es igual a $g'(0)$.

La noción de diferencial para funcionales, queda descrita mediante la siguiente definición.

Definición 2.4 (Derivada de Fréchet).

Si J es un funcional de X en \mathbb{R} , diremos que J tiene una derivada de Fréchet en $x_0 \in X$ si existe un funcional lineal continuo Φ tal que,

$$\Phi(h) + \epsilon \|h\| = J(x_0 + h) - J(x_0)$$

y $|\epsilon| \rightarrow 0$ cuando $\|h\| \rightarrow 0$. A Φ se le llama la variación o derivada de Fréchet de J y se denota por $DJ(x_0)$.

Existe cierta analogía entre el concepto de diferencial y derivada del cálculo de variable real y el concepto de derivada de Gâteaux y de Fréchet del cálculo de variaciones. Es posible probar que la derivada de Fréchet es única, sin embargo, antes de enunciar propiamente este resultado, probaremos un lema.

Lema 2.2.

Si Φ es un funcional lineal continuo definido sobre X tal que $\Phi(h)/\|h\| \rightarrow 0$ cuando $\|h\| \rightarrow 0$ entonces $\Phi \equiv 0$.

Demostración. Supongamos que existe $h_0 \neq 0$ en X tal que $\Phi(h_0) \neq 0$. Consideremos ahora la sucesión $h_n = h_0/n$ con $n \in \mathbb{N}$, así,

$$\|h_n\| = \frac{1}{n} \|h_0\| \rightarrow 0 \quad \text{cuando } n \rightarrow \infty$$

y por otro lado,

$$\frac{\Phi(h_n)}{\|h_n\|} = \frac{\Phi(\frac{h_0}{n})}{\frac{1}{n} \|h_0\|} = \frac{\frac{1}{n} \Phi(h_0)}{\frac{1}{n} \|h_0\|} = \frac{\Phi(h_0)}{\|h_0\|} = \lambda \neq 0 \quad \text{pues } h_0 \neq 0$$

entonces, existe el límite de $\Phi(h_n)/\|h_n\|$ cuando $n \rightarrow \infty$ y además,

$$\lim_{n \rightarrow \infty} \frac{\Phi(h_n)}{\|h_n\|} = \lambda \neq 0$$

lo cual contradice nuestra hipótesis, por lo tanto $\Phi \equiv 0$

□

Ahora sí, estamos en posición de probar el siguiente teorema.

Teorema 2.3 (Unicidad de la derivada de Fréchet).

Si existe la derivada de Fréchet de un funcional es única.

Demostración. Sea J un funcional para el cual existen dos derivadas de Fréchet en algún elemento del dominio de J . Sean estas derivadas Φ_1 y Φ_2 , entonces de acuerdo con la definición, se satisface que,

$$\Phi_1(h) + \epsilon_1 \|h_n\| = \Phi_2(h) + \epsilon_2 \|h_n\|$$

de donde se sigue que, si $g(h) = \Phi_1(h) - \Phi_2(h)$, entonces, $g(h) = \|h_n\| (\epsilon_2 - \epsilon_1)$, además, por su definición, g es un funcional lineal continuo tal que $g(h)/\|h_n\| = \epsilon_2 - \epsilon_1 \rightarrow 0$ cuando $\|h_n\| \rightarrow 0$ pues $\epsilon_1, \epsilon_2 \rightarrow 0$ cuando $\|h_n\| \rightarrow 0$ (de acuerdo a la definición (2.4)). Así, concluimos por el lema (2.2) que $g \equiv 0$ y por lo tanto $\Phi_1 \equiv \Phi_2$.

□

A continuación, se enuncian un par de resultados que relacionan los dos conceptos de derivada (el de Gateaux y el de Fréchet) que hemos definido para funcionales.

Teorema 2.4.

Si el funcional J definido sobre el espacio lineal normado X tiene derivada de Fréchet en $x_0 \in X$, entonces también tiene derivada de Gâteaux en x_0 en toda dirección $h \in X$ y además $D_h J(x_0) = DJ(x_0)(h)$.

Demostración. Como J es Fréchet-diferenciable en x_0 , entonces,

$$J(x_0 + th) - J(x_0) = DJ(x_0)(th) + \epsilon \|th\|$$

así,

$$\frac{J(x_0 + th) - J(x_0)}{t} = \frac{1}{t} DJ(x_0)(th) + \epsilon \frac{|t|}{t} \|h\| = DJ(x_0)(h) + \epsilon \frac{|t|}{t} \|h\|$$

pero $\|th\| \rightarrow 0$ si $t \rightarrow 0$, entonces, $\epsilon \rightarrow 0$ si $t \rightarrow 0$ y como $-1 \leq |t|/t \leq 1$ concluimos que $\epsilon(|t|/t) \|h\| \rightarrow 0$ cuando $t \rightarrow 0$. Por lo tanto, podemos escribir

$$\lim_{t \rightarrow 0} \frac{J(x_0 + th) - J(x_0)}{t} = DJ(x_0)(h)$$

de esta manera concluimos que $D_h J(x_0) = DJ(x_0)(h)$. □

Un resultado análogo al precedente pero en dirección contraria se enuncia a continuación.

Teorema 2.5.

Sea J un funcional definido sobre el espacio lineal normado X para el cual existe la derivada de Gâteaux en una vecindad de $x_0 \in X$ en toda dirección $h \in X$. Si la derivada de Gâteaux $D_h J$ es continua en dicha vecindad y lineal en h entonces J tiene derivada de Fréchet en x_0 con $DJ(x_0)(h) = D_h J(x_0)$.

Demostración. Sea $V_\delta(x_0)$ la vecindad donde la derivada de Gâteaux de J existe en toda dirección $h \in X$ y es continua. Consideremos $h \in X$ tal que $\|h\| = 1$, entonces $x = x_0 + th \in V_\delta(x_0)$ con $|t| < \delta$. Por hipótesis la función $f(t) = J(x_0 + th)$ tiene derivada para $|t| < \delta$ ya que $f'(t) = D_h J(x_0 + th)$. Por otro lado, por hipótesis, también tenemos que $D_h J$ es continuo para $x \in V_\delta(x_0)$, entonces $f'(t)$ es una función continua para $|t| < \delta$.

Por el teorema del valor medio, existe ξ con $0 < \xi < t$ tal que $f(t) = f(0) + f'(\xi)t$, es decir,

$$J(x) = J(x_0) + tD_h J(x_0 + \xi h) = J(x_0) + tD_h J(x_0) + t(D_h J(x_0 + \xi h) - D_h J(x_0)),$$

y usando finalmente la hipótesis de linealidad en h de $D_h J(x_0)$, concluimos

$$J(x) = J(x_0) + D_{th} J(x_0) + \epsilon t$$

donde ϵ no depende de la dirección h y tiene la propiedad de tender a cero si t tiende a cero. □

La extensión natural del concepto de extremo para un funcional queda establecida por la siguiente definición.

Definición 2.5 (Extremo de un funcional).

Sea J un funcional definido sobre un espacio lineal normado X . Diremos que J tiene un extremo local en $x_0 \in X$ si $J(x) - J(x_0)$ no cambia de signo para todo $x \in X$ perteneciente a alguna vecindad de x_0 para el cual J esté definido.

En la definición anterior, si $J(x) - J(x_0) \geq 0$ se dice que J alcanza un mínimo local en x_0 y si $J(x) - J(x_0) \leq 0$ se dice que J tiene un máximo local en x_0 .

Como es de esperarse, existe una relación estrecha entre los extremos de un funcional y sus derivadas.

Teorema 2.6.

Sea J un funcional definido sobre el espacio lineal normado X . Si J tiene un extremo local en $x_0 \in X$, entonces $D_h J(x_0) = 0$ en todas las direcciones h donde la derivada de Gâteaux esté definida. En particular si J tiene una derivada de Fréchet en x_0 , se tiene que $DJ(x_0) \equiv 0$.

Demostración. Consideremos $h \in X$ una dirección donde la derivada de Gâteaux existe, entonces la función $f(t) = J(x_0 + th)$ está definida para t en alguna vecindad del 0 y tiene un extremo en $t = 0$, por lo tanto $f'(0) = 0$, es decir $D_h J(x_0) = 0$. Si J tuviera una derivada de Fréchet en x_0 , entonces por el teorema (2.4) como la derivada de Gâteaux en toda dirección admisible es cero, se concluye que $DJ(x_0) \equiv 0$. □

2.2. Ecuación de Euler-Lagrange

Antes de entrar en materia, probaremos un par de lemas que serán de gran utilidad.

Lema 2.7.

Si $\alpha \in C[a, b]$ y si $\int_a^b \alpha(x)h'(x) dx = 0$ para toda $h \in C^1[a, b]$ que cumple con $h(a) = h(b) = 0$ entonces $\alpha(x) = c$ para toda $x \in [a, b]$ con $c \in \mathbb{R}$ constante.

Demostración. Consideremos la constante c y la función h definidas como sigue:

$$c = \frac{1}{b-a} \int_a^b \alpha(x) dx \quad \text{y} \quad h(x) = \int_a^x (\alpha(\xi) - c) d\xi.$$

Como $\alpha(x)$ es continua en $[a, b]$, entonces por el teorema fundamental del cálculo $h(x)$ es diferenciable con primera derivada continua y además $h(a) = h(b) = 0$.

Así,

$$\begin{aligned} \int_a^b [\alpha(x) - c]h'(x) dx &= \int_a^b \alpha(x)h'(x) dx - c \int_a^b h'(x) dx \\ &= \int_a^b \alpha(x)h'(x) dx - c[h(b) - h(a)] \\ &= \int_a^b \alpha(x)h'(x) dx \\ &= 0 \quad (\text{por hipótesis}). \end{aligned}$$

Por otro lado,

$$\int_a^b [\alpha(x) - c]h'(x) dx = \int_a^b [\alpha(x) - c]^2 dx,$$

de donde concluimos que

$$\int_a^b [\alpha(x) - c]^2 dx = 0.$$

Como además se tiene que $[\alpha(x) - c]^2 \geq 0$ y continua en $[a, b]$, tenemos que $\alpha(x) - c = 0$ y por lo tanto $\alpha(x) = c$ para todo $x \in [a, b]$. \square

Lema 2.8.

Si $\alpha, \beta \in C[a, b]$ y si

$$\int_a^b [\alpha(x)h(x) + \beta(x)h'(x)] dx = 0$$

para toda $h \in C^1[a, b]$ que cumple con $h(a) = h(b) = 0$, entonces β es diferenciable y $\beta'(x) = \alpha(x)$ para toda $x \in [a, b]$.

Demostración. Consideremos

$$A(x) = \int_0^x a(\xi) d\xi.$$

Entonces, integrando por partes tenemos que

$$\int_a^b \alpha(x)h(x) dx = \alpha(x)h(x)|_a^b - \int_a^b A(x)h'(x) dx = - \int_a^b A(x)h'(x) dx$$

ya que h se anula en a y b . De esta manera,

$$\begin{aligned} \int_a^b [\alpha(x)h(x) + \beta(x)h'(x)] dx &= \int_a^b \alpha(x)h(x) dx + \int_a^b \beta(x)h'(x) dx \\ &= - \int_a^b A(x)h'(x) dx + \int_a^b \beta(x)h'(x) dx \\ &= \int_a^b [\beta(x) - A(x)]h'(x) dx \\ &= 0. \end{aligned}$$

Ahora, $\beta - A \in C[a, b]$ pues $\beta, A \in C[a, b]$, entonces por el lema anterior, $\beta(x) - A(x) = cte$ para toda $x \in [a, b]$, es decir que $\beta(x) = A(x) + cte$, de donde, de acuerdo con la definición de $A(x)$ se obtiene que β es diferenciable y $\beta'(x) = A'(x) = \alpha(x)$ para toda $x \in [a, b]$.

□

Deduciremos la ecuación de Euler-Lagrange de un problema variacional. Para ello debemos considerar funcionales de la forma:

$$J(\mathbf{y}) = \int_a^b F(x, \mathbf{y}, \mathbf{y}') dx \quad (2.1)$$

donde J está definido para curvas en \mathbb{R}^n continuamente diferenciables con parámetro en $[a, b]$ y norma $\|\cdot\|_1$ ². De esta manera,

$$\mathbf{y}(x) = \begin{pmatrix} y_1(x) \\ y_2(x) \\ \vdots \\ y_n(x) \end{pmatrix} \in \mathbb{R}^n \quad \text{con} \quad x \in [a, b]$$

y $y_i \in C^1[a, b]$ ($i = 1, \dots, n$). Supondremos que la función $F(x, y_1, \dots, y_n, z_1, \dots, z_n)$, involucrada en la ecuación (2.1), es continuamente diferenciable en $[a, b] \times \mathbb{R}^{2n}$. Es decir, sus derivadas parciales $F_{y_1}, \dots, F_{y_n}, F_{z_1}, \dots, F_{z_n}$ también son funciones continuas en $[a, b] \times \mathbb{R}^{2n}$.

El problema variacional del que se deduce la ecuación de Euler-Lagrange consiste en buscar condiciones necesarias para los extremos de funcionales con la forma descrita mediante la ecuación (2.1) y que además satisfacen las condiciones de frontera

$$\mathbf{y}(a) = \mathbf{a} \in \mathbb{R}^n \quad \text{y} \quad \mathbf{y}(b) = \mathbf{b} \in \mathbb{R}^n. \quad (2.2)$$

Con la finalidad de encontrar dichas condiciones para los extremos de un funcional como J , primero calcularemos su derivada de Fréchet. Consideremos \mathbf{h} una curva en \mathbb{R}^n para la cual J está definido, entonces,

$$J(\mathbf{y} + \mathbf{h}) - J(\mathbf{y}) = \int_a^b F(x, \mathbf{y} + \mathbf{h}, \mathbf{y}' + \mathbf{h}') - F(x, \mathbf{y}, \mathbf{y}') dx. \quad (2.3)$$

Si usamos el teorema de Taylor para $F(x, y_1 + h_1, \dots, y_n + h_n, y'_1 + h'_1, \dots, y'_n + h'_n)$ (pensando fijo a x) alrededor del punto $(y_1, \dots, y_n, y'_1, \dots, y'_n)$ obtenemos que

$$F(x, \mathbf{y} + \mathbf{h}, \mathbf{y}' + \mathbf{h}') = F(x, \mathbf{y}, \mathbf{y}') + \sum_{i=1}^n F_{y_i}(x, \mathbf{y}, \mathbf{y}')h_i + F_{z_i}(x, \mathbf{y}, \mathbf{y}')h'_i + s(x, \mathbf{h}), \quad (2.4)$$

²Recordemos que $\|\mathbf{y}\|_1 = \max_{a \leq x \leq b} \|\mathbf{y}(x)\| + \max_{a \leq x \leq b} \|\mathbf{y}'(x)\|$.

donde $s(x, \mathbf{h})$ denota la suma de los términos no lineales en $h_1, \dots, h_n, h'_1, \dots, h'_n$.

Si sustituimos la expresión (2.4) en (2.3), obtenemos que

$$J(\mathbf{y} + \mathbf{h}) - J(\mathbf{y}) = \int_a^b \sum_{i=1}^n F_{y_i} h_i + F_{z_i} h'_i dx + \int_a^b s(x, \mathbf{h}) dx, \quad (2.5)$$

donde las derivadas parciales F_{y_i}, F_{z_i} ($i = 1, \dots, n$) están evaluadas en $(x, \mathbf{y}, \mathbf{y}')$, h_i ($i = 1, \dots, n$) en x y s en (x, \mathbf{h}) .

Si definimos a la función $\Phi(\mathbf{h})$ y a la constante ϵ como sigue,

$$\Phi(\mathbf{h}) = \int_a^b \sum_{i=1}^n F_{y_i} h_i + F_{z_i} h'_i dx \quad \text{y} \quad \epsilon = \frac{1}{\|\mathbf{h}\|_1} \int_a^b s(x, \mathbf{h}) dx,$$

tenemos que

$$J(\mathbf{y} + \mathbf{h}) - J(\mathbf{y}) = \Phi(\mathbf{h}) + \|\mathbf{h}\|_1 \epsilon. \quad (2.6)$$

Veremos ahora que Φ es un funcional lineal. Tomemos \mathbf{h}_1 y \mathbf{h}_2 curvas en el dominio de J y $c \in \mathbb{R}$ entonces

$$\Phi(\mathbf{h}_1 + c\mathbf{h}_2) = \int_a^b \sum_{i=1}^n F_{y_i}(h_{1i} + ch_{2i}) + F_{z_i}(h'_{1i} + ch'_{2i}) dx,$$

si factorizamos c y agrupamos términos, obtenemos

$$\begin{aligned} \Phi(\mathbf{h}_1 + c\mathbf{h}_2) &= \int_a^b \sum_{i=1}^n F_{y_i} h_{1i} + F_{z_i} h'_{1i} dx + c \int_a^b \sum_{i=1}^n F_{y_i} h_{2i} + F_{z_i} h'_{2i} dx \\ &= \Phi(\mathbf{h}_1) + c\Phi(\mathbf{h}_2). \end{aligned}$$

Es decir Φ es lineal. Ahora demostraremos la continuidad de Φ en 0,

$$\begin{aligned} |\Phi(\mathbf{h}) - \Phi(0)| &= |\Phi(\mathbf{h})| \leq \int_a^b \left| \sum_{i=1}^n F_{y_i} h_i + F_{z_i} h'_i \right| dx \\ &\leq \int_a^b \sum_{i=1}^n |F_{y_i}| |h_i| + |F_{z_i}| |h'_i| dx \\ &\leq \int_a^b \sum_{i=1}^n |F_{y_i}| \|\mathbf{h}\|_1 + |F_{z_i}| \|\mathbf{h}\|_1 dx \\ &= \left(\int_a^b \sum_{i=1}^n |F_{y_i}| + |F_{z_i}| dx \right) \|\mathbf{h}\|_1, \end{aligned}$$

donde la última expresión tiende a cero cuando $\|\mathbf{h}\|_1 \rightarrow 0$ ya que el término entre paréntesis es una constante. Concluimos que Φ es un funcional continuo en 0 y por el lema (2.1), como además Φ es lineal, concluimos que Φ es continuo para toda \mathbf{h} en el dominio de J .

Así, hemos probado que Φ es un funcional lineal continuo en el dominio de J . Probaremos ahora que $|\epsilon| \rightarrow 0$ cuando $\|\mathbf{h}\|_1 \rightarrow 0$.

De la desigualdad del triángulo tenemos

$$0 \leq |\epsilon| \leq \frac{1}{\|\mathbf{h}\|_1} \int_a^b |s(x, \mathbf{h})| dx.$$

Si nombramos $r_1(x, \mathbf{h})$ a la suma de los valores absolutos de cada término de $s(x, \mathbf{h})$, tenemos

$$\int_a^b |s(x, \mathbf{h})| dx \leq \int_a^b r_1(x, \mathbf{h}) dx.$$

Como cada sumando de $r_1(x, \mathbf{h})$ involucra términos no lineales en $|h_i|$ y $|h'_i|$ ($i = 1, \dots, n$) y además $|h_i|, |h'_i| < \|\mathbf{h}\|_1$ ($i = 1, \dots, n$) podemos intercambiar en cada sumando de $r_1(x, \mathbf{h})$ a $|h_i|$ y $|h'_i|$ por $\|\mathbf{h}\|_1$. Obtenemos así una nueva cota para $|\epsilon|$,

$$0 \leq |\epsilon| \leq \frac{1}{\|\mathbf{h}\|_1} \int_a^b r_2(x, \mathbf{h}) dx,$$

donde $r_2(x, \mathbf{h})$ es simplemente $r_1(x, \mathbf{h})$ pero en vez de tener los términos no lineales en $|h_i|$ y $|h'_i|$ ($i = 1, \dots, n$), tiene potencias mayores o iguales a dos de $\|\mathbf{h}\|_1$. Así, podemos obtener una nueva cota,

$$0 \leq |\epsilon| \leq \int_a^b r_3(x, \mathbf{h}) dx,$$

donde $r_3(x, \mathbf{h}) = r_2(x, \mathbf{h})/\|\mathbf{h}\|_1$. Es decir, $r_3(x, \mathbf{h})$ es, para cada x , una suma de constantes (las derivadas parciales de F evaluadas en $(x, \mathbf{y}, \mathbf{y}')$) multiplicadas por una potencias mayores o iguales a uno de $\|\mathbf{h}\|_1$. De este modo, cuando $\|\mathbf{h}\|_1 \rightarrow 0$ cada sumando de $r_3(x, \mathbf{h})$ tiende también a cero y por tanto $r_3(x, \mathbf{h}) \rightarrow 0$. Como no hay problemas de convergencia con la integral (es respecto de x), tenemos que,

$$\int_a^b r_3(x, \mathbf{h}) dx \rightarrow 0 \quad \text{cuando} \quad \|\mathbf{h}\|_1 \rightarrow 0.$$

Finalmente,

$$|\epsilon| \rightarrow 0 \quad \text{cuando} \quad \|\mathbf{h}\|_1 \rightarrow 0. \quad (2.7)$$

Podemos ahora concluir en virtud de la ecuación (2.6) que como Φ es un funcional lineal continuo en el dominio de J y $|\epsilon|$ satisface (2.7)

$$DJ(\mathbf{y})(\mathbf{h}) = \Phi(\mathbf{h}).$$

Una condición necesaria según el teorema (2.6) para los extremos de un funcional es que $DJ(\mathbf{y})(\mathbf{h}) = 0$, es decir, de acuerdo con la definición de $\Phi(\mathbf{h})$,

$$\int_a^b \sum_{i=1}^n F_{y_i} h_i + F_{z_i} h'_i dx = 0. \quad (2.8)$$

Como \mathbf{h} es arbitraria (siempre que J esté definido en \mathbf{h}), podemos escogerla de modo que todas sus coordenadas sean cero excepto la k -ésima y que además se satisfagan las condiciones de frontera (2.2). Entonces la condición (2.8) se escribe como,

$$\int_a^b F_{y_k} h_k + F_{z_k} h'_k dx = 0.$$

Como por hipótesis F_{y_k} y F_{z_k} son funciones continuas en $[a, b]$ podemos hacer uso del lema (2.8) para concluir que un extremo de J satisface la ecuación

$$F_{y_k} - \frac{d}{dx} F_{z_k} = 0.$$

Esta ecuación es conocida como la **ecuación de Euler-Lagrange**. Como k es arbitraria, tenemos que, de hecho en un extremo de J se satisface el sistema de ecuaciones de Euler-Lagrange,

$$F_{y_i} - \frac{d}{dx} F_{z_i} = 0 \quad (i = 1, \dots, n). \quad (2.9)$$

Cada ecuación del sistema (2.9) es de segundo orden, por tanto su solución depende de $2n$ constantes arbitrarias que son determinadas por las condiciones de frontera (2.2). Finalmente podemos escribir este resultado en forma de teorema.

Teorema 2.9.

Un extremo de un funcional que puede ser descrito mediante la ecuación (2.1), satisface el sistema de ecuaciones de Euler-Lagrange (2.9).

2.3. Coordenadas canónicas

En la sección anterior se obtuvo el sistema de ecuaciones de Euler-Lagrange (2.9) asociado a un funcional cuya forma general está descrita por la ecuación (2.1). Como el sistema (2.9) está compuesto por n ecuaciones de segundo orden, éste puede llevarse a un sistema de $2n$ ecuaciones de primer orden. Para ello definiremos,

$$p_i = F_{z_i}(x, y_1, \dots, y_n, z_1, \dots, z_n) \quad (i = 1, \dots, n). \quad (2.10)$$

Supondremos que

$$\frac{\partial(p_1, \dots, p_n)}{\partial(z_1, \dots, z_n)} \neq 0.$$

Luego, mediante el teorema de la función inversa, es posible invertir localmente el sistema (2.10) y expresar z_1, \dots, z_n en términos de $x, y_1, \dots, y_n, p_1, \dots, p_n$. De este modo, en adelante todo el desarrollo tendrá carácter local.

Ahora consideraremos la función $H(x, y_1, \dots, y_n, p_1, \dots, p_n)$ que se encuentra relacionada con $F(x, y_1, \dots, y_n, z_1, \dots, z_n)$ mediante la siguiente ecuación,

$$H = \sum_{i=1}^n z_i p_i - F. \quad (2.11)$$

Por supuesto, pensando a z_i ($i = 1, \dots, n$) como función de $x, y_1, \dots, y_n, p_1, \dots, p_n$. A la función H se le conoce como **Hamiltoniano** y a las variables $x, y_1, \dots, y_n, p_1, \dots, p_n$, H se les conoce como coordenadas canónicas.

Haciendo uso de la definición de p_i ($i = 1, \dots, n$) veremos ahora algunas relaciones que cumple H .

$$\frac{\partial H}{\partial x} = -\frac{\partial F}{\partial x} + \sum_{i=1}^n \frac{\partial z_i}{\partial x} p_i - \sum_{i=1}^n F_{z_i} \frac{\partial z_i}{\partial x} = -\frac{\partial F}{\partial x}, \quad (2.12)$$

$$\frac{\partial H}{\partial y_i} = -\frac{\partial F}{\partial y_i} + \sum_{i=1}^n \frac{\partial z_i}{\partial y_i} p_i - \sum_{i=1}^n F_{z_i} \frac{\partial z_i}{\partial y_i} = -\frac{\partial F}{\partial y_i}, \quad (2.13)$$

$$\frac{\partial H}{\partial p_i} = z_i + \sum_{i=1}^n \frac{\partial z_i}{\partial p_i} p_i - \sum_{i=1}^n F_{z_i} \frac{\partial z_i}{\partial p_i} = z_i. \quad (2.14)$$

Por otro lado, si hacemos,

$$z_i = \frac{dy_i}{dx} = y'_i \quad (i = 1, \dots, n),$$

podemos escribir el sistema de n ecuaciones (2.9), como el siguiente sistema de $2n$ ecuaciones:

$$\begin{aligned} \frac{dy_i}{dx} &= z_i \\ F_{y_i} - \frac{d}{dx} F_{z_i} &= 0 \quad (i = 1, \dots, n). \end{aligned} \quad (2.15)$$

En este sistema de ecuaciones, las derivadas parciales F_{y_i} y F_{z_i} ($i=1, \dots, n$) están evaluadas en $(x, \mathbf{y}, \mathbf{z})$.

Si ahora sustituimos la ecuaciones (2.12), (2.13) y (2.14) en el sistema de ecuaciones (2.15) obtenemos un nuevo sistema de ecuaciones,

$$\begin{aligned} \frac{dy_i}{dx} &= \frac{\partial H}{\partial p_i} \\ \frac{dp_i}{dx} &= -\frac{\partial H}{\partial y_i} \quad (i = 1, \dots, n). \end{aligned} \quad (2.16)$$

Este sistema de $2n$ ecuaciones diferenciales de primer orden es equivalente al sistema de ecuaciones (2.15) y es llamado **sistema canónico** de ecuaciones de Euler-Lagrange para el funcional (2.1).

Por último, enunciaremos una observación que ahora resulta inmediata.

Observación 2.10.

El sistema de ecuaciones canónicas (2.16) es el sistema de ecuaciones de Euler-Lagrange asociado al funcional,

$$A(\mathbf{y}, \mathbf{p}) = \int_a^b \mathbf{p} \cdot \mathbf{y}' - H(x, \mathbf{y}, \mathbf{p}) dx, \quad (2.17)$$

con las condiciones de frontera (2.2).

2.3.1. Transformada de Legendre

Consideremos primero una función $f(x, y)$ de dos variables continuamente diferenciable y que la diferencial total de f se expresa como:

$$df = udx + vdy \quad (2.18)$$

con

$$u = \frac{\partial f}{\partial x} \quad y \quad v = \frac{\partial f}{\partial y}.$$

Lo que se desea hacer es cambiar las coordenadas de descripción x y y por las variables independientes x y v . A la función que transforma las coordenadas (x, y) en (x, v) se le conoce como transformada de Legendre. Se hace necesario pues describir la dinámica en términos de las nuevas variables (en diferenciales respecto de las nuevas variables), para ello introducimos una función que depende de x y v ,

$$g(x, v) = vx - f$$

entonces,

$$dg = vdy + ydv - df$$

y por (2.18),

$$dg = ydv + udx.$$

Esta última expresión hace una descripción en términos de diferenciales de v y x como se desea. Por supuesto,

$$y = \frac{\partial g}{\partial v} \quad y \quad u = \frac{\partial g}{\partial x}.$$

Una aplicación directa de la transformada de Legendre es que se pueden encontrar las ecuaciones canónicas transformando a las coordenadas $(x, \mathbf{y}, \mathbf{z})$ en las coordenadas $(x, \mathbf{y}, \mathbf{p})$ (las componentes de \mathbf{p} están definidas por las ecuaciones (2.10)). De manera análoga a como se hizo antes es necesario introducir una función de las nuevas variables, ésta es a la que se conoce como Hamiltoniano,

$$H(x, \mathbf{y}, \mathbf{p}) = \mathbf{z} \cdot \mathbf{p} - F(x, \mathbf{y}, \mathbf{z})^{**}$$

así, si calculamos la diferencial de H , obtenemos por una parte,

$$dH = \sum_{i=1}^n \frac{\partial H}{\partial y_i} dy_i + \sum_{i=1}^n \frac{\partial H}{\partial p_i} dp_i + \frac{\partial H}{\partial x} dx$$

y por la otra,

$$dH = \sum_{i=1}^n z_i dp_i + \sum_{i=1}^n p_i dz_i - \sum_{i=1}^n \frac{\partial F}{\partial y_i} dy_i - \sum_{i=1}^n \frac{\partial H}{\partial z_i} dz_i - \frac{\partial F}{\partial x} dx.$$

De estas dos últimas expresiones, haciendo uso de las ecuaciones (2.10) y (2.15) e identificando términos, obtenemos las ecuaciones canónicas (2.16).

** Esta expresión es la misma que (2.11).

2.4. Transformaciones canónicas

Las transformaciones canónicas son aquellas transformaciones de las coordenadas para las cuales el sistema canónico de ecuaciones de Euler-Lagrange asociado a un funcional de la forma (2.1) se transforma en un nuevo sistema canónico al considerar las ecuaciones originales con las variables transformadas, es decir tiene la forma (2.16).

Definición 2.6 (Transformación canónica).

La transformación,

$$(\mathbf{y}, \mathbf{p}) \longrightarrow (\mathbf{s}, \mathbf{r}) \quad (2.19)$$

es canónica, si las ecuaciones de Euler-Lagrange asociadas al funcional $A^*(\mathbf{s}, \mathbf{r}) = A(\mathbf{y}(\mathbf{s}, \mathbf{r}), \mathbf{p}(\mathbf{s}, \mathbf{r}))$ se pueden escribir como

$$\begin{aligned} \frac{ds_i}{dx} &= \frac{\partial H^*}{\partial r_i} \\ -\frac{dr_i}{dx} &= \frac{\partial H^*}{\partial s_i} \quad (i = 1, \dots, n), \end{aligned} \quad (2.20)$$

para alguna función $H^* = H^*(x, \mathbf{s}, \mathbf{r})$.

Lo que ahora queremos encontrar son condiciones que garanticen que una transformación es canónica. De acuerdo con la observación (2.10), el sistema de ecuaciones de Euler-Lagrange asociado al funcional,

$$J^*(\mathbf{s}, \mathbf{r}) = \int_a^b \mathbf{r} \cdot \mathbf{s}' - H^*(x, \mathbf{s}, \mathbf{r}) dx,$$

es justo el sistema (2.20). Así, una transformación canónica es una transformación de variables que hace que las ecuaciones de Euler-Lagrange asociadas a los funcionales $A^*(\mathbf{s}, \mathbf{r})$ y $J^*(\mathbf{s}, \mathbf{r})$ sean las mismas.

Estudiaremos pues, bajo qué condiciones dos funcionales que pueden ser descritos por la ecuación (2.1) tienen asociado el mismo sistema de ecuaciones de Euler-Lagrange. El siguiente lema nos ofrece un resultado al respecto.

Lema 2.11.

Si J y J' son funcionales que tienen la forma general de la ecuación (2.1) con integrandos $F(x, \mathbf{y}, \mathbf{y}')$ y $F'(x, \mathbf{y}, \mathbf{y}')$ respectivamente entonces J y J' tiene asociado el mismo sistema de ecuaciones de Euler-Lagrange si

$$F'(x, \mathbf{y}, \mathbf{z}) = F(x, \mathbf{y}, \mathbf{z}) + \Psi(x, \mathbf{y}, \mathbf{z}),$$

donde

$$\Psi(x, \mathbf{y}, \mathbf{z}) = \frac{\partial G}{\partial x} + \sum_{i=1}^n \frac{\partial G}{\partial y_i} z_i,$$

siendo $G = G(x, \mathbf{y})$ una función continuamente diferenciable dos veces en $[a, b] \times \mathbb{R}^n$.

Demostración. Calculemos,

$$\frac{\partial \Psi}{\partial y_i}(x, \mathbf{y}, \mathbf{y}') = \frac{\partial G}{\partial y_i \partial x} + \sum_{k=1}^n \frac{\partial G}{\partial y_i \partial y_k} y'_k$$

y además,

$$\frac{d}{dx} \left(\frac{\partial \Psi}{\partial z_i}(x, \mathbf{y}, \mathbf{y}') \right) = \frac{d}{dx} \left(\frac{\partial G}{\partial y_i}(x, \mathbf{y}) \right) = \frac{\partial G}{\partial x \partial y_i} + \sum_{k=1}^n \frac{\partial G}{\partial y_k \partial y_i} y'_k$$

entonces,

$$\frac{\partial \Psi}{\partial y_i} - \frac{d}{dx} \left(\frac{\partial \Psi}{\partial z_i} \right) = 0,$$

por lo tanto J y J' tienen asociado el mismo sistema de ecuaciones de Euler-Lagrange. \square

La hipótesis del lema anterior, implica también que los integrandos de los funcionales J y J' difieren por una derivada total, ya que si $\mathbf{y} = \mathbf{y}(x)$ entonces,

$$\Psi(x, \mathbf{y}, \mathbf{y}') = \frac{dG}{dx}(x, \mathbf{y}).$$

De esta manera, una condición que garantiza que $A^*(\mathbf{s}, \mathbf{r})$ y $J^*(\mathbf{s}, \mathbf{r})$ tienen asociado el mismo sistema de ecuaciones de Euler-Lagrange es la siguiente igualdad,

$$\mathbf{p} \cdot \mathbf{y}' - H = \mathbf{r} \cdot \mathbf{s}' - H^* + \frac{dG}{dx}. \quad (2.21)$$

Por supuesto pensando el lado izquierdo de esta igualdad como función de \mathbf{r} y \mathbf{s} . Así, si la transformación (2.19) es tal que existe una función G que satisface la condición (2.21), entonces tal transformación es canónica. A la función G se le conoce como **función generadora de la transformación canónica** o **función generatriz**. A continuación justificaremos el porque se le da este nombre a G .

