Kurmyshev

Fundamentos de Métodos Matemáticos para Física e Ingeniería

LIMUSA
NORIEGA EDITORES

GRUPO NORIEGA EDITORES
BALDERAS 95, MÉXICO, D.F.
C.P. 06040
(5) 8503-80-50
01(800) 7-06-91-00
(5) 512-29-03
limusa@noriega.com.mx
www.noriega.com.mx

CANIEM NÚM. 121

PRIMERA EDICIÓN
HECHO EN MÉXICO
ISBN 968-18-6366-6

Contenido

Prefacio

La complejidad de los problemas y sistemas a tratar en las ciencias exactas y naturales requiere de manera indispensable la aplicación de conceptos y métodos de las Matemáticas cada vez más complicados y poderosos. La aplicación de las Matemáticas, una herramienta de investigación científica bella y poderosa, a un problema de la Física o la Ingeniería consiste principalmente en tres fases de igual importancia:

1. Modelo. La elaboración de un modelo matemático consiste en presentar la relación entre los parámetros sustanciales (que se consideran como variables dependientes e independientes) de un proceso, sistema o fenómeno físico o de ingeniería en la forma de un sistema de ecuaciones matemáticas.

2. Resolución. El tratamiento del modelo por medio de métodos matemáticos nos lleva a una solución del problema dado en forma de una relación funcional entre variables dependientes e independientes del modelo.

3. Interpretación. La interpretación de la solución matemática en términos físicos que permite describir el comportamiento del sistema o fenómeno físico o de ingeniería bajo consideración.

Para lograr el éxito en cada una de las tres fases se requiere de intuición, de una forma de pensar matemática y del manejo de los métodos matemáticos. Sin embargo, en nuestros tiempos de "computación total", existe la creencia bastante común de que los problemas matemáticos en su mayoría se resuelven por computadora, es decir, a menudo se desprecia el conocimiento de los conceptos y métodos generales de las Matemáticas que se confunden con la habilidad de manejar, casi a ciegas, los programas computacionales ya hechos para las aplicaciones particulares. Es, por tanto, el objetivo y propósito principal de este libro, que el estudiante se familiarice profundamente con los conceptos matemáticos.

El motivo particular para la elaboración de los "Fundamentos de Métodos Matemáticos" surgió principalmente debido a la siguiente razón. La experiencia de investigador y docente de uno de los autores (E. Kurmyshev) a nivel de posgrado en física en varias materias (Métodos Matemáticos, Mecánica Cuántica, Mecánica Estadística, etc.), junto con las pláticas entre los colegas de trabajo y estudiantes, demostraron varias deficiencias e incongruencias que se mencionan a continuación. Ante todo, hay que reconocer que una parte de los estudiantes de posgrado son ingenieros por formación que en realidad tienen un nivel de estudios en Matemáticas un tanto bajo y conocimientos fragmentarios en la materia. Sin embargo, los programas de estudios en doctorado están elaborados con base en la suposición de un nivel bastante alto que podría permitir los estudios a fondo de materias como mecánica cuántica, electromagnetismo, óptica, estado sólido y mecánica estadística que requieren métodos matemáticos avanzados. Además, varias veces me encontré con una situación un tanto extraña pero real, que dentro de un programa de estudios de doctorado, los métodos matemáticos se imparten en uno o dos semestres y, a veces, al mismo tiempo o después de las materias que los requieren. A pesar de que existen varios libros de texto excelentes en matemáticas y con una amplia cobertura de temas, la mayoría de éstos están escritos en inglés o algún otro idioma extranjero que, por supuesto, dificulta el aprendizaje a los estudiantes de habla española y aún más dentro de un tiempo muy limitado. Por tanto, surgió la necesidad de corregir las incongruencias existentes.

Como requisitos para escribir los "Fundamentos de Métodos Matemáticos" nos ajustamos a los siguientes lineamientos: 1) la selección de temas está apegada al programa de estudios de Doctorado en Física (Óptica) del Centro de Investigaciones en Óptica, A.C., y cubre las necesidades fundamentales en métodos matemáticos de varias materias como óptica clásica, mecánica clásica, electromagnetismo, las bases de teoría de ondas y oscilaciones, mecánica cuántica y mecánica estadística; 2) los "Fundamentos de Métodos Matemáticos" deben ser compactos, bastante rigurosos y con alta cultura matemática; 3) están dirigidos a los estudiantes del programa de posgrado en Física, pero pueden usarse como material de apoyo en ramas afines de ingeniería. El libro se pensó como autoconsistente y autosuficiente a su nivel.

El texto pretende abarcar los temas en matemáticas más indispensables hasta para estudios de Doctorado en Física (Óptica) y áreas afines. Dado que a la materia de matemáticas se le dedica un solo semestre en muchos de los programas de estudios de Doctorado en Física, sólo cambiando la selección de temas y su nivel es prudente ofrecer a los estudiantes un libro de texto que sea consistente y conciso a la vez, dejando una parte del desarrollo matemático para los estudiantes y no sobrecargando el texto con ejemplos. El material de un curso matemático demasiado *"machacado"*, que es bastante común en muchos textos recientes, pierde el sabor de las matemáticas, diluye su esencia y en lugar de ayudar a los estudiantes a descubrir y desarrollar sus habilidades matemáticas puede convertirse en una serie sin fin de ejemplos y ejercicios aburridos. El propósito de los "Fundamentos de Métodos Matemáticos para Física e Ingeniería" es dar a los estudiantes bases sólidas en métodos matemáticos, despertar su mente y enseñarles cierta cultura de estudios y aplicaciones de métodos matemáticos en su trabajo profesional. Se piensa que al alcanzar las metas propuestas por el texto, el estudiante se sentirá mucho más cómodo y con más confianza en sí mismo al estudiar otras materias y, en caso de que surja la necesidad, los conocimientos adquiridos por medio de este libro le permitirán avanzar de forma autodidacta en los estudios de otros métodos matemáticos, apoyándose en libros de texto más completos y avanzados.

Aprovechando el hecho de que ya hay muchos libros que contienen una gran cantidad de problemas de diferentes niveles sobre todos los temas que contiene este libro, los problemas y ejercicios de éste no son numerosos, sino escogidos para ser altamente instructivos; no son únicamente para ejercitar al estudiante, sino para enriquecer y ampliar el conocimiento conceptual sobre el tema correspondiente.

Este libro fue pensado como un repaso conciso, autoconsistente y autosuficiente, una consulta para los que ya saben bastante sobre los métodos matemáticos y/o un curso intensivo (de uno o dos semestres) para los que quieren o necesitan elevar de manera adecuada el nivel de sus conocimientos en la materia.

La bibliografía es muy escasa e incluye básicamente los libros que se usaron para dar forma y contenido al texto.