Notemos que, además de x , G depende de \mathbf{s} y \mathbf{r} . Como la transformación de coordenadas define $2n$ variables, en general G dependerá de las $2n$ variables que se toman como independientes en dicha transformación y de x . Entonces, tenemos cuatro posibles elecciones de variables independientes; \mathbf{y} y \mathbf{s} , \mathbf{y} y \mathbf{r} , \mathbf{p} y \mathbf{s} , \mathbf{p} y \mathbf{r} .

Si G dependiera de \mathbf{y} y \mathbf{s} , es decir, si $G = G_1(x, \mathbf{y}, \mathbf{s})$ la ecuación (2.21) se podría escribir simplemente como

$$\frac{dG_1}{dx} = \mathbf{p} \cdot \mathbf{y}' - \mathbf{r} \cdot \mathbf{s}' + (H^* - H) \quad (2.22)$$

encontrando así

$$p_i = \frac{\partial G_1}{\partial y_i}, \quad r_i = -\frac{\partial G_1}{\partial s_i}, \quad H^* - H = \frac{\partial G_1}{\partial x} \quad (i = 1, \dots, n). \quad (2.23)$$

El primer sistema de ecuaciones, que involucra las variables x , y_i , p_i y s_i ($i = 1, \dots, n$), puede ser resuelto para s en términos de x , y y \mathbf{p} obteniendo así una parte de la transformación. El segundo sistema completa la transformación expresando las \mathbf{r} en función de x , y y \mathbf{p} . La última ecuación provee una relación entre el Hamiltoniano H y el nuevo Hamiltoniano H^* .

Si las variables independientes de G fueran y y \mathbf{r} , como,

$$r_i = -\frac{\partial G_1}{\partial s_i} \quad (i = 1, \dots, n),$$

la función $G = G_2(x, y, \mathbf{r})$ puede encontrarse al hacer la transformación de Legendre $(x, y, s) \rightarrow (x, y, \mathbf{r})$ para expresar la ecuación (2.21) en términos de diferenciales respecto de y_i y r_i ($i = 1, \dots, n$). Obtenemos así que

$$G_2(x, y, \mathbf{r}) = G_1(x, y, s) + \mathbf{r} \cdot s.$$

Si despejamos $G_1(x, y, s)$ de la ecuación anterior sustituimos la expresión encontrada en la igualdad (2.22) obtenemos

$$\mathbf{p} \cdot \mathbf{y}' - H = \mathbf{r} \cdot \mathbf{s}' - H^* + \frac{d}{dx} (G_2(x, y, \mathbf{r}) - \mathbf{r} \cdot s),$$

de donde

$$\frac{dG_2}{dx} = \mathbf{p} \cdot \mathbf{y}' + \mathbf{r}' \cdot s + (H^* - H) \quad (2.24)$$

de donde identificamos, de manera análoga a como se hizo antes, los siguiente términos,

$$p_i = \frac{\partial G_2}{\partial y_i}, \quad s_i = \frac{\partial G_2}{\partial r_i}, \quad H^* - H = \frac{\partial G_2}{\partial x} \quad (i = 1, \dots, n). \quad (2.25)$$

El primero de estos sistemas, puede ser resuelto para \mathbf{r} en términos de x , y y \mathbf{p} pudiendo entonces completar la transformación mediante el segundo sistema, expresando s en términos de x , y y \mathbf{p} . La última ecuación relaciona H con H^* .

Si las variables independientes fueran ahora \mathbf{p} y s , de nueva cuenta, como

$$p_i = \frac{\partial G_1}{\partial y_i} \quad (i = 1, \dots, n),$$

se podría relacionar $G = G_3(x, \mathbf{p}, s)$ con $G_1(x, y, s)$ mediante la transformación de Legendre $(x, y, s) \rightarrow (x, \mathbf{p}, s)$ encontrando que

$$G_3(x, \mathbf{p}, s) = G_1(x, y, s) - \mathbf{y} \cdot \mathbf{p}.$$

Si sustituimos la expresión que resulta para $G_1(x, y, s)$ de la ecuación anterior en (2.22) se obtiene

$$\frac{dG_3}{dx} = -s' \cdot \mathbf{r} - \mathbf{y} \cdot \mathbf{p}' + (H^* - H) \quad (2.26)$$

de donde se sigue

$$y_i = -\frac{\partial G_3}{\partial p_i}, \quad r_i = -\frac{\partial G_3}{\partial s_i}, \quad H^* - H = \frac{\partial G_3}{\partial x} \quad (i = 1, \dots, n). \quad (2.27)$$

Como en los casos anteriores el primer sistema puede ser resuelto para \mathbf{s} en términos de x , \mathbf{y} y \mathbf{p} entonces el segundo sistema puede ser resuelto para \mathbf{r} en términos de x , \mathbf{y} y \mathbf{p} .

Finalmente cuando \mathbf{p} y \mathbf{r} fueran tomadas como variables independientes, tendríamos que en virtud de la ecuación,

$$p_i = \frac{\partial G_2}{\partial y_i} \quad (i = 1, \dots, n),$$

podríamos relacionar $G = G_4(x, \mathbf{p}, \mathbf{r})$ con $G_2(x, \mathbf{y}, \mathbf{r})$ haciendo la transformación de Legendre $(x, \mathbf{y}, \mathbf{r}) \rightarrow (x, \mathbf{p}, \mathbf{r})$. Obteniendo así,

$$G_4(x, \mathbf{p}, \mathbf{r}) = G_2(x, \mathbf{y}, \mathbf{r}) - \mathbf{p} \cdot \mathbf{y}.$$

Al despejar $G_2(x, \mathbf{y}, \mathbf{r})$ de la ecuación anterior, obtenemos una nueva expresión para ésta y al substituir dicha expresión en (2.24) se obtiene

$$\frac{dG_4}{dx} = \mathbf{s} \cdot \mathbf{r}' - \mathbf{y} \cdot \mathbf{p}' + (H^* - H), \quad (2.28)$$

que al identificar términos en ambos lados de la igualdad se obtienen la ecuaciones, que como antes se ha verificado, definen la transformación de coordenadas que ha de ser canónica. Estas ecuaciones son,

$$y_i = -\frac{\partial G_4}{\partial p_i}, \quad s_i = \frac{\partial G_4}{\partial r_i}, \quad H^* - H = \frac{\partial G_4}{\partial x} \quad (i = 1, \dots, n). \quad (2.29)$$

2.4.1. Ejemplos de transformaciones canónicas

Como primer ejemplo, consideraremos una función generatriz del segundo tipo, es decir, $G_2(x, \mathbf{y}, \mathbf{r})$ pero con la forma particular descrita por la siguiente ecuación

$$G_2 = \mathbf{y} \cdot \mathbf{r} \quad (2.30)$$

entonces, las ecuaciones (2.25) quedan escritas como

$$\begin{aligned} p_i &= \frac{\partial G_2}{\partial y_i} = r_i \\ s_i &= \frac{\partial G_2}{\partial r_i} = y_i \\ H^* &= H. \end{aligned}$$

Entonces la función generatriz (2.30) genera la transformación identidad, es decir que deja las coordenadas sin transformar.

Una función generatriz un poco más general es

$$G_2 = \sum_i f_i(t, \mathbf{y}) r_i \quad (2.31)$$

donde $\{f_i\}_{i \in I}$ es cualquier conjunto de funciones, así, de acuerdo con las ecuaciones (2.25)

$$s_i = \frac{\partial G_2}{\partial r_i} = f_i(t, \mathbf{y}).$$

Este tipo de transformaciones se llaman **transformaciones de puntos**, ya que las s_i dependen solo de las y_i y del tiempo, es decir, transforman solo las coordenadas y no los momentos. Lo que probamos es que un cambio de coordenadas (sin los momentos) es una transformación canónica.

Si consideramos ahora una función generatriz del primer tipo con la forma

$$G_1(x, \mathbf{y}, \mathbf{s}) = \mathbf{y} \cdot \mathbf{s} \quad (2.32)$$

entonces la ecuaciones de transformación (2.23) quedan

$$\begin{aligned} p_i &= \frac{\partial G_1}{\partial y_i} = s_i \\ r_i &= -\frac{\partial G_1}{\partial s_i} = -y_i \\ H^* &= H. \end{aligned}$$

Esta transformación de coordenadas es canónica ya que la ecuación (2.32) nos proporciona la función generatriz de dicha transformación. Esta transformación es ilustrativa, ya que al intercambiar los momentos con las coordenadas, se hace patente su independencia y por tanto la necesidad de ambas variables en la formulación Hamiltoniana.

Ahora consideraremos un caso que será de utilidad en el próximo capítulo. Si una función generatriz del tercer tipo tiene la forma

$$G_3(x, \mathbf{p}, \mathbf{s}) = -\mathbf{p} \cdot \mathbf{s} \quad (2.33)$$

entonces, las ecuaciones (2.27)

$$\begin{aligned} y_i &= -\frac{\partial G_3}{\partial p_i} = s_i \\ r_i &= -\frac{\partial G_3}{\partial s_i} = p_i \\ H^* &= H \end{aligned}$$

generan la transformación identidad.

Una pregunta que es posible hacernos es si será posible encontrar una función generatriz que avance en el tiempo las coordenadas generalizadas, es decir, una transformación que nos permita obtener de manera discreta los valores de

\mathbf{p} y \mathbf{q} . Esto es, que si conocemos el valor de (\mathbf{p}, \mathbf{q}) para un tiempo t_0 , podamos vía una transformación canónica encontrar los valores de (\mathbf{p}, \mathbf{q}) para un tiempo posterior. De ser así, obtendríamos los valores de la solución exacta en un conjunto discreto de valores del tiempo.

Esta idea no resulta extraña, pues si consideramos a una función en su serie de Taylor

$$p(t) = p(0) + \dot{p}(0)t + \frac{1}{2}\ddot{p}(0)t^2 + \dots$$

y además si

$$\dot{p}(0) = \frac{\partial H}{\partial q}(p(0), q(0)) \quad \text{y} \quad p(t) = \frac{\partial G}{\partial q}(p(0), q(t))$$

es decir, $p(0)$ es una coordenada de un sistema Hamiltoniano y $p(t)$ se obtiene al transformar coordenadas vía la función generatriz G , buscamos G tal que

$$p(t) = \frac{\partial G}{\partial q}(p(0), q(t)) + \frac{\partial H}{\partial q}\left(\frac{\partial G}{\partial q}, q\right)(0, t)t + \dots$$

algo similar puede hacerse con q para obtener una ecuación para G , de esta manera parece posible encontrar la función generatriz G con la propiedad de que sólo avanza en el tiempo la posición del sistema.

En el capítulo siguiente veremos como encontrar dicha función generatriz para poder así encontrar un modo de integrar un sistema Hamiltoniano discretamente. De hecho, en ello se basa la obtención de cierto tipo de integradores simplécticos y que son el objeto de esta tesis.

2.5. Notación clásica

En esta sección daremos sentido físico a la teoría que hemos venido desarrollando y de este modo, introduciremos la notación clásica para este tipo de problemas.

Consideremos un sistema de N partículas con coordenadas $\mathbf{x}_i(t)$ ($i = 1, \dots, n$) en \mathbb{R}^3 . Nos gustaría que las trayectorias de las partículas fueran extremos de algún funcional, es decir que quedaran descritas mediante las ecuaciones de Euler-Lagrange para un funcional de la forma

$$I(\mathbf{x}_1, \dots, \mathbf{x}_N) = \int_{t_0}^{t_1} L(\mathbf{x}_1, \dots, \mathbf{x}_N, \dot{\mathbf{x}}_1, \dots, \dot{\mathbf{x}}_N) dt.$$

Esto significaría que las ecuaciones de Euler-Lagrange

$$\frac{d}{dt}L_{\dot{\mathbf{x}}_i} = L_{\mathbf{x}_i} \quad (i = 1, \dots, N)$$

deberían ser las ecuaciones de movimiento de Newton

$$m_i \ddot{\mathbf{x}}_i = F_i(t, \mathbf{x}_1, \dots, \mathbf{x}_N) \quad (i = 1, \dots, N),$$

que si identificamos término a término obtenemos para $i = 1, \dots, N$ que

$$L_{\dot{\mathbf{x}}_i} = m_i \dot{\mathbf{x}}_i \quad \text{y} \quad F_i = L_{\mathbf{x}_i}.$$

Si ahora integramos la primera ecuación respecto de $\dot{\mathbf{x}}_i$ obtenemos

$$L = \frac{1}{2} \sum_1^N m_i |\dot{\mathbf{x}}_i|^2 - U(t, \mathbf{x}_1, \dots, \mathbf{x}_N)$$

y si derivamos esta última expresión respecto de \mathbf{x}_i tenemos

$$F_i = -U_{\mathbf{x}_i}.$$

En física, se define la energía cinética del sistema como

$$T = \frac{1}{2} \sum_1^N m_i |\dot{\mathbf{x}}_i|^2$$

y la energía potencial U tal que $F_i = -U_{\mathbf{x}_i}$. Así concluimos que las ecuaciones de Newton son las ecuaciones de Euler para el funcional

$$\int_{t_0}^{t_1} (T - U) dt.$$

En general, puede ser más conveniente pensar en otro tipo de coordenadas que determinen con precisión la posición del sistema y no sólo en coordenadas cartesianas; éstas pueden ser por ejemplo coordenadas esféricas o cilíndricas dependiendo de las simetrías del problema o incluso de cantidades muy particulares del sistema que se estudie. Estas coordenadas se denominan **coordenadas Lagrangianas** y usualmente se denotan por q_1, \dots, q_n y se suponen linealmente independientes. Por supuesto, el vector de coordenadas Lagrangianas $\mathbf{q} = (q_1, \dots, q_n)$ está relacionado con la posición de las partículas mediante una transformación de coordenadas, es decir, $\mathbf{x} = C(\mathbf{q})$ que cuando las partículas se mueven significa que $\dot{\mathbf{x}} = JC(\mathbf{q}) \cdot \dot{\mathbf{q}}$. Es decir, la energía cinética queda escrita

$$T = \frac{1}{2} \sum_{i,j=1}^n a_{ij}(\mathbf{q}) \dot{q}_i \dot{q}_j$$

donde los coeficientes $a_{ij}(\mathbf{q})$ involucran a m_i y a $JC(\mathbf{q})$. Entonces,

$$L = T - U = \frac{1}{2} \sum_{i,j=1}^n a_{ij}(\mathbf{q}) \dot{q}_i \dot{q}_j - U(t, \mathbf{q}).$$

El **principio de Hamilton** o **principio de mínima acción** establece que las trayectorias de las partículas en las coordenadas Lagrangianas son los puntos críticos del funcional de acción

$$I(\mathbf{q}) = \int_{t_0}^{t_1} L(t, \mathbf{q}, \dot{\mathbf{q}}) dt \quad (2.34)$$

con $\mathbf{q}(t_0)$ y $\mathbf{q}(t_1)$ dados. A la función L se le llama **Lagrangiana** en mecánica y se supone que es continuamente diferenciable. De esta manera, las ecuaciones de movimiento para (2.34) son

$$\frac{d}{dt}L_{\dot{q}_i} = L_{q_i} \quad (i = 1, \dots, n).$$

Como hicimos antes, si definimos

$$p_i = L_{z_i}(t, q_1, \dots, q_n, z_1, \dots, z_n) \quad (i = 1, \dots, n) \quad (2.35)$$

y suponemos que la matriz $(L_{z_i z_j})$ es invertible, entonces si hacemos

$$H(\mathbf{p}, \mathbf{q}) = \mathbf{p} \cdot \mathbf{z} - L(t, \mathbf{q}, \mathbf{z})$$

por supuesto, considerando $\mathbf{z} = \mathbf{z}(t, \mathbf{p}, \mathbf{q})$ y gracias a la definición de \mathbf{p} dada en (2.35), las ecuaciones de Euler-Lagrange se transforman en el sistema canónico

$$\dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}}, \quad \dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{q}} \quad (2.36)$$

con $\mathbf{q}(t_0)$ y $\mathbf{q}(t_1)$ dados. Esta es la notación clásica para el sistema canónico de ecuaciones de Euler-Lagrange y suele también llamarse **sistema Hamiltoniano** o **ecuaciones de movimiento de Hamilton**. A las coordenadas \mathbf{p} y \mathbf{q} se les llama **coordenadas generalizadas** y al espacio cartesiano formado por las coordenadas generalizadas se le llama **espacio fase**.

De acuerdo con la observación (2.10), este sistema es el sistema de Euler para el funcional

$$A(\mathbf{p}, \mathbf{q}) = \int_{t_0}^{t_1} \mathbf{p} \cdot \dot{\mathbf{q}} - H(t, \mathbf{p}, \mathbf{q}) dt.$$

A este funcional se le llama **acción generalizada**.

La **energía del sistema** se define como

$$E(t, \mathbf{q}, \dot{\mathbf{q}}) = \sum_{i=1}^n \dot{q}_i L_{z_i}(t, \mathbf{q}, \dot{\mathbf{q}}) - L(t, \mathbf{q}, \dot{\mathbf{q}})$$

que en el ejemplo de las N partículas que habíamos tomado, se escribe

$$E = \sum_{i,j=1}^n a_{ij}(\mathbf{q}) \dot{q}_i \dot{q}_j - L = 2T - (T - U) = T + U.$$

Además como $L_{z_i} = p_i$ y $\dot{q}_i = z_i$ se tiene que

$$E(t, \mathbf{q}, \dot{\mathbf{q}}) = H(t, \mathbf{p}, \mathbf{q}),$$

es decir,

$$H = T + U.$$

Finalmente escribiremos los resultados más importantes de la sección anterior haciendo uso de esta nueva notación que hemos introducido.

La definición (2.6) puede ser escrita como sigue

Definición 2.7 (Transformación canónica).

La transformación,

$$(\mathbf{p}_0, \mathbf{q}_0) \longrightarrow (\mathbf{p}, \mathbf{q}) \quad (2.37)$$

es canónica, si las ecuaciones de Euler-Lagrange asociadas al funcional

$A^*(\mathbf{p}, \mathbf{q}) = A(\mathbf{p}_0(\mathbf{p}, \mathbf{q}), \mathbf{q}_0(\mathbf{p}, \mathbf{q}))$ se pueden escribir como:

$$\dot{\mathbf{q}} = \frac{\partial H^*}{\partial \mathbf{p}}, \quad \dot{\mathbf{p}} = -\frac{\partial H^*}{\partial \mathbf{q}} \quad (2.38)$$

para alguna función $H^* = H^*(t, \mathbf{p}, \mathbf{q})$.

Se escriben las coordenadas $(\mathbf{p}_0, \mathbf{q}_0)$ como coordenadas iniciales porque pueden ser pensadas como la posición del sistema al tiempo cero y las coordenadas finales como (\mathbf{p}, \mathbf{q}) que sería la posición del sistema al tiempo t . Probamos también el siguiente teorema.

Teorema 2.12.

La transformación de coordenadas (2.37) es canónica si existe una función G que depende sólo de $(t, \mathbf{q}_0, \mathbf{q})$ ó $(t, \mathbf{q}_0, \mathbf{p})$ ó $(t, \mathbf{p}_0, \mathbf{q})$ ó $(t, \mathbf{p}_0, \mathbf{p})$ y una función $H^*(t, \mathbf{p}, \mathbf{q})$ tales que

$$\frac{dG}{dt} = \mathbf{p}_0 \cdot \dot{\mathbf{q}}_0 - \mathbf{q} \cdot \dot{\mathbf{p}} + (H^* - H) \quad (2.39)$$

donde $H = H(t, \mathbf{p}_0, \mathbf{q}_0)$ es el Hamiltoniano que satisface las ecuaciones (2.36).

Como habíamos dicho, a G se le llama **función generatriz** o **función generadora de la transformación canónica**.

Probamos que para las distintas elecciones de variables independientes en G las respectivas funciones $G_1(t, \mathbf{q}_0, \mathbf{q})$, $G_2(t, \mathbf{q}_0, \mathbf{p})$, $G_3(t, \mathbf{p}_0, \mathbf{q})$ y $G_4(t, \mathbf{p}_0, \mathbf{p})$ generan las siguientes transformaciones

$$\mathbf{p}_0 = \frac{\partial G_1}{\partial \mathbf{q}_0}, \quad \mathbf{p} = -\frac{\partial G_1}{\partial \mathbf{q}}, \quad H^* - H = \frac{\partial G_1}{\partial t} \quad (2.40)$$

$$\mathbf{p}_0 = \frac{\partial G_2}{\partial \mathbf{q}_0}, \quad \mathbf{q} = \frac{\partial G_2}{\partial \mathbf{p}}, \quad H^* - H = \frac{\partial G_2}{\partial t} \quad (2.41)$$

$$\mathbf{q}_0 = -\frac{\partial G_3}{\partial \mathbf{p}_0}, \quad \mathbf{p} = -\frac{\partial G_3}{\partial \mathbf{q}}, \quad H^* - H = \frac{\partial G_3}{\partial t} \quad (2.42)$$

$$\mathbf{q}_0 = -\frac{\partial G_4}{\partial \mathbf{p}_0}, \quad \mathbf{q} = \frac{\partial G_4}{\partial \mathbf{p}}, \quad H^* - H = \frac{\partial G_4}{\partial t}. \quad (2.43)$$

Por otro lado, si una función G satisface alguno de los grupos de ecuaciones (2.40), (2.41), (2.42) ó (2.43) se tiene que entonces también satisface (2.39). Es decir que si queremos encontrar una transformación canónica basta encontrar una función que satisfaga alguno de los grupos de ecuaciones mencionados.

Vimos también algunos ejemplos particulares, uno de ellos fue considerar $G = G_3(t, \mathbf{p}_0, \mathbf{q})$ con la forma particular de

$$G_3(t, \mathbf{p}_0, \mathbf{q}) = -\mathbf{p}_0 \cdot \mathbf{q} \quad (2.44)$$

y demostramos que la transformación (canónica) que genera es la identidad.

Esta es la notación clásica que se sigue en la formulación Hamiltoniana de la mecánica y que por convención seguiremos nosotros también a lo largo de la tesis.

Capítulo 3

Integración simpléctica

En este capítulo estudiaremos, por un lado, la estructura geométrica del flujo de un sistema Hamiltoniano y por el otro, el problema de integrar numéricamente el sistema conservando dicha estructura. Por comodidad usaremos la notación vectorial clásica desarrollada en el capítulo precedente, es decir, que si $\mathbf{p} = (p_1, p_2, \dots, p_n)^t$ y $\mathbf{q} = (q_1, q_2, \dots, q_n)^t$ son funciones del tiempo, entonces un sistema Hamiltoniano, puede ser escrito simplemente por

$$\frac{\partial H}{\partial \mathbf{p}} = \dot{\mathbf{q}}, \quad -\frac{\partial H}{\partial \mathbf{q}} = \dot{\mathbf{p}}. \quad (3.1)$$

Donde $H(t, \mathbf{p}, \mathbf{q})$ es la función Hamiltoniana.

3.1. Simplecticidad

En esta sección estudiaremos algunas propiedades del flujo de un sistema Hamiltoniano. Para ello es necesario definir simplecticidad.

Recordemos que dada una matriz cuadrada A de $k \times k$ con entradas en \mathbb{R} define una transformación lineal de $\mathbb{R}^k \rightarrow \mathbb{R}^k$ denotada por L_A y definida por

$$L_A(\mathbf{x}) = A\mathbf{x}.$$

En adelante conservaremos esta notación para referirnos a la transformación lineal asociada a una matriz $A \in M_k(\mathbb{R})$ ¹

Definición 3.1 (Transformación simpléctica).

Sea $A \in M_{2n}(\mathbb{R})$, entonces la transformación lineal $L_A : \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$ es una transformación simpléctica si

$$A^t J A = J \quad \text{con} \quad J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$$

donde I es la matriz identidad de dimensión n y 0 es la matriz de ceros también de dimensión n . A la matriz J se le llama **matriz simpléctica**.

¹ Este conjunto denota al conjunto de matrices de $k \times k$ con entradas en \mathbb{R} .

Esta definición resulta útil pues para saber si una transformación lineal es simpléctica lo único que se debe hacer son dos productos de matrices. Sin embargo la siguiente definición equivalente nos permite apreciar mejor la estructura geométrica de una transformación simpléctica.

Definición 3.2 (Transformación simpléctica).

Sea $A \in M_{2n}(\mathbb{R})$, entonces la transformación lineal $L_A : \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$ es una transformación simpléctica si $w(A\xi, A\eta) = w(\xi, \eta)$ para toda $\xi, \eta \in \mathbb{R}^{2n}$ donde

$$w(\xi, \eta) = \xi^t J \eta \quad (3.2)$$

y J es la matriz simpléctica de dimensión $2n$.

La prueba de la equivalencia de las definiciones anteriores es muy sencilla, si A satisface la condición de la definición (3.1), entonces

$$w(A\xi, A\eta) = \xi^t A^t J A \eta = \xi^t J \eta,$$

es decir, A satisface la definición (3.2) e inversamente si A satisface que

$$w(A\xi, A\eta) = w(\xi, \eta) \quad \text{para toda } \xi, \eta \in \mathbb{R}^{2n}$$

entonces la igualdad anterior también se tiene y además es posible escribirla como

$$\xi^t (A^t J A - J) \eta = 0 \quad \text{para toda } \xi, \eta \in \mathbb{R}^{2n}$$

lo que implica que $J = A^t J A$, es decir que A satisface la definición (3.1).

Es posible dar una interpretación geométrica a la simplecticidad de una transformación lineal. La siguiente observación establece dicha interpretación.

Observación 3.1.

Si $\xi, \eta \in \mathbb{R}^{2n} = \{(p, q) : p, q \in \mathbb{R}^n\}$ y P es el paralelogramo definido mediante la expresión

$$P = \{t\xi + s\eta : 0 \leq t \leq 1, 0 \leq s \leq 1\},$$

entonces la simplecticidad de una transformación lineal L_A implica que la suma de las áreas de las proyecciones de P sobre los planos (p_i, q_i) ($i = 1, \dots, n$) se preservan bajo la transformación.

Demostración. Consideremos a $\xi, \eta \in \mathbb{R}^{2n}$ como

$$\xi = \begin{pmatrix} \xi^p \\ \xi^q \end{pmatrix} \quad \text{y} \quad \eta = \begin{pmatrix} \eta^p \\ \eta^q \end{pmatrix}$$

con $\eta^p, \eta^q, \xi^p, \xi^q \in \mathbb{R}^n$ entonces, se tiene que

$$\begin{aligned} w(\xi, \eta) &= \xi^t J \eta \\ &= ((\xi^p)^t \ (\xi^q)^t) \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} \eta^p \\ \eta^q \end{pmatrix} \\ &= (\xi^p)^t \eta^q - (\xi^q)^t \eta^p \\ &= \sum_{i=1}^n (\xi_i^p \eta_i^q - \xi_i^q \eta_i^p) \\ &= \sum_{i=1}^n \det \begin{pmatrix} \xi_i^p & \eta_i^p \\ \xi_i^q & \eta_i^q \end{pmatrix}, \end{aligned}$$

es decir que $w(\xi, \eta)$ es la suma de las áreas de las proyecciones de P sobre los planos (p_i, q_i) , luego, de la definición (3.2) se desprende que dicha suma permanece invariante bajo la transformación A . En otras palabras, la suma de las áreas de las proyecciones de P sobre los planos (p_i, q_i) es la misma que la suma de las áreas de las proyecciones para $A(P)$. \square

El siguiente lema establece propiedades sencillas pero útiles de la función $w(\xi, \eta)$. Esta función es conocida como el **área simpléctica** del paralelogramo, en \mathbb{R}^{2n} , generado por los vectores ξ y η en \mathbb{R}^{2n} .

Lema 3.2.

La función $w : \mathbb{R}^{2n} \times \mathbb{R}^{2n} \rightarrow \mathbb{R}$ definida por la ecuación (3.2) satisface para cualesquiera $\xi, \eta, \nu, \zeta \in \mathbb{R}^{2n}$ los siguientes tres incisos

(1) $w(\xi, \eta)$ es bilineal

(2) $-w(\xi, \eta) = w(\eta, \xi)$

(3) $w(\xi + \nu, \eta + \zeta) = w(\xi, \eta) + w(\nu, \zeta) + w(\xi, \zeta) + w(\nu, \eta)$

Demostración. El inciso uno se sigue inmediatamente de la distributividad del producto de matrices sobre la suma de vectores y de la factorización de escalares del producto de matrices por vectores. Para la demostración del inciso dos considere

$$-w(\xi, \eta) = \xi^t (-J) \eta = \xi^t \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} \eta = (\xi^q)^t \eta^p - (\xi^p)^t \eta^q \quad (3.3)$$

$$w(\eta, \xi) = \eta^t J \xi = \eta^t \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \xi = (\xi^q)^t \eta^p - (\xi^p)^t \eta^q \quad (3.4)$$

así, de las ecuaciones (3.3) y (3.4) se concluye el segundo inciso. La demostración del tercer inciso queda escrita en las siguientes líneas

$$\begin{aligned} w(\xi + \nu, \eta + \zeta) &= (\xi + \nu)^t J (\eta + \zeta) = (\xi^t + \nu^t) J (\eta + \zeta) = (\xi^t J + \nu^t J) (\eta + \zeta) \\ &= \xi^t J \eta + \xi^t J \zeta + \nu^t J \eta + \nu^t J \zeta = w(\xi, \eta) + w(\nu, \zeta) + w(\xi, \zeta) + w(\nu, \eta) \end{aligned}$$

\square

Corolario 3.3.

Si $\xi, \eta, \nu, \zeta \in \mathbb{R}^{2n}$ son tales que $\xi = a\zeta$ y $\nu = b\eta$ para algunos números $a, b \in \mathbb{R}$ entonces $w(\xi + \nu, \eta + \zeta) = w(\xi, \eta) + w(\nu, \zeta)$.

Demostración. Veamos pues, por el inciso (3) del lema (3.2) tenemos que

$$\begin{aligned} w(\xi + \nu, \eta + \zeta) &= w(\xi, \eta) + w(\nu, \zeta) + w(\xi, \zeta) + w(\nu, \eta) \\ &= w(\xi, \eta) + w(\nu, \zeta) + w(a\zeta, \zeta) + w(b\eta, \eta) \end{aligned}$$

y por el inciso (1) se transforma en

$$= w(\xi, \eta) + w(\nu, \zeta) + aw(\zeta, \zeta) + bw(\eta, \eta)$$

pero de la definición (3.2) se sigue inmediatamente que para cualquier $\rho \in \mathbb{R}^{2n}$ $w(\rho, \rho) = 0$ y por lo tanto

$$w(\xi + \nu, \eta + \zeta) = w(\xi, \eta) + w(\nu, \zeta).$$

□

Más adelante veremos que la estructura geométrica de las transformaciones simplécticas, demostrada en la observación (3.1) resulta muy importante. Ahora extenderemos las definiciones (3.1) y (3.2) para funciones no necesariamente lineales, esto se puede hacer sólo para aquellas funciones que pueden ser aproximadas localmente por una función lineal, es decir para funciones diferenciables.

Definición 3.3 (Función simpléctica).

Una función diferenciable $g : U \rightarrow \mathbb{R}^{2n}$ (donde $U \subset \mathbb{R}^{2n}$ es un conjunto abierto) es simpléctica si la transformación lineal asociada a la matriz jacobiana $g'(y)$ es simpléctica. Es decir si $L_{g'(y)}$ es una transformación simpléctica.

Para ilustrar la importancia de la simplecticidad de la matriz jacobiana de una función diferenciable $g(y)$ consideremos $M \subset U$ una variedad de dimensión 2. Supongamos que M puede ser parametrizada por una función $\psi : K \rightarrow \mathbb{R}^{2n}$ continuamente diferenciable e invertible con $K \subset \mathbb{R}^2$ compacto, así $\psi(K) = M$. De este modo M puede ser visto como el límite de la unión de pequeños paralelogramos generados por

$$\frac{\partial \psi}{\partial s}(s, t) \quad \text{y} \quad \frac{\partial \psi}{\partial t}(s, t).$$

Para cada uno de estos paralelogramos consideremos la suma de las áreas de sus proyecciones sobre los planos (p_i, q_i) y después sumemos sobre todos los paralelogramos de M , en el límite obtenemos la expresión

$$\Omega(M) = \iint_K w \left(\frac{\partial \psi}{\partial s}(s, t), \frac{\partial \psi}{\partial t}(s, t) \right) ds dt. \quad (3.5)$$

Esta expresión es el **área simpléctica** de la variedad 2-dimensional M . La fórmula de cambio de variables para integrales dobles implica que $\Omega(M)$ no depende de la parametrización de M , como lo establece la siguiente proposición.

Proposición 3.4.

La expresión (3.5) del área simplectica de una variedad M 2-dimensional no depende de la parametrización ψ .