Y, por fin, a los lectores les pido el gran favor de apuntar pero no contar mis errores (¡no los tiene aquel que no hace nada!). Tal vez mi error principal fue empezar el trabajo de escribir este libro, distrayéndome de la investigación científica que indudablemente regala muchos momentos de satisfacción. Justamente valorando el esfuerzo de la *tribu* llamada estudiantes sentí la necesidad de apoyar sus estudios con mi experiencia personal. El resultado está aquí y a ustedes se les otorga el pleno derecho de juzgarlo.

Se les agradece a Raymundo Mendoza Arce la elaboración de la portada y a Víctor Ayala Ramírez su asesoría en cuanto a la edición del texto en LaTeX.

Dedicatoria. A mi familia y a mis estudiantes, que me dan la razón para vivir una vida completa y feliz.

E.V. Kurmyshev

Capítulo 1

Ecuaciones diferenciales ordinarias

1.1. Clasificación y origen de las ecuaciones diferenciales

Una **ecuación diferencial** es una ecuación que contiene derivadas. La ecuación de la forma

$$F\left(x, y, y', y'', \ldots, y^{(n)}\right) = 0 \tag{1.1}$$

en donde $y = y(x)$ es función de la variable independiente x que se busca y $y^{(n)}$ es la n-ésima derivada con respecto a x, se llama **ecuación diferencial ordinaria de orden n-ésimo**. El **orden** de una ecuación diferencial corresponde al orden de la derivada mayor que interviene en la ecuación, mientras que el **grado** de la ecuación diferencial denota el grado del exponente de la derivada de mayor orden (se implica que la ecuación está escrita en forma polinomial con respecto a la derivada mayor). Presentamos ejemplos de ecuaciones diferenciales:

$$y' = 5 + x \tag{1.2}$$

$$(y'')^2 + (y')^3 + 3y = x^2 \tag{1.3}$$

$$y'' + 3y' + 2y = 0 \tag{1.4}$$

$$y''' + 2(y'')^2 + y' = \cos x \tag{1.5}$$

$$z_{xx} + z_{yy} = x^2 + y \tag{1.6}$$

en donde $y' = dy/dx$ es la derivada ordinaria de la función $y(x)$ con respecto a la variable independiente x, $z_{xx} = \partial^2 z/\partial x^2$ es la derivada parcial de segundo orden. Las ecuaciones (1.2), (1.3), (1.4) y (1.5) son diferenciales ordinarias (EDO), mientras que la ecuación (1.6) es un ejemplo de ecuación diferencial parcial (EDP). La ecuación (1.2) es de orden 1 y grado 1, EDO (1,1); la ecuación (1.3) es de orden 2 y grado 2, EDO (2,2); la ecuación (1.4) es de orden 2 y grado

1, EDO (2,1); la ecuación (1.5) es de orden 3 y grado 1, EDO (3,1); la ecuación (1.6) es EDP (2,1).

Las EDO pueden surgir a partir de

a) problemas geométricos,

b) problemas de ciencias naturales,

c) primitivas de EDO.

Cualquier función $y = y(x)$ que, al sustituirla en la ecuación (1.1), transforma la ecuación diferencial en identidad, recibe el nombre de solución de esta ecuación. A menudo, la solución se da en forma implícita, $\Phi(x, y) = 0$. Una **primitiva** es una relación funcional entre variables x y y que contiene n constantes arbitrarias $C_1, C_2, \ldots, C_n$:

$$\Phi(x, y, C_1, C_2, \ldots, C_n) = 0 \tag{1.7}$$

En general, una primitiva que contenga n **constantes arbitrarias esenciales** (que no se pueden sustituir por un número menor de constantes) se puede reducir a una EDO de orden n, excluyendo las constantes $C_1, C_2, \ldots, C_n$ del sistema de ecuaciones obtenidas derivando la primitiva n veces con respecto a la variable independiente:

$$\Phi(x, y, C_1, C_2, \ldots, C_n) = 0, \quad \frac{d\Phi}{dx} = 0, \quad \ldots, \quad \frac{d^n\Phi}{dx^n} = 0 \tag{1.8}$$

Ejemplo 1.1. *Derivando la primitiva $y = x^2 + C$ con respecto a x, excluimos la constante C y obtenemos la EDO correspondiente, $y' = 2x$.*

Ejemplo 1.2. *La primitiva $y = C_1x^2 + C_2$ contiene dos constantes esenciales. Derivando una vez se excluye la constante C_2 y se tiene $y' = 2C_1x$; derivando nuevamente, se tiene $y'' = 2C_1$. Resolviendo esta ecuación para la constante C_1 y sustituyendo en la ecuación de la primera derivación, resulta la EDO $y' = y''x$ de segundo orden.*

1.2. Soluciones de ecuaciones diferenciales ordinarias

Resolver una EDO de orden n implica encontrar una relación entre variables x y y que contenga n constantes arbitrarias independientes y que satisfaga la EDO, es decir, obtener una primitiva. Por ejemplo, la ecuación $y''' = 0$ tiene como su solución o primitiva la función $y = Ax^2 + Bx + C$. La primitiva de una ecuación diferencial se denomina normalmente la **solución general** de la ecuación. Una **solución particular** de una EDO se obtiene de la primitiva dando valores determinados a las constantes arbitrarias. Así, soluciones particulares de la ecuación anterior son: $y = 0$ para $A = B = C = 0$; $y = 2x + 1$ para $A = 0$, $B = 2$, $C = 1$; $y = 3x^2$ para $A = 3$, $B = C = 0$. Geométricamente, la primitiva es la ecuación de una familia de curvas y una solución particular es la ecuación de una de las curvas.

Por regla, la solución de una EDO de orden n está sujeta a **condiciones iniciales**, que son valores de la función dependiente y de sus $(n-1)$ primeras derivadas en un punto dado x_0: $y(x_0) = a_0$, $y'(x_0) = a_1$, $y''(x_0) = a_2$, $\ldots$, $y^{(n-1)}(x_0) = a_{n-1}$, en donde $a_0, \ldots, a_{n-1}$ son constantes. Al resolver una EDO con condiciones iniciales, se encuentra uno ante los **problemas de existencia y unicidad** de la solución, es decir: ¿bajo qué condiciones se tiene al menos una solución? Y, ¿bajo qué condiciones la EDO tiene sólo una solución? En el caso de una EDO (1,1), la respuesta se da (sin demostración) por el siguiente teorema de existencia y unicidad.

Teorema 1.1. *Dada la EDO de primer orden y primer grado:* $y' = g(x,y)$. *Si* $g(x,y)$ *es continua en un dominio* $R = \{(x,y)\}$ *y* $\partial g/\partial y$ *existe y es continua en* R, *entonces existe una primitiva (una familia infinita de soluciones),* $f(x,y,C) = 0$, *siendo* C *una constante arbitraria, tal que* $y' = g(x,y)$. *Además, la EDO con condiciones iniciales tiene sólo una solución tal que para cualquier punto* (x_0,y_0) *se satisfacen las condiciones iniciales:* $y(x_0) = y_0$, $y'(x_0) = g(x_0,y_0)$.