Demostración. Considere dos parametrizaciones C^1 -invertibles

$$\psi_1(s, t) : K_1 \rightarrow M \quad \text{y} \quad \psi_2(u, v) : K_2 \rightarrow M$$

con $K_1, K_2 \subset \mathbb{R}^2$ compactos. Entonces si consideramos el cambio de variables $h : K_1 \rightarrow K_2$ dado por

$$h(s, t) = \psi_2^{-1} \circ \psi_1(s, t)$$

tenemos que $h(s, t) = (h_1(s, t), h_2(s, t)) = (u, v)$. El teorema de cambio de variables² para integrales dobles establece que

$$\begin{aligned} \iint_{K_1} w \left(\frac{\partial \psi_2}{\partial u}(h(s, t)), \frac{\partial \psi_2}{\partial v}(h(s, t)) \right) |\Delta_h(s, t)| ds dt = \\ \iint_{K_2} w \left(\frac{\partial \psi_2}{\partial u}(u, v), \frac{\partial \psi_2}{\partial v}(u, v) \right) du dv \end{aligned} \quad (3.6)$$

ya que $h(K_1) = K_2$ y donde

$$\Delta_h(s, t) = \det \begin{pmatrix} \frac{\partial h_1}{\partial s} & \frac{\partial h_1}{\partial t} \\ \frac{\partial h_2}{\partial s} & \frac{\partial h_2}{\partial t} \end{pmatrix}.$$

De la definición de $h(s, t)$, se sigue que $\psi_1(s, t) = \psi_2 \circ h(s, t)$ y por lo tanto, si aplicamos la regla de la cadena obtenemos que

$$\frac{\partial \psi_1}{\partial s}(s, t) = \frac{\partial \psi_2}{\partial u}(h(s, t)) \frac{\partial h_1}{\partial s}(s, t) + \frac{\partial \psi_2}{\partial v}(h(s, t)) \frac{\partial h_2}{\partial s}(s, t) \quad (3.7)$$

$$\frac{\partial \psi_1}{\partial t}(s, t) = \frac{\partial \psi_2}{\partial u}(h(s, t)) \frac{\partial h_1}{\partial t}(s, t) + \frac{\partial \psi_2}{\partial v}(h(s, t)) \frac{\partial h_2}{\partial t}(s, t). \quad (3.8)$$

Supongamos que $\Delta_h(s, t) > 0$, entonces el integrando del lado derecho de la ecuación (3.6) se puede escribir como

$$\begin{aligned} w \left(\frac{\partial \psi_2}{\partial u}(h), \frac{\partial \psi_2}{\partial v}(h) \right) \Delta_h \\ = \frac{\partial h_1}{\partial s} \frac{\partial h_2}{\partial t} w \left(\frac{\partial \psi_2}{\partial u}(h), \frac{\partial \psi_2}{\partial v}(h) \right) - \frac{\partial h_1}{\partial t} \frac{\partial h_2}{\partial s} w \left(\frac{\partial \psi_2}{\partial u}(h), \frac{\partial \psi_2}{\partial v}(h) \right) \end{aligned}$$

y como w es una función bilineal según el inciso (1) del lema (3.2),

$$\begin{aligned} w \left(\frac{\partial \psi_2}{\partial u}(h), \frac{\partial \psi_2}{\partial v}(h) \right) \Delta_h \\ = w \left(\frac{\partial \psi_2}{\partial u}(h) \frac{\partial h_1}{\partial s}, \frac{\partial \psi_2}{\partial v}(h) \frac{\partial h_2}{\partial t} \right) - w \left(\frac{\partial \psi_2}{\partial u}(h) \frac{\partial h_1}{\partial t}, \frac{\partial \psi_2}{\partial v}(h) \frac{\partial h_2}{\partial s} \right) \end{aligned}$$

²Ver S. Lang [9].

que utilizando el segundo inciso del mismo lema se transforma en

$$= w \left(\frac{\partial \psi_2}{\partial u}(h) \frac{\partial h_1}{\partial s}, \frac{\partial \psi_2}{\partial v}(h) \frac{\partial h_2}{\partial t} \right) + w \left(\frac{\partial \psi_2}{\partial v}(h) \frac{\partial h_2}{\partial s}, \frac{\partial \psi_2}{\partial u}(h) \frac{\partial h_1}{\partial t} \right)$$

y de nueva cuenta en virtud de la bilinealidad tenemos

$$= w \left(\frac{\partial \psi_2}{\partial u}(h) \frac{\partial h_1}{\partial s}, \frac{\partial \psi_2}{\partial v}(h) \frac{\partial h_2}{\partial t} \right) + w \left(\frac{\partial \psi_2}{\partial v}(h) \frac{\partial h_2}{\partial s}, \frac{\partial \psi_2}{\partial u}(h) \frac{\partial h_1}{\partial t} \right).$$

Debido a que

$$w \left(\frac{\partial \psi_2}{\partial u}(h) \frac{\partial h_1}{\partial s}, \frac{\partial \psi_2}{\partial u}(h) \frac{\partial h_1}{\partial t} \right) = \frac{\partial h_1}{\partial s} \frac{\partial h_1}{\partial t} w \left(\frac{\partial \psi_2}{\partial u}(h), \frac{\partial \psi_2}{\partial u}(h) \right) = 0$$

$$w \left(\frac{\partial \psi_2}{\partial v}(h) \frac{\partial h_2}{\partial t}, \frac{\partial \psi_2}{\partial v}(h) \frac{\partial h_2}{\partial s} \right) = \frac{\partial h_2}{\partial t} \frac{\partial h_2}{\partial s} w \left(\frac{\partial \psi_2}{\partial v}(h), \frac{\partial \psi_2}{\partial v}(h) \right) = 0$$

entonces gracias al corolario (3.3) podemos concluir que

$$w \left(\frac{\partial \psi_2}{\partial u}(h), \frac{\partial \psi_2}{\partial v}(h) \right) \Delta_h =$$

$$w \left(\frac{\partial \psi_2}{\partial u}(h) \frac{\partial h_1}{\partial s} + \frac{\partial \psi_2}{\partial v}(h) \frac{\partial h_2}{\partial s}, \frac{\partial \psi_2}{\partial v}(h) \frac{\partial h_2}{\partial t} + \frac{\partial \psi_2}{\partial u}(h) \frac{\partial h_1}{\partial t} \right)$$

donde todas las funciones en esta expresión están evaluadas en (s, t) . Así, de acuerdo con las ecuaciones (3.7) y (3.8) concluimos que

$$w \left(\frac{\partial \psi_2}{\partial u}(h(s, t)), \frac{\partial \psi_2}{\partial v}(h(s, t)) \right) \Delta_h(s, t) = w \left(\frac{\partial \psi_1}{\partial s}(s, t) \frac{\partial \psi_2}{\partial s}(s, t) \right)$$

y si sustituimos la expresión anterior en (3.6) finalmente tenemos

$$\iint_{K_1} w \left(\frac{\partial \psi_1}{\partial s}(s, t), \frac{\partial \psi_1}{\partial t}(s, t) \right) ds dt = \iint_{K_2} w \left(\frac{\partial \psi_2}{\partial u}(u, v), \frac{\partial \psi_2}{\partial v}(u, v) \right) du dv$$

es decir, $\Omega(M)$ no depende de la parametrización que se elija. El caso en el que $\Delta_h(s, t) < 0$ es totalmente análogo. \square

Una vez probada la independencia de la parametrización en la expresión (3.5) del área simpléctica estamos en posición de probar un lema que extiende el resultado visto en la observación (3.1) para transformaciones lineales simplécticas a funciones (no necesariamente lineales) simplécticas.

Lema 3.5.

Si $g : U \rightarrow \mathbb{R}^{2n}$ es una función simpléctica en U , entonces

$$\Omega(M) = \Omega(g(M))$$

para todas las variedades $M \subset U$ de dimensión 2 para las cuales existe una función continuamente diferenciable $\psi : K \rightarrow \mathbb{R}^{2n}$ ($K \subset \mathbb{R}^2$ compacto) tal que $\psi(K) = M$.

Demostración. la variedad $g(M)$ puede ser parametrizada por $g \circ \psi$, entonces,

$$\Omega(g(M)) = \iint_K w \left(\frac{\partial(g \circ \psi)}{\partial s}(s, t), \frac{\partial(g \circ \psi)}{\partial t}(s, t) \right) ds dt.$$

Sin embargo, tenemos que $(g \circ \psi)'(s, t) = g'(\psi(s, t))\psi'(s, t)$ entonces,

$$\begin{aligned} \frac{\partial(g \circ \psi)}{\partial s}(s, t) &= (g \circ \psi)'(s, t)(1, 0)^t = g'(\psi(s, t))\psi'(s, t)(1, 0)^t \\ &= g'(\psi(s, t)) \frac{\partial \psi}{\partial s}(s, t) \end{aligned}$$

y también,

$$\begin{aligned} \frac{\partial(g \circ \psi)}{\partial t}(s, t) &= (g \circ \psi)'(s, t)(0, 1)^t = g'(\psi(s, t))\psi'(s, t)(0, 1)^t \\ &= g'(\psi(s, t)) \frac{\partial \psi}{\partial t}(s, t) \end{aligned}$$

por lo que

$$\begin{aligned} w \left(\frac{\partial(g \circ \psi)}{\partial s}(s, t), \frac{\partial(g \circ \psi)}{\partial t}(s, t) \right) &= w \left(g'(\psi(s, t)) \frac{\partial \psi}{\partial s}(s, t), g'(\psi(s, t)) \frac{\partial \psi}{\partial t}(s, t) \right) \\ &= w \left(\frac{\partial \psi}{\partial s}(s, t), \frac{\partial \psi}{\partial t}(s, t) \right) \end{aligned}$$

pues g , al ser una función simpléctica, satisface la definición (3.2). Si sustituimos esta última expresión en la integral para $\Omega(g(M))$ concluimos que $\Omega(g(M)) = \Omega(M)$. \square

Corolario 3.6.

Si $g : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}^2$ es una función simpléctica en U , entonces g preserva áreas.

Demostración. Tomemos $n = 1$ en el lema anterior, entonces $M \subset \mathbb{R}^2$. Escojamos $K = M$ con ψ la función identidad, entonces

$$\Omega(M) = \iint_M ds dt$$

representa el área de M y es conservada por g . \square

3.1.1. Transformaciones canónicas y simplecticidad

En el capítulo anterior estudiamos transformaciones de las coordenadas generalizadas para las cuales la forma canónica de las ecuaciones de Euler-Lagrange se conservaba. Es decir transformaciones para las cuales al escribir las ecuaciones de movimiento de Hamilton para las coordenadas transformadas se obtenían las mismas ecuaciones pero con un Hamiltoniano diferente, de hecho, si la función Hamiltoniana no depende explícitamente del tiempo, entonces la función

generatriz tampoco, pudiendo de este modo concluir, de acuerdo con las ecuaciones (2.40), (2.41), (2.42) y (2.43) que $H = H^*$. Es decir que cuando H no depende explícitamente del tiempo entonces una transformación canónica deja invariantes las ecuaciones del movimiento.

Si H no depende de t , el sistema (3.1) puede escribirse como

$$\dot{\mathbf{y}} = J^{-1} \nabla H(\mathbf{y}) \quad (3.9)$$

donde $\mathbf{y} = (\mathbf{p}, \mathbf{q})$.

Con esta notación, se tiene que $\mathbf{z} = \psi(\mathbf{y})$ es una transformación canónica de coordenadas si las ecuaciones de movimiento expresadas en términos de \mathbf{z} se escriben como

$$\dot{\mathbf{z}} = J^{-1} \nabla H^*(\mathbf{z}) \quad (3.10)$$

con $H^*(\mathbf{z}) = H(\psi^{-1}(\mathbf{z})) = H(\mathbf{y})$.

Entonces para este caso, puede demostrarse que una transformación canónica de coordenadas es equivalente a una transformación simpléctica (no necesariamente lineal) de las coordenadas.

Teorema 3.7.

Sea $\psi : U \rightarrow V$ un cambio de coordenadas tal que ψ y ψ^{-1} son funciones continuamente diferenciables. Entonces ψ es simpléctica si y solo si ψ es una transformación canónica de coordenadas.

Demostración. Por la regla de la cadena, si $\mathbf{z} = \psi(\mathbf{y})$ y considerando $H^*(\mathbf{z}) = H(\mathbf{y})$ obtenemos

$$\dot{\mathbf{z}} = \psi'(\mathbf{y})\dot{\mathbf{y}} \quad \text{y} \quad \psi'(\mathbf{y})^t \nabla H^*(\mathbf{z}) = \nabla H(\mathbf{y})$$

entonces el sistema (3.9) en las nuevas variables se escribe como

$$\dot{\mathbf{z}} = \psi'(\mathbf{y})J^{-1}\psi'(\mathbf{y})^t \nabla H^*(\mathbf{z}) \quad (3.11)$$

y es equivalente al sistema (3.10) si

$$\psi'(\mathbf{y})J^{-1}\psi'(\mathbf{y})^t = J^{-1} \quad (3.12)$$

si multiplicamos por la derecha esta igualdad por $\psi'(\mathbf{y})^{-t}$ y por la izquierda por $\psi'(\mathbf{y})^{-1}$ y tomamos inversos obtenemos $J = \psi'(\mathbf{y})^t J \psi'(\mathbf{y})$ lo que muestra que la ecuación (3.12) es equivalente a la simplecticidad de ψ . Es decir que hemos probado que si ψ es simpléctica entonces es una transformación canónica. Si ahora ψ es una transformación canónica, entonces de (3.11) y (3.10) se obtiene (3.12) lo que significa que ψ es simpléctica. \square

Reconstrucción de una transformación simpléctica

Una implicación importante del teorema (3.7) es que dada una función generatriz de una transformación canónica de coordenadas del sistema (3.9) es posible reconstruir una transformación simpléctica vía las ecuaciones (2.40), (2.41), (2.42) y (2.43) dependiendo de las variables que se elijan como independientes.

3.2. Integrales primeras

La idea de una integral primera es la de una cantidad que permanece constante a lo largo de las trayectorias que solucionan un cierto sistema de ecuaciones diferenciales. La siguiente definición la establece formalmente.

Definición 3.4 (Integral primera).

Una función $I : \mathbb{R}^k \rightarrow \mathbb{R}$ es una integral primera del sistema de k ecuaciones diferenciales $\dot{y} = f(y)$ si

$$\frac{d}{dt}(I(y(t))) = 0$$

donde $y(t)$ es cualquier solución del sistema.

La definición anterior implica que $I(y(t)) = I(y_0) = cte$. En lo sucesivo usaremos indistintamente los términos **integral primera**, **cantidad conservada**, **invariante** o también **constante de movimiento**.

Es posible encontrar una condición necesaria y suficiente para que una función Φ sea una integral primera, para ello es necesario definir el corchete de Poisson.

Definición 3.5 (Corchete de Poisson).

Si $\Phi_1(x, y)$ y $\Phi_2(x, y)$ son dos funciones reales continuamente diferenciables definidas sobre algún abierto $U \subset \mathbb{R}^{2n}$ el corchete de Poisson se define mediante la siguiente expresión

$$[\Phi_1, \Phi_2] = \frac{\partial \Phi_1}{\partial y} \cdot \frac{\partial \Phi_2}{\partial x} - \frac{\partial \Phi_1}{\partial x} \cdot \frac{\partial \Phi_2}{\partial y}.$$

La condición que buscamos queda establecida en el siguiente teorema

Teorema 3.8 (Condición de Frobenius).

La función $\Phi : \mathbb{R}^{2n} \rightarrow \mathbb{R}$ es una integral primera del sistema (3.1) si $[\Phi, H] = 0$ (considerando a H como función sólo de (p, q)).

Demostración. Debemos calcular la derivada total respecto de t de $\Phi(p, q)$, esto es

$$\begin{aligned} \frac{d\Phi}{dt}(p, q) &= \sum_{i=1}^n \frac{\partial \Phi}{\partial p_i} \dot{p}_i + \sum_{i=1}^n \frac{\partial \Phi}{\partial q_i} \dot{q}_i \\ &= \frac{\partial \Phi}{\partial p} \cdot \dot{p} - \frac{\partial \Phi}{\partial q} \cdot \dot{q} \\ &= \frac{\partial \Phi}{\partial q} \cdot \frac{\partial H}{\partial p} - \frac{\partial \Phi}{\partial p} \cdot \frac{\partial H}{\partial q} \\ &= [\Phi, H] = 0 \end{aligned}$$

por lo tanto Φ satisface la definición (3.4). □

Corolario 3.9.

Si en el sistema (3.1) H no depende explícitamente del tiempo entonces es una integral primera de dicho sistema. Es decir que en el sistema (3.9) H es una integral primera.

Demostración. De acuerdo con el teorema anterior, lo único que debe verificarse es que $[H, H] = 0$ lo cual es inmediata de la definición (3.8). \square

El corolario (3.9) tiene como consecuencia que en un sistema de N partículas la energía total $H = T + U$ se conserva siempre que U no dependa de t^{**} .

Hemos visto ya que un función simpléctica tiene ciertas características geométricas como la conservación de área demostrada en el corolario (3.6). Ahora veremos que los sistemas Hamiltonianos tienen flujo simpléctico y por ende su geometría tiene los rasgos que ya hemos empezado a estudiar. Primero recordemos la definición de flujo para un sistema Hamiltoniano.

Definición 3.6 (Flujo Hamiltoniano).

La función $\varphi_t : U \rightarrow \mathbb{R}^{2n}$ es el flujo del sistema Hamiltoniano (3.1) si

$$\varphi_t(\mathbf{p}_0, \mathbf{q}_0) = (\mathbf{p}(t, \mathbf{p}_0, \mathbf{q}_0), \mathbf{q}(t, \mathbf{p}_0, \mathbf{q}_0))$$

donde $(\mathbf{p}(t, \mathbf{p}_0, \mathbf{q}_0), \mathbf{q}(t, \mathbf{p}_0, \mathbf{q}_0))$ es la solución del sistema Hamiltoniano al tiempo t para la condición inicial $\mathbf{p}(0) = \mathbf{p}_0$ y $\mathbf{q}(0) = \mathbf{q}_0$

Así, la definición anterior, indica que el flujo de un sistema Hamiltoniano es la función que avanza la solución del sistema al tiempo t . Una característica del flujo queda establecida en el siguiente lema.

Lema 3.10.

Si φ_t es el flujo del sistema (3.9), entonces $\partial\varphi_t/\partial\mathbf{y}_0$ es solución a la ecuación variacional

$$\dot{\Psi} = J^{-1}\nabla^2 H(\varphi_t(\mathbf{y}_0))\Psi \quad (3.13)$$

donde $\mathbf{y}_0 = (\mathbf{p}_0, \mathbf{q}_0)$.

Demostración. Comenzaremos derivando respecto del tiempo a $\partial\varphi_t/\partial\mathbf{y}_0$ y utilizando las ecuaciones del sistema (3.9)

$$\begin{aligned} \frac{d}{dt} \left(\frac{\partial\varphi_t}{\partial\mathbf{y}_0} \right) &= \frac{d}{dt} \left(\frac{\partial\mathbf{y}}{\partial\mathbf{y}_0} \right) = \frac{\partial\dot{\mathbf{y}}}{\partial\mathbf{y}_0} = \frac{\partial}{\partial\mathbf{y}_0} (J^{-1}\nabla H(\mathbf{y})) \\ &= J^{-1} \frac{\partial}{\partial\mathbf{y}_0} (\nabla H(\mathbf{y})) = J^{-1}\nabla^2 H(\mathbf{y}) \frac{\partial\mathbf{y}}{\partial\mathbf{y}_0} = J^{-1}\nabla^2 H(\mathbf{y}) \frac{\partial\varphi_t}{\partial\mathbf{y}_0} \end{aligned}$$

es decir

$$\frac{d}{dt} \left(\frac{\partial\varphi_t}{\partial\mathbf{y}_0} \right) = J^{-1}\nabla^2 H(\varphi_t(\mathbf{y}_0)) \frac{\partial\varphi_t}{\partial\mathbf{y}_0}$$

por lo tanto queda este lema demostrado. \square

**Recordemos el ejemplo de la sección 2.5.

El siguiente resultado fue demostrado por Poincaré y establece que el flujo de un sistema Hamiltoniano es simpléctico.

Teorema 3.11 (Poincaré).

Sea $H(\mathbf{p}, \mathbf{q})$ una función continuamente diferenciable dos veces sobre $U \subset \mathbb{R}^{2n}$. Entonces para cada t el flujo φ_t del sistema (3.9) es una transformación simpléctica en donde esté definida.

Demostración. Calcularemos

$$\frac{d}{dt} \left(\left(\frac{\partial \varphi_t}{\partial \mathbf{y}_0} \right)^t J \left(\frac{\partial \varphi_t}{\partial \mathbf{y}_0} \right) \right) = \frac{d}{dt} \left(\frac{\partial \varphi_t}{\partial \mathbf{y}_0} \right)^t J \left(\frac{\partial \varphi_t}{\partial \mathbf{y}_0} \right) + \left(\frac{\partial \varphi_t}{\partial \mathbf{y}_0} \right)^t J \frac{d}{dt} \left(\frac{\partial \varphi_t}{\partial \mathbf{y}_0} \right)$$

que de acuerdo con el lema (3.10) y porque $\nabla^2 H(\mathbf{y})$ es una matriz simétrica (por hipótesis $H(\mathbf{y})$ es C^2) se transforma en

$$\begin{aligned} \frac{d}{dt} \left(\left(\frac{\partial \varphi_t}{\partial \mathbf{y}_0} \right)^t J \left(\frac{\partial \varphi_t}{\partial \mathbf{y}_0} \right) \right) = \\ \left(\frac{\partial \varphi_t}{\partial \mathbf{y}_0} \right)^t \nabla^2 H(\varphi_t(\mathbf{y}_0)) J^{-t} J \left(\frac{\partial \varphi_t}{\partial \mathbf{y}_0} \right) + \left(\frac{\partial \varphi_t}{\partial \mathbf{y}_0} \right)^t \nabla^2 H(\varphi_t(\mathbf{y}_0)) \left(\frac{\partial \varphi_t}{\partial \mathbf{y}_0} \right) \end{aligned}$$

y como $J^{-t} J = -I$ se tiene que

$$\frac{d}{dt} \left(\left(\frac{\partial \varphi_t}{\partial \mathbf{y}_0} \right)^t J \left(\frac{\partial \varphi_t}{\partial \mathbf{y}_0} \right) \right) = 0 \quad (3.14)$$

además, es claro que

$$\left(\frac{\partial \varphi_t}{\partial \mathbf{y}_0} \right)^t J \left(\frac{\partial \varphi_t}{\partial \mathbf{y}_0} \right) = J \quad \text{en} \quad t = 0$$

pues $\varphi_0 = \mathbf{y}_0$ y por tanto $\partial \varphi_t / \partial \mathbf{y}_0 = I$. De este modo concluimos, en virtud de (3.14) que

$$\left(\frac{\partial \varphi_t}{\partial \mathbf{y}_0} \right)^t J \left(\frac{\partial \varphi_t}{\partial \mathbf{y}_0} \right) = J \quad \text{para} \quad t \geq 0$$

y por lo tanto φ_t es simpléctica. \square

Como el teorema anterior establece que el flujo φ_t del sistema (3.9) es simpléctico, entonces si consideramos una variedad $M \subset \mathbb{R}^{2n}$ de dimensión dos totalmente inmersa en el plano (p_i, q_i) y tomamos $K = M_i$ visto como conjunto de \mathbb{R}^2 y ψ como la función identidad de K en M_i , tenemos que por el lema (3.5) que

$$\Omega(M_i) = \iint_{M_i} dq_i dp_i$$

se conserva bajo φ_t . De hecho, como M_i está totalmente contenida en el plano (p_i, q_i) , entonces la integral

$$\iint_{M_i} dq_j dp_j = 0 \quad \text{si} \quad i \neq j.$$

Entonces, podemos concluir que

$$\Omega(M_i) = \iint_{M_i} \sum_{k=1}^n dq_k dp_k \quad (3.15)$$

se preserve bajo φ_t . Más aún, como $\Omega(M_i)$ se preserve para $i = 1, \dots, n$, se tiene que $\sum_{i=1}^n \Omega(M_i)$ también se preserve. Por otro lado, podemos pensar a las M_i como las proyecciones de una superficie $S \subset \mathbb{R}^{2n}$ sobre los planos (p_i, q_i) para obtener que

$$\iint_S \sum_{k=1}^n dq_k dp_k \quad (3.16)$$

se preserve bajo φ_t . Esto pues la integral (3.16) es igual a $\sum_{i=1}^n \Omega(M_i)$ considerando cada $\Omega(M_i)$ según la expresión (3.15). La integral (3.16) es conocida como la primera de las **integrales invariantes de Poincaré**. Por otro lado, la integral (3.16) representa lo que se conoce como el **área simpléctica** de S . Así, usualmente se hace referencia a este hecho simplemente como la preservación del área del flujo Hamiltoniano.

Se puede demostrar de manera análoga pero más complicada que la forma general de la **integral invariante de Poincaré** dada por

$$\int \cdots \int_S \sum_{i_1 \leq \dots \leq i_k} dp_{i_1} \cdots dp_{i_k} dq_{i_1} \cdots dq_{i_k}$$

es una cantidad conservada por el sistema (3.1) para cualquier superficie $S \subset \mathbb{R}^{2n}$ $2k$ -dimensional. Cuando $k = n$ se tiene la conservación de volumen del espacio fase, este resultado es conocido como el **teorema de Liouville**.

3.3. Integradores simplécticos

Los integradores simplécticos comenzaron a estudiarse por DeVogelaere en 1956, por Ruth en 1983 y por Feng Kang en 1985. Si se considera el sistema

$$\frac{dy}{dt} = f(t, y) \quad t_0 \leq t \quad \text{con condición inicial} \quad y(t_0) = y_0, \quad (3.17)$$

entonces integrarlo simplecticamente significa que el flujo numérico del sistema (el que se obtiene mediante el integrador) sea una transformación simpléctica.

Definición 3.7 (Integrador simpléctico).

El método numérico de un paso³

$$\mathbf{y}_{k+1} = \mathbf{y}_k + h\Phi(t_k, \mathbf{y}_k) \quad (3.18)$$

es simpléctico si la transformación $\Lambda_h(\mathbf{y}_0) = \mathbf{y}_0 + h\Phi(t_0, \mathbf{y}_0) = \mathbf{y}_1$ es simpléctica para cualquier elección de \mathbf{y}_0 .

La definición (1.4) dada en el capítulo 1 establece lo que entenderemos como el orden del método.

Uno de los primeros intentos por encontrar un método simpléctico es el llamado método de **Euler implícito** que, para un sistema como (3.17), se escribe como

$$\mathbf{y}_{n+1} = \mathbf{y}_n + hf(t, \mathbf{y}_{n+1}). \quad (3.19)$$

Sin embargo aunque este método conserva mejor la estructura simpléctica de un sistema como (3.9) no es propiamente simpléctico.

Algunos de los integradores simplécticos clásicos para el sistema (3.9) son por ejemplo el método de **Euler simpléctico** definido por las ecuaciones

$$\mathbf{p}_{n+1} = \mathbf{p}_n - h \frac{\partial H}{\partial \mathbf{q}}(\mathbf{p}_n, \mathbf{q}_{n+1}), \quad \mathbf{q}_{n+1} = \mathbf{q}_n - h \frac{\partial H}{\partial \mathbf{p}}(\mathbf{p}_n, \mathbf{q}_{n+1})$$

este es un método de orden 1. Una modificación inmediata de este es el **método del punto intermedio**

$$\mathbf{y}_{n+1} = \mathbf{y}_n + hJ^{-1}\nabla H((\mathbf{y}_{n+1} + \mathbf{y}_n)/2)$$

Otro tipo de esquemas de integración fueron desarrollados para Hamiltonianos separables, es decir para Hamiltonianos H tales que

$$H(\mathbf{p}, \mathbf{q}) = T(\mathbf{p}) + U(\mathbf{q}).$$

Para estos casos, el Hamiltoniano anterior es considerado como suma de dos Hamiltonianos, una dependiente sólo de \mathbf{p} y el otro de \mathbf{q} . Los correspondientes sistemas son

$$\begin{array}{l} \dot{\mathbf{p}} = 0 \\ \dot{\mathbf{q}} = T_{\mathbf{p}}(\mathbf{p}) \end{array} \quad \text{y} \quad \begin{array}{l} \dot{\mathbf{p}} = -U_{\mathbf{q}}(\mathbf{q}) \\ \dot{\mathbf{q}} = 0 \end{array}$$

que pueden ser resueltos sin problema para obtener que

$$\begin{array}{l} \mathbf{p}(t) = \mathbf{p}_0 \\ \mathbf{q}(t) = \mathbf{q}_0 + tT_{\mathbf{p}}(\mathbf{p}_0) \end{array} \quad \text{y} \quad \begin{array}{l} \dot{\mathbf{p}} = \mathbf{p}_0 - tU_{\mathbf{q}}(\mathbf{q}_0) \\ \dot{\mathbf{q}} = \mathbf{q}_0 \end{array}$$

ahora bien, si denotamos por φ_t^T y φ_t^U los flujos correspondientes a los sistemas anteriores, tenemos que el método de Euler simpléctico es justo la composición $\varphi_h^U \circ \varphi_h^T$.

³Esta es la definición general de método de un paso dada en (1.2) sólo que generalizada a sistemas de ecuaciones.

Otro método simpléctico basado en esta idea es el de **Störmer-Verlet**, éste se forma con la composición $\varphi_{h/2}^U \circ \varphi_h^T \circ \varphi_{h/2}^U$. Un rasgo interesante de este modo de proceder es que la propia definición del método establece que es simpléctico, pues sabemos según el teorema de Poincaré (3.11) que el flujo de un sistema Hamiltoniano es simpléctico y obviamente la composición de transformaciones simplécticas es simpléctica.

Una idea un poco más general de este esquema de integración es considerar como método a

$$\Lambda_h = \varphi_{b_m h}^U \circ \varphi_{a_m h}^T \circ \cdots \circ \varphi_{b_1 h}^U \circ \varphi_{a_1 h}^T$$

donde, por supuesto $\sum_{i=1}^m a_i = \sum_{i=1}^m b_i = 1$. Se puede demostrar que si $\sum_{i=1}^m a_i^{p+1} = \sum_{i=1}^m b_i^{p+1} = 0$ entonces el método definido por Λ_h de al menos orden p . Este tipo de esquemas fueron comenzados a estudiar por Ruth en 1983 para desarrollar métodos de orden 4 y sistemáticamente fueron estudiados por Yoshida en 1990 y por Suzuki en 1990 y 1992.

Otra tendencia a sido la de "simplectificar" los métodos clásicos que ha dado como resultado por ejemplo el método de **Runge-Kutta simpléctico** de orden R el cual se obtiene al imponer la siguiente condición al método clásico dado por (1.3)

$$a_r q_{r,s} + a_s q_{s,r} = a_r a_s \quad \text{con} \quad r, s = 1, \dots, R.$$

Sin embargo, nosotros nos concentraremos en el desarrollo de los integradores simplécticos basados en funciones generatrices.

3.4. Integradores basados en función generatriz

En esta sección buscaremos un integrador del sistema (3.1) cuando H no depende explícitamente del tiempo basado en funciones generatrices. La estrategia, es compaginar el paso de integración con el avance en el tiempo del flujo del sistema. Para lograr dicho efecto, buscaremos una transformación de coordenadas que transforme las coordenadas generalizadas en un tiempo inicial t_0 en las coordenadas generalizadas al tiempo t . Esto es posible ya que el teorema de Poincaré (3.11) establece que el flujo del sistema, como función de t , es una función simpléctica y entonces vía el teorema (3.7) es una transformación canónica de coordenadas así, para encontrar dicha transformación buscaremos su función generatriz G . De este modo obtendremos nuevos valores de las coordenadas generalizadas que por ser obtenidos mediante una transformación canónica dejan invariantes las ecuaciones de movimiento de Hamilton, es decir, satisfacen el sistema original (3.1).

Supongamos que $(\mathbf{p}_0, \mathbf{q}_0)$ son las condiciones iniciales del sistema (3.1), es decir $\mathbf{p}(t_0) = \mathbf{p}_0$ y $\mathbf{q}(t_0) = \mathbf{q}_0$. Para obtener el valor de (\mathbf{p}, \mathbf{q}) en un tiempo posterior, debemos transformar canónicamente las coordenadas $(\mathbf{p}_0, \mathbf{q}_0)$. Para ello, será conveniente considerar una función generatriz del tercer tipo, es decir $G = G(\mathbf{p}_0, \mathbf{q})$. Ésta genera la transformación de coordenadas dada por las

siguientes ecuaciones ⁴

$$\mathbf{p} = -\frac{\partial G}{\partial \mathbf{q}} \quad (3.20)$$

$$\mathbf{q}_0 = -\frac{\partial G}{\partial \mathbf{p}_0}. \quad (3.21)$$

Si expandemos la función generatriz como

$$G(\mathbf{p}_0, \mathbf{q}) = \sum_{m=0}^{\infty} \frac{h^m}{m!} K_m(\mathbf{p}_0, \mathbf{q}) \quad (3.22)$$

donde $h = t - t_0$ es el paso de integración y $K_0 = -\mathbf{p}_0^t \cdot \mathbf{q}^5$ es, según vimos en la sección 2.5, la función que genera la transformación identidad. La elección de K_0 de esta manera se debe a que el flujo $\varphi_t(\mathbf{p}_0, \mathbf{q}_0)$ del sistema (3.1) satisface que $\lim_{t \rightarrow t_0} \varphi_t = \varphi_{t_0}$ pero $\varphi_{t_0}(\mathbf{p}_0, \mathbf{q}_0) = (\mathbf{p}_0, \mathbf{q}_0)$ es decir que el límite del flujo cuando $t \rightarrow t_0$ es la transformación identidad. En otras palabras $(\mathbf{p}, \mathbf{q}) \rightarrow (\mathbf{p}_0, \mathbf{q}_0)$ cuando $t \rightarrow t_0$. Si además,

$$\mathbf{p} = \mathbf{p}_0 + \sum_{m=1}^{\infty} \frac{h^m}{m!} \mathbf{p}_m(\mathbf{p}_0, \mathbf{q}), \quad (3.23)$$

entonces, si sustituimos la ecuaciones (3.22) y (3.23) en (3.20) e igualamos los coeficientes de las series tenemos que

$$\mathbf{p}_m = -\frac{\partial K_m}{\partial \mathbf{q}}(\mathbf{p}_0, \mathbf{q}), \quad m \in \mathbb{N}. \quad (3.24)$$

Por otro lado, tenemos que la derivada total de \mathbf{p} respecto del tiempo esta dada por

$$\dot{\mathbf{p}} = \frac{\partial \mathbf{p}}{\partial t} + \frac{\partial \mathbf{p}}{\partial \mathbf{q}} \cdot \dot{\mathbf{q}}. \quad (3.25)$$

Utilizando la expresión (3.23) para \mathbf{p} y sustituyendo en la ecuación (3.25) obtenemos

$$\dot{\mathbf{p}} = \sum_{m=1}^{\infty} \frac{h^{m-1}}{(m-1)!} \mathbf{p}_m + \sum_{m=1}^{\infty} \frac{h^m}{m!} \frac{\partial \mathbf{p}_m}{\partial \mathbf{q}} \cdot \dot{\mathbf{q}},$$

Así, si sustituimos la expresión para $\dot{\mathbf{p}}$ del sistema (3.1) tenemos

$$\sum_{m=1}^{\infty} \frac{h^{m-1}}{(m-1)!} \mathbf{p}_m + \sum_{m=1}^{\infty} \frac{h^m}{m!} \frac{\partial \mathbf{p}_m}{\partial \mathbf{q}} \cdot \dot{\mathbf{q}} = -\frac{\partial H}{\partial \mathbf{q}}(\mathbf{p}_0 + \Delta \mathbf{p}, \mathbf{q}), \quad (3.26)$$

donde

$$\Delta \mathbf{p} = \sum_{m=1}^{\infty} \frac{h^m}{m!} \mathbf{p}_m(\mathbf{p}_0, \mathbf{q}) = \mathbf{p} - \mathbf{p}_0. \quad (3.27)$$

⁴La transformación de coordenadas para este caso está dada por las ecuaciones (2.42).

⁵Donde \cdot denota el producto de matrices. En adelante se conservara esta notación.

Ahora bien, utilizando la expresión correspondiente a $\dot{\mathbf{q}}$ del sistema (3.1), podemos expandir en serie de Taylor⁶ cada entrada de $\dot{\mathbf{q}}$ alrededor de \mathbf{p}_0 . Así, obtenemos

$$\dot{\mathbf{q}} = \left[\sum_{k=0}^{\infty} \frac{1}{k!} \left(\Delta \mathbf{p}^t \cdot \frac{\partial}{\partial \mathbf{p}} \right)^k \right] \frac{\partial H}{\partial \mathbf{p}}(\mathbf{p}_0, \mathbf{q}). \quad (3.28)$$

Lo mismo podemos hacer con el lado derecho de la ecuación (3.26) para obtener una expresión de $\dot{\mathbf{p}}$. Obtenemos así

$$\dot{\mathbf{p}} = \left[\sum_{k=0}^{\infty} \frac{1}{k!} \left(\Delta \mathbf{p}^t \cdot \frac{\partial}{\partial \mathbf{p}} \right)^k \right] \frac{\partial H}{\partial \mathbf{q}}(\mathbf{p}_0, \mathbf{q}). \quad (3.29)$$

Si definimos al operador M mediante la ecuación

$$M = \sum_{k=0}^{\infty} \frac{1}{k!} \left(\Delta \mathbf{p}^t \cdot \frac{\partial}{\partial \mathbf{p}} \right)^k \quad (3.30)$$

y sustituimos (3.28) y (3.29) en (3.26), obtenemos

$$\sum_{m=0}^{\infty} \frac{h^m}{m!} \mathbf{P}_{m+1} = \left(\sum_{m=1}^{\infty} \frac{h^m}{m!} \frac{\partial^2 K_m}{\partial \mathbf{q}^2} \cdot M \frac{\partial H}{\partial \mathbf{p}} \right) - M \frac{\partial H}{\partial \mathbf{q}}. \quad (3.31)$$

Donde todas las derivadas parciales de H están evaluadas en $(\mathbf{p}_0, \mathbf{q})$.

El siguiente paso es, para $m \in \mathbb{N} \cup \{0\}$, igualar el coeficiente de h^m de ambos lados de la igualdad anterior y así encontrar una expresión de \mathbf{P}_{m+1} . Como el operador M involucra potencias de h de orden arbitrario, debemos obtener de él sólo una parte que involucre al menos todas las potencias de h de grado menor o igual a m . Para ello trabajaremos sobre la ecuación (3.30) la cual define a M .

De acuerdo con el teorema multinomial⁷, tenemos que para $k \in \mathbb{N} \cup \{0\}$

$$\begin{aligned} \left(\Delta \mathbf{p}^t \cdot \frac{\partial}{\partial \mathbf{p}} \right)^k &= \left(\Delta p_1 \frac{\partial}{\partial p_1} + \dots + \Delta p_n \frac{\partial}{\partial p_n} \right)^k \\ &= \sum_{\mathbf{r} \in N_n(k)} \frac{k!}{\mathbf{r}!} \left(\Delta p_1 \frac{\partial}{\partial p_1} \right)^{r_1} \dots \left(\Delta p_n \frac{\partial}{\partial p_n} \right)^{r_n}. \end{aligned} \quad (3.32)$$

Si $k \in \mathbb{N} \cup \{0\}$ y $d \in \mathbb{N}$, definimos

$$N_d(k) = \{ \mathbf{v} = (v_1, \dots, v_d) \in (\mathbb{N} \cup \{0\})^d : v_1 + \dots + v_d = k \} \quad (3.33)$$

y si $\mathbf{v} \in N_d(k)$ entonces, definimos también

$$\mathbf{v}! = v_1! \dots v_d!. \quad (3.34)$$

⁶Ver S. Lang [9].

⁷Ver M. H. Protter [10].

De esta manera, si $\mathbf{r} \in N_n(k)$ debemos calcular $(\Delta \mathbf{p}_i)^{r_i}$ que de acuerdo con la definición de $\Delta \mathbf{p}$ dada por (3.27) se tiene que

$$(\Delta p_i)^{r_i} = \left(\sum_{l=1}^{\infty} \frac{h^l}{l!} p_{li} \right)^{r_i} \quad (i = 1, \dots, n) \quad (3.35)$$

donde p_{li} es la i -ésima componente del vector \mathbf{p}_l .

Podemos observar que cada Δp_i ($i = 1, \dots, n$) tiene potencias de h mayores o iguales a 1. Entonces, cuando $m \neq 0$ para encontrar el coeficiente de h^m no es necesario considerar la suma en (3.35) hasta infinito, basta con sumar hasta m , es decir, que para $\mathbf{r} \in N_n(k)$ sólo tenemos que calcular

$$({}^m \Delta p_i)^{r_i} = \left(\sum_{l=1}^m \frac{h^l}{l!} p_{li} \right)^{r_i} \quad (i = 1, \dots, n)$$

que de nuevo por el teorema multinomial podemos escribir para $i = 1, \dots, n$, como

$$\begin{aligned} ({}^m \Delta p_i)^{r_i} &= \sum_{\mathbf{x}^i \in N_m(r_i)} \frac{r_i!}{\mathbf{x}^i!} \left(h \frac{p_{1i}}{1!} \right)^{x_1^i} \dots \left(h^m \frac{p_{mi}}{m!} \right)^{x_m^i} \\ &= \sum_{\mathbf{x}^i \in N_m(r_i)} \frac{r_i!}{\mathbf{x}^i!} h^{\sum_{j=1}^m j x_j^i} \left(\frac{p_{1i}}{1!} \right)^{x_1^i} \dots \left(\frac{p_{mi}}{m!} \right)^{x_m^i}. \end{aligned} \quad (3.36)$$

Así entonces, la ecuación (3.36) permite obtener para $m \in \mathbb{N}$ y $\mathbf{r} \in N_n(k)$ que

$$\begin{aligned} \prod_{i=1}^n ({}^m \Delta p_i)^{r_i} &= \\ \sum_{\mathbf{x}^1 \in N_m(r_1)} \dots \sum_{\mathbf{x}^n \in N_m(r_n)} \frac{\mathbf{r}!}{\mathbf{x}^1! \dots \mathbf{x}^n!} h^{\sum_{i=1}^n \sum_{j=1}^m j x_j^i} \prod_{i=1}^n \left(\frac{p_{1i}}{1!} \right)^{x_1^i} \dots \left(\frac{p_{mi}}{m!} \right)^{x_m^i}. \end{aligned} \quad (3.37)$$

Utilizando esta última igualdad, podemos obtener en virtud de (3.32) una expresión para

$$\left(\Delta \mathbf{p}^L \cdot \frac{\partial}{\partial \mathbf{p}} \right)^k \quad k \in \mathbb{N} \cup \{0\} \quad (3.38)$$

la cual contenga todos los términos que involucran potencias de h menores o iguales a m con $m \in \mathbb{N}$ (puede contener potencias de h mayores a m , dependiendo de k , pero al menos contiene todas las que son menores o iguales a m). Entonces, para $m \in \mathbb{N}$, si

$$({}^m \Delta \mathbf{p})^L = ({}^m \Delta p_1, \dots, {}^m \Delta p_n),$$

la expresión buscada para $k \in \mathbb{N} \cup \{0\}$ es

$$\begin{aligned} \left(({}^m \Delta \mathbf{p})^t \cdot \frac{\partial}{\partial \mathbf{p}} \right)^k &= \left({}^m \Delta p_1 \frac{\partial}{\partial p_1} + \dots + {}^m \Delta p_n \frac{\partial}{\partial p_n} \right)^k \\ &= \sum_{\mathbf{r} \in N_n(k)} \frac{k!}{\mathbf{r}!} \left({}^m \Delta p_1 \frac{\partial}{\partial p_1} \right)^{r_1} \dots \left({}^m \Delta p_n \frac{\partial}{\partial p_n} \right)^{r_n} \\ &= \sum_{\mathbf{r} \in N_n(k)} \frac{k!}{\mathbf{r}!} \prod_{i=1}^n \left(\frac{\partial}{\partial p_i} \right)^{r_i} \prod_{i=1}^n ({}^m \Delta p_i)^{r_i}, \end{aligned}$$

o bien, sustituyendo (3.37) en la relación anterior, obtenemos para $m \in \mathbb{N}$ que

$$\begin{aligned} \left(({}^m \Delta \mathbf{p})^t \cdot \frac{\partial}{\partial \mathbf{p}} \right)^k &= \sum_{\mathbf{r} \in N_n(k)} k! \prod_{i=1}^n \left(\frac{\partial}{\partial p_i} \right)^{r_i} \\ &\quad \sum_{\mathbf{x}^1 \in N_m(r_1)} \dots \sum_{\mathbf{x}^n \in N_m(r_n)} \frac{1}{\mathbf{x}^1! \dots \mathbf{x}^n!} h^{\sum_{i=1}^n \sum_{j=1}^m j x_j^i} \prod_{i=1}^n \left(\frac{p_{1i}}{1!} \right)^{x_1^i} \dots \left(\frac{p_{mi}}{m!} \right)^{x_m^i}. \end{aligned} \quad (3.39)$$

Como ya habíamos observado, en el lado derecho de la ecuación (3.31), el operador M tiene potencias de h de orden arbitrario. Dado que estamos interesados en los términos que involucran h^m con $m \in \mathbb{N} \cup \{0\}$, debemos obtener una expresión para el operador M de la cual podamos asegurar que contiene todos las potencias de h menores o iguales a m . Para el caso $m \in \mathbb{N}$, la expresión (3.39) es válida y contiene las potencias de h que buscamos, entonces podemos pensar en sustituir (3.38) por (3.39) en (3.30) y obtener así una parte del operador M que precisamente contiene al menos las potencias de h deseadas (puede contener potencias de h de orden mayor a m). Sin embargo esta expresión, que es

$$\sum_{k=0}^{\infty} \frac{1}{k!} \left(({}^m \Delta \mathbf{p})^t \cdot \frac{\partial}{\partial \mathbf{p}} \right)^k,$$

contiene potencias de h de todos los ordenes aun cuando hemos restringido la suma de $\Delta \mathbf{p}$ hasta m debido a que k puede tomar cualquier valor en $\mathbb{N} \cup \{0\}$. De hecho, basta con que $k \in \{0, \dots, m\}$ para generar las potencias de h menores o iguales a m ya que si k es mayor a m , el grado menor de h será mayor o igual a $m + 1$. Así pues, la parte de M que contiene las potencias de h que son a lo más de grado m cuando $m \in \mathbb{N}$ se obtiene al variar $k \in \{0, \dots, m\}$ y sustituir (3.38) por (3.39) en (3.30) siendo

$$\begin{aligned} {}^m M &= \sum_{k=0}^m \sum_{\mathbf{r} \in N_n(k)} \prod_{i=1}^n \left(\frac{\partial}{\partial p_i} \right)^{r_i} \\ &\quad \sum_{\mathbf{x}^1 \in N_m(r_1)} \dots \sum_{\mathbf{x}^n \in N_m(r_n)} \frac{h^{\sum_{i=1}^n \sum_{j=1}^m j x_j^i}}{\mathbf{x}^1! \dots \mathbf{x}^n!} \prod_{i=1}^n \prod_{j=1}^m \left(\frac{p_{ji}}{j!} \right)^{x_j^i}. \end{aligned} \quad (3.40)$$

Cuando $m = 0$ buscamos el coeficiente de h^0 es decir el término libre. Éste es fácil de hallar, ya que como ninguna Δp_i ($i = 1, \dots, n$) tiene términos libres, la única forma en la cual M genera dichos términos es, de acuerdo con su definición dada por (3.30), cuando $k = 0$ obteniendo así que

$${}^0M = 1 \quad (3.41)$$

es el término libre. De este modo, podemos encontrar la parte del lado derecho de la ecuación (3.31) que contiene al menos todas las potencias de h que son menores o iguales a m con $m \in \mathbb{N} \cup \{0\}$. Esta expresión es:

$$\left(\sum_{s=1}^m \frac{h^s}{s!} \frac{\partial^2 K_s}{\partial q^2} \cdot {}^{m-s}M \frac{\partial H}{\partial p} \right) - {}^mM \frac{\partial H}{\partial q}. \quad (3.42)$$

Donde todas las parciales en esta expresión están evaluadas en (p_0, q) . Notamos además que en la expresión anterior, la suma dentro del paréntesis no es infinita como en (3.31) sino que está restringida hasta m , pues sólo queremos generar potencias de h a lo más de grado m . Así, cada potencia h^s con $s = 1, \dots, m$ está multiplicada por los términos que involucra el operador ${}^{m-s}M$ (que contiene todas las potencias de h de grado menor o igual a $m - s$), de este modo el producto tiene todos los términos que involucran h^k con $k = 1, \dots, m$. El otro término también tiene, como ya se ha dicho, todas las potencias de h de grado menor o igual a m . Por tanto, la expresión (3.42) es una parte del lado derecho de la ecuación (3.31) que contiene todas las potencias de h de grado menor o igual a m . Así, si hemos de encontrar el coeficiente de h^m debemos restringir la suma en mM a solo los términos que tienen h^m , para ello definamos el siguiente conjunto

$$\Upsilon(m, r) = \left\{ \mathbf{x} = (x^1, \dots, x^n) \in N_m(r_1) \times \dots \times N_m(r_n) : \sum_{i=1}^n \sum_{j=1}^m j x_j^i = m \right\}, \quad (3.43)$$

donde

$$r = (r_1, \dots, r_n) \in (\mathbb{N} \cup \{0\})^n, \quad k \in \mathbb{N} \cup \{0\} \quad \text{y} \quad m \in \mathbb{N}.$$

Hecho lo anterior, podemos escribir para $m \in \mathbb{N}$ precisamente los términos que tienen h^m en el operador mM . Estos quedan determinados mediante la siguiente expresión

$$[{}^m]M = h^m C(m), \quad (3.44)$$

donde,

$$C(m) = \sum_{k=0}^m \sum_{r \in N_n(k)} \prod_{i=1}^n \left(\frac{\partial}{\partial p_i} \right)^{r_i} \sum_{\mathbf{x} \in \Upsilon(m, r)} \frac{1}{x^1! \dots x^n!} \prod_{i=1}^n \prod_{j=1}^m \left(\frac{p_{ji}}{j!} \right)^{x_j^i}. \quad (3.45)$$

En el caso $m = 0$, tenemos que

$$[{}^0]M = {}^0M = 1,$$

de donde, si

$${}^{(0)}M = h^0 C(0), \tag{3.46}$$

entonces

$$C(0) = 1. \tag{3.47}$$

Las ecuaciones (3.44) y (3.46) nos permiten escribir, de acuerdo con (3.31), la siguiente igualdad para $m \in \mathbb{N} \cup \{0\}$

$$\frac{h^m}{m!} P_{m+1} = \left(\sum_{s=1}^m \frac{h^s}{s!} \frac{\partial^2 K_s}{\partial \mathbf{q}^2} \cdot {}^{(m-s)}M \frac{\partial H}{\partial \mathbf{p}} \right) - {}^{(m)}M \frac{\partial H}{\partial \mathbf{q}}.$$

En donde todas las derivadas parciales en la ecuación anterior están evaluadas en $(\mathbf{p}_0, \mathbf{q})$. Si despejamos P_{m+1} de la última ecuación y la sustituimos por su expresión dada en (3.24), encontramos que para $m \in \mathbb{N} \cup \{0\}$ se tiene que

$$\frac{\partial K_{m+1}}{\partial \mathbf{q}} = -m! \left(\sum_{s=1}^m \frac{1}{s!} \frac{\partial^2 K_s}{\partial \mathbf{q}^2} \cdot C(m-s) \frac{\partial H}{\partial \mathbf{p}} \right) + m! C(m) \frac{\partial H}{\partial \mathbf{q}}. \tag{3.48}$$

Donde todas las derivadas parciales en la ecuación anterior están evaluadas en $(\mathbf{p}_0, \mathbf{q})$.

La ecuación (3.48) es la ecuación maestra para obtener K_s cuando $s \in \mathbb{N}$ (K_0 la conocemos desde el principio). De este modo, podemos encontrar la función generatriz $G(\mathbf{p}_0, \mathbf{q})$ de la ecuación (3.22) hasta el orden que se desee.

Calcularemos K_1 , K_2 y K_3 para ejemplificar el procedimiento a seguir si se desea encontrar cualquier K_s con $s \in \mathbb{N}$ y así aproximar G a cualquier orden.

Antes de comenzar los cálculos haremos una observación sobre el conjunto $\Upsilon(m, \mathbf{r})$ que es útil cuando se desea calcular K_s con $s \in \mathbb{N} \setminus \{1\}$.

Observación 3.12.

$\Upsilon(m, \mathbf{0}) = \emptyset$ para $m \in \mathbb{N}$.

Demostración. Es claro que el único vector $\mathbf{x} = (x^1, \dots, x^n) \in N_m(0) \times \dots \times N_m(0)$ es $\mathbf{x} = \mathbf{0}$, pues para cada $i \in \{1, \dots, n\}$ como $x^i \in N_m(0)$ debe pasar que $\sum_{j=1}^m x_j^i = 0$, lo cual sucede sólo si $x_j^i = 0$ para toda $j \in \{1, \dots, m\}$ es decir si $x^i = \mathbf{0}$. Entonces, de acuerdo con (3.43), tenemos que $\Upsilon(m, \mathbf{0}) = \emptyset$ pues $\sum_{i=1}^n \sum_{j=1}^m j x_j^i = 0 \neq m$. \square

Comenzaremos calculando K_1 , para ello fijaremos $m = 0$ en (3.48) obteniendo así

$$\frac{\partial K_1}{\partial \mathbf{q}} = C(0) \frac{\partial H}{\partial \mathbf{q}},$$

esto debido a que $m! = 0! = 1$ y a que la suma entre paréntesis en (3.48) no tiene sentido cuando $m = 0$. Además, de acuerdo con (3.47), $C_0 = 1$, entonces tenemos que

$$\frac{\partial K_1}{\partial \mathbf{q}} = \frac{\partial H}{\partial \mathbf{q}}.$$

Esta igualdad puede escribirse también como el sistema de ecuaciones

$$\frac{\partial K_1}{\partial q_i} = \frac{\partial H}{\partial q_i} \quad (i = 1, \dots, n).$$

Podemos integrar cada ecuación de este sistema de la siguiente manera

$$\int \frac{\partial K_1}{\partial q_i} dq_i = \int \frac{\partial H}{\partial q_i} dq_i \quad (i = 1, \dots, n),$$

obteniendo así que

$$K_1 = H. \quad (3.49)$$

Ahora obtendremos K_2 . Con ese fin, debemos fijar $m = 1$ en nuestra ecuación maestra (3.48). Obtenemos así que

$$\frac{\partial K_2}{\partial \mathbf{q}} = - \left(\frac{\partial^2 K_1}{\partial \mathbf{q}^2} \cdot C(0) \frac{\partial H}{\partial \mathbf{p}} \right) + C(1) \frac{\partial H}{\partial \mathbf{q}}. \quad (3.50)$$

Como $C(0) = 1$, solo falta calcular $C(1)$. De acuerdo con (3.45), tenemos

$$C(1) = \sum_{k=0}^1 \sum_{\mathbf{r} \in N_n(k)} \prod_{i=1}^n \left(\frac{\partial}{\partial p_i} \right)^{r_i} \sum_{\mathbf{x} \in \Upsilon(1, \mathbf{r})} \frac{1}{x^1! \cdots x^n!} \prod_{i=1}^n \prod_{j=1}^1 \left(\frac{p_{ji}}{j!} \right)^{x_j^i}.$$

En la igualdad anterior, tenemos que cuando $k = 0$ el único vector $\mathbf{r} \in N_n(0)$ es $\mathbf{r} = \mathbf{0}$ de acuerdo con la definición de $N_n(k)$ dada en (3.33). Así pues, tenemos ahora que describir $\Upsilon(1, \mathbf{0})$, pero la observación (3.12) nos dice que $\Upsilon(1, \mathbf{0}) = \emptyset$, entonces $C(1)$ se reduce a

$$C(1) = \sum_{\mathbf{r} \in N_n(1)} \prod_{i=1}^n \left(\frac{\partial}{\partial p_i} \right)^{r_i} \sum_{\mathbf{x} \in \Upsilon(1, \mathbf{r})} \frac{1}{x^1! \cdots x^n!} \prod_{i=1}^n \left(\frac{p_{1i}}{1!} \right)^{x_i^1}. \quad (3.51)$$

Por otro lado, si $\mathbf{r} \in N_n(1)$ se tiene que sus entradas suman 1, esto quiere decir que todas son cero excepto una. Es fácil ver, de acuerdo con su definición, que

$$N_n(1) = \{ \mathbf{v} = (v_1, \dots, v_n) \in \{0, 1\}^n : \exists i \text{ con } v_i = 1 \text{ y } v_j = 0 \forall j \neq i \}. \quad (3.52)$$

Entonces para cada vector $\mathbf{r} \in N_n(1)$ debemos describir $\Upsilon(\mathbf{r}, 1)$. Tomemos un elemento típico en $N_n(1)$, sea $\mathbf{r} = (0, \dots, r_i = 1, \dots, 0)$ tal elemento. Entonces

$$\begin{aligned} \Upsilon(1, \mathbf{r}) &= \left\{ \mathbf{x} \in N_1(0) \times \cdots \times N_1(r_i = 1) \times \cdots \times N_1(0) : \sum_{i=1}^n \sum_{j=1}^1 j x_j^i = 1 \right\} \\ &= \left\{ \mathbf{x} \in N_1(0) \times \cdots \times N_1(r_i = 1) \times \cdots \times N_1(0) : \sum_{i=1}^n x_1^i = 1 \right\}. \end{aligned}$$

Pero como $\mathbf{x} \in N_1(0) \times \cdots \times N_1(r_i = 1) \times \cdots \times N_1(0)$ entonces $x_j^i = 0$ siempre que $j \neq i$ y $x_1^i = 1$, entonces

$$\Upsilon(1, \mathbf{r}) = \{(0, \dots, x_1^i = 1, \dots, 0)\}. \quad (3.53)$$

Así, si sustituimos las ecuaciones (3.52) y (3.53) simultáneamente en (3.51), podemos escribir

$$C(1) = \sum_{i=1}^n \frac{\partial}{\partial p_i} p_{1i}. \quad (3.54)$$

Entonces, si sustituimos las expresiones de $C(0)$ y $C(1)$ en (3.50), obtenemos

$$\frac{\partial K_2}{\partial \mathbf{q}} = -\frac{\partial^2 K_1}{\partial \mathbf{q}^2} \cdot \frac{\partial H}{\partial \mathbf{p}} + \left[\sum_{i=1}^n \frac{\partial}{\partial p_i} p_{1i} \right] \frac{\partial H}{\partial \mathbf{q}}.$$

Esta igualdad también se puede escribir en forma de sistema como

$$\frac{\partial K_2}{\partial q_i} = -\frac{\partial}{\partial q} \left(\frac{\partial K_1}{\partial q_i} \right) \cdot \frac{\partial H}{\partial \mathbf{p}} + \left[\sum_{j=1}^n \frac{\partial}{\partial p_j} p_{1j} \right] \frac{\partial H}{\partial q_i} \quad (i = 1, \dots, n).$$

Para cada $i \in \{1, \dots, n\}$ la igualdad anterior se transforma en

$$\begin{aligned} \frac{\partial K_2}{\partial q_i} &= -\sum_{j=1}^n \frac{\partial^2 K_1}{\partial q_j \partial q_i} \frac{\partial H}{\partial p_j} - \sum_{j=1}^n \frac{\partial^2 H}{\partial p_j \partial q_i} \frac{\partial K_1}{\partial q_j} \\ &= -\sum_{j=1}^n \frac{\partial^2 K_1}{\partial q_j \partial q_i} \frac{\partial H}{\partial p_j} + \frac{\partial^2 H}{\partial p_j \partial q_i} \frac{\partial K_1}{\partial q_j} \end{aligned}$$

esto por la definición de p_1 dada en (3.24). Además si suponemos que K_1 es continuamente diferenciable dos veces, entonces

$$\begin{aligned} &= -\sum_{j=1}^n \frac{\partial}{\partial q_i} \left(\frac{\partial K_1}{\partial q_j} \frac{\partial H}{\partial p_j} \right) \\ &= -\frac{\partial}{\partial q_i} \sum_{j=1}^n \frac{\partial K_1}{\partial q_j} \frac{\partial H}{\partial p_j}. \end{aligned}$$

Obteniendo por fin que

$$\frac{\partial K_2}{\partial q_i} = -\frac{\partial}{\partial q_i} \left(\left(\frac{\partial K_1}{\partial \mathbf{q}} \right)^t \cdot \frac{\partial H}{\partial \mathbf{p}} \right) \quad (i = 1, \dots, n).$$

Si integramos el sistema anterior tenemos

$$\int \frac{\partial K_2}{\partial q_i} dq_i = -\int \frac{\partial}{\partial q_i} \left(\left(\frac{\partial K_1}{\partial \mathbf{q}} \right)^t \cdot \frac{\partial H}{\partial \mathbf{p}} \right) dq_i \quad (i = 1, \dots, n),$$

de donde finalmente obtenemos la ecuación para K_2

$$K_2 = -\left(\frac{\partial K_1}{\partial \mathbf{q}} \right)^t \cdot \frac{\partial H}{\partial \mathbf{p}}. \quad (3.55)$$

Comenzaremos ahora el cálculo de K_3 como hemos venido haciendo pero ahora fijando $m = 2$. Así, nuestra ecuación maestra (3.48) queda de la siguiente forma

$$\begin{aligned} \frac{\partial K_3}{\partial \mathbf{q}} &= -2! \left(\sum_{s=1}^2 \frac{1}{s!} \frac{\partial^2 K_s}{\partial \mathbf{q}^2} \cdot C(2-s) \frac{\partial H}{\partial \mathbf{p}} \right) + 2! C(2) \frac{\partial H}{\partial \mathbf{q}} \\ &= -2 \left(\frac{\partial^2 K_1}{\partial \mathbf{q}^2} \cdot C(1) \frac{\partial H}{\partial \mathbf{p}} + \frac{1}{2} \frac{\partial^2 K_2}{\partial \mathbf{q}^2} \cdot C(0) \frac{\partial H}{\partial \mathbf{p}} \right) + 2C(2) \frac{\partial H}{\partial \mathbf{q}}. \end{aligned} \quad (3.56)$$

De la ecuación anterior, sólo desconocemos $C(2)$, por ende, lo único que resta es encontrar una expresión para $C(2)$. De acuerdo con (3.45), podemos ver que

$$C(2) = \sum_{k=0}^2 \sum_{\mathbf{r} \in N_n(k)} \prod_{i=1}^n \left(\frac{\partial}{\partial p_i} \right)^{r_i} \sum_{\mathbf{x} \in \Upsilon(2, \mathbf{r})} \frac{1}{\mathbf{x}^1! \cdots \mathbf{x}^n!} \prod_{i=1}^n \prod_{j=1}^2 \left(\frac{p_{ji}}{j!} \right)^{x_j^i},$$

teniendo que cuando $k = 0$ el único vector $\mathbf{r} \in N_n(0)$ es $\mathbf{r} = \mathbf{0}$. Así las cosas, debemos describir $\Upsilon(2, \mathbf{0})$, pero la observación (3.12) implica que $\Upsilon(2, \mathbf{0}) = \emptyset$. Entonces podemos escribir

$$C(2) = \sum_{k=1}^2 \sum_{\mathbf{r} \in N_n(k)} \prod_{i=1}^n \left(\frac{\partial}{\partial p_i} \right)^{r_i} \sum_{\mathbf{x} \in \Upsilon(2, \mathbf{r})} \frac{1}{\mathbf{x}^1! \cdots \mathbf{x}^n!} \prod_{i=1}^n \left(\frac{p_{1i}}{1!} \right)^{x_1^i} \left(\frac{p_{2i}}{2!} \right)^{x_2^i}. \quad (3.57)$$

Cuando $k = 1$, tenemos que sumar sobre los vectores \mathbf{r} en $N_n(1)$. Utilizando la expresión (3.52) de $N_n(1)$, tomamos \mathbf{r} un elemento típico en $N_n(1)$ y debemos describir $\Upsilon(2, \mathbf{r})$. Entonces tomando en cuenta (3.52) y la definición de $\Upsilon(2, \mathbf{r})$, tenemos

$$\begin{aligned} \Upsilon(2, \mathbf{r}) &= \left\{ \mathbf{x} \in N_2(0) \times \cdots \times N_2(r_i = 1) \times \cdots \times N_2(0) : \sum_{i=1}^n \sum_{j=1}^2 j x_j^i = 2 \right\} \\ &= \left\{ \mathbf{x} \in N_2(0) \times \cdots \times N_2(r_i = 1) \times \cdots \times N_2(0) : \sum_{i=1}^n x_1^i + 2x_2^i = 2 \right\} \end{aligned}$$

y como $\mathbf{x}^j \in N_2(0)$ para $j \neq i$ entonces $\mathbf{x}^j = \mathbf{0}$ si $j \neq i$, así se tiene que

$$= \left\{ \mathbf{x} \in N_2(0) \times \cdots \times N_2(r_i = 1) \times \cdots \times N_2(0) : x_1^i + 2x_2^i = 2 \right\}$$

además, como $\mathbf{x}^i \in N_2(1)$, tenemos que $x_1^i + x_2^i = 1$, es decir que $x_1^i = 1$ ó $x_2^i = 1$; así concluimos que

$$= \left\{ \mathbf{x} \in N_2(0) \times \cdots \times N_2(r_i = 1) \times \cdots \times N_2(0) : x_2^i = 1 \right\}.$$

Es decir que finalmente

$$\Upsilon(2, \mathbf{r}) = \{(\mathbf{0}, \dots, (0, x_2^i = 1), \dots, \mathbf{0})\}. \quad (3.58)$$

Entonces, para $k = 1$ podemos sustituir simultáneamente las ecuaciones (3.52) y (3.58) en (3.57) para obtener

$$C(2) = \sum_{i=1}^n \frac{\partial}{\partial p_i} \frac{p_{2i}}{2!} + \sum_{\mathbf{r} \in N_n(2)} \prod_{i=1}^n \left(\frac{\partial}{\partial p_i} \right)^{r_i} \sum_{\mathbf{x} \in \Upsilon(2, \mathbf{r})} \frac{1}{\mathbf{x}^1! \cdots \mathbf{x}^n!} \prod_{i=1}^n \left(\frac{p_{1i}}{1!} \right)^{x_i} \left(\frac{p_{2i}}{2!} \right)^{x_i} \quad (3.59)$$

Ahora, encontraremos una mejor expresión para la suma sobre los índices $\mathbf{r} \in N_n(2)$. El primer paso es encontrar una mejor expresión de $N_n(2)$. Como los vectores en $N_n(2)$ son aquellos con entradas en los enteros no negativos y que su suma es 2, entonces la única forma de que ello suceda es que dos de sus entradas sean 1 y las demás 0 ó que una de sus entradas sea 2 y las demás 0. Así,

$$N_n(2) = N_n^1(2) \cup N_n^2(2), \quad (3.60)$$

siendo

$$N_n^1(2) = \{(v_1, \dots, v_n) \in \{0, 1\}^n : \exists i, j, i < j \text{ con } v_i = v_j = 1 \text{ y } v_k = 0 \text{ si } k \neq i, j\} \quad (3.61)$$

$$N_n^2(2) = \{(v_1, \dots, v_n) \in \{0, 1\}^n : \exists i \text{ con } v_i = 2 \text{ y } v_k = 0 \text{ si } k \neq i\}. \quad (3.62)$$

Es claro además que en (3.60) la unión es disjunta, es decir $N_n^1(2) \cap N_n^2(2) = \emptyset$, por lo tanto la suma sobre los vectores en $N_n(2)$ se puede escribir como dos sumas, una sobre los vectores en $N_n^1(2)$ y la otra sobre los vectores en $N_n^2(2)$. De esta manera, lo siguiente es tomar un vector $\mathbf{r} \in N_n^1(2)$ y describir $\Upsilon(\mathbf{r}, 2)$. Entonces, si para $i, j \in \{1, \dots, n\}$ con $i < j$, hacemos

$$N_2(i, j) = N_2(0) \times \cdots \times N_2(r_i = 1) \times \cdots \times N_2(r_j = 1) \times \cdots \times N_2(0)$$

tenemos

$$\begin{aligned} \Upsilon(2, \mathbf{r}) &= \left\{ \mathbf{x} \in N_2(i, j) : \sum_{i=1}^n \sum_{j=1}^2 j x_j^i = 2 \right\} \\ &= \left\{ \mathbf{x} \in N_2(i, j) : \sum_{i=1}^n x_1^i + 2x_2^i = 2 \right\} \end{aligned}$$

y como $\mathbf{x}^k \in N_2(0)$ para $k \neq i, j$ entonces $\mathbf{x}^k = 0$ si $k \neq i, j$, de este modo

$$= \left\{ \mathbf{x} \in N_2(i, j) : x_1^i + 2x_2^i + x_1^j + 2x_2^j = 2 \right\}$$

además, como $\mathbf{x}^i, \mathbf{x}^j \in N_2(1)$, tenemos que $x_1^i + x_2^i = x_1^j + x_2^j = 1$, es decir $x_1^i = 1$ ó $x_2^i = 1$ y $x_1^j = 1$ ó $x_2^j = 1$. Así, concluimos que

$$= \left\{ \mathbf{x} \in N_2(i, j) : x_1^i = x_1^j = 1 \right\},$$

de donde finalmente

$$\Upsilon(2, \mathbf{r}) = \{(0, \dots, (x_1^i = 1, 0), \dots, (x_1^j = 1, 0), \dots, 0)\}. \quad (3.63)$$