Una EDO (1,1), $y' = g(x,y)$, asocia con cada punto (x_0,y_0) de la región $R = \{(x,y)\}$ una dirección $m = y'(x_0,y_0) = g(x_0,y_0)$. La dirección en cada punto es la de la tangente a la curva de la familia, $f(x,y,C) = 0$, que pasa por este punto. La región R con la dirección en cada uno de sus puntos se llama **campo de direcciones** de la ecuación diferencial $y' = g(x,y)$ y se representa generalmente por medio de un sistema de flechas con un ángulo de inclinación $\tan\alpha = g(x,y)$. Las curvas $K = g(x,y)$, en cuyos puntos la inclinación del campo tiene un valor constante K, se llaman **isoclinas**.

1.3. Ecuaciones diferenciales de primer orden y primer grado

Una EDO (1,1) se puede escribir en una de las dos formas:

$$M(x,y)dx + N(x,y)dy = 0 \tag{1.9}$$

$$y' = g(x,y) \tag{1.10}$$

Ejemplo 1.3. *La ecuación* $y' = 5+x$ *se puede escribir así:* $(5+x)\,dx - dy = 0$, *donde* $M = 5+x$, $N = -1$. *La ecuación* $y' = (y+x)/(y-x)$ *se puede escribir así:* $(y+x)\,dx + (x-y)\,dy = 0$, *donde* $M = y+x$, $N = x-y$.

Si $M(x,y)dx + N(x,y)dy$ es una diferencial completa de una función $\mu(x,y)$, de modo que $d\mu(x,y) = Mdx + Ndy$, la ecuación (1.9) se llama una **ecuación diferencial exacta** y $\mu(x,y) = C$ es su primitiva o solución general.

Ejemplo 1.4. *Sea* $\mu(x,y) = x^3y^2$, *entonces, la diferencial completa de esta función es* $d\mu(x,y) = 3x^2y^2dx + 2x^3ydy$. *La ecuación diferencial exacta es* $3x^2y^2dx + 2x^3ydy = 0$ *y, por lo tanto, la solución general de ésta es* $x^3y^2 = C$.

Nótese que la ecuación diferencial a su vez se puede escribir como $y' = \frac{-3x^2y^2}{2x^3y} = -\frac{3y}{2x}$.

La diferencial completa de una función $\mu(x, y)$ es

$$d\mu(x, y) = \frac{\partial \mu}{\partial x}dx + \frac{\partial \mu}{\partial y}dy \tag{1.11}$$

y para las segundas derivadas parciales se cumple la igualdad $\frac{\partial^2 \mu}{\partial y \partial x} = \frac{\partial^2 \mu}{\partial x \partial y}$. Por lo tanto, el criterio de que la forma diferencial $M(x, y)dx + N(x, y)dy$ es una diferencial completa se da por la igualdad

$$\frac{\partial M(x, y)}{\partial y} = \frac{\partial N(x, y)}{\partial x} \tag{1.12}$$

Si la ecuación (1.9) no es exacta, sin embargo se encuentra una función $\xi(x, y)$ tal que

$$\xi(x, y)(Mdx + Ndy) = d\mu(x, y) \tag{1.13}$$

la función $\xi(x, y)$ se llama un **factor integrante** y $\mu(x, y) = C$ es la primitiva de la ecuación (1.9).

Ejemplo 1.5. *La ecuación* $3ydx + 2xdy = 0$ *no es exacta, pero al multiplicarla por* x^2y, *se tiene la ecuación* $3x^2y^2dx + 2x^3ydy = 0$, *que es exacta con la primitiva (su solución general)* $x^3y^2 = C$.

Si la ecuación (1.9) no es exacta y no se encuentra rápidamente un factor integrante, es posible que mediante un cambio de una o de las dos variables se obtenga una ecuación en la que se pueda hallar un factor integrante. Aunque no se puede dar una regla general para hallar un factor integrante o una transformación, se pueden considerar tres agrupaciones importantes, que son en cierto modo convencionales, para obtener las soluciones:
1) Ecuaciones del tipo de separación de variables y reducibles a ellas,
2) ecuaciones diferenciales exactas y reducibles a ellas,
3) EDO (1,1) lineales, de la forma $y' + P(x)y = Q(x)$.

Ejemplo 1.6. *La ecuación diferencial* $xdy - ydx = 0$ *se puede clasificar en cualquiera de los grupos: 1) las variables se separan multiplicando la ecuación por el factor integrante* $1/xy$, *así que* $dy/y - dx/x = 0$, *de donde* $\ln y - \ln x = \ln C$, *o sea,* $y/x = C$; *2) por medio del factor integrante* $1/x^2$, *la ecuación se convierte en exacta,* $(xdy - ydx)/x^2 = 0$, *de donde la primitiva es* $y/x = C$; *por medio del factor integrante* $1/y^2$, *la ecuación se convierte en exacta,* $(xdy - ydx)/y^2 = 0$ *con la primitiva* $-x/y = C_1$ *o* $y/x = -1/C_1 = C$; *3) escrita en la forma* $y' - y/x = 0$, *es la ecuación lineal de primer orden.*

Llama la atención el hecho de que la forma de la primitiva no es única. La primitiva en el ejemplo anterior puede ponerse en las siguientes formas: a) $\ln y - \ln x = \ln C$, b) $y/x = C$, c) $y = Cx$, d) $x/y = K$. Aceptando cualquiera de estas formas, se pueden perder ciertas soluciones particulares.

1.3.1. Separación de variables

Las EDO (1,1) de la forma

$$y' = f(x)\, g(y) \quad \text{o} \quad X(x)dx + Y(y)dy = 0 \tag{1.14}$$

se llaman ecuaciones diferenciales con **variables separables**. Dividiendo la primera ecuación por $g(y)$ y multiplicando por dx, obtenemos $dy/g(y) = f(x)\, dx$. Integrando la ecuación (1.14), se obtiene la solución general en la forma

$$\int \frac{dy}{g(y)} = \int f(x)\, dx + C \quad \text{o} \quad \int X(x)dx + \int Y(y)dy = C \tag{1.15}$$

Se halla rápidamente un factor integrante para la ecuación de la forma

$$f_1(x)g_2(y)dx + f_2(x)g_1(y)dy = 0 \tag{1.16}$$

Al multiplicarla por el factor integrante $\xi(x,y) = \frac{1}{g_2(y)f_2(x)}$, la ecuación resulta en la forma con variables separadas

$$\frac{f_1(x)}{f_2(x)}dx + \frac{g_1(y)}{g_2(y)}dy = 0 \tag{1.17}$$

cuya primitiva se obtiene simplemente integrando:

$$\int \frac{f_1(x)}{f_2(x)}dx + \int \frac{g_1(y)}{g_2(y)}dy = C \tag{1.18}$$

Ejemplo 1.7. *La ecuación diferencial $(x-1)^2 y dx + x^2(y+1)dy = 0$ que no es exacta, multiplicada por el factor integrante $\frac{1}{yx^2}$ resulta en $\frac{(x-1)^2}{x^2}dx + \frac{(y+1)}{y}dy = 0$. Integrando se obtiene la solución general, $x - 2\ln x - \frac{1}{x} + y + \ln y = C$.*