Luego, podemos sustituir simultáneamente las ecuaciones (3.61) y (3.63) en (3.59) para obtener

$$\begin{aligned} C(2) &= \sum_{i=1}^n \frac{\partial}{\partial p_i} \frac{p_{2i}}{2!} + \sum_{\substack{i,j=1 \\ i < j}}^n \frac{\partial}{\partial p_i \partial p_j} \frac{p_{1i} p_{1j}}{1! 1!} \\ &+ \sum_{\mathbf{r} \in N_n^2(2)} \prod_{i=1}^n \left(\frac{\partial}{\partial p_i} \right)^{r_i} \sum_{\mathbf{x} \in \Upsilon(2, \mathbf{r})} \frac{1}{\mathbf{x}^1! \cdots \mathbf{x}^n!} \prod_{i=1}^n \left(\frac{p_{1i}}{1!} \right)^{x_i^1} \left(\frac{p_{2i}}{2!} \right)^{x_i^2}. \end{aligned} \quad (3.64)$$

Basta describir $\Upsilon(2, \mathbf{r})$ cuando $\mathbf{r} \in N_n^2(2)$, supongamos pues que \mathbf{r} pertenece a $N_n^2(2)$, entonces de acuerdo con (3.43) y (3.62)

$$\begin{aligned} \Upsilon(2, \mathbf{r}) &= \left\{ \mathbf{x} \in N_2(0) \times \cdots \times N_2(r_i = 2) \times \cdots \times N_2(0) : \sum_{i=1}^n \sum_{j=1}^2 j x_j^i = 2 \right\} \\ &= \left\{ \mathbf{x} \in N_2(0) \times \cdots \times N_2(r_i = 2) \times \cdots \times N_2(0) : \sum_{i=1}^n x_1^i + 2x_2^i = 2 \right\} \end{aligned}$$

y como $\mathbf{x}^j \in N_2(0)$ para $j \neq i$ entonces $\mathbf{x}^j = \mathbf{0}$, así

$$= \{ \mathbf{x} \in N_2(0) \times \cdots \times N_2(r_i = 2) \times \cdots \times N_2(0) : x_1^i + 2x_2^i = 2 \}$$

además, como $\mathbf{x}^i \in N_2(2)$, tenemos que $x_1^i + x_2^i = 2$, es decir $x_1^i = 1$ y $x_2^i = 1$ ó $x_1^i = 2$ ó $x_2^i = 2$. Concluimos que

$$= \{ \mathbf{x} \in N_2(0) \times \cdots \times N_2(r_i = 2) \times \cdots \times N_2(0) : x_1^i = 2 \}.$$

Es decir que finalmente

$$\Upsilon(2, \mathbf{r}) = \{(0, \dots, (x_1^i = 2, 0), \dots, 0)\}. \quad (3.65)$$

Si sustituimos simultáneamente las ecuaciones (3.62) y (3.65) en (3.64) obtenemos

$$C(2) = \sum_{i=1}^n \frac{\partial}{\partial p_i} \frac{p_{2i}}{2!} + \sum_{\substack{i,j=1 \\ i < j}}^n \frac{\partial}{\partial p_i \partial p_j} \frac{p_{1i} p_{1j}}{1! 1!} + \sum_{i=1}^n \frac{1}{2!} \frac{\partial^2}{\partial p_i^2} \left(\frac{p_{1i}}{1!} \right)^2. \quad (3.66)$$

Ahora, como las ecuaciones (3.47), (3.54) y (3.66) nos dan los valores de $C(0)$, $C(1)$ y $C(2)$ respectivamente, podemos escribir en vez de (3.56)

$$\begin{aligned} \frac{\partial K_3}{\partial \mathbf{q}} &= -2 \left(\frac{\partial^2 K_1}{\partial \mathbf{q}^2} \cdot \left[\sum_{i=1}^n \frac{\partial}{\partial p_i} p_{1i} \right] \frac{\partial H}{\partial \mathbf{p}} + \frac{1}{2} \frac{\partial^2 K_2}{\partial \mathbf{q}^2} \cdot \frac{\partial H}{\partial \mathbf{p}} \right) \\ &+ 2 \left[\sum_{i=1}^n \frac{\partial}{\partial p_i} \frac{p_{2i}}{2} + \sum_{\substack{i,j=1 \\ i < j}}^n \frac{\partial}{\partial p_i \partial p_j} p_{1i} p_{1j} + \sum_{i=1}^n \frac{1}{2} \frac{\partial^2}{\partial p_i^2} p_{1i}^2 \right] \frac{\partial H}{\partial \mathbf{q}} \end{aligned}$$

la cual también podemos escribir, para $s = 1, \dots, n$, en forma de sistema como

$$\begin{aligned} \frac{\partial K_3}{\partial q_s} = & -2 \frac{\partial}{\partial q} \left(\frac{\partial K_1}{\partial q_s} \right)^t \cdot \left[\sum_{i=1}^n \frac{\partial}{\partial p_i} p_{1i} \right] \frac{\partial H}{\partial p} - \frac{\partial}{\partial q} \left(\frac{\partial K_2}{\partial q_s} \right)^t \cdot \frac{\partial H}{\partial p} \\ & + \left[\sum_{i=1}^n \frac{\partial}{\partial p_i} p_{2i} + 2 \sum_{\substack{i,j=1 \\ i < j}}^n \frac{\partial}{\partial p_i \partial p_j} p_{1i} p_{1j} + \sum_{i=1}^n \frac{\partial^2}{\partial p_i^2} p_{1i}^2 \right] \frac{\partial H}{\partial q_s} \end{aligned}$$

que sustituyendo (3.24) se transforma en

$$\begin{aligned} = & 2 \sum_{j=1}^n \frac{\partial^2 K_1}{\partial q_j \partial q_s} \sum_{i=1}^n \frac{\partial^2 H}{\partial p_i \partial p_j} \frac{\partial K_1}{\partial q_i} - \sum_{i=1}^n \frac{\partial^2 K_2}{\partial q_i \partial q_s} \frac{\partial H}{\partial p_i} \\ & - \sum_{i=1}^n \frac{\partial^2 H}{\partial p_i \partial q_s} \frac{\partial K_2}{\partial q_i} + 2 \sum_{\substack{i,j=1 \\ i < j}}^n \frac{\partial^3 H}{\partial p_i \partial p_j \partial q_s} \frac{\partial K_1}{\partial q_i} \frac{\partial K_1}{\partial q_j} + \sum_{i=1}^n \frac{\partial^3 H}{\partial p_i^2 \partial q_s} \left(\frac{\partial K_1}{\partial q_i} \right)^2 \end{aligned}$$

o agrupando términos podemos escribir

$$\begin{aligned} \frac{\partial K_3}{\partial q_s} = & 2 \sum_{i,j=1}^n \frac{\partial^2 K_1}{\partial q_j \partial q_s} \frac{\partial^2 H}{\partial p_i \partial p_j} \frac{\partial K_1}{\partial q_i} - \sum_{i=1}^n \left(\frac{\partial^2 K_2}{\partial q_i \partial q_s} \frac{\partial H}{\partial p_i} + \frac{\partial^2 H}{\partial p_i \partial q_s} \frac{\partial K_2}{\partial q_i} \right) \\ & + 2 \sum_{\substack{i,j=1 \\ i < j}}^n \frac{\partial^3 H}{\partial p_i \partial p_j \partial q_s} \frac{\partial K_1}{\partial q_i} \frac{\partial K_1}{\partial q_j} + \sum_{i=1}^n \frac{\partial^3 H}{\partial p_i^2 \partial q_s} \left(\frac{\partial K_1}{\partial q_i} \right)^2 \end{aligned}$$

además si suponemos que H es continuamente diferenciable dos veces, entonces

$$2 \sum_{\substack{i,j=1 \\ i < j}}^n \frac{\partial^3 H}{\partial p_i \partial p_j \partial q_s} \frac{\partial K_1}{\partial q_i} \frac{\partial K_1}{\partial q_j} = \sum_{\substack{i,j=1 \\ i \neq j}}^n \frac{\partial^3 H}{\partial p_i \partial p_j \partial q_s} \frac{\partial K_1}{\partial q_i} \frac{\partial K_1}{\partial q_j},$$

luego,

$$\begin{aligned} & 2 \sum_{\substack{i,j=1 \\ i < j}}^n \frac{\partial^3 H}{\partial p_i \partial p_j \partial q_s} \frac{\partial K_1}{\partial q_i} \frac{\partial K_1}{\partial q_j} + \sum_{i=1}^n \frac{\partial^3 H}{\partial p_i^2 \partial q_s} \left(\frac{\partial K_1}{\partial q_i} \right)^2 \\ & = \sum_{\substack{i,j=1 \\ i \neq j}}^n \frac{\partial^3 H}{\partial p_i \partial p_j \partial q_s} \frac{\partial K_1}{\partial q_i} \frac{\partial K_1}{\partial q_j} + \sum_{i=1}^n \frac{\partial^3 H}{\partial p_i^2 \partial q_s} \left(\frac{\partial K_1}{\partial q_i} \right)^2 \\ & = \sum_{i,j=1}^n \frac{\partial^3 H}{\partial p_i \partial p_j \partial q_s} \frac{\partial K_1}{\partial q_i} \frac{\partial K_1}{\partial q_j}, \end{aligned}$$

y por tanto

$$\begin{aligned}
 \frac{\partial K_3}{\partial q_s} &= 2 \sum_{i,j=1}^n \frac{\partial^2 K_1}{\partial q_j \partial q_s} \frac{\partial^2 H}{\partial p_i \partial p_j} \frac{\partial K_1}{\partial q_i} + \sum_{i,j=1}^n \frac{\partial^3 H}{\partial p_i \partial p_j \partial q_s} \frac{\partial K_1}{\partial q_i} \frac{\partial K_1}{\partial q_j} \\
 &\quad - \sum_{i=1}^n \left(\frac{\partial^2 K_2}{\partial q_i \partial q_s} \frac{\partial H}{\partial p_i} + \frac{\partial^2 H}{\partial p_i \partial q_s} \frac{\partial K_2}{\partial q_i} \right) \\
 &= \sum_{i,j=1}^n \frac{\partial}{\partial q_s} \left(\frac{\partial K_1}{\partial q_i} \frac{\partial K_1}{\partial q_j} \frac{\partial^2 H}{\partial p_i \partial p_j} \right) - \sum_{i=1}^n \frac{\partial}{\partial q_s} \left(\frac{\partial K_2}{\partial q_i} \frac{\partial H}{\partial p_i} \right) \\
 &= \frac{\partial}{\partial q_s} \sum_{i,j=1}^n \frac{\partial K_1}{\partial q_i} \frac{\partial K_1}{\partial q_j} \frac{\partial^2 H}{\partial p_i \partial p_j} - \frac{\partial}{\partial q_s} \sum_{i=1}^n \frac{\partial K_2}{\partial q_i} \frac{\partial H}{\partial p_i} \\
 &= \frac{\partial}{\partial q_s} \sum_{i,j=1}^n \frac{\partial K_1}{\partial q_i} \frac{\partial K_1}{\partial q_j} \frac{\partial^2 H}{\partial p_i \partial p_j} - \frac{\partial}{\partial q_s} \left(\left(\frac{\partial K_2}{\partial \mathbf{q}} \right)^t \cdot \frac{\partial H}{\partial \mathbf{p}} \right).
 \end{aligned}$$

Si integramos este sistema para cada $s \in \{1, \dots, n\}$ con respecto de q_s , tenemos que

$$\begin{aligned}
 \int \frac{\partial K_3}{\partial q_s} dq_s &= - \int \frac{\partial}{\partial q_s} \left(\left(\frac{\partial K_2}{\partial \mathbf{q}} \right)^t \cdot \frac{\partial H}{\partial \mathbf{p}} \right) dq_s \\
 &\quad + \int \frac{\partial}{\partial q_s} \sum_{i,j=1}^n \frac{\partial K_1}{\partial q_i} \frac{\partial K_1}{\partial q_j} \frac{\partial^2 H}{\partial p_i \partial p_j} dq_s
 \end{aligned}$$

obteniendo finalmente

$$K_3 = - \left(\frac{\partial K_2}{\partial \mathbf{q}} \right)^t \cdot \frac{\partial H}{\partial \mathbf{p}} + \sum_{i,j=1}^n \frac{\partial K_1}{\partial q_i} \frac{\partial K_1}{\partial q_j} \frac{\partial^2 H}{\partial p_i \partial p_j}. \quad (3.67)$$

En las ecuaciones (3.49), (3.55) y (3.67) todas las funciones que aparecen están evaluadas en $(\mathbf{p}_0, \mathbf{q})$.

De manera análoga, pueden calcularse la funciones K_m para cualquier entero positivo y así aproximar $G(\mathbf{p}_0, \mathbf{q})$ al orden que se desee. A continuación, escribimos K_4 , K_5 y K_6 .

$$K_4 = - \left(\frac{\partial K_3}{\partial \mathbf{q}} \right)^t \cdot \frac{\partial H}{\partial \mathbf{p}} + 3 \sum_{i,j=1}^n \frac{\partial K_1}{\partial q_i} \frac{\partial K_2}{\partial q_j} \frac{\partial^2 H}{\partial p_i \partial p_j} - \sum_{i,j,k=1}^n \frac{\partial K_1}{\partial q_i} \frac{\partial K_1}{\partial q_j} \frac{\partial K_1}{\partial q_k} \frac{\partial^3 H}{\partial p_i \partial p_j \partial p_k}, \quad (3.68)$$

$$\begin{aligned}
 K_5 &= - \left(\frac{\partial K_4}{\partial \mathbf{q}} \right)^t \cdot \frac{\partial H}{\partial \mathbf{p}} + 3 \sum_{i,j=1}^n \frac{\partial K_2}{\partial q_i} \frac{\partial K_2}{\partial q_j} \frac{\partial^2 H}{\partial p_i \partial p_j} - 6 \sum_{i,j,k=1}^n \frac{\partial K_1}{\partial q_i} \frac{\partial K_1}{\partial q_j} \frac{\partial K_1}{\partial q_k} \frac{\partial^3 H}{\partial p_i^3} \\
 &\quad + 4 \sum_{i,j=1}^n \frac{\partial K_1}{\partial q_i} \frac{\partial K_3}{\partial q_j} \frac{\partial^2 H}{\partial p_i \partial p_j} + \sum_{i,j,k,l=1}^n \frac{\partial K_1}{\partial q_i} \frac{\partial K_1}{\partial q_j} \frac{\partial K_1}{\partial q_k} \frac{\partial K_1}{\partial q_l} \frac{\partial^4 H}{\partial p_i \partial p_j \partial p_k \partial p_l},
 \end{aligned} \quad (3.69)$$

$$\begin{aligned}
K_6 = & - \left(\frac{\partial K_5}{\partial \mathbf{q}} \right)^t \cdot \frac{\partial H}{\partial \mathbf{p}} + 10 \sum_{i,j=1}^n \frac{\partial K_2}{\partial q_i} \frac{\partial K_3}{\partial q_j} \frac{\partial^2 H}{\partial p_i \partial p_j} - 10 \sum_{i,j,k=1}^n \frac{\partial K_1}{\partial q_i} \frac{\partial K_1}{\partial q_j} \frac{\partial K_3}{\partial q_k} \frac{\partial^3 H}{\partial p_i^3} \\
& + 5 \sum_{i,j=1}^n \frac{\partial K_1}{\partial q_i} \frac{\partial K_4}{\partial q_j} \frac{\partial^2 H}{\partial p_i \partial p_j} + 10 \sum_{i,j,k,l=1}^n \frac{\partial K_1}{\partial q_i} \frac{\partial K_1}{\partial q_j} \frac{\partial K_1}{\partial q_k} \frac{\partial K_2}{\partial q_l} \frac{\partial^4 H}{\partial p_i \partial p_j \partial p_k \partial p_l} \\
& - 15 \sum_{i,j,k=1}^n \frac{\partial K_1}{\partial q_i} \frac{\partial K_2}{\partial q_j} \frac{\partial K_2}{\partial q_k} \frac{\partial^3 H}{\partial p_i^3} \\
& - \sum_{i,j,k,l,m=1}^n \frac{\partial K_1}{\partial q_i} \frac{\partial K_1}{\partial q_j} \frac{\partial K_1}{\partial q_k} \frac{\partial K_1}{\partial q_l} \frac{\partial K_1}{\partial q_m} \frac{\partial^5 H}{\partial p_i \partial p_j \partial p_k \partial p_l \partial p_m}. \tag{3.70}
\end{aligned}$$

Utilizando estas expresiones para la funciones K_m con $m \in \mathbb{N}$, como ya habíamos dicho, se puede aproximar a la función generatriz G . De este modo las ecuaciones que definen la transformación canónica (3.20) y (3.21) pueden ser utilizadas de la siguiente manera, la ecuación (3.21) puede ser escrita como

$$\mathbf{q} = \mathbf{q}_0 + \sum_{m=1}^{\infty} \frac{h^m}{m!} \frac{\partial K_m}{\partial \mathbf{p}_0}(\mathbf{p}_0, \mathbf{q}). \tag{3.71}$$

Esta ecuación es implícita para \mathbf{q} pero cuando es resuelta, entonces (3.20) es explícita para \mathbf{p} . De este modo hemos resuelto el sistema (3.9) de manera exacta. Sin embargo, en la práctica el cálculo exacto de la serie (3.71) no es posible en general, así el método numérico que genera este procedimiento es limitar la expansión de G dada en (3.22) hasta un entero positivo k para obtener la expresión

$$\mathbf{q} = \mathbf{q}_0 + \sum_{m=1}^k \frac{h^m}{m!} \frac{\partial K_m}{\partial \mathbf{p}_0}(\mathbf{p}_0, \mathbf{q}). \tag{3.72}$$

Las ecuaciones (3.71) y (3.72) pueden ser resueltas iterativamente sustituyendo en el lado derecho \mathbf{q} por \mathbf{q}_0 para obtener una aproximación \mathbf{q}_1 de \mathbf{q} así, podemos repetir el proceso hasta una $k \in \mathbb{N}$ tal que $|\mathbf{q}_k - \mathbf{q}_{k+1}| < \epsilon$ con ϵ tan pequeña como se requiera. Otro modo de hacerlo es encontrar la raíz de la función

$$f(\mathbf{q}) = \mathbf{q}_0 - \mathbf{q} + \sum_{m=1}^{\omega} \frac{h^m}{m!} \frac{\partial K_m}{\partial \mathbf{p}_0}(\mathbf{p}_0, \mathbf{q}), \tag{3.73}$$

con $\omega = \infty$ ó $\omega = k$ según sea el caso, por el método de Newton-Raphson.

3.4.1. Orden y symplecticidad del Método

Cuando en la expresión de G se considera la serie completa y es posible resolver la ecuación (3.71) para \mathbf{q} entonces las ecuaciones (3.71) y (3.20) definen un integrador exacto del sistema (3.9). El caso en que se considera la serie (3.22) truncada hasta k entonces el error del método es justo el error que se comete

al aproximar en serie de Taylor a G hasta orden k , pues este error es el mismo que se comete en las ecuaciones (3.72) y (3.20) es decir, que de acuerdo con la definición (1.4) tenemos un método de orden k ya que para cada t se tiene que si expandimos en serie de Taylor cada entrada de $\mathbf{y}(t) = (\mathbf{p}(t), \mathbf{q}(t))$ hasta orden k alrededor de $t + h$ obtenemos que

$$\mathbf{y}(t+h) - \mathbf{y}(t) = \sum_{m=1}^k h^m \frac{d^m}{dt} \mathbf{y}(t) + O(h^{k+1}). \quad (3.74)$$

Por otro lado, en concordancia con la definición del método de un paso dada en (3.18) tenemos que

$$\Phi(t, \mathbf{y}(t)) = \frac{1}{h} \begin{pmatrix} - \sum_{m=1}^k \frac{h^m}{m!} \frac{\partial K_m}{\partial \mathbf{q}} \\ \sum_{m=1}^k \frac{h^m}{m!} \frac{\partial K_m}{\partial \mathbf{p}} \end{pmatrix} (\mathbf{p}_0, \mathbf{q}(t))$$

pero la primer entrada del vector en la expresión anterior es igual a $\mathbf{p} - \mathbf{p}_0 + O(h^{k+1})$ de acuerdo con la ecuación (3.23) y la segunda entrada es igual a $\mathbf{q} - \mathbf{q}_0 + O(h^{k+1})$ de acuerdo con la ecuación (3.71). Entonces obtenemos la expresión

$$\Phi(t, \mathbf{y}(t)) = \frac{1}{h} \begin{pmatrix} \mathbf{p}(t) - \mathbf{p}_0 \\ \mathbf{q}(t) - \mathbf{q}_0 \end{pmatrix} + O(h^{k+1}) \quad (3.75)$$

Si expandimos en serie de Taylor las funciones $\mathbf{p}_0 = \mathbf{p}(t_0)$ y $\mathbf{q}_0 = \mathbf{q}(t_0)$ alrededor de t , obtenemos que

$$\begin{aligned} \mathbf{p}_0 &= \sum_{m=0}^k h^m \frac{d^m}{dt} \mathbf{p}(t) + O(h^{k+1}) \\ \mathbf{q}_0 &= \sum_{m=0}^k h^m \frac{d^m}{dt} \mathbf{q}(t) + O(h^{k+1}). \end{aligned}$$

Así, la expresión (3.75) se transforma en

$$\Phi(t, \mathbf{y}(t)) = \sum_{m=1}^k h^{m-1} \frac{d^m}{dt} \mathbf{y}(t) + O(h^k) \quad (3.76)$$

y entonces el error de discretización en un punto dado por (1.2) se escribe como

$$\begin{aligned} L(t, h) &= \frac{1}{h} [\mathbf{y}(t+h) - \mathbf{y}(t)] - \Phi(t, \mathbf{y}(t)) \\ &= \sum_{m=1}^k h^{m-1} \frac{d^m}{dt} \mathbf{y}(t) - \sum_{m=1}^k h^{m-1} \frac{d^m}{dt} \mathbf{y}(t) + O(h^k) \\ &= O(h^k) \end{aligned}$$

gracias a las expresiones (3.76) y (3.74).

Como la igualdad anterior se tiene para toda t , se sigue de la definición (1.3) que $L(h) = O(h^k)$ es decir que según la definición de orden (1.4) el método es de orden k .

Si además se utiliza el método de Newton-Raphson para resolver cualquiera de las ecuaciones (3.71) ó (3.72) se tiene que considerar el error con el cual se aproxima la solución de dichas ecuaciones.

La simplecticidad del método definido por las ecuaciones (3.71) y (3.20) se sigue de que son obtenidas mediante la transformación canónica generada por G y entonces según el teorema (3.7) dicha transformación es simpléctica y por lo tanto de acuerdo con la definición (3.7) el método es simpléctico.

Una observación interesante es que el método de orden 1 es justo el método que se conoce como Euler simpléctico.

Capítulo 4

Simulaciones Numéricas

En este capítulo implementaremos los métodos numéricos que hemos estudiado tanto en el capítulo 1 como en el capítulo 3 para algunos sistemas Hamiltonianos clásicos con la finalidad de comparar la exactitud de los integradores en la preservación de algunas cantidades para tiempos de integración grandes. El primer sistema que estudiaremos representa matemáticamente el movimiento de un oscilador armónico, el segundo de ellos es el que surge al modelar el movimiento de un péndulo sin fricción¹ y el último será el que se conoce como la cadena de Toda.

La primer razón para elegir los problemas del oscilador, del péndulo y de la cadena de Toda es que conocemos el comportamiento explícito del sistema, por ejemplo, la elección del problema del oscilador armónico se debe a que es un sistema Hamiltoniano con una característica muy especial: sin importar el valor de la energía del sistema, el período de la solución permanece constante. El péndulo se escogió porque es un sistema clásicamente Hamiltoniano (entre otras cosas no tiene la particularidad mencionada para el oscilador) y la cadena de Toda porque el Hamiltoniano del sistema involucra la evaluación de exponenciales, lo que lo hace numéricamente inestable, es decir, los errores numéricos cometidos crecen exponencialmente. Un último motivo para la elección de estos sistemas es que el oscilador armónico y el péndulo ideal representan problemas con un grado de libertad, es decir que dependen solo de una posición y un momento generalizados y la cadena de Toda que hemos elegido puede transformarse (más adelante veremos como) en un problema con dos grados de libertad.

Todos los sistemas que hemos elegido son sistemas para los cuales el Hamiltoniano involucrado no depende explícitamente del tiempo y por el corolario (3.9) el propio Hamiltoniano (la energía del sistema) es una integral primera. Adicionalmente, para la cadena de Toda se conoce otra cantidad conservada. En las siguientes secciones al integrar estos sistemas, más que en la solución numérica del sistema estaremos interesados en la preservación de dichas cantidades.

Los métodos clásicos de un paso que se eligieron para integrar los sistemas

¹Péndulo ideal.

que hemos mencionado arriba son los métodos de: Euler (EU), Euler implícito (EUI) y Runge-Kutta de orden 4 (RK4), y de los métodos multipaso se eligieron los métodos de; Adams-Bashforth de orden 4 (AB4) y Adams-Moulton de orden 4 (M4). Para éstos últimos utilizaremos como arrancador el método de Runge-Kutta de orden 4.

De los métodos simplécticos estudiados en el capítulo 3 implementaremos sólo los de orden 1 (S1), orden 2 (S2) y orden 3 (S3).

Una vez elegidos los integradores que implementaremos, para determinar que tan precisos son en la conservación de la energía, calcularemos dos cantidades en cada paso de integración:

- (1) El error relativo de la energía promedio respecto del valor exacto de la energía.
- (2) El error relativo de la energía máxima respecto del valor exacto de la energía.

Más precisamente, si el sistema que integraremos es (3.9)

$$\dot{\mathbf{y}} = J^{-1} \nabla H(\mathbf{y})$$

y deseamos analizar el comportamiento numérico de una integral primera $I(\mathbf{y})$, debemos definir $I_k = I(\mathbf{y}_k)$ donde \mathbf{y}_k es la aproximación numérica a la solución exacta en el k -ésimo paso de integración, y entonces, si

$$IM_n = \max_{1 \leq k \leq n} \{I_k\} \quad \text{y} \quad IP_n = \frac{1}{n} \sum_{k=1}^n I_k,$$

calcular en el n -ésimo paso de integración las cantidades

$$(1) \quad I_{prom} = \left| \frac{I_0 - IP_n}{IP_n} \right| \quad (4.1)$$

$$(2) \quad I_{max} = \left| \frac{I_0 - IM_n}{IM_n} \right| \quad (4.2)$$

para $n \geq 1$.

En el análisis de las cantidades conservadas para cada sistema que estudiaremos en las próximas secciones, obtendremos los valores de las integrales primeras en los pasos de integración que son potencias de 2 y graficaremos el logaritmo en base 10 del tiempo contra el logaritmo en base 10 de las cantidades asociadas a las constantes de movimiento dadas en (4.1) y (4.2). Los resultados son presentados de este modo, por dos razones, la primera para que las gráficas sean más claras y la segunda para que al tomar los logaritmos en base 10 del tiempo y de (4.1) y (4.2) se obtenga una escala más o menos lineal que permita, a simple vista, darse cuenta de la precisión con la que el integrador numérico preserva las constantes de movimiento.

4.1. El oscilador armónico

En esta sección tomaremos el problema de describir el movimiento de una masa atada por un resorte a una pared²,

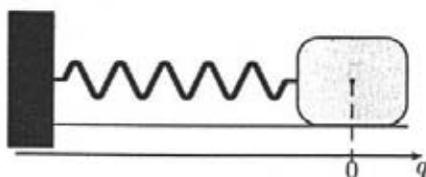


Figura 4.1: Oscilador armónico

Si q denota la posición de la masa en el tiempo, las ecuaciones de movimiento que consideraremos para este problema quedan dadas por

$$\dot{q} = p \quad (4.3)$$

$$\dot{p} = -9q \quad (4.4)$$

esto implica que

$$H(p, q) = \frac{1}{2}p^2 + \frac{9}{2}q^2$$

y por tanto su plano fase está descrito por las curvas de nivel de H . En la figura 4.2 es posible apreciar que dichas curvas son elipses sin importar las condiciones iniciales.

4.1.1. Resultados numéricos

Consideraremos las ecuaciones (4.3) y (4.4) con condiciones iniciales $p(0) = 3$ y $q(0) = 2$ e implementaremos los métodos clásicos de Euler (EU), Euler implícito (EUI) Runge-Kutta de orden 4 (RK4), Adams-Bashforth de orden 4 (AB4) y Adams-Moulton de orden 4 (M4). Para estos dos últimos utilizaremos como arrancador al método de RK4. También los simplécticos vistos en el capítulo anterior considerando solo los de orden 1 (S1), orden 2 (S2) y orden 3 (S3). Una primer observación es que los integradores de Euler implícito y el simpléctico de orden uno son exactamente el mismo para este caso, entonces solo haremos el análisis para S1 en lugar de EUI.

A continuación se muestran las gráficas de las cantidades (4.1) y (4.2) asociadas a la energía del sistema obtenidas por los diferentes integradores con distintos pasos de integración h en escalas logarítmicas.

²Ver R. L. Devaney [1].