1.3.2. Ecuaciones diferenciales homogéneas y reducción a separación de variables

Una función $f(x,y)$ es homogénea de grado n si para un número entero n se cumple la igualdad

$$f(\lambda x, \lambda y) = \lambda^n f(x,y) \tag{1.19}$$

Ejemplo 1.8. *La función $f(x,y) = x^3 - x^2 y$ es de tercer grado. La $f(x,y) = \exp(y/x) + \tan(x/y)$ es de grado 0. La función $f(x,y) = x^3 - x^2 \operatorname{sen} y$ no es homogénea.*

Una EDO (1,1), $M(x,y)dx + N(x,y)dy = 0$, es una **ecuación diferencial homogénea** si tanto $M(x,y)$ como $N(x,y)$ son funciones homogéneas del mismo grado n.

Ejemplo 1.9. *La EDO (1,1) $x\ln(y/x)\, dx + (y^2/x) \operatorname{arcsen}(y/x)\, dy = 0$ es la ecuación diferencial homogénea, dado que tiene funciones $M(x,y)$ y $N(x,y)$ homogéneas de primer grado.*

Proposición 1.1. *La transformación*

$$y = vx, \ \text{con } dy = xdv + vdx \tag{1.20}$$

reduce cualquier EDO (1,1) homogénea, $M(x,y)dx + N(x,y)dy = 0$, a una ecuación del tipo de variables separables.

Demostración. *Dado que $M(x,y)$ y $N(x,y)$ son homogéneas de grado n y considerando $\lambda = \frac{1}{x}$, tenemos*

$$M(\lambda x, \lambda y) = \lambda^n M(x,y) \quad \Rightarrow \quad M(1, \frac{y}{x}) = \frac{1}{x^n} M(x,y)$$

$$N(\lambda x, \lambda y) = \lambda^n N(x,y) \quad \Rightarrow \quad N(1, \frac{y}{x}) = \frac{1}{x^n} N(x,y)$$

Empleando la notación $N_1(\frac{y}{x}) \equiv N(1, \frac{y}{x})$, $M_1(\frac{y}{x}) \equiv M(1, \frac{y}{x})$, reescribimos la ecuación inicial como

$$x^n \left[M_1\left(\frac{y}{x}\right) dx + N_1\left(\frac{y}{x}\right) dy \right] = 0 \quad \Rightarrow M_1\left(\frac{y}{x}\right) dx + N_1\left(\frac{y}{x}\right) dy = 0$$

Aplicando la transformación $y = vx$, $dy = xdv + vdx$ a la ecuación anterior se tiene

$$\left[M_1(v) + vN_1(v)\right] dx + N_1(v) x dv = 0$$

que tiene la forma

$$f(v)dx + xN_1(v)dv = 0 \tag{1.21}$$

Al multiplicar por el factor integrante

$$\frac{1}{f(v)x} = \frac{1}{x\left[M_1(v) + vN_1(v)\right]} \tag{1.22}$$

resulta en la ecuación

$$\frac{dx}{x} + \frac{N_1(v)}{M_1(v) + vN_1(v)} dv = 0 \tag{1.23}$$

cuya solución general se obtiene integrando,

$$\int \frac{dx}{x} + \int \frac{N_1(v)}{M_1(v) + vN_1(v)} dv = C \tag{1.24}$$

Quedando completa la demostración. ■

Problema. *Sean $f_1(x,y)$ y $f_2(x,y)$ dos funciones homogéneas de grado k y m, respectivamente. ¿Qué se puede decir sobre la homogeneidad de las funciones $g(x,y) = f_1(x,y) f_2(x,y)$ y $h(x,y) = \alpha f_1(x,y) + \beta f_2(x,y)$, en donde α y β son constantes?*

Problema. *Sea $f(x,y)$ una función homogénea de grado n. Demuestre la igualdad (el teorema de Euler),*

$$xf_x + yf_y = nf(x,y)$$

en donde f_x y f_y son derivadas parciales.

1.3.3. Ecuaciones con $M(x,y)$ y $N(x,y)$ lineales, pero no homogéneas

La ecuación no homogénea

$$(a_1x + b_1y + c_1)dx + (a_2x + b_2y + c_2)dy = 0 \tag{1.25}$$

se reduce por una transformación lineal de variables a la forma en la que las variables son separables. Las funciones $g_1(x,y) \equiv a_1x+b_1y+c_1 = 0$ y $g_2(x,y) \equiv a_2x+b_2y+c_2 = 0$ definen dos rectas que en el caso de que $c_1 = c_2 = 0$ pasan por el origen del sistema de coordenadas. Notando que con $c_1 = c_2 = 0$ la ecuación (1.25) será homogénea de grado uno, encontramos el punto de intersección (h,k) de las rectas $g_1(x,y) = 0$ y $g_2(x,y) = 0$ y movemos allí el origen del nuevo sistema de coordenadas, quedando la transformación lineal $x = X+h$, $y = Y+k$. En las nuevas coordenadas las ecuaciones de dichas rectas son

$$a_1(X+h) + b_1(Y+k) + c_1 = a_1X + b_1Y + a_1h + b_1k + c_1 = 0$$

$$a_2(X+h) + b_2(Y+k) + c_2 = a_2X + b_2Y + a_2h + b_2k + c_2 = 0$$

y el punto de intersección se define a condición de que $X = Y = 0$, es decir, que $a_1h + b_1k + c_1 = 0$ y $a_2h + b_2k + c_2 = 0$. Dado que $a_1b_2 - a_2b_1 \neq 0$, de las últimas dos ecuaciones se obtienen unívocamente (h,k). Por la transformación lineal $x = X+h$, $y = Y+k$, la ecuación (1.25) se reduce a la ecuación homogénea de grado uno

$$(a_1X + b_1Y)\,dX + (a_2X + b_2Y)\,dY = 0 \tag{1.26}$$

que a su vez se reduce a una ecuación del tipo de separación de variables.

Cuando $a_1b_2 - a_2b_1 = 0$, la ecuación (1.25) se reduce por la transformación

$$a_1x + b_1y = t, \quad dy = (dt - a_1dx)/b_1$$

a la forma $P(t)\,dx + Q(t)\,dt = 0$ en la que las variables son separables.