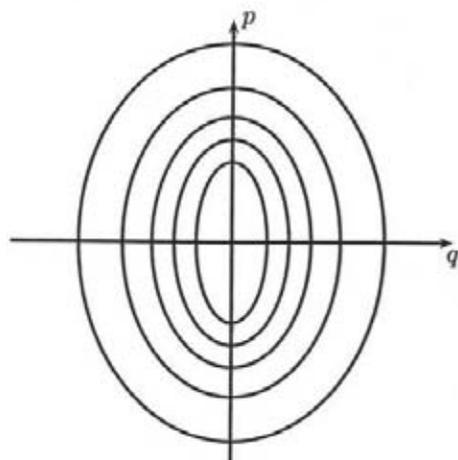
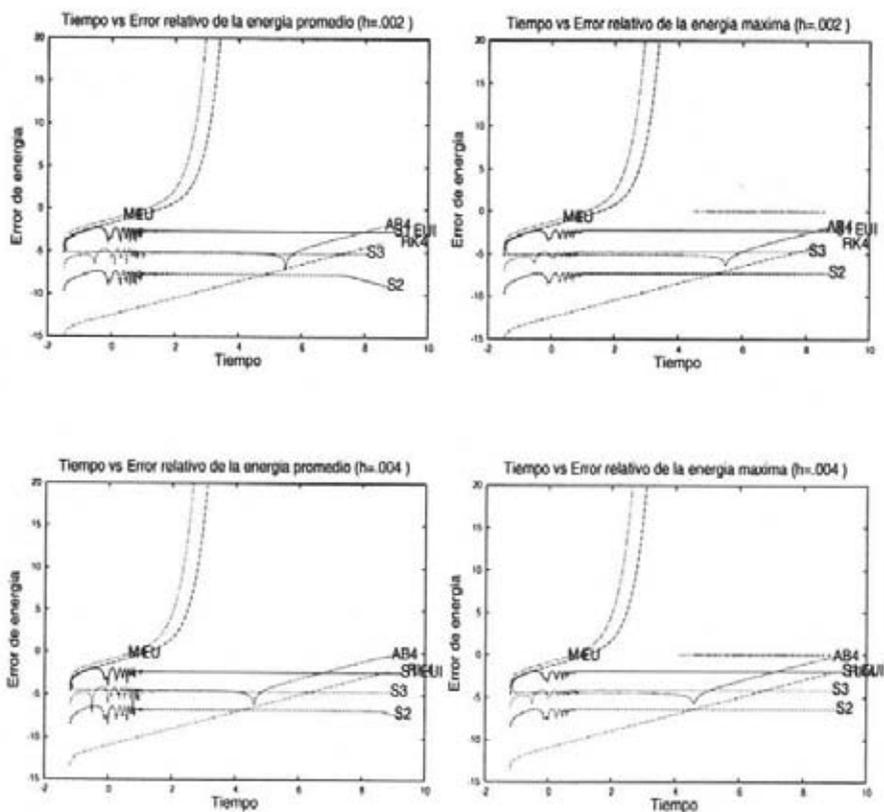
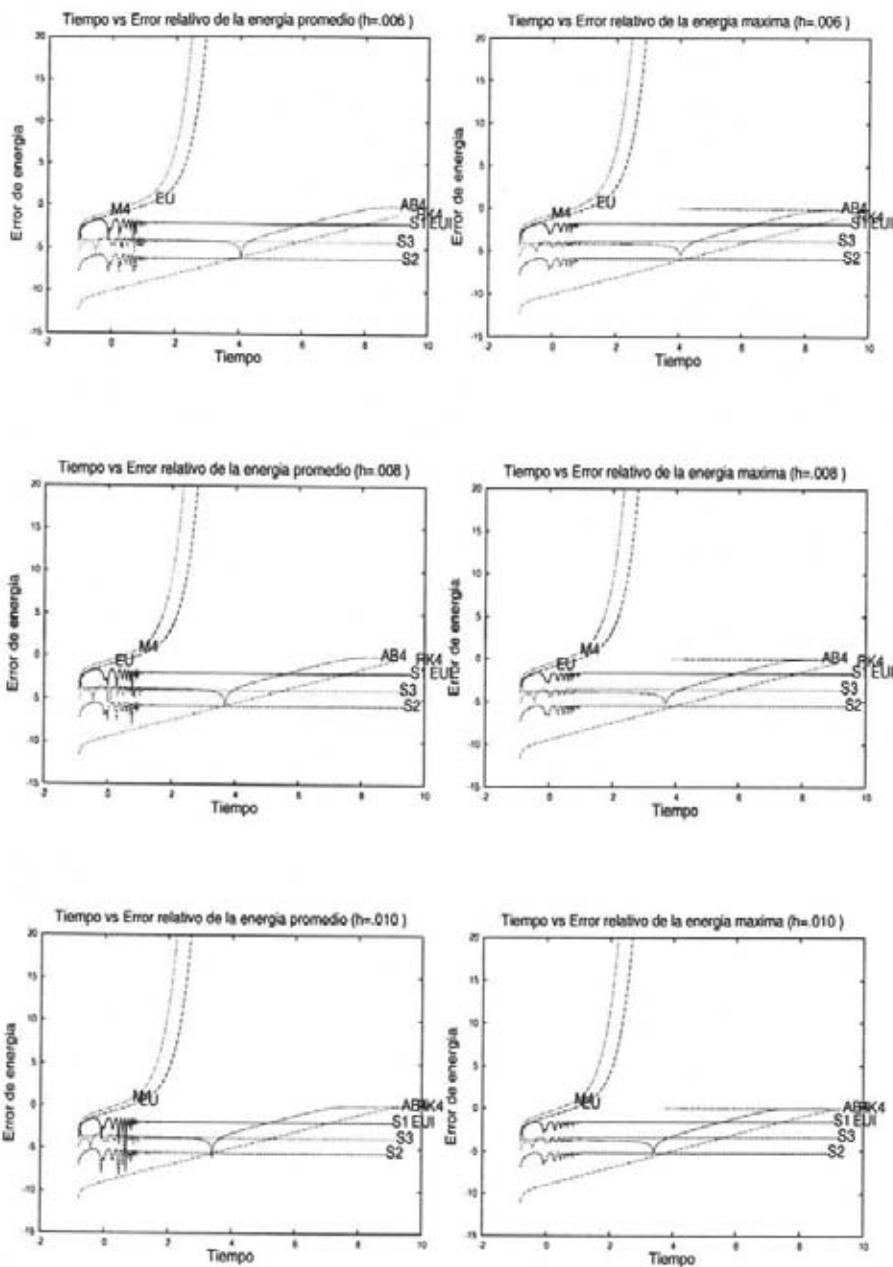
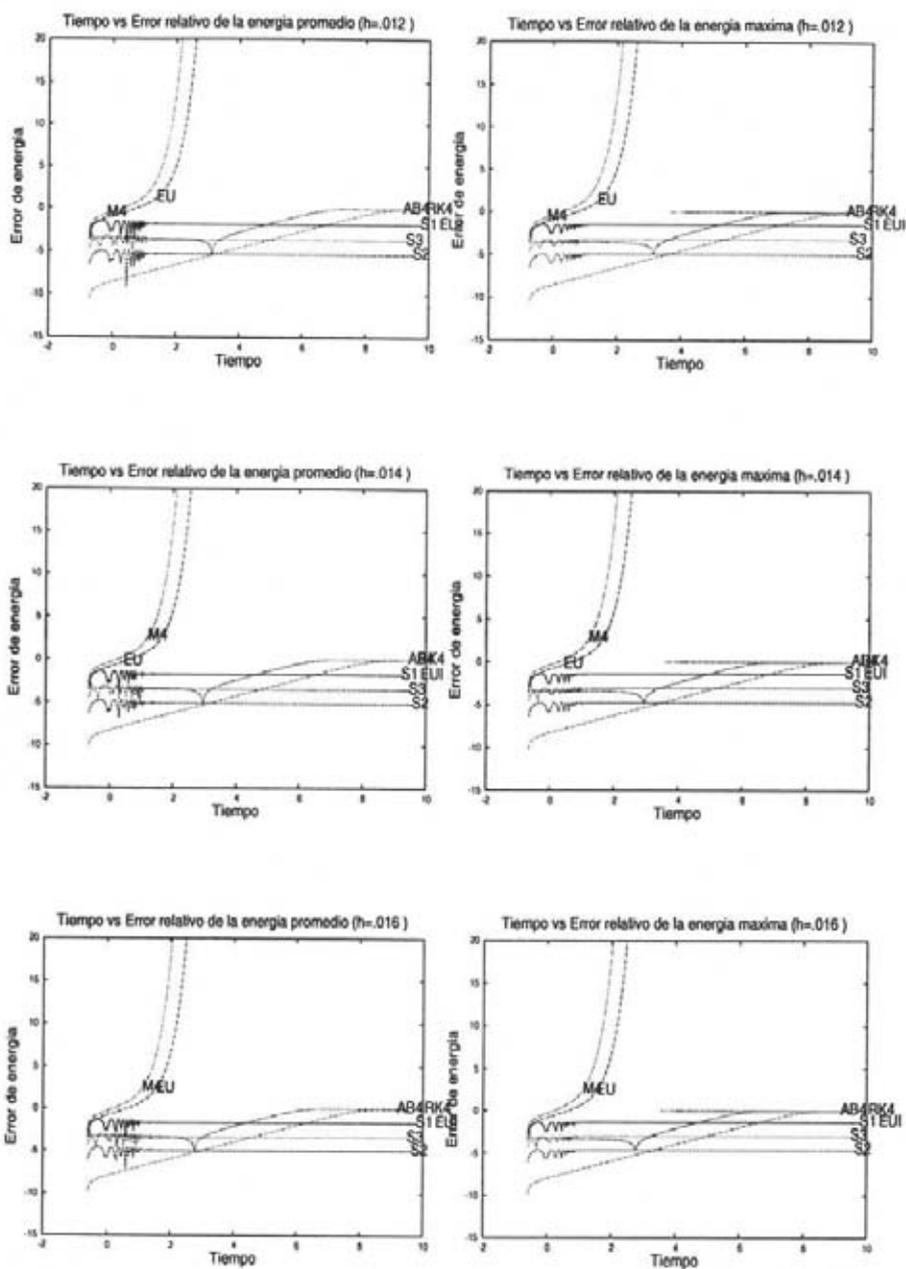
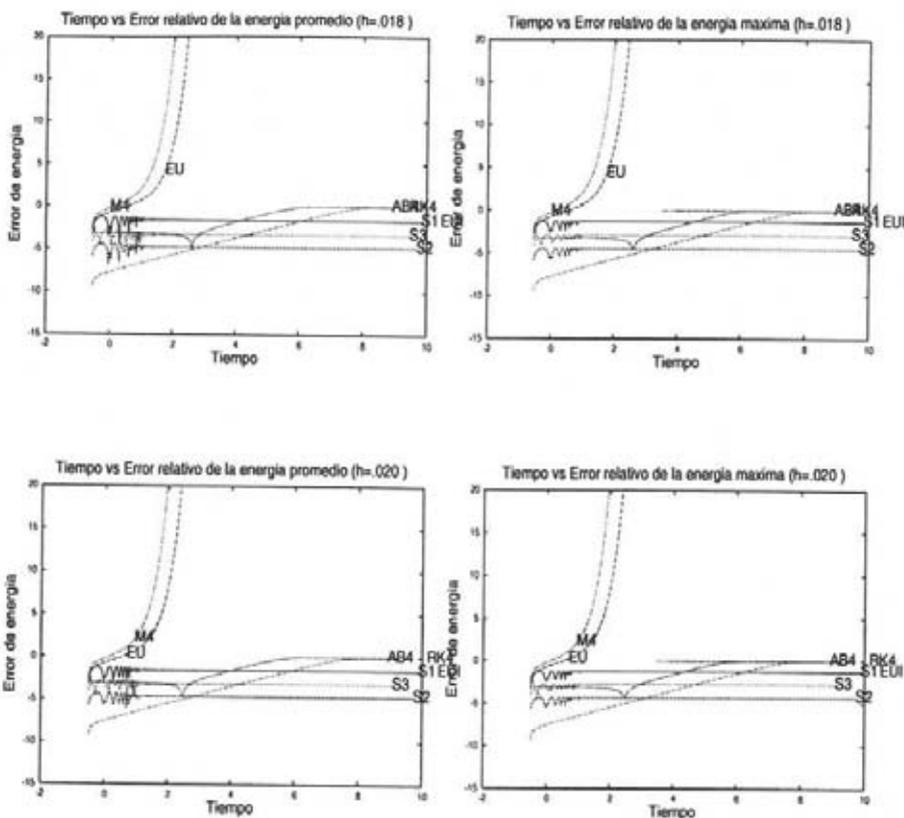


Figura 4.2: Plano fase del oscilador armónico

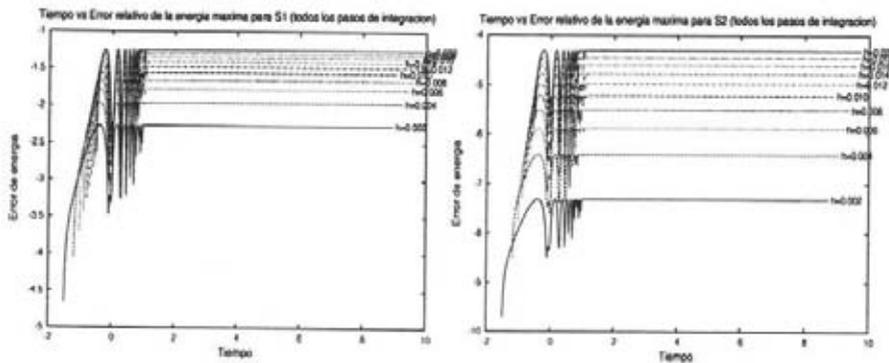


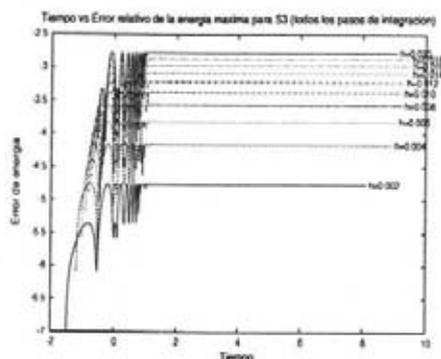






Estas gráficas muestran que tanto la energía promedio como la máxima se preservan bastante bien por los métodos simplécticos. De hecho, podemos apreciar que los métodos simplécticos tienden a un valor constante cuando integramos por tiempos muy largos y que dicho valor depende del paso de integración h , como lo muestran las siguientes gráficas.





Resulta mucho mejor observar el comportamiento del error relativo de la energía máxima para los distintos pasos de integración y los diferentes métodos simplécticos pues de este modo podemos asegurar que globalmente la energía es bien conservada.

El valor constante al que tienden los errores relativos de la energía máxima para los distintos integradores simplécticos y todos los pasos de integración queda esquematizado en el cuadro (4.1).

h	S1	S2	S3
0.002	0.0024	9.1×10^{-10}	6.3×10^{-06}
0.004	0.0048	5.4×10^{-08}	2.5×10^{-05}
0.006	0.0072	4.3×10^{-07}	5.6×10^{-05}
0.008	0.0097	1.3×10^{-06}	9.9×10^{-05}
0.010	0.0122	2.6×10^{-06}	0.0001
0.012	0.0147	4.7×10^{-06}	0.0002
0.014	0.0172	7.4×10^{-06}	0.0002
0.016	0.0197	1.1×10^{-05}	0.0003
0.018	0.0223	1.5×10^{-05}	0.0004
0.020	0.0249	1.5×10^{-05}	0.0005

Cuadro 4.1: Valor estable del error relativo de la energía máxima para el oscilador armónico.

Del cuadro (4.1) podemos observar que para todos los métodos simplécticos el error relativo disminuye al disminuir el paso de integración, y que además el error también disminuye al aumentar el orden del integrador simpléctico, exceptuando el cambio de S2 a S3. Más aún, si ajustamos una curva del tipo $y = ax^b$ con y representando el error relativo de la energía máxima y x los distintos pasos de integración obtenemos que los datos del cuadro (4.1) se ajustan a una curva con los parámetros dados en el cuadro (4.2). Luego, obtenemos fórmulas que ponen al error relativo de la energía máxima como función del paso de

integración, esto es

$$\text{Error relativo} = a \times (\text{paso de integración})^b.$$

El cuadro (4.2), nos dice que al disminuir el paso de integración, para S1 el error relativo disminuye linealmente, para S2 cúbicamente y para S3 cuadráticamente. Ésta situación solo se da en éste caso, pues como veremos en las siguientes secciones, en general obtendremos un comportamiento del error relativo que disminuye de forma lineal para S1, cuadrático para S2 y cúbico para S3 al disminuir el paso de integración. Una observación importante es que los

	S1	S2	S3
a	2.9043	5.8720	4.0016
b	1.0119	2.9919	1.9894

Cuadro 4.2: Parámetros de ajuste de la curva $y = ax^b$.

integradores simplécticos preservan la energía de manera muy constante para tiempos largos de integración. Estas simulaciones se integraron hasta el tiempo 10^{10} lo que representa de acuerdo con los pasos de integración escogidos que iteramos los métodos alrededor de cien mil millones de veces.

Los métodos clásicos EU y M4 fracasaron muy rápidamente, antes del tiempo 100 su energía ya no tenía ningún dígito de precisión sin embargo los métodos AB4 y RK4 comenzaron con errores relativos más o menos similares a los errores de los métodos simplécticos. Por ejemplo, AB4 mantiene su energía como S2 hasta el tiempo 10,000 y después el error cometido crece muy rápidamente. Los errores para RK4 crecen de manera constante teniendo mejor precisión que los métodos simplécticos hasta antes del tiempo 10,000 alcanzando la precisión de S2 un poco antes del tiempo 1,000,000. Un hecho remarcable es que para AB4 y RK4 la energía en realidad no crece, sino que decrece tendiendo a cero, para así obtener en el error relativo 1 y por ende en las gráficas vemos que su energía tiende a 1 ($\log_{10}(1) = 0$).

4.2. El péndulo ideal

El problema que simularemos numéricamente en esta sección es el de un péndulo de masa igual a 1 y de longitud 1 sin fricción.

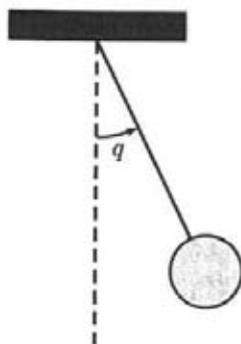


Figura 4.3: Péndulo ideal

Si q es el ángulo entre el péndulo y la vertical, entonces el sistema Hamiltoniano que integraremos es

$$\dot{q} = p \quad (4.5)$$

$$\dot{p} = -9 \operatorname{sen}(q) \quad (4.6)$$

siendo así

$$H(p, q) = \frac{1}{2}p^2 - 9 \cos(q).$$

Para este sistema, sus soluciones se ven en el plano fase como las curvas de nivel de H , de hecho como es bien sabido existen esencialmente dos tipos de soluciones, unas que corresponden al péndulo oscilando alrededor de un punto de estabilidad como 0 y π y otras que corresponden al péndulo girando sin cesar. En la siguiente gráfica del plano fase se aprecian las primeras como curvas cerradas y las segundas como curvas oscilatorias.

de hecho sabemos que la curva separatriz³ pasa por el punto $(\pi, 0)$ y

$$H(\pi, 0) = 9$$

³Esta curva separa a las soluciones cerradas de las oscilatorias.

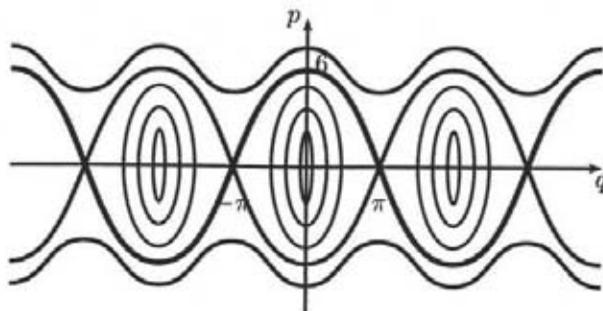


Figura 4.4: Plano fase del péndulo ideal

entonces si queremos encontrar el valor de p para el cual la condición inicial $(p, 0)$ hace que el sistema genere como solución a la separatriz del plano fase debemos resolver para p la ecuación

$$9 = \frac{1}{2}p^2 - 9,$$

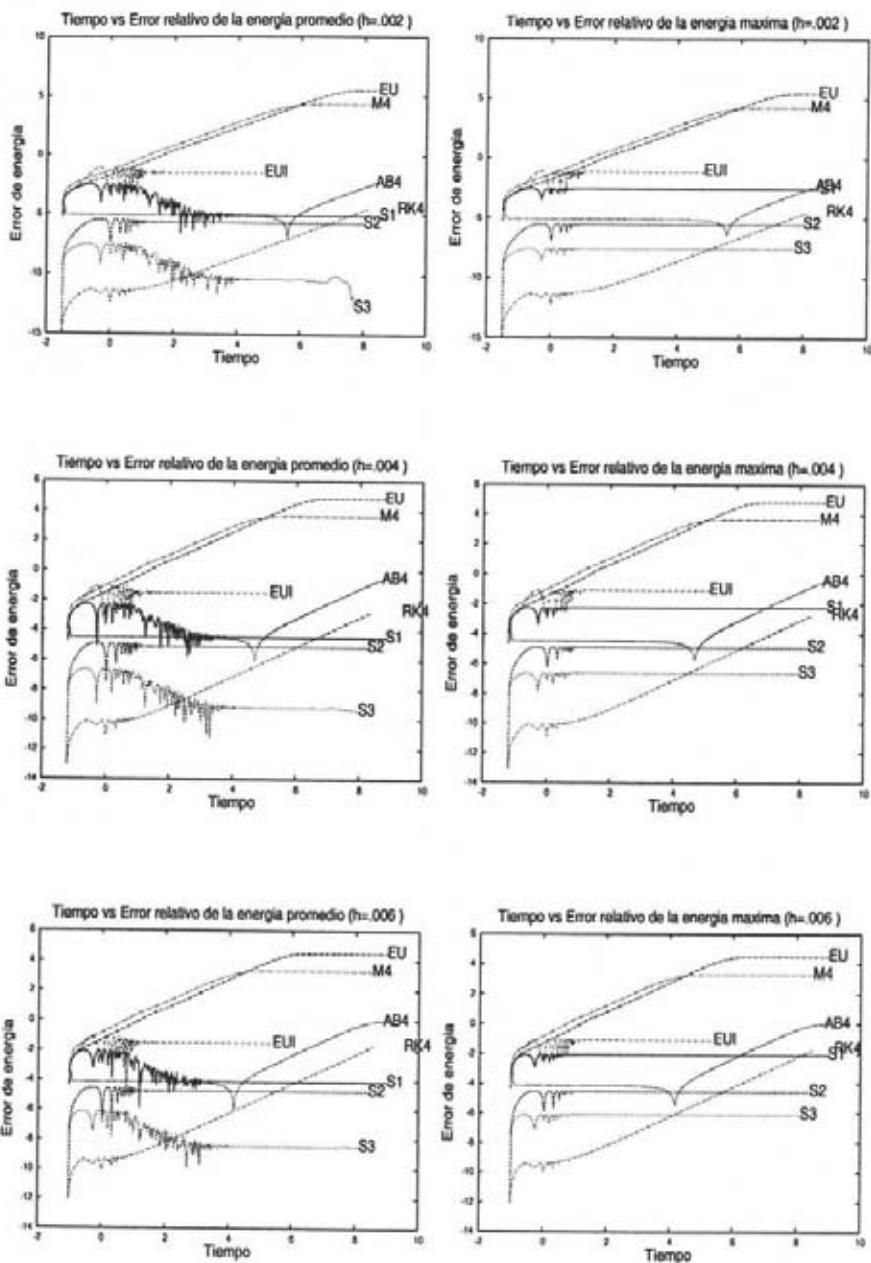
de donde obtenemos que $p = 6$. En la figura 4.4 la curva separatriz es la que pasa por el punto $(0, 6)$ y está indicada con color más obscuro.

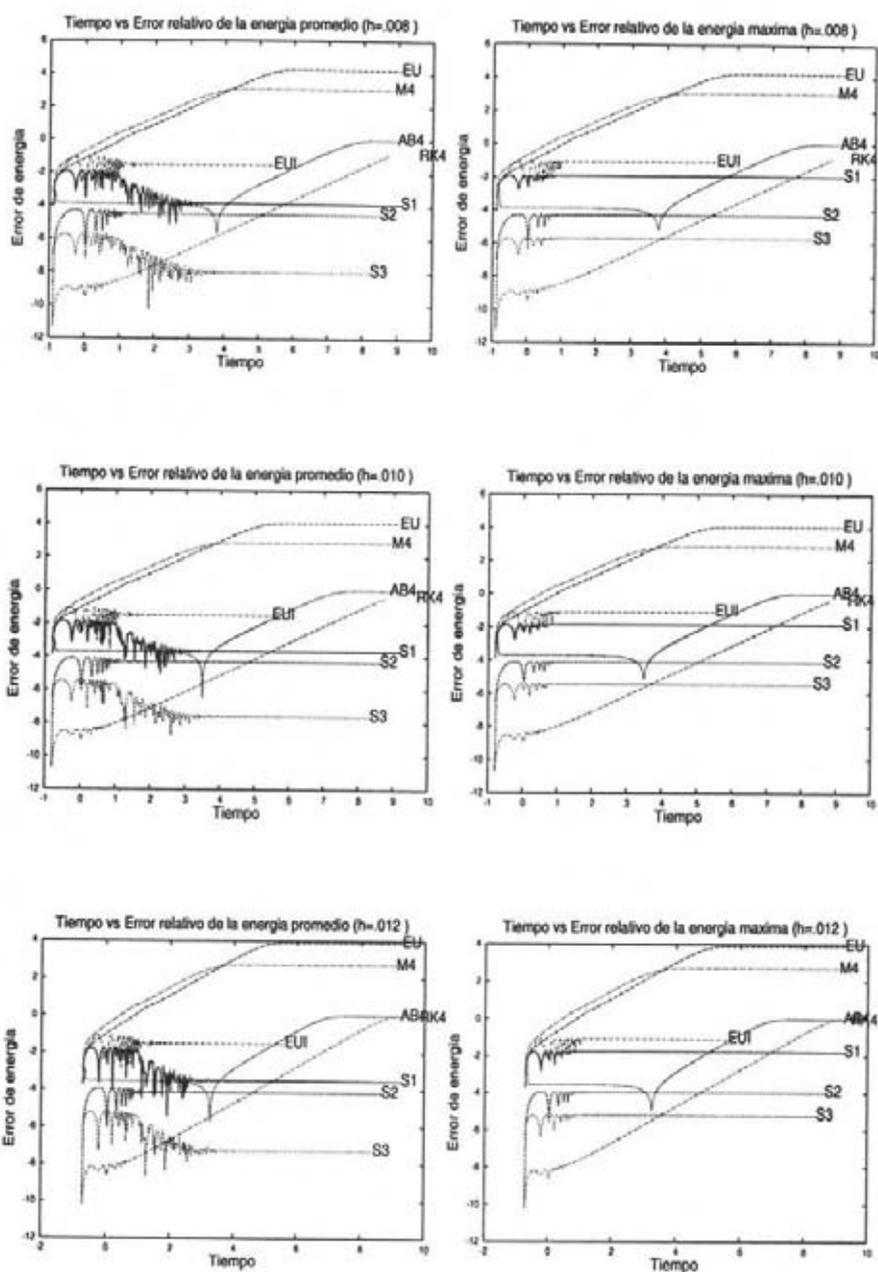
4.2.1. Resultados numéricos

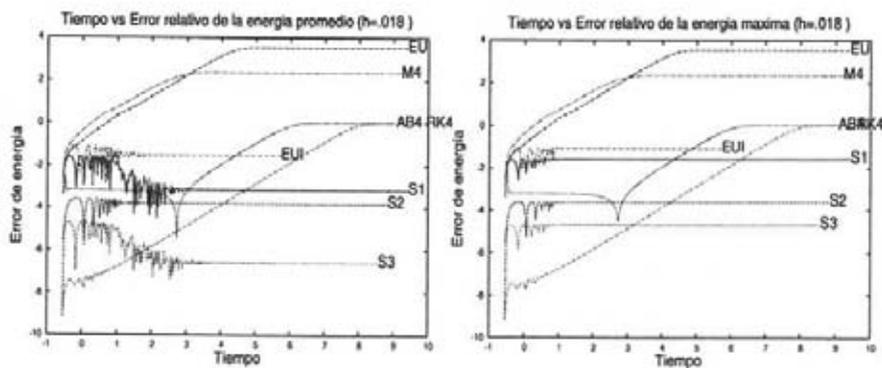
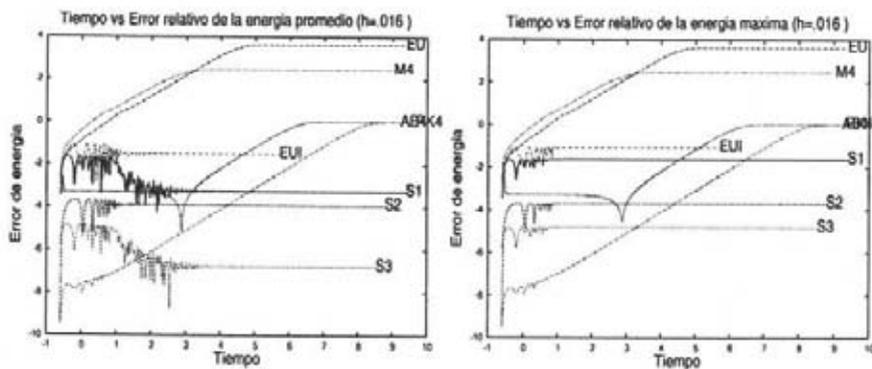
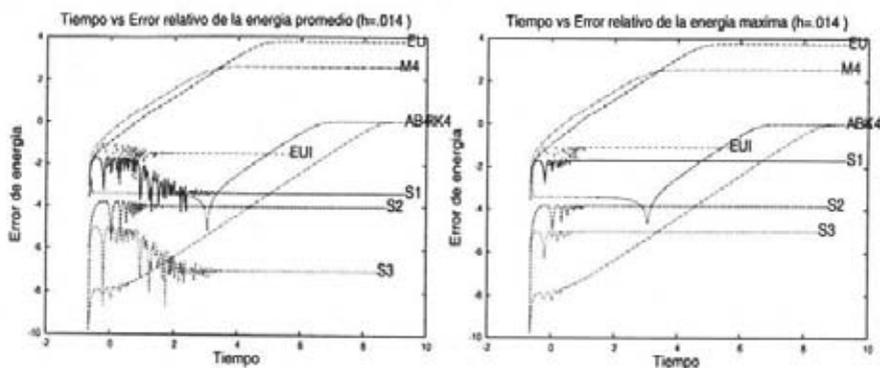
Las simulaciones numéricas que haremos a continuación serán con dos distintos tipos de condiciones iniciales, unas que representen el movimiento oscilatorio del péndulo alrededor de un punto de equilibrio y las otras que representen al péndulo girando.

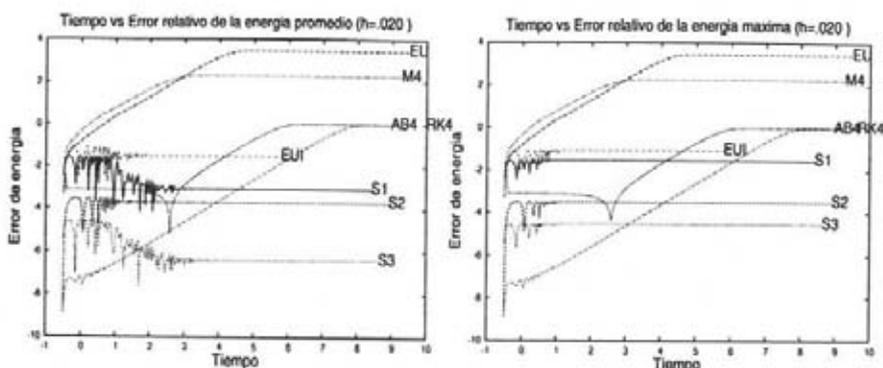
$$\begin{array}{l} p(0) = 3 \\ q(0) = 0 \end{array} \quad \text{y} \quad \begin{array}{l} p(0) = 6.001 \\ q(0) = 0. \end{array}$$

Como se hizo antes se presentan los resultados de las simulaciones numéricas de los errores relativos de la energía promedio y de la energía máxima para los métodos numéricos que ya hemos elegido y con los mismos pasos de integración. Para las primeras condiciones iniciales, tenemos las siguientes gráficas:

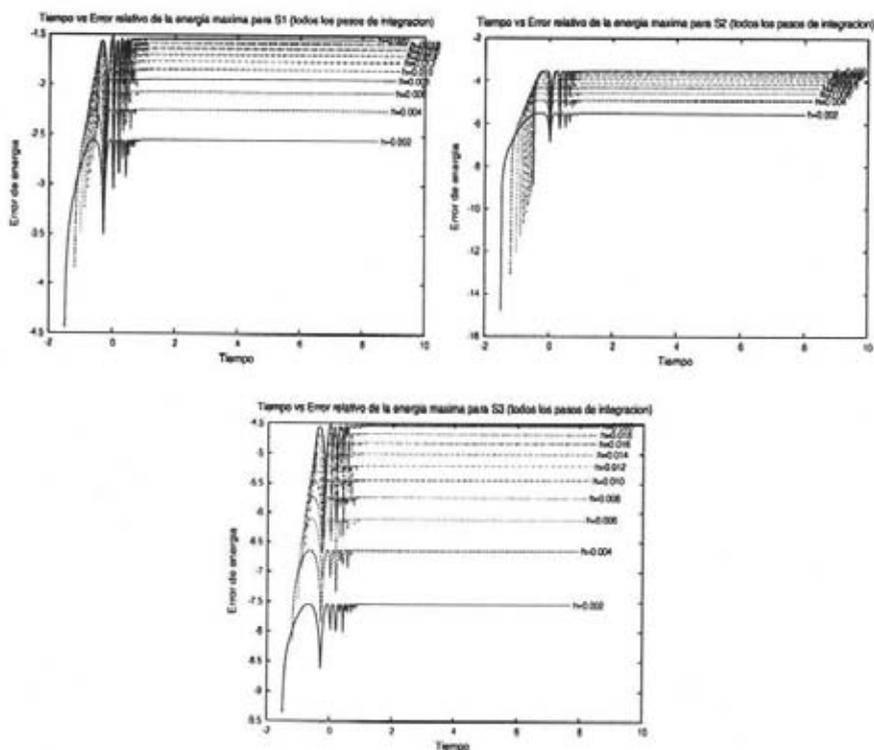








Primero analizaremos los integradores simplécticos. Es posible observar que tanto los errores relativos de la energía promedio como de la máxima tienden a un valor constante, es decir que para tiempos largos de integración (10^{10}) tanto la energía promedio como la máxima tienden a un valor constante que depende del tamaño del paso de integración como se aprecia en las siguientes gráficas.



Los valores límite de los errores relativos para la energía máxima están presentados en el cuadro (4.3). Si ajustamos nuevamente una curva de la forma

h	S1	S2	S3
0.002	0.0028	3.0×10^{-06}	2.8×10^{-08}
0.004	0.0056	1.2×10^{-05}	2.2×10^{-07}
0.006	0.0085	2.7×10^{-05}	7.7×10^{-07}
0.008	0.0113	4.8×10^{-05}	1.8×10^{-06}
0.010	0.0142	7.5×10^{-05}	3.6×10^{-06}
0.012	0.0171	0.00010	6.2×10^{-06}
0.014	0.0201	0.00014	9.9×10^{-06}
0.016	0.0230	0.00019	1.4×10^{-05}
0.018	0.0260	0.00024	2.1×10^{-05}
0.020	0.0290	0.00030	2.9×10^{-05}

Cuadro 4.3: Valor estable del error relativo de la energía máxima para el péndulo.

$y = ax^b$ con y el valor al que tienden los errores relativos de la energía máxima para los métodos S1, S2 y S3 y x el paso de integración obtenemos que las curvas tienen los parámetros dados en el cuadro (4.4). La curva $y = ax^b$ nos dice que

	S1	S2	S3
a	1.5131	.7763	3.7452
b	1.0119	2.0048	3.0081

Cuadro 4.4: Parámetros de ajuste de la curva $y = ax^b$.

el error relativo de la energía máxima está dado, aproximadamente, en función del paso de integración de la siguiente manera

$$\text{Error relativo} = a \times (\text{paso de integración})^b$$

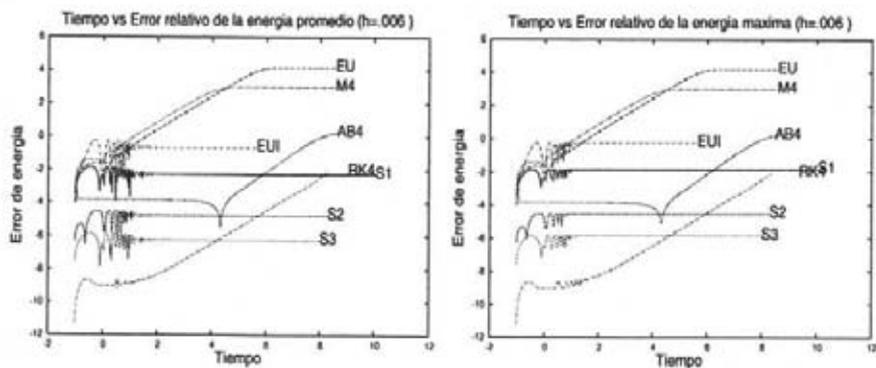
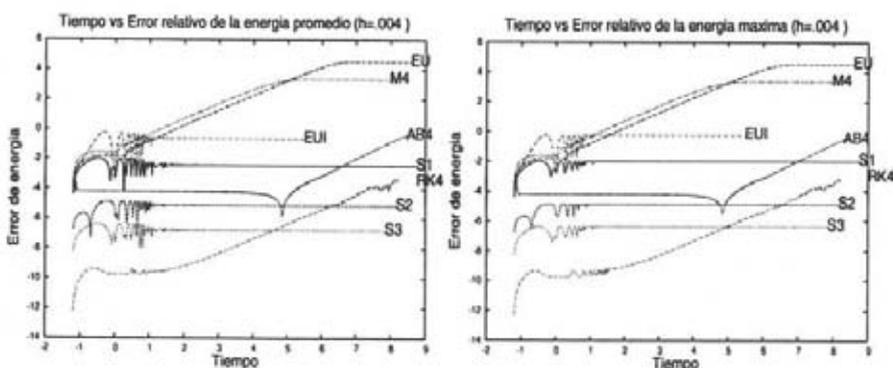
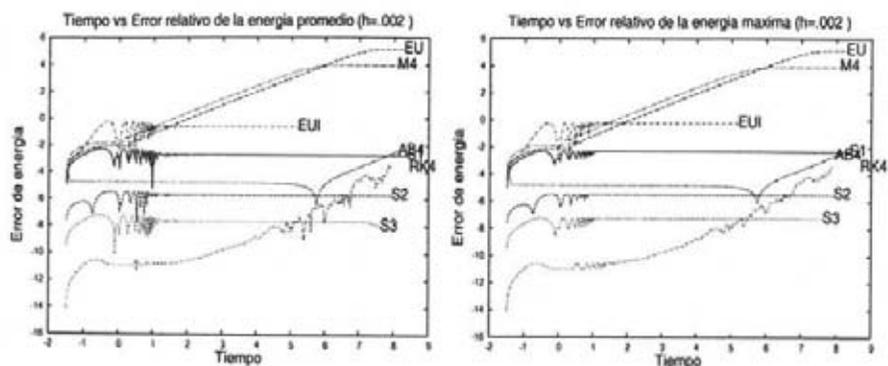
y de acuerdo con el cuadro (4.4) podemos concluir que el máximo error relativo de la energía decrece linealmente respecto del paso de integración para S1, decrece cuadráticamente para S2 y cúbicamente para S3. Para este problema existe una relación estrecha entre el orden del método y los errores relativos para la energía máxima, de hecho, se aprecia que los errores en la energía máxima disminuyen con el orden del método.

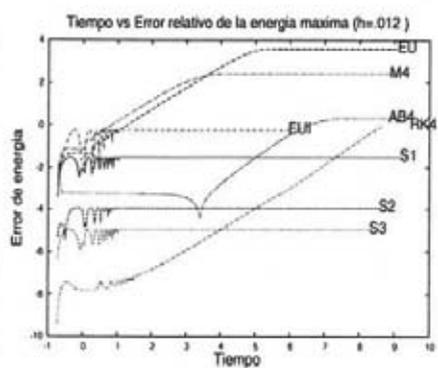
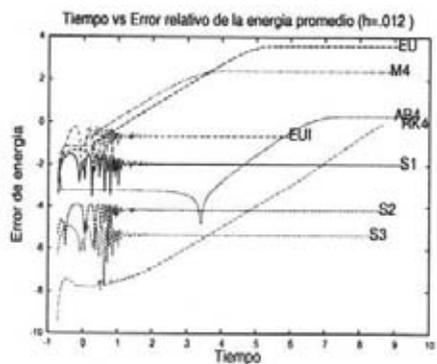
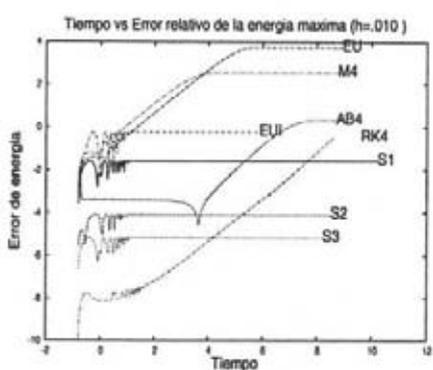
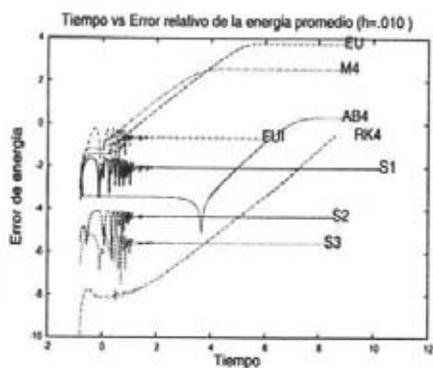
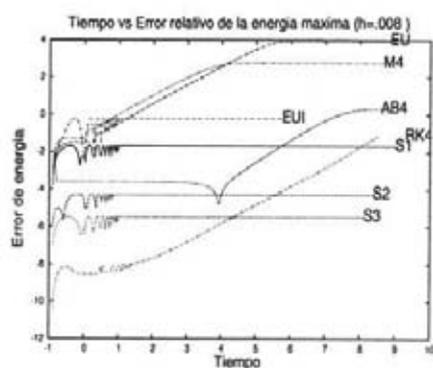
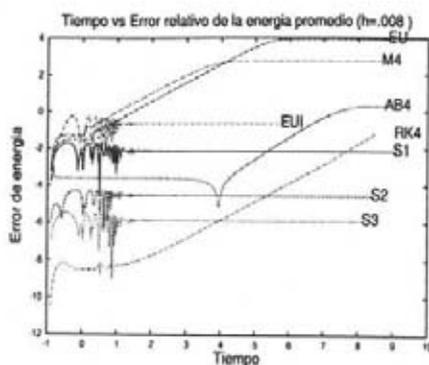
Para las condiciones iniciales

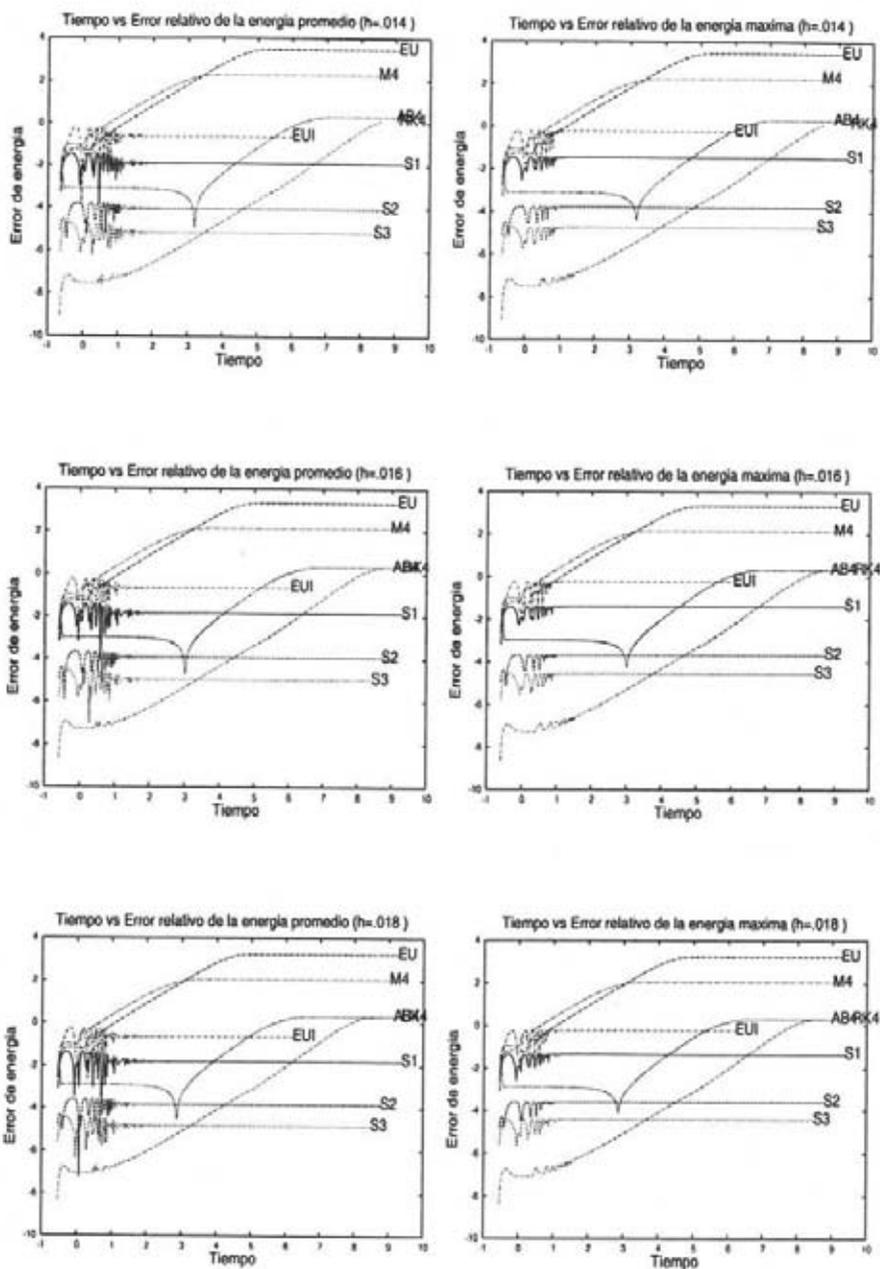
$$p(0) = 6.001$$

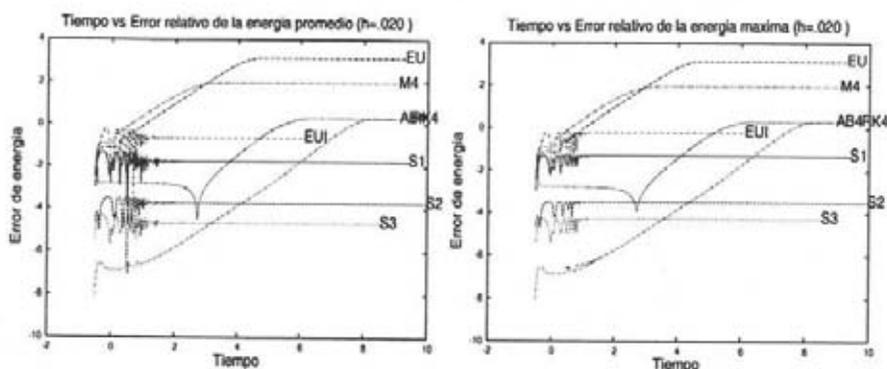
$$q(0) = 0$$

obtuvimos los siguientes resultados



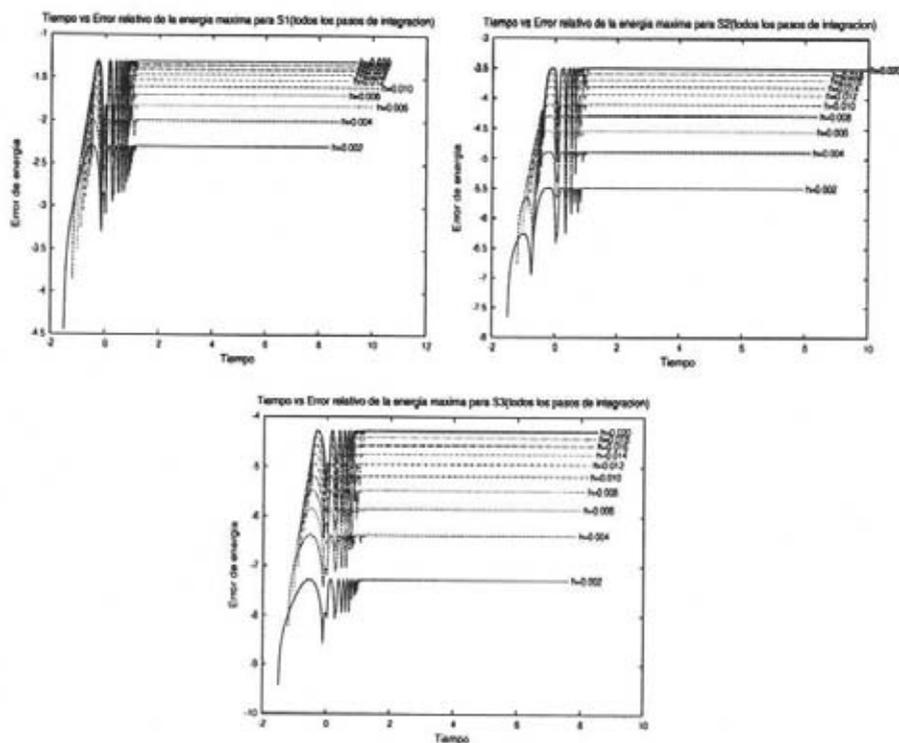






De las gráficas anteriores es posible observar que las simulaciones para las dos distintas condiciones iniciales son cualitativamente las mismas.

Veamos, si graficamos los distintos errores relativos de la energía máxima para los integradores simplécticos con todos los pasos de integración, obtenemos las siguientes gráficas



que como antes, apreciamos de igual manera que al crecer el tiempo de integración el error relativo se estabiliza en un valor. En el cuadro (4.5) se presentan los valores donde se estabilizan los errores relativos para los métodos simplécticos.

h	S1	S2	S3
0.002	0.0050	3.2×10^{-06}	5.2×10^{-08}
0.004	0.0101	1.2×10^{-05}	4.2×10^{-07}
0.006	0.0151	2.9×10^{-05}	1.4×10^{-06}
0.008	0.0201	5.1×10^{-05}	3.3×10^{-06}
0.010	0.0251	8.0×10^{-05}	6.5×10^{-06}
0.012	0.0300	0.00011	1.1×10^{-05}
0.014	0.0349	0.00015	1.7×10^{-05}
0.016	0.0398	0.00020	2.6×10^{-05}
0.018	0.0447	0.00025	3.8×10^{-05}
0.020	0.0495	0.00031	5.2×10^{-05}

Cuadro 4.5: Valor estable del error relativo de la energía máxima para el péndulo.

Si de nueva cuenta ajustamos a los datos presentados en el cuadro (4.5) la curva $y = ax^b$ obtenemos los parámetros presentados en el cuadro (4.6). Igual

	S1	S2	S3
a	2.3826	0.7791	6.4572
b	0.9890	1.9929	2.9964

Cuadro 4.6: Parámetros de ajuste de la curva $y = ax^b$.

que antes la precisión de los métodos en la energía máxima va conforme al orden del método.

Para los integradores clásicos los errores relativos crecen rápidamente, siendo RK4 y AB4 los mejores. Para RK4, el error relativo de la energía máxima es menor que el de S3 para tiempos menores que 1,000, sin embargo crece alcanzando la precisión de S2 en, aproximadamente, el tiempo 10,000 y S1 en 1,000,000. AB4 comienza con errores relativos para la energía máxima muy similares a los de S2 y se preserva de este modo hasta el tiempo 1,000, después de este tiempo crece más o menos con la velocidad con la que crecen los errores de RK4 para finalmente estabilizarse, junto con RK4, en 1. Esto significa que la energía tiende a cero para estos métodos obteniendo así que los errores relativos tienden a 1 (como en el oscilador armónico). El integrador EUI se comporta bastante bien para tiempos largos, ya que al igual que los simplécticos tiende a un valor constante, sin embargo, su precisión es peor que la de S1. Finalmente los integradores EU y M4 desde el comienzo tienen errores relativos mayores que cualquier otro integrador que hemos probado y desde este punto su precisión empeora todavía más conforme avanza el tiempo de integración.

4.3. La cadena de Toda

Comenzaremos esta sección explicando como surge este sistema Hamiltoniano y algunos antecedentes sobre su estudio.

4.3.1. Antecedentes y desarrollo

La ecuación de movimiento para una cadena unidimensional de osciladores, con interacción del vecino más cercano, está determinada por la segunda ley de Newton, esto es:

$$m \frac{d^2 y_n}{dt^2} = f(\Delta y_{n-1}) - f(\Delta y_n), \quad n \in \mathbb{Z} \quad (4.7)$$

donde y_n es el desplazamiento de la n -ésima partícula medida desde un sistema de referencia fijo, y

$$\begin{aligned} \Delta y_n &= y_{n+1} - y_n, & n \in \mathbb{Z} \\ f(x) &= -U'(x) \end{aligned}$$

con $U(x)$ la energía potencial almacenada en un resorte cuando su longitud es aumentada en una cantidad x .

El tratamiento de oscilaciones en cadenas no lineales⁴ comenzó a estudiarse muy seriamente en los inicios de 1950 cuando Fermi, Pasta y Ulam (FPU) estudiaron numéricamente la equipartición de la energía, esto es, en una cadena lineal de osciladores los modos normales de oscilación son mutuamente independientes, por lo que no hay intercambio de energía a través de estos modos. Esto pensaron Fermi, Pasta y Ulam; si una interacción no lineal era introducida, el flujo de energía tomaría lugar a través de los modos lineales, finalmente establecieron la equipartición de la energía, característica de los sistemas ergódicos, y quisieron verificar esta afirmación a través de experimentos numéricos. Sin embargo, contrario a lo que ellos pensaban, solamente ocurrió un poco de equipartición de energía y encontraron que el estado del sistema regresaba periódicamente al estado inicial, éste fenómeno de recurrencia dependía de la magnitud del término no lineal, la energía dada al modo inicial, la longitud de la cadena y las condiciones iniciales como un todo.

El modelo de FPU se puede considerar como la primera aproximación no lineal a la fuerza de interacción entre partículas cercanas de la forma:

$$f(r) = a(e^{-br} - 1), \quad ab > 0 \quad (4.8)$$

entonces la ecuación de movimiento (4.7), resulta ser

$$m \frac{d^2 y_n}{dt^2} = a(e^{-b(y_n - y_{n-1})} - e^{-b(y_{n+1} - y_n)}), \quad n \in \mathbb{Z}. \quad (4.9)$$

⁴Las cadenas no lineales se tienen cuando $U'(x)$ es no lineal.

La cadena unidimensional de osciladores con la ecuación de movimiento anterior es conocida como **cadena exponencial** o **cadena de Toda** ⁵.

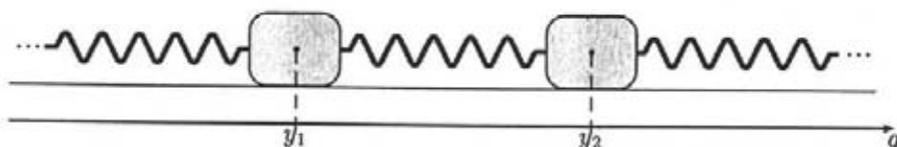


Figura 4.5: Cadena de Toda

Las ecuaciones de movimiento (4.9) son un ejemplo no trivial de una cadena no lineal en la que la equipartición de la energía no tiene lugar ⁶ y cuya solución fue encontrada por M. Toda:

$$y_n(t) = \frac{1}{b} \log \frac{1 + e^{2(kn - k - \beta t)}}{1 + e^{2(kn - \beta t)}} + cte, \quad n \in \mathbb{Z} \quad (4.10)$$

donde

$$\beta = \sqrt{\frac{ab}{m}} \sinh(k), \quad ab > 0, \quad k = cte.$$

La solución (4.10) es el frente de onda de una onda oscilatoria o solitón.

Si se considera una cadena de Toda finita con condiciones de frontera periódicas y masas unitarias, entonces se tendrá el siguiente sistema

$$\dot{p}_n = e^{-(y_n - y_{n+1})} - e^{-(y_{n+1} - y_n)}, \quad n = 1, \dots, N,$$

con condiciones

$$p_n = \frac{dy_n}{dt}$$

$$y_0 = y_{N+1}$$

cuyo Hamiltoniano es

$$H = \frac{1}{2} \sum_{n=0}^N p_n^2 + \sum_{n=0}^N (e^{y_n - y_{n+1}} - 1). \quad (4.11)$$

Para una cadena de tres partículas ($N = 3$), se tiene el Hamiltoniano

$$H = \frac{2}{2} (p_1^2 + p_2^2 + p_3^2) + e^{y_1 - y_2} + e^{y_2 - y_3} + e^{y_3 - y_1} - 3. \quad (4.12)$$

⁵Esta cadena fue propuesta por primera vez por M. Toda en 1967. Ver M. Toda [13, 12].

⁶Ésta es la paradoja de Fermi, Pasta y Ulam. La razón por la cual el sistema FPU no exhibió equipartición de energía se debió a que es una aproximación de un sistema totalmente integrable.

Consideremos el cambio de variables dado por

$$\begin{aligned}\xi &= Ay \\ \eta &= Ap\end{aligned}$$

donde

$$A = \begin{pmatrix} \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}.$$

Este cambio de variables corresponde a una rotación de cuerpo rígido del sistema coordinado original (\mathbf{y}, \mathbf{p}) a través de los ángulos e Euler, de hecho los ángulos son $\phi = \pi/6$, $\tau = \pi/4$ y $\theta \approx 0.9557$. Luego el Hamiltoniano (4.12) en las nuevas variables (ξ, η) queda

$$H = \frac{2}{2}(\eta_1^2 + \eta_2^2 + \eta_3^2) + e^{\frac{\sqrt{2}}{2}\xi_2 + \frac{\sqrt{6}}{2}\xi_1} + e^{\frac{\sqrt{2}}{2}\xi_2 - \frac{\sqrt{6}}{2}\xi_1} + e^{\frac{\sqrt{6}}{2}\xi_2} - 3. \quad (4.13)$$

Aquí ξ_3 al ser coordenada cíclica no aparece y $\sqrt{3}\eta_3 = p_1 + p_2 + p_3 = cte$, es decir, se han reducido los grados de libertad del sistema (4.12) a dos. La ecuaciones de movimiento son

$$\begin{aligned}\dot{\xi}_1 &= \frac{\partial H}{\partial \eta_1} = \eta_1 \\ \dot{\eta}_1 &= -\frac{\partial H}{\partial \xi_1} = \frac{\sqrt{6}}{2}(-e^{\frac{\sqrt{2}}{2}\xi_2 + \frac{\sqrt{6}}{2}\xi_1} + e^{\frac{\sqrt{2}}{2}\xi_2 - \frac{\sqrt{6}}{2}\xi_1}) \\ \dot{\xi}_2 &= \frac{\partial H}{\partial \eta_2} = \eta_2 \\ \dot{\eta}_2 &= -\frac{\partial H}{\partial \xi_2} = -\frac{\sqrt{2}}{2}(-e^{\frac{\sqrt{2}}{2}\xi_2 + \frac{\sqrt{6}}{2}\xi_1} + e^{\frac{\sqrt{2}}{2}\xi_2 - \frac{\sqrt{6}}{2}\xi_1}) + \sqrt{2}e^{-\sqrt{2}\xi_2},\end{aligned}$$

si además definimos $\xi_1 = 2\sqrt{2}q_1$, $\xi_2 = 2\sqrt{2}q_2$ y $t = \tau/\sqrt{3}$ entonces el sistema anterior se transforma en

$$\begin{aligned}\frac{dq_1}{d\tau} &= \frac{1}{2\sqrt{6}}\eta_1 = p_1 \\ \frac{dp_1}{d\tau} &= \frac{1}{2\sqrt{6}}\frac{d\eta_1}{d\tau} = \frac{1}{4\sqrt{3}}(-e^{2q_2+2\sqrt{3}q_1} + e^{2q_2-2\sqrt{3}q_1}) \\ \frac{dq_2}{d\tau} &= \frac{1}{2\sqrt{6}}\eta_2 = p_2 \\ \frac{dp_2}{d\tau} &= \frac{1}{2\sqrt{6}}\frac{d\eta_2}{d\tau} = -\frac{1}{12}(-e^{2q_2+2\sqrt{3}q_1} + e^{2q_2-2\sqrt{3}q_1}) + \frac{1}{6}e^{-4q_2}\end{aligned}$$

encontrando finalmente que el nuevo Hamiltoniano es

$$H = \frac{1}{2}(p_1^2 + p_2^2) + \frac{1}{24}(e^{2q_2+2\sqrt{3}q_1} + e^{2q_2-2\sqrt{3}q_1} + e^{-4q_2}) - \frac{1}{8}. \quad (4.14)$$

Ford estudió este problema y mostró que el sistema anterior tiene planos de Poincaré que están cubiertos enteramente por curvas cerradas y suaves. Es decir, que el sistema es completamente integrable. De hecho, las cantidades conservadas son:

$$E = \frac{1}{2}(p_1^2 + p_2^2) + \frac{1}{24}(e^{2q_2+2\sqrt{3}q_1} + e^{2q_2-2\sqrt{3}q_1} + e^{-4q_2}) - \frac{1}{8} \quad (4.15)$$

la energía mecánica del sistema y

$$F = 8p_1(p_1^2 - 3p_2^2) + (p_1 + \sqrt{3}p_2)e^{2q_2-2\sqrt{3}q_1} + (p_1 - \sqrt{3}p_2)e^{2q_2+2\sqrt{3}q_1} - 2p_1e^{-4q_2} \quad (4.16)$$

un momento angular generalizado.

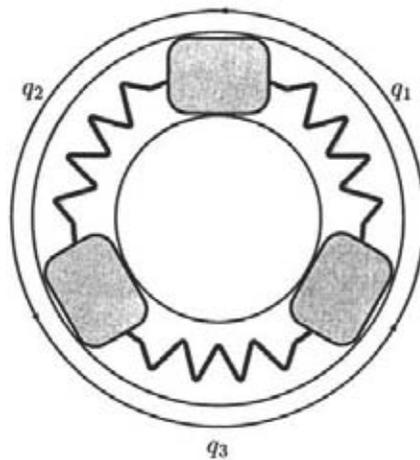


Figura 4.6: Cadena de Toda finita con condiciones de frontera periódicas

4.3.2. Implementación numérica de los integradores simplécticos

En esta sección implementaremos los métodos simplécticos desarrollados en el capítulo 3 al sistema generado por el Hamiltoniano (4.14). Si definimos las funciones

$$e_1(\mathbf{q}) = e^{2q_2+2\sqrt{3}q_1}$$

$$e_2(\mathbf{q}) = e^{2q_2-2\sqrt{3}q_1}$$

$$e_3(\mathbf{q}) = e^{-4q_2}$$

entonces el sistema (4.14) puede ser escrito como

$$H = \frac{1}{2}(p_1^2 + p_2^2) + \frac{1}{24}(e_1 + e_2 + e_3) - \frac{1}{8}. \quad (4.17)$$

En el desarrollo subsecuente usaremos la siguiente notación

$$\begin{aligned} E_1 &= e_1 + e_2 \\ E_2 &= e_1 - e_2 \\ E_3 &= e_1 + e_2 - 2e_3 \\ E_4 &= e_1 + e_2 + 4e_3 \\ E_5 &= e_1 + e_2 - 8e_3 \end{aligned}$$

y si $r = \sqrt{3}$ las constantes

$$\begin{aligned} R_1 &= 2r \\ R_2 &= \frac{1}{24} \\ R_3 &= R_1 R_2 = \frac{r}{12} \\ R_4 &= 2R_3 = \frac{r}{6} \\ R_5 &= 2R_2 = \frac{1}{12} \end{aligned}$$

además $R_1 R_3 = 1/2$ y $R_1 R_4 = 1$.

Como $K_0 = -(p_1 q_1 + p_2 q_2)$, de acuerdo con las ecuaciones (3.49) y (3.55), tenemos que

$$K_1 = \frac{1}{2}(p_1^2 + p_2^2) + R_2(e_1 + e_2 + e_3) - \frac{1}{8} \quad (4.18)$$

y que

$$K_2 = - \left(\frac{\partial K_1}{\partial \mathbf{q}} \right)^t \cdot \frac{\partial H}{\partial \mathbf{p}} = - \left(\frac{\partial H}{\partial \mathbf{q}} \right)^t \cdot \frac{\partial H}{\partial \mathbf{p}}$$

concluyendo que

$$K_2 = (R_3 E_2, 2R_2 E_3)^t (p_1, p_2) = -(R_3 E_2 p_1 + R_5 E_3 p_2). \quad (4.19)$$

Para calcular K_3 es necesario calcular $\partial K_2 / \partial \mathbf{q}$ para encontrar

$$- \frac{\partial K_2}{\partial \mathbf{q}} = \left(\frac{1}{2} E_1 p_1 + R_4 E_2 p_2, R_4 E_2 p_1 + 2R_5 E_4 p_2 \right)$$

y por tanto

$$- \left(\frac{\partial K_2}{\partial \mathbf{q}} \right)^t \cdot \frac{\partial H}{\partial \mathbf{p}} = \frac{1}{2} E_1 p_1^2 + 4R_3 E_2 p_1 p_2 + 4R_2 E_4 p_2^2$$

por otra parte, si

$$S = \sum_{i,j=1}^2 \frac{\partial K_1}{\partial q_i} \frac{\partial K_1}{\partial q_j} \frac{\partial^2 H}{\partial p_i \partial p_j}$$

entonces como $\partial^2 H / \partial p_i \partial p_j = 0$ obtenemos que

$$\begin{aligned} S &= \left(\frac{\partial K_1}{\partial q_1} \right)^2 \frac{\partial^2 H}{\partial p_1^2} + \left(\frac{\partial K_1}{\partial q_2} \right)^2 \frac{\partial^2 H}{\partial p_2^2} \\ &= (R_3 E_2)^2 + (R_5 E_3)^2 \end{aligned}$$

y por tanto

$$K_3 = \frac{1}{2} E_1 p_1^2 + 4R_3 E_2 p_1 p_2 + 4R_2 E_4 p_2^2 + (R_3 E_2)^2 + (R_5 E_3)^2 \quad (4.20)$$

entonces, de acuerdo con la ecuación (3.73), las nuevas coordenadas $\mathbf{q} = (q_1, q_2)$ son las raíces de la ecuación

$$f(\mathbf{q}) = 0$$

con

$$f(\mathbf{q}) = \mathbf{q}_0 + \begin{pmatrix} -q_1 + hp_1 - \frac{h^2}{2} R_3 E_2 + \frac{h^3}{3} (R_1 R_3 E_1 p_1 + R_4 E_2 p_2) \\ -q_2 + hp_2 - \frac{h^2}{2} R_2 E_3 + \frac{h^3}{3} (R_4 E_2 p_1 + 4R_2 E_4 p_2) \end{pmatrix}. \quad (4.21)$$

Para encontrar las raíces de la ecuación anterior utilizaremos el método de Newton-Raphson, de acuerdo con éste una aproximación \mathbf{q}_n de la raíz está dada por

$$\mathbf{q}_n = \mathbf{q}_{n-1} - (Jf(\mathbf{q}_{n-1}))^{-1} f(\mathbf{q}_{n-1})$$

donde $Jf(\mathbf{q})$ representa la matriz jacobiana de la transformación (4.21), es decir, $Jf(\mathbf{q})$ es

$$\begin{pmatrix} -1 - \frac{h^2}{4} E_1 + \frac{h^3}{3} (r p_1 E_2 + E_1 p_2) & -h^2 R_3 E_2 + \frac{h^3}{3} (E_1 p_1 + 2R_4 E_2 p_2) \\ -h^2 R_3 E_2 + \frac{h^3}{3} (E_1 p_1 + 4R_3 E_2 p_2) & -1 - h^2 R_2 E_4 + \frac{h^3}{3} (2R_4 E_2 p_1 + 4R_5 E_5 p_2) \end{pmatrix}$$

de esta manera el método de Newton se arranca con $\mathbf{q}_1 = \mathbf{q}_0$. La precisión que utilizaremos al aproximar \mathbf{q} será de 10^{-6} , pues se efectuaron pruebas tomando precisiones al aproximar \mathbf{q} de hasta 10^{-12} sin obtener una mejoría en los errores relativos de la energía mecánica y el momento angular generalizado. Una vez obtenida esta aproximación entonces de acuerdo con la ecuación (3.20), tenemos una expresión explícita para \mathbf{p} .