1.3.4. Ecuaciones de la forma $f(xy)ydx + g(xy)xdy = 0$

La transformación $xy = z$, reduce una ecuación de este tipo a la forma de separación de variables. Sea $xy = z$. Sustituyendo $y = \frac{z}{x}$ y $dy = \frac{xdz - zdx}{x^2}$ en la ecuación original, se tiene

$$\left(\frac{z}{x}\right) f(z)dx + xg(z)\left(\frac{xdz - zdx}{x^2}\right) = 0 \tag{1.27}$$

de donde se obtiene la ecuación en la que las variables son separables:

$$z\,(f(z) - g\,(z))\,dx + xg(z)dz = 0 \Rightarrow \frac{1}{x}dx + \frac{g(z)}{z\,(f(z) - g(z))}dz = 0 \tag{1.28}$$

1.3.5. Ecuaciones diferenciales exactas y reducción a ellas

Una ecuación diferencial de la forma de la ecuación (1.9),

$$M(x,y)dx + N(x,y)dy = 0$$

se llama exacta si existe una función $\mu(x,y)$ tal que

$$d\mu(x,y) = Mdx + Ndy \tag{1.29}$$

en cuyo caso $\mu(x,y) = C$ es la primitiva de la ecuación.

La condición suficiente y necesaria de que la forma diferencial $Mdx + Ndy$ sea la diferencial completa de una función $\mu(x,y)$ se da por la ecuación (1.12),

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$$

En este caso,

$$M(x,y) = \frac{\partial \mu}{\partial x}, \quad N(x,y) = \frac{\partial \mu}{\partial y} \tag{1.30}$$

Ejemplo 1.10. *Sea $(x^2 - y)dx + (-x + y^2)dy = 0$, entonces $\frac{\partial(x^2-y)}{\partial y} = -1$, y $\frac{\partial(-x+y^2)}{\partial x} = -1$, por lo que la ecuación es exacta.*

La integral $\int_A^B d\mu(x,y) = \mu(x,y) - \mu(x_a, y_a)$ no depende de la trayectoria, donde $A(x_a, y_a)$ y $B(x,y)$ son dos puntos cualesquiera en el plano xy. Por lo tanto, integrando con respecto a x y siendo y constante, tenemos

$$\mu(x,y) = \int^x \frac{\partial \mu}{\partial x} dx + \varphi(y) = \int^x M(x,y)\,dx + \varphi(y) \tag{1.31}$$

donde la función $\varphi(y)$ es la "constante" de integración con respecto a x y que aún no está determinada. Luego,

$$\frac{\partial \mu}{\partial y} = \frac{\partial}{\partial y}\left[\int^x M(x,y)dx + \varphi(y)\right] = \frac{\partial}{\partial y}\left[\int^x M(x,y)dx\right] + \varphi'(y) = N(x,y) \tag{1.32}$$

De la última ecuación podemos encontrar $\varphi(y)$, integrando con respecto a y y tratando x como una constante:

$$\varphi(y) = \int^y N(x,y)dy - \int^y \frac{\partial}{\partial y}\left[\int^x M(x,y)dx\right] dy$$

Sustituyendo la $\varphi(y)$ en la ecuación (1.31), encontraremos la primitiva de la ecuación diferencial exacta:

$$\mu(x,y) = \int^x M(x,y)\,dx + \int^y N(x,y)dy - \int^y \frac{\partial}{\partial y}\left[\int^x M(x,y)dx\right] dy = C \tag{1.33}$$

Ejemplo 1.11. *La ecuación* $(2x^3+3y)dx+(3x+y+1)dy=0$ *es exacta, pues* $\frac{\partial M}{\partial y}=\frac{\partial N}{\partial x}=3$. *Luego,*

$$M(x,y)=\frac{\partial\mu}{\partial x}=2x^3+3y, \qquad N(x,y)=\frac{\partial\mu}{\partial y}=3x+y+1$$

así que

$$\mu(x,y)=\int^x (2x^3+3y)\,dx+\varphi(y)=\frac{x^4}{2}+3xy+\varphi(y)$$

$$N(x,y)=\frac{\partial\mu}{\partial y}=3x+\varphi'(y)=3x+y+1$$

de donde $\varphi'(y)=y+1$, *que integrando resulta en* $\varphi(y)=\frac{y^2}{2}+y$. *Finalmente, la primitiva de la ecuación es*

$$\mu(x,y)=\frac{x^4}{2}+3xy+\frac{y^2}{2}+y=C$$

Si la ecuación (1.9) no es exacta, se busca algún factor integrante.

Proposición 1.2. *Si* $\frac{M_y-N_x}{N}=f(x)$, *función sólo de* x, *entonces* $\exp\left[\int f(x)dx\right]$ *es un factor integrante de la ecuación (1.9).*

Demostración. *Verificamos la condición necesaria y suficiente. El factor* $\xi(x)=\exp\left[\int f(x)dx\right]$ *es una función sólo de* x. *Entonces,*

$$\frac{\partial}{\partial y}\left(\xi(x)\,M(x,y)\right)=\xi M_y$$

$$\frac{\partial}{\partial x}\left(\xi(x)\,N(x,y)\right)=\xi N_x+\xi fN=\xi N_x+\xi\left(M_y-N_x\right)=\xi M_y$$

Esto completa la demostración. ■

Proposición 1.3. *Si* $\frac{M_y-N_x}{M}=-g(y)$, *función sólo de* y, *entonces* $\exp\left[\int g(y)dy\right]$ *es un factor integrante de la ecuación (1.9).*

Demostración. *El factor* $\xi(y)=\exp\left[\int g(y)dy\right]$ *es función sólo de* y. *Al multiplicar por el factor integrante,* $d\mu(x,y)=\xi\ Mdx+\xi\ Ndy$, *se tiene que cumplir la condición de que* $\frac{\partial(\xi\ M)}{\partial y}=\frac{\partial(\xi\ N)}{\partial x}$. *Diferenciando,*

$$\frac{\partial(\xi\ M)}{\partial y}=\xi M_y+\xi gM=\xi M_y+\xi\left(N_x-M_y\right)=\xi N_x$$

$$\frac{\partial(\xi\ N)}{\partial x}=\xi N_x$$

Por lo tanto, queda demostrado que el factor es integrante. ■

Proposición 1.4. *Si la ecuación (1.9) es homogénea y* $Mx+Ny \neq 0$*, entonces,* $(Mx+Ny)^{-1}$ *es un factor integrante.*

Demostración. *Sea* $\xi(x,y) = (Mx+Ny)^{-1}$. *Se tiene*

$$\frac{\partial(\xi\ M)}{\partial y} = \frac{\partial}{\partial y}\frac{M}{Mx+Ny} = \frac{yM_yN - yN_yM - MN}{(Mx+Ny)^2}$$

$$\frac{\partial(\xi\ N)}{\partial x} = \frac{\partial}{\partial x}\frac{N}{Mx+Ny} = \frac{xN_xM - xM_xN - MN}{(Mx+Ny)^2}$$

Entonces, usando el teorema de Euler de una función $H(x,y)$ *homogénea de grado* n*:*

$$xH_x + yH_y = nH$$

se tiene

$$\frac{\partial(\xi\ M)}{\partial y} - \frac{\partial(\xi\ N)}{\partial x} = \frac{(yM_y + xM_x)\,N - (yN_y + xN_x)\,M}{(Mx+Ny)^2} = 0$$

y queda demostrado que el factor es integrante. ■

Nótese que una EDO (1,1) homogénea siempre puede ser resuelta reduciéndola a la ecuación del tipo de separación de variables por medio de la transformación de variables, ecuación (1.20). Además, puesto que $M\ x + N\ y \neq 0$, ésta se resuelve reduciéndola a la exacta por medio de un factor integrante.

A veces un factor integrante se halla por inspección, agrupando convenientemente los términos de la ecuación y reconociendo un cierto grupo de términos como parte de una diferencial exacta. En la siguiente tabla se presentan algunos factores integrantes de utilidad.

Grupo de términos	**Factor integrante**	**Primitiva de ecuación**
$-ydx + xdy$	$\frac{1}{x^2}$	$\frac{y}{x} = C$
$-ydx + xdy$	$\frac{1}{y^2}$	$-\frac{x}{y} = C$
$-ydx + xdy$	$\frac{1}{xy}$	$\ln\frac{y}{x} = C$
$-ydx + xdy$	$\frac{1}{x^2+y^2}$	$\arctan\frac{y}{x} = C$
$ydx + xdy$	$\frac{1}{xy}$	$\ln xy = C$
$xdx + ydy$	$\frac{1}{x^2+y^2}$	$\frac{1}{2}\ln(x^2+y^2) = C$

1.3.6. Ecuaciones diferenciales lineales y reducibles a lineales

La ecuación

$$y' + P(x)y = Q(x) \tag{1.34}$$

se llama **ecuación diferencial lineal de primer orden**, siendo una forma lineal tanto en la variable dependiente como en su derivada.

Ejemplo 1.12. *La ecuación* $y' + 3x^2y = \operatorname{sen} x$*, es lineal, mientras que* $y' + 3x^2y^2 = \operatorname{sen} x$*, no lo es.*

La ecuación (1.34) escrita en diferenciales es $(yP - Q)dx + dy = 0$. Luego, $M(x, y) = yP(x) - Q(x)$ y $N(x, y) = 1$ resultando en que $\frac{M_y - N_x}{N} = P(x)$ es sólo función de x. Por lo tanto, según la Proposición 18, un factor integrante de la ecuación (1.34) es $\xi(x) = e^{\int P(x)dx}$. La ecuación diferencial, multiplicada por el factor integrante, resulta en

$$e^{\int P(x)dx}(y' + yP(x)) = \frac{d}{dx}\left[ye^{\int P(x)dx}\right] = Q(x)e^{\int P(x)dx}$$

y la primitiva de la ecuación (1.34) es

$$ye^{\int P(x)dx} = \int Q(x)e^{\int P(x)dx}dx + C \tag{1.35}$$

1.3.7. Ecuación de Bernoulli

La ecuación diferencial

$$y' + P(x)y = Q(x)y^a \tag{1.36}$$

donde a es cualquier número, se llama **ecuación de Bernoulli**. Para $a = 0$ y $a = 1$, la ecuación es lineal; de lo contrario no. Una forma equivalente es $y^{-a}y' + y^{1-a}P(x) = Q(x)$. Haciendo el cambio de variable

$$z = y^{1-a} \tag{1.37}$$

se tiene que $z' = (y^{1-a})' = (1-a)y^{-a}y'$. Al sustituir en la ecuación (1.36), se tiene

$$z' + (1-a)zP(x) = (1-a)Q(x) \tag{1.38}$$

que es una ecuación lineal.

El factor integrante $e^{(1-a)\int P(x)dx}$, aplicado a la ecuación (1.38), lleva a la solución

$$z = y^{1-a} = e^{(a-1)\int P(x)dx}\left[\int (1-a)Q(x)e^{(1-a)\int P(x)dx}dx\right] + Ce^{(a-1)\int P(x)dx} \tag{1.39}$$

1.3.8. Ecuación de Riccati

La ecuación

$$y' + P(x)y + Q(x)y^2 + R(x) = 0 \tag{1.40}$$

donde $P(x)$, $Q(x)$ y $R(x)$ son funciones de x, es conocida como **ecuación de Riccati generalizada**. La ecuación de Riccati original es $y' + Ay^2 = Bx^m$, donde A y B son constantes y m es un número natural. En caso de $R(x) = 0$, la ecuación (1.40) es de Bernoulli con $a = 2$. Si $Q(x) = 0$, se tiene una ecuación lineal.

En general, la ecuación de Riccati generalizada no es integrable en cuadraturas. Sin embargo, cuando se conoce cualquier solución particular, la solución general de ésta se obtiene fácilmente reduciéndola a la ecuación de Bernoulli, y esta última es reducida a una ecuación lineal.

Sea y_1 una solución particular de (1.40). Buscamos la solución general de la ecuación (1.40) en la forma

$$y = y_1 + z \tag{1.41}$$

Sustituyendo (1.41) en (1.40), se tiene

$$z' + Q(x)\left[2y_1 z + z^2\right] + P(x)z = 0 \quad \text{o} \quad z' + \tilde{P}(x)z = \tilde{Q}(x)z^2 \tag{1.42}$$

que es la ecuación de Bernoulli, ecuación (1.36), con respecto a la incógnita z, donde $\tilde{Q}(x) = -Q(x)$, $\tilde{P}(x) = P(x) + 2y_1 Q(x)$ y $a = 2$. El cambio de variable,

$$v = \frac{1}{z}$$

reduce la última ecuación a una ecuación lineal (véase la sección anterior).

Proposición 1.5. *Sean y_1, y_2, y_3 soluciones particulares de (1.40) y sea y su solución general. Entonces,*

$$u = \frac{1}{y - y_1},\ u_1 = \frac{1}{y_2 - y_1},\ u_2 = \frac{1}{y_3 - y_1}$$

son soluciones de la misma ecuación diferencial lineal y debido a esto

$$\frac{u - u_1}{u_2 - u_1} = C \quad \Rightarrow \quad \frac{y - y_2}{y - y_1} = C\,\frac{y_3 - y_2}{y_3 - y_1}$$

donde C es una constante.

La última fórmula muestra que la solución general de la ecuación de Riccati se puede expresar de la manera racional en términos de cualquiera de tres soluciones particulares distintas; además, se tiene que la solución general es una función racional de la constante de integración. Por tanto, cualquier función del tipo

$$y = \frac{Cf_1 + f_2}{Cf_3 + f_4} \tag{1.43}$$

satisface una ecuación de Riccati, ecuación (1.40), donde C es una constante arbitraria y f_1, f_2, f_3 y f_4 son funciones dadas de x. Lo último se demuestra eliminando C de las expresiones para y y su derivada y'.

Con $Q(x) = 0$ la ecuación de Riccati se transforma en una ecuación lineal. Si $Q(x) \neq 0$, la ecuación de Riccati es reducible a una EDO (2,1) lineal. Haciendo el cambio de variable dependiente,

$$y = \frac{v}{Q(x)} \tag{1.44}$$

en la ecuación (1.40), resulta

$$v' + v^2 + v\left[P - \frac{Q'}{Q}\right] + R\,Q = 0$$

Luego, sustituyendo $v = (\ln u)' = \frac{u'}{u}$, se tiene

$$\frac{u''u - u'^2}{u^2} + \left(\frac{u'}{u}\right)^2 + \frac{u'}{u}\left[P - \frac{Q'}{Q}\right] + R\,Q = 0$$

y finalmente

$$u'' + u'\left[P - \frac{Q'}{Q}\right] + uR\,Q = 0 \tag{1.45}$$

que es la ecuación diferencial lineal de segundo orden.

Ejemplo 1.13. *La ecuación $y' - y^2 + 2xy - (x^2 + 1) = 0$ es de Riccati. Es fácil ver que $y = x$ es una solución particular. Haciendo la sustitución $y = x + z$, se tiene*

$$1 + z' - (x^2 + 2xz + z^2) + 2x(x + z) - (x^2 + 1) = 0$$
$$z' = z^2$$

La última ecuación se resuelve fácilmente por separación de variables,

$$\frac{dz}{dx} = z^2 \quad \Rightarrow \quad \int \frac{dz}{z^2} = \int dx \quad \Rightarrow \quad x = -\frac{1}{z} + C$$

También puede resolverse como ecuación de Bernoulli realizando la sustitución $v = \frac{1}{z}$, $z = \frac{1}{v}$, $z' = -\frac{v'}{v^2}$,

$$-\frac{v'}{v^2} = \frac{1}{v^2} \quad \Rightarrow \quad v' = -1 \quad \Rightarrow \quad v = -x + C \quad \Rightarrow \quad -x + C = \frac{1}{z}$$

Así, $z = \frac{1}{C-x}$, luego, como $y = x + z$, se tiene finalmente la solución general

$$y = \frac{Cx - x^2 + 1}{C - x}$$

Ejemplo 1.14. *La ecuación $y' - y^2 + \frac{y}{x} + \frac{1}{x^2} = 0$ es de Riccati. Se puede comprobar fácilmente que $y_1 = \frac{1}{x}$ es una solución particular. Sustituyendo $y = y_1 + z = \frac{1}{x} + z$ en la ecuación, se tiene*

$$\left(\frac{1}{x}\right)' + z' - \left(\frac{1}{x^2} + \frac{2z}{x} + z^2\right) + \frac{1}{x^2} + \frac{z}{x} + \frac{1}{x^2} = 0$$

o bien, reacomodando los términos se tiene la ecuación de Bernoulli con $a = 2$:

$$z' - \frac{z}{x} = z^2$$

El cambio de la variable dependiente, $u = z^{-1}$, pues $z = u^{-1}$ y $z' = -u^{-2}u'$, reduce la ecuación anterior a la ecuación lineal

$$u' + \frac{1}{x}u = -1$$

cuyo factor integrante es

$$\xi(x) = e^{\int \frac{1}{x}dx} = e^{\ln x} = x$$

Al multiplicar la ecuación diferencial por el factor integrante se tiene

$$xu' + u = -x, \quad \int d(xu) = -\int x dx, \quad xu = -\frac{x^2}{2} + C$$

de donde $u = -\frac{x}{2} + \frac{C}{x}$. *Ahora, como* $z = u^{-1}$, *tenemos que*

$$z = \frac{2x}{2C - x^2}$$

Finalmente, como $y = \frac{1}{x} + z$, *se tiene que la solución general es*

$$y = \frac{x^2 + 2C}{x(2C - x^2)}$$

1.4. Ecuaciones de primer orden y de grado superior

Una ecuación diferencial de primer orden tiene la forma $f(x, y, y') = 0$, o bien, $f(x, y, p) = 0$, donde sustituimos y' por p, $y' = p$. Por ejemplo, la ecuación $p^2 - 3px + 2y = 0$, es una EDO (1,2). La ecuación diferencial ordinaria de primer orden de grado n se puede escribir en la forma

$$p^n + P_1(x, y)p^{n-1} + P_2(x, y)p^{n-2} + \cdots + P_{n-1}(x, y)p + P_n(x, y) = 0 \quad (1.46)$$

Al resolver estas ecuaciones, EDO (1,n), pueden presentarse varios casos.

1.4.1. Ecuación resuelta con respecto a y'

Supongamos que la ecuación (1.46), considerada como un polinomio en p, se puede resolver con respecto a p, es decir, esta ecuación polinomial se puede presentar en la forma de n factores

$$(p - F_1(x, y))(p - F_2(x, y)) \quad \cdots \quad (p - F_n(x, y)) = 0 \quad (1.47)$$

donde las raíces de la ecuación (1.46), $F_i(x, y)$, $i = 1, 2, \ldots, n$, son funciones de x y y. Igualando a cero cada factor y resolviendo las n ecuaciones diferenciales resultantes de primer orden y primer grado

$$\frac{dy}{dx} = F_1(x, y), \quad \frac{dy}{dx} = F_2(x, y), \quad \ldots, \quad \frac{dy}{dx} = F_n(x, y)$$

se obtienen las primitivas correspondientes

$$f_1(x, y, C) = 0, \quad f_2(x, y, C) = 0, \quad \ldots, \quad f_n(x, y, C) = 0 \quad (1.48)$$

Cada una de estas primitivas anula uno de los factores en la ecuación (1.47) y, por esto satisface tanto la ecuación (1.47) como la ecuación (1.46). Por lo tanto, cada una de las primitivas (1.48) puede llamarse la primitiva parcial de las ecuaciones (1.47, 1.46). La primitiva de (1.47) y, por tanto, de (1.46), es el producto

$$f_1(x,y,C)f_2(x,y,C) \quad \cdots \quad f_n(x,y,C)=0 \tag{1.49}$$

de n soluciones, ecuación (1.48).

Ejemplo 1.15. *Es fácil ver que la EDO*

$$p^2-2x\,(y+1)\,p+4x^2y=0, \quad \text{o bien,} \quad (p-2x)\,(p-2xy)=0$$

tiene las siguientes primitivas parciales

$$\begin{aligned} y &= x^2+C, \quad \text{o bien,} \quad f_1\,(x,y,C)\equiv y-x^2-C=0 \\ y &= C\exp\left(x^2\right), \quad \text{o bien,} \quad f_2\,(x,y,C)\equiv y-C\exp\left(x^2\right)=0 \end{aligned}$$

que son soluciones generales de las EDO (1,1), $p-2x=0$ *y* $p-2xy=0$, *respectivamente. Entonces, la primitiva completa es*

$$f_1\,(x,y,C)\,f_2\,(x,y,C)\equiv\left(y-x^2-C\right)\left(y-C\exp\left(x^2\right)\right)=0$$

1.4.2. Ecuación resuelta con respecto a y

Supongamos que la EDO de primer orden se resuelve con respecto a y,

$$y=f(x,p) \tag{1.50}$$

Derivando la ecuación (1.50) con respecto a x, hallamos la ecuación que, con respecto a p y dp/dx, es una EDO (1,1):

$$\frac{dy}{dx}=p=\frac{\partial f}{\partial x}+\frac{\partial f}{\partial p}\frac{dp}{dx} \quad \Rightarrow \quad \frac{dp}{dx}=F\,(x,p) \tag{1.51}$$

donde $F\,(x,p)=\left(p-\frac{\partial f}{\partial x}\right)/\frac{\partial f}{\partial p}$. Resuélvase la ecuación (1.51) obteniendo la primitiva $\phi(x,p,C)=0$. Después, obténgase la primitiva $\Phi(x,y,C)=0$ excluyendo p entre las dos ecuaciones,

$$y=f(x,p) \quad \text{y} \quad \phi(x,p,C)=0$$

cuando sea posible, o bien se quedan x y y expresadas como funciones del parámetro p.

1.4.3. Ecuación resuelta con respecto a x

Cuando la ecuación (1.46) se resuelve con respecto a x,

$$x=f(y,p) \tag{1.52}$$

derivando la ecuación (1.52) con respecto a y, se tiene una EDO (1,1) con respecto a p y dp/dy:

$$\frac{dx}{dy} = \frac{1}{p} = \frac{\partial f}{\partial y} + \frac{\partial f}{\partial p}\frac{dp}{dy} \quad \Rightarrow \quad \frac{dp}{dy} = F(y,p) \tag{1.53}$$

donde $F(x,p) = \left(\frac{1}{p} - \frac{\partial f}{\partial y}\right) / \frac{\partial f}{\partial p}$. Hallando la primitiva $\phi(y,p,C) = 0$ de la ecuación (1.53) y excluyendo p, cuando sea posible, de las dos ecuaciones,

$$x = f(y,p) \quad \text{y} \quad \phi(y,p,C) = 0$$

encontramos la primitiva $\Phi(x,y,C) = 0$.

Ejemplo 1.16. *Considérese la ecuación*

$$y = 3px + 6p^2y^2$$

Resolviendo con respecto a x, se tiene

$$3x = \frac{y}{p} - 6py^2$$

Luego, derivando con respecto a y,

$$(1 + 6p^2y)\left(2p + y\frac{dp}{dy}\right) = 0$$

Igualando a cero el segundo factor se deduce que $py^2 = C$. Despejando p y sustituyendo en la ecuación diferencial original se obtiene la primitiva $y^3 = 3Cx + 6C^2$. Nótese que al igualar a cero el primer factor se tiene solamente la ecuación algebraica, $1 + 6p^2y = 0$, con respecto a p. Despejando p y sustituyendo en la ecuación diferencial original se obtiene la solución singular que no contiene ninguna constante arbitraria, esto es, $24y^3 + 9x^2 = 0$.

1.4.4. Ecuación de Clairaut

La ecuación de la forma

$$y = px + f(p) \tag{1.54}$$

se conoce como **ecuación de Clairaut**. Su primitiva es

$$y = Cx + f(C) \tag{1.55}$$

que se comprueba de la siguiente manera. Derivando la ecuación (1.54), tenemos $p = p + x\frac{dp}{dx} + \frac{df}{dp}\frac{dp}{dx}$, de donde $\frac{dp}{dx}\left(x + \frac{df}{dp}\right) = 0$. Una solución es $\frac{dp}{dx} = 0$, de donde $p = C$. Obténgase la primitiva $y = Cx + f(C)$ sustituyendo $p = C$ en la ecuación (1.54). La ecuación $\frac{df(p)}{dp} = -x$ es algebraica y no contiene derivada de p con respecto a x. Despejando p como función de x y sustituyendo en la ecuación diferencial original se obtiene una solución singular que no contiene ninguna constante arbitraria.

Ejemplo 1.17. *La ecuación* $y = px + \sqrt{4+p^2}$ *tiene la primitiva* $y = Cx + \sqrt{4+C^2}$. *Despejando* p *de la ecuación* $-x = d\left(\sqrt{4+p^2}\right)/dp = p/\sqrt{4+p^2}$ *se tiene* $p = \pm 2x/\sqrt{1-x^2}$, *y sustituyendo en la ecuación diferencial original se obtiene la solución singular,* $y = 2\left(1 \pm x^2\right)/\sqrt{1-x^2}$.

Ejemplo 1.18. *La ecuación* $y = px + p^2$ *tiene la primitiva* $y = Cx + C^2$. *La solución singular es* $y = -x^2/4$.

1.5. EDO de orden n

1.5.1. EDO (n,1) lineales

Se dice que una ecuación diferencial de orden n es lineal si tiene la forma

$$P_0(x)y^{(n)} + P_1(x)y^{(n-1)} + \cdots + P_{n-1}(x)y^{(1)} + P_n(x)y = Q(x) \tag{1.56}$$

donde $P_0(x) \neq 0, P_1(x), \ldots, P_n(x)$, $Q(x)$ son funciones de x o constantes y $y^{(n)} = \frac{d^n y}{dx^n}$. Introduciendo el operador diferencial del orden n,

$$L^{(n)} = P_0(x)\frac{d^n}{dx^n} + P_1(x)\frac{d^{n-1}}{dx^{n-1}} + \cdots + P_n(x) \tag{1.57}$$

podemos escribir la ecuación (1.56) en forma compacta,

$$L^{(n)}[y] = Q(x) \tag{1.58}$$

La característica distintiva de esta ecuación es su linealidad con respecto a la función desconocida $y(x)$ y sus derivadas, es decir, la ecuación es de primer grado con respecto a éstas.

Si $Q(x) \equiv 0$, queda la ecuación

$$L^{(n)}[y] = 0 \tag{1.59}$$

que se llama **homogénea**; en el caso de $Q(x) \neq 0$, la ecuación (1.56) es **no homogénea**.

Ejemplo 1.19. *La ecuación diferencial*

$$x^2\frac{d^3y}{dx^3} - 3x\frac{d^2y}{dx^2} + 5\frac{dy}{dx} + xy = \cos x$$

es lineal de tercer orden no homogénea, mientras que la siguiente ecuación lineal es de segundo orden homogénea,

$$3x\frac{d^2y}{dx^2} - 5\frac{dy}{dx} + x^2y = 0$$

Las ecuaciones diferenciales lineales homogéneas poseen una importante propiedad, llamada