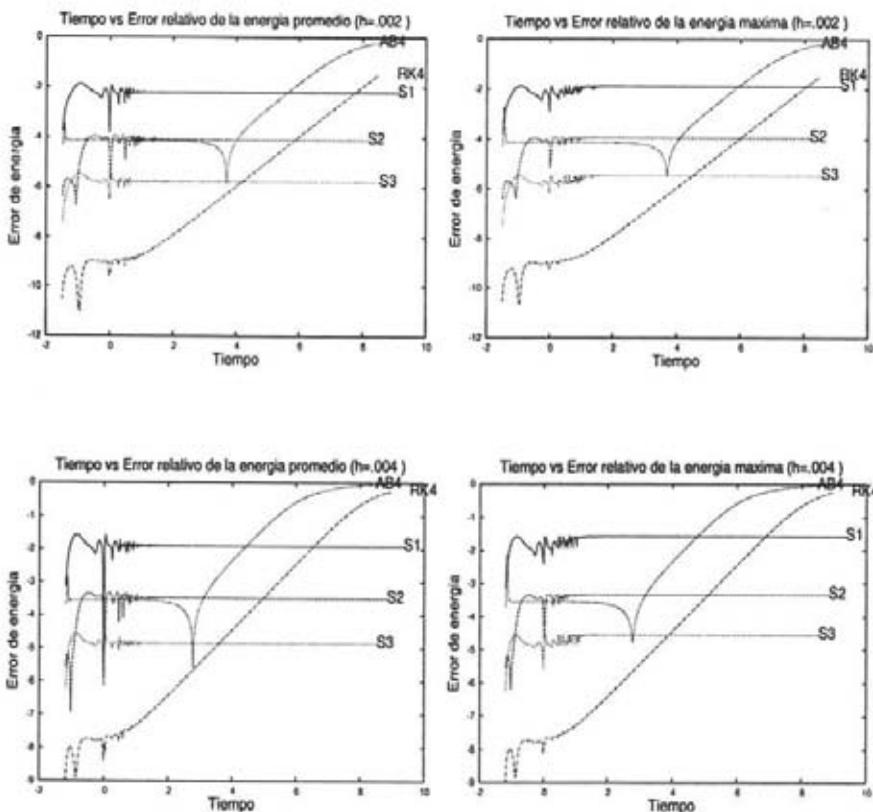
4.3.3. Resultados numéricos

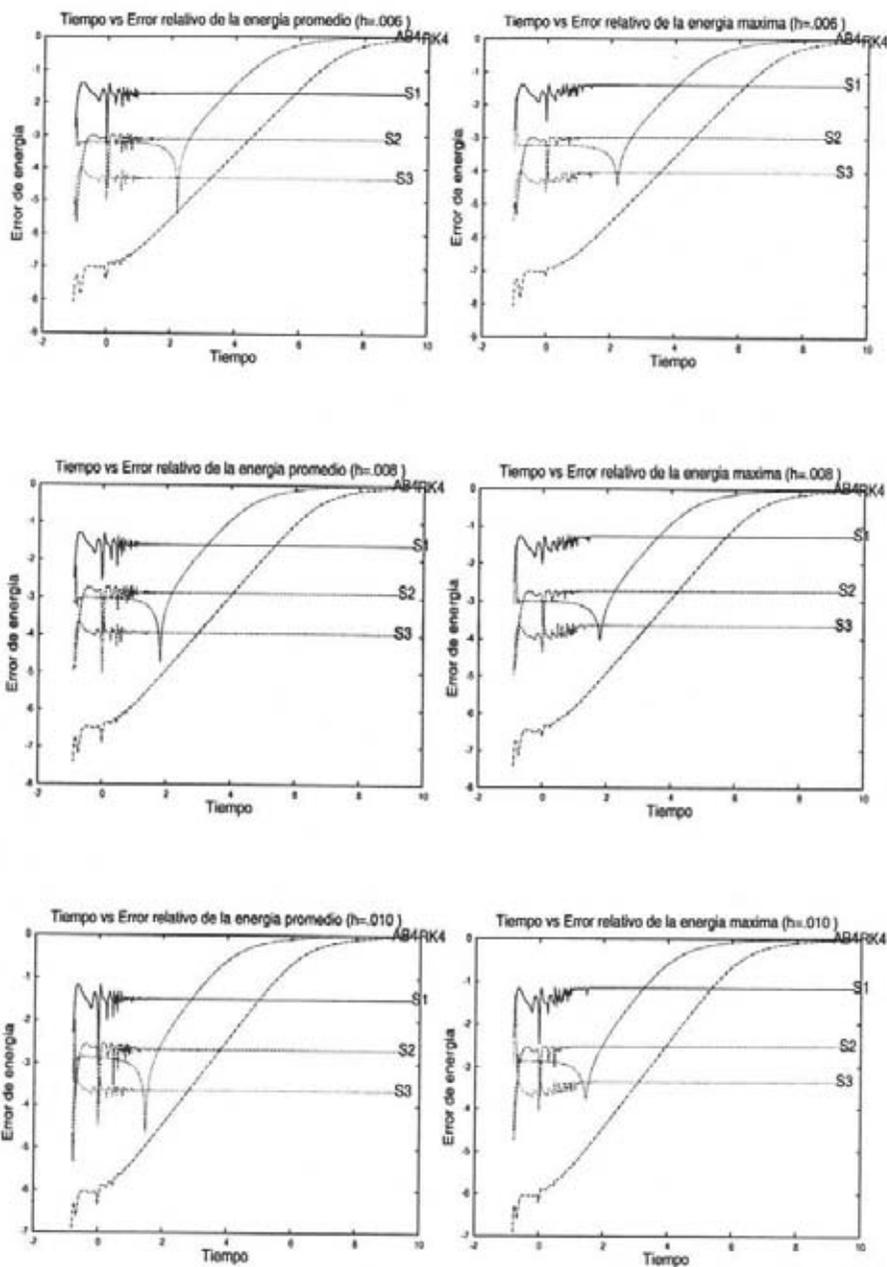
Las simulaciones numéricas efectuadas se hicieron para la cadena de Toda finita de tres partículas, es decir, se efectuaron simulaciones para las ecuaciones de movimiento asociadas al Hamiltoniano (4.14) con condiciones iniciales

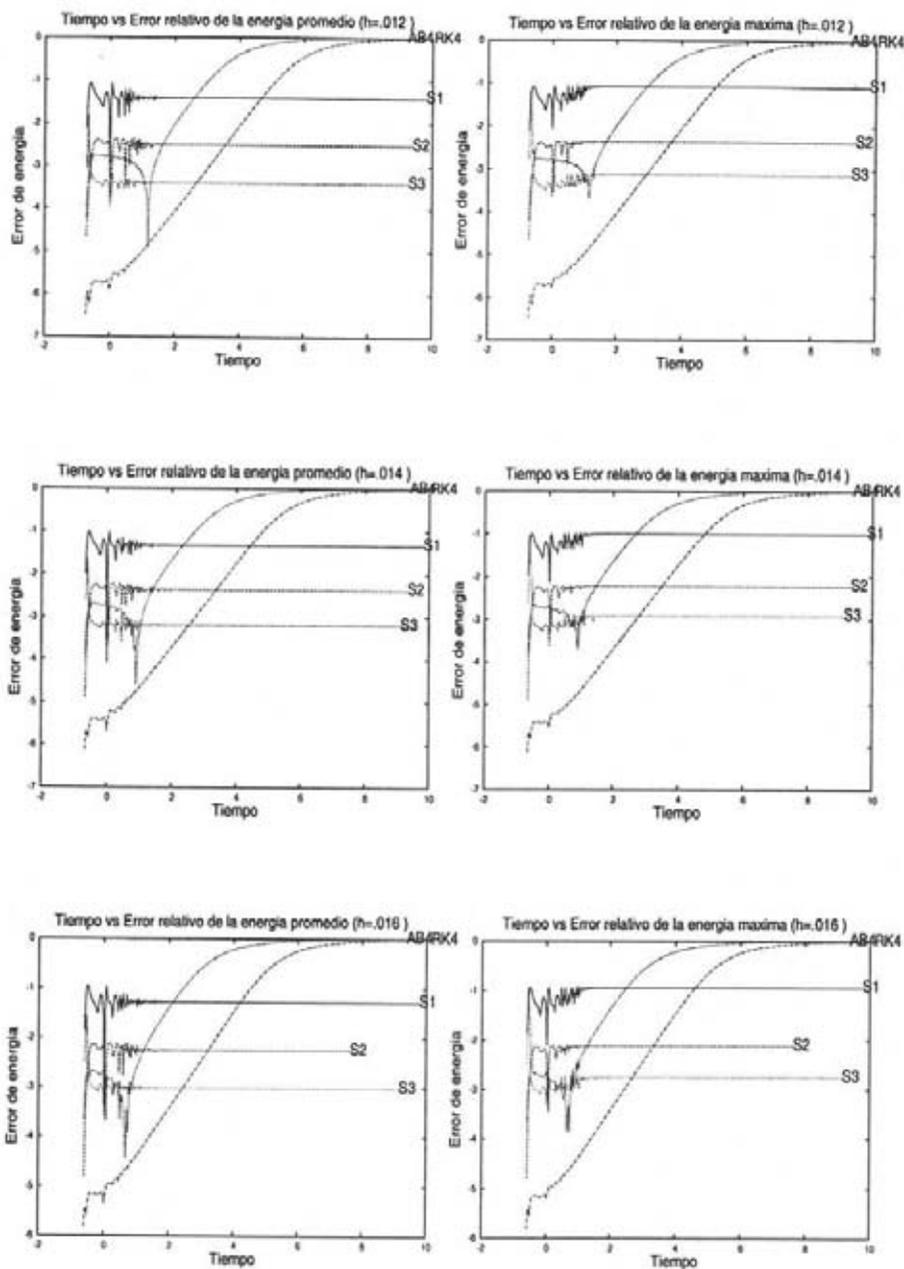
$$\mathbf{p}(0) = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad \text{y} \quad \mathbf{q}(0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

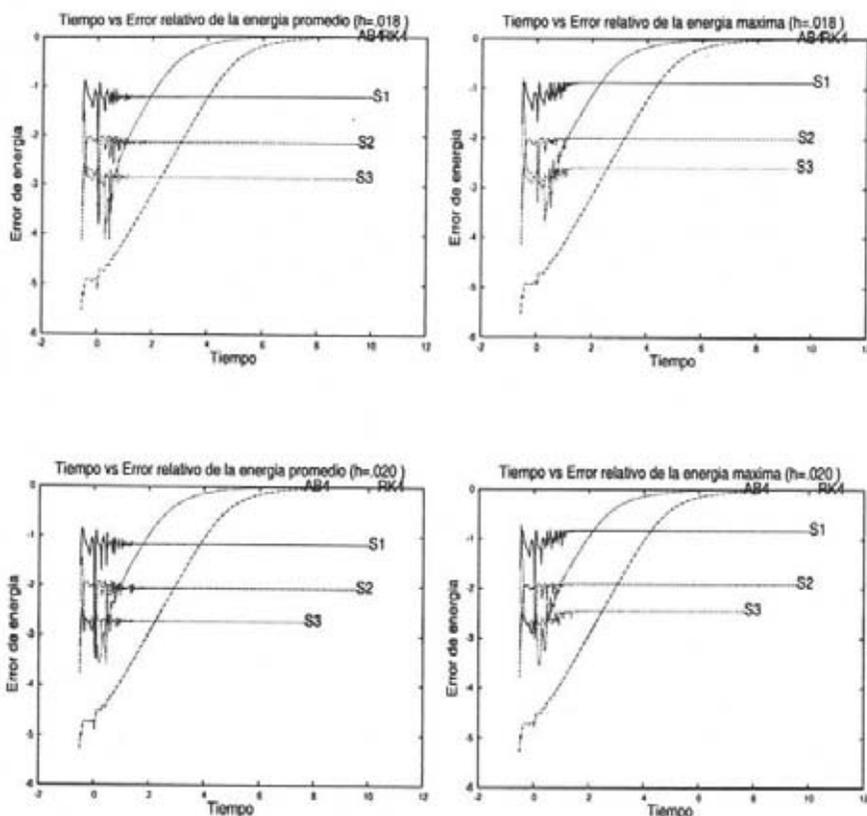
Las simulaciones se hicieron para distintos integradores. En las simulaciones del péndulo ideal y del oscilador armónico, los mejores integradores clásicos fueron RK4 y AB4, razón por la cual los hemos elegido para implementarlos junto con los integradores simplécticos S1, S2 y S3 con los pasos de integración h que hemos venido utilizando.

A continuación se presentan las gráficas de los errores relativos para la energía mecánica y el momento angular generalizado para los integradores elegidos con los distintos pasos de integración.

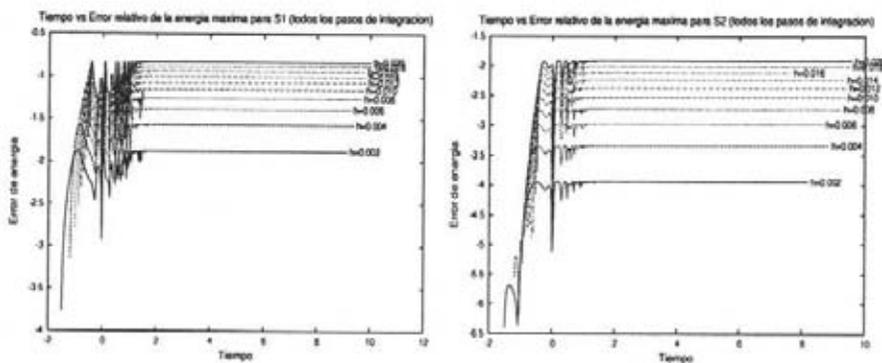


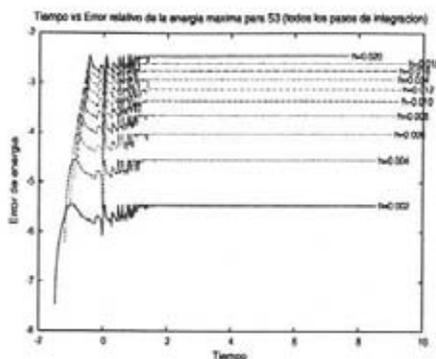






Para los métodos simplécticos podemos realizar un resumen de los errores relativos de la energía máxima cometidos en cada caso mediante el cuadro (4.7) debido a que los errores relativos de los integradores simplécticos tienden a un valor constante como lo muestran las siguientes gráficas.





h	S1	S2	S3
0.002	0.0131	0.00011	3.4×10^{-06}
0.004	0.0266	0.00046	2.7×10^{-05}
0.006	0.0404	0.00104	6.3×10^{-05}
0.008	0.0546	0.00187	0.00022
0.010	0.0691	0.00295	0.00043
0.012	0.0840	0.00429	0.00075
0.014	0.0993	0.00588	0.00119
0.016	0.1149	0.00773	0.00178
0.018	0.1310	0.00985	0.00253
0.020	0.1476	0.01224	0.00348

Cuadro 4.7: Valor estable del error relativo de la energía máxima para la cadena de Toda.

Podemos ver que al aumentar el orden del método obtenemos una mejoría en el error relativo de la energía máxima. Si ajustamos una curva de la forma $y = ax^b$ a los datos del cuadro (4.7) obtenemos para cada integrador simpléctico los parámetros dados en el cuadro (4.8). De este modo obtenemos los parámetros

	S1	S2	S3
a	8.7785	34.1546	435.3281
b	1.0491	2.0300	3.0000

Cuadro 4.8: Parámetros de ajuste de la curva $y = ax^b$.

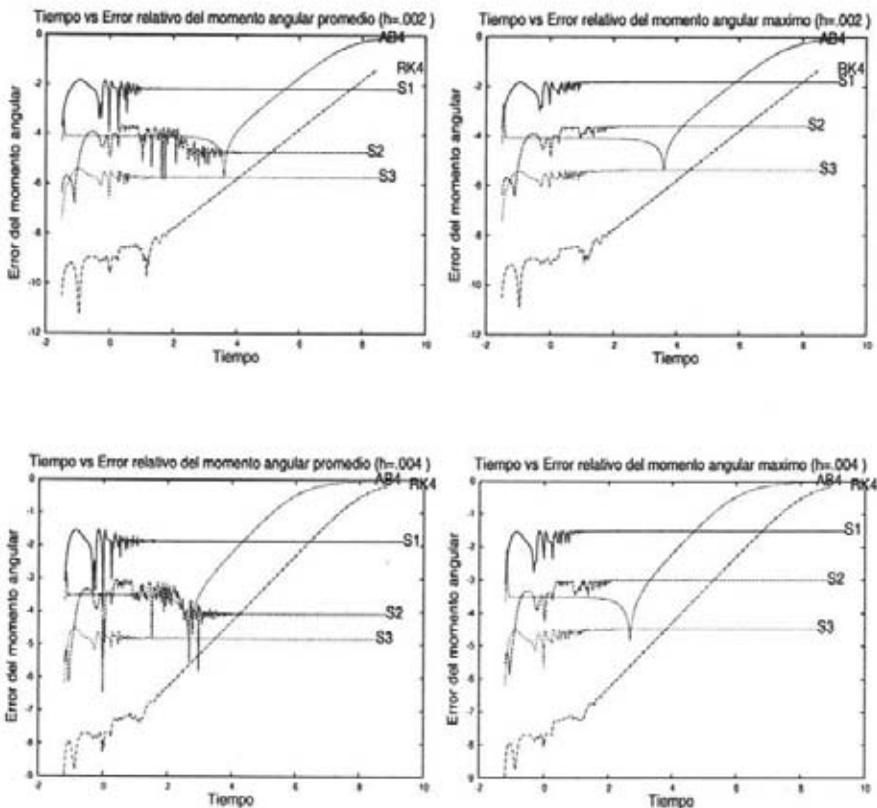
de la ecuación

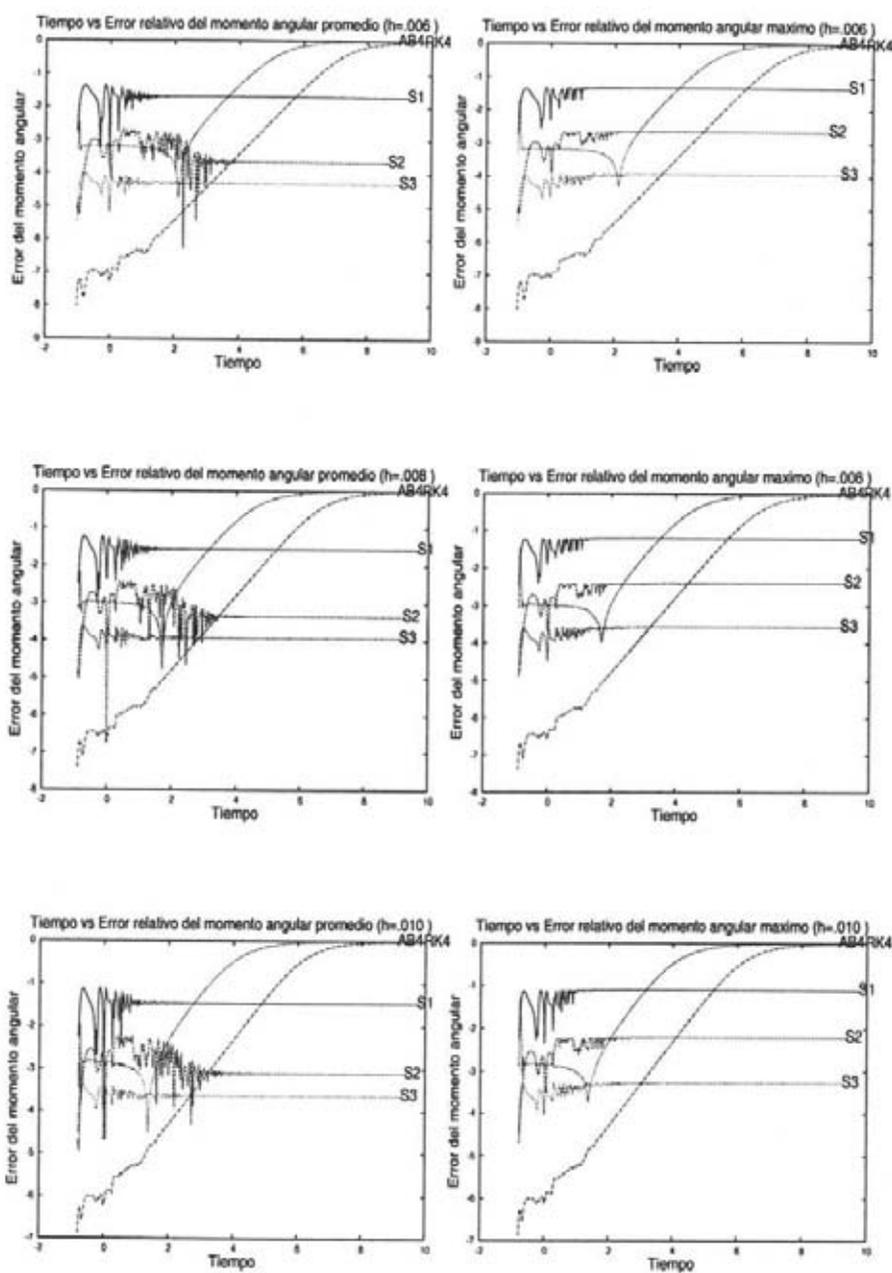
$$\text{Error relativo} = a \times (\text{paso de integración})^b$$

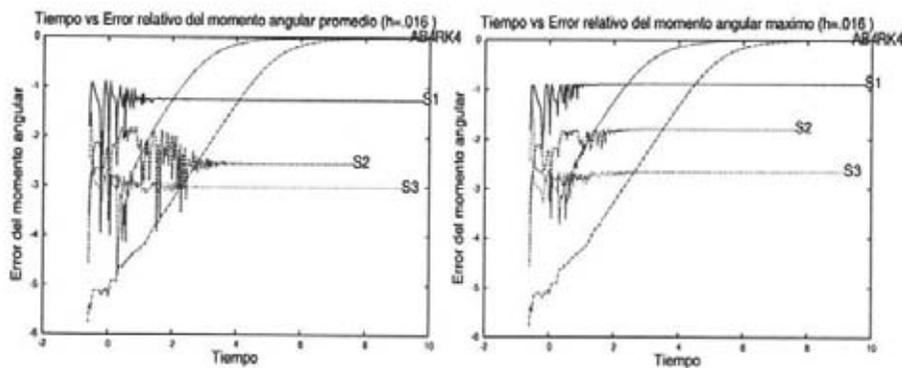
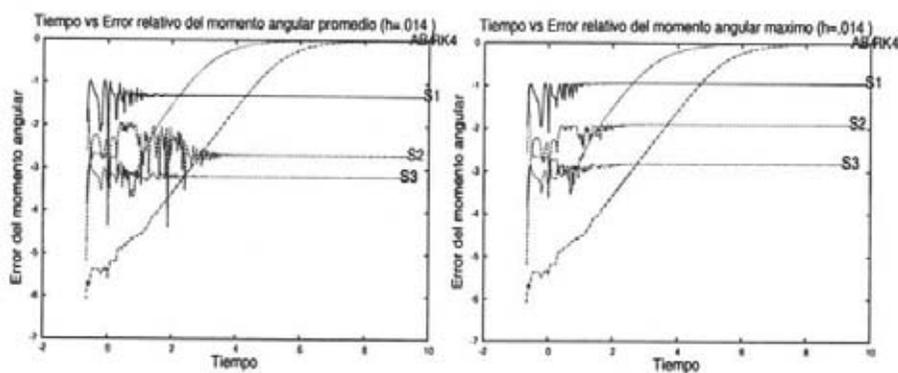
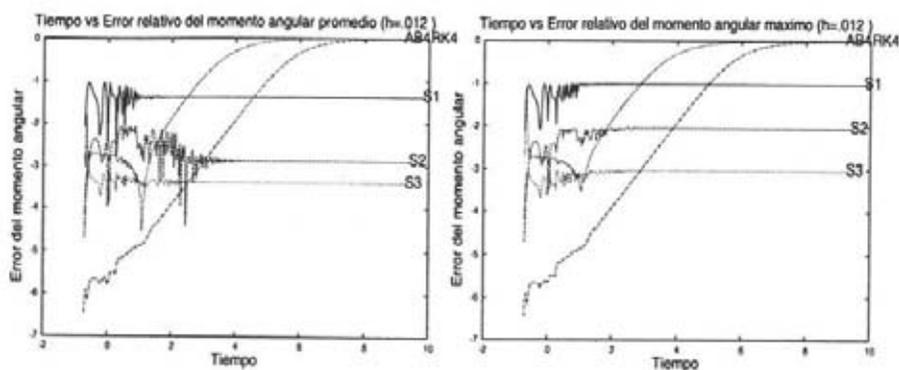
para S1, S2 y S3. Los parámetros dados en el cuadro (4.8) nos indican una estrecha relación entre el orden del método simpléctico y la velocidad con la

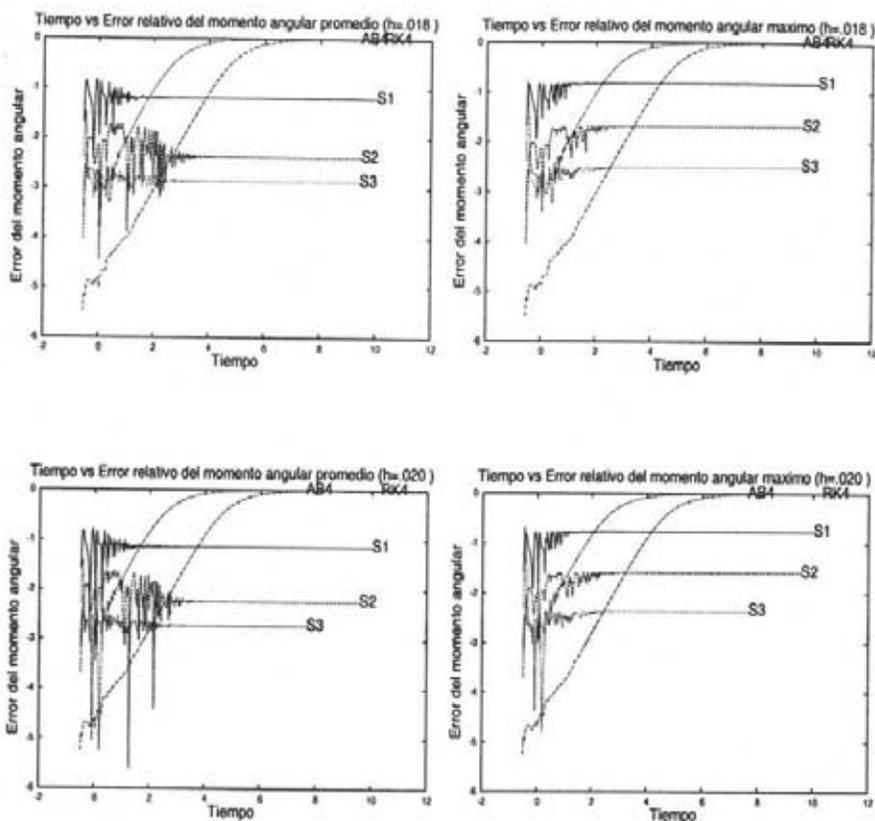
que mejora el error relativo de la energía máxima. Por ejemplo para S1 se tiene que el error relativo mejora con la velocidad con la que hacemos más pequeño el paso de integración h , para S2 mejora cuadráticamente y para S3 cubicamente. Los métodos clásicos resultan ser bastante ineficientes, ya que por ejemplo el integrador RK4 aunque comienza mejor que los simplécticos, con un error relativo de la energía máxima aproximadamente 10^{-5} se tiene que al tiempo 1,000 ya había alcanzado la precisión de S3 y para el tiempo 10,000 ya era menos preciso que S1. El integrador AB4 comienza más o menos con la precisión de S3 pero para el tiempo 100 ya ha alcanzado a S1. En el error relativo de la energía promedio se comportan mejor, teniendo que RK4 permanece por debajo de S3 hasta el tiempo 10,000 y alcanza a S1 en el tiempo 100 millones. AB4 comete el mismo error relativo para la energía promedio que S2 hasta el tiempo 10,000 y para el tiempo 100,000 ya es menos eficiente que S1. Como hasta ahora a sucedido, conforme se aumenta el paso de integración se aprecia un comportamiento más uniforme de los integradores y que los métodos RK4 y AB4 en realidad pierden energía ya que sus errores relativos tienden a 1.

Veamos ahora las gráficas del momento angular generalizado

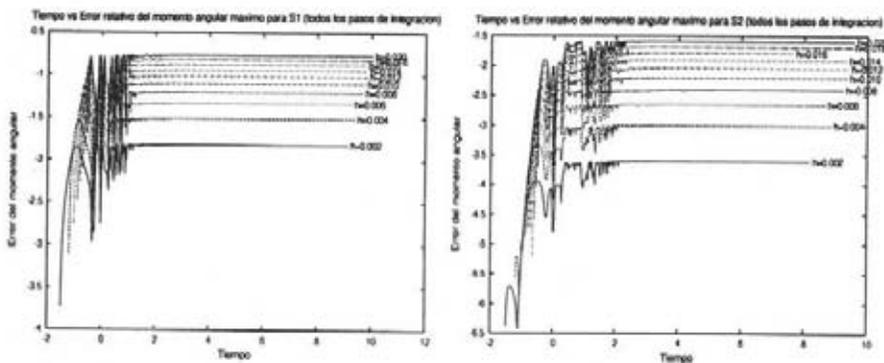


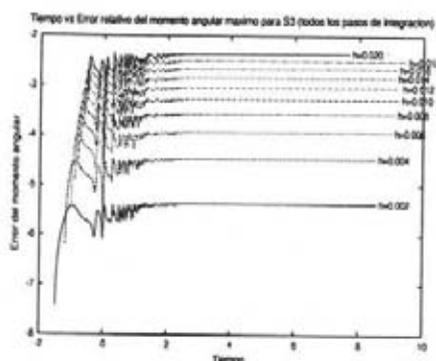






En este caso los errores relativos asociados al promedio y al máximo del momento angular generalizado se comportan esencialmente igual. Sin embargo el error relativo del máximo del momento angular es más representativo ya que podemos ofrecer cotas para los errores. El error relativo del máximo del momento angular, al igual que en la energía, tiende a estabilizarse en un valor como podemos apreciar en las siguientes gráficas





Los valores donde se estabilizan los errores relativos del máximo del momento angular generalizado son presentados en el cuadro (4.9). Como hicimos para

h	S1	S2	S3
0.002	0.0151	0.00024	4.2×10^{-06}
0.004	0.0306	0.00098	3.4×10^{-05}
0.006	0.0465	0.00222	0.00011
0.008	0.0627	0.00398	0.00027
0.010	0.0794	0.00624	0.00053
0.012	0.0964	0.00902	0.00092
0.014	0.1140	0.01233	0.00147
0.016	0.1320	0.01617	0.00220
0.018	0.1505	0.02055	0.00315
0.020	0.1698	0.02547	0.00436

Cuadro 4.9: Valor estable del error relativo del momento angular máximo para la cadena de Toda.

los errores relativos de la energía máxima, ajustaremos a los datos del cuadro (4.9) una curva con los mismo parámetros. Obtenemos así los parámetros dados en el cuadro (4.10) para cada integrador. Igual que antes, la misma relación

	S1	S2	S3
a	10.0663	65.8579	557.13078
b	1.0487	2.0109	3.0084

Cuadro 4.10: Parámetros de ajuste de la curva $y = ax^b$.

estrecha, entre orden del método y precisión en los errores relativos del momento angular generalizado máximo, existe. Es decir, los errores relativos disminuyen lineal, cuadráticamente o cúbicamente (según sea el caso) conforme se reduce el tamaño del paso de integración. La única diferencia es que, como se aprecia en

los cuadros (4.7) y (4.9) los métodos simplécticos son un poco menos precisos en la conservación del momento angular generalizado que en la conservación de la energía mecánica y a su vez tienen un comportamiento menos regular.

El comportamiento de los métodos clásicos resulta esencialmente igual que en el caso de la energía, RK4 comienza con un error relativo del momento angular de aproximadamente 10^{-10} pero rápidamente este error crece hasta alcanzar la precisión de S3 en el tiempo 1,000,000. AB4 se mantiene cualitativamente igual que S2 hasta el tiempo 10,000 y después crece rápidamente su error hasta alcanzar a S1 en el tiempo 10 millones. Una observación interesante es que los métodos clásicos al igual que hacían con la energía también van perdiendo momento angular pues el logaritmo en base de 10 de los errores relativos del momento angular máximo tienden a 0 lo que significa que los errores relativos tienden a 1 y por ende el momento angular tiende a 0.

La ventaja que ofrecen los métodos simplécticos sobre los clásicos es que su comportamiento se mantiene mucho más constante para tiempos largos de integración mientras que los clásicos aumentan el error cometido rápidamente.

Conclusiones

En el capítulo 4 se efectuaron distintas simulaciones numéricas para los sistemas Hamiltonianos del oscilador armónico, del péndulo ideal y de la cadena de Toda. Obtuvimos así una serie de conclusiones alusivas a las diferencias entre los métodos clásicos y los métodos simplécticos. También analizamos diferencias entre los métodos simplécticos de distinto orden en cada caso.

De todos los resultados y observaciones presentadas en el último capítulo cabe destacar algunos hechos importantes. En el caso del péndulo, obtuvimos resultados cualitativamente iguales en el error relativo de la energía mecánica para los distintos tipos de soluciones que puede tener este sistema, las soluciones que se aprecian como curvas cerradas en el plano fase y las oscilatorias, es decir para los métodos simplécticos de orden 1, 2 y 3 tanto el error relativo promedio como el máximo se estabilizan en un valor que depende del paso de integración. Además, también observamos que los valores de los errores relativos disminuyen de manera lineal para el simpléctico de orden 1, cuadrática para el de orden 2 y cúbica para el de orden 3. Así, la manera en que se reducen los errores relativos está íntimamente relacionada con el orden de cada método simpléctico. Los resultados numéricos obtenidos para el oscilador no difieren de los obtenidos para el péndulo exceptuando el hecho de que en este caso el simpléctico de orden 2 es mejor que el simpléctico de orden 3. De hecho la forma en la que se reducen los errores relativos de la energía al disminuir el paso de integración parece estar invertida, pues es cúbica para el simpléctico de orden 2 y cuadrática para el de orden 3. El comportamiento constante de los errores relativos para tiempos de integración grandes no cambia en ninguno de estos dos ejemplos. De hecho, para la cadena de Toda obtenemos los mismos resultados, la energía mecánica es conservada de manera muy regular para tiempos grandes de integración y ésta precisión está relacionada con el tamaño del paso de integración de la misma manera que antes, es decir que vuelve a ser lineal, cuadrática y cúbica para los integradores simplécticos de orden 1, 2 y 3 respectivamente. Éste es un hecho remarcable, ya que la cadena de Toda involucra la evaluación constante de exponenciales y la implementación del método de Newton-Raphson multivariado lo que implica una fuente de error adicional a las propias del método. No sólo se obtuvieron estos resultados para la energía mecánica de la cadena de Toda sino también para el momento angular generalizado se obtuvieron exactamente los mismo resultados cualitativos.

Los métodos clásicos demostraron no ser muy eficientes en la conservación de

integrales primeras pues todos sin excepción tienen un comportamiento creciente en sus errores relativos. Los mejores integradores clásicos, desde el punto de vista simpléctico, son de acuerdo con las observaciones hechas, el método de Runge-Kutta de orden 4 y el de Adams-Bashforth también de orden 4. Estos integradores demostraron tener errores relativos más pequeños que los obtenidos para los métodos simplécticos cuando el tiempo de integración es no mayor a 10,000, pues después de este tiempo su precisión comienza a ser menor. Luego el método de Runge-Kutta es comparable con el simpléctico de orden 3 y el método de Adams-Bashforth con el de orden 2 siempre que el tiempo sea, aproximadamente, menor a 10,000.

Una desventaja de los métodos simplécticos es que se debe implementar el método para cada sistema de ecuaciones que se desee integrar, pues al requerir las derivadas parciales del Hamiltoniano involucrado es necesario calcularlas en cada caso. Por supuesto se podrían interpolar las derivadas mediante la evaluación del Hamiltoniano en algunos puntos para así no requerir de sus derivadas explícitamente. Sin embargo habría que considerar el error generado al efectuar dicha interpolación. Otra posible desventaja es que al conservar la estructura simpléctica, es decir la preservación del área y de integrales primeras en general, es probable que no se conserven exactamente las trayectorias de la solución, es decir, puede preservar el área pero no la figura que contiene dicha área.

Los métodos simplécticos basados en funciones generatrices estudiados en la sección (3.4) han demostrado ser bastante eficientes en la preservación de integrales primeras, de hecho, de acuerdo con los resultados obtenidos en esta tesis, si se deseara que el error relativo de una integral primera de un sistema Hamiltoniano fuera $O((\text{paso de integración})^n)$, el método simpléctico que debería implementarse es el de orden n . Una observación importante es que estos métodos suponen que el Hamiltoniano H , es una función armónica y que su serie de Taylor decae exponencialmente, luego no podemos esperar que estos métodos simplécticos funcionen para cualquier sistema Hamiltoniano, pero sí para una gran variedad de problemas.

Bibliografía

- [1] Paul Blanchard, Rober L. Devaney, and Glen R. Hall. *Differential Equations*, pages 438–442. Books/Cole Publishing Company, 1998.
- [2] P. J. Channell and C. Scovel. Symplectic integration of Hamiltonian systems. *Nonlinearity*, 3:231–259, 1990.
- [3] Lars Eldén and Linde Wittmeyer-Koch. *Numerical Analysis. An Introduction*, pages 114–126, 97–100. Academic Press, Inc., 1990.
- [4] I. M. Gelfand and S. V. Fomin. *Calculus of Variations*, pages 4–24, 34–36, 67–79. Prentice-Hall, Inc., 1963.
- [5] Herbert Goldstein. *Classical Mechanics*, pages 237–250, 266–268. Addison-Wesley Series in Advanced Physics. Addison-Wesley Publishing Company, 1965.
- [6] Gene H. Golub and James M. Ortega. *Scientific Computing and Differential Equations. An Introduction to Numerical Methods*, pages 22–51. Academic Press, Inc., 1992.
- [7] Ernst Hairer, Christian Lubich, and Gerhard Wanner. *Geometric Numerical Integration. Structure-Preserving Algorithms for ordinary Differential Equations*, pages 170–190. Number 31 in Springer Series in Computational Mathematics. Springer-Verlag Berlin Heidelberg New York, 2002.
- [8] Jorge Ize. *Calculo de Variaciones*, pages 55–72, 155–166. IIMAS-FENOMECC, UNAM, second edition, 2002.
- [9] Serge Lang. *Calculus of Several Variables*, pages 121–129, 271–290. Addison-Wesley Publishing Company, 1973.
- [10] M. H. Protter and C. B. Morrey. *A First Course in Real Analysis*, pages 179–181. Undergraduate Texts in Mathematics. Springer-Verlag Berlin Heidelberg New York, 1977.
- [11] T. J. Stuchi. Symplectic integrators revisited. *Brazilian Journal of Physics*, 32(4):958–979, December 2002.

-
- [12] M. Toda. Vibration of a chain with nonlinear interaction. *J. Phys. Soc. Japan*, 22:431–436, 1967.
- [13] M. Toda. Waves in nonlinear lattices. *Prog. Theor. Phys. Suppl*, 45:174–200, 1970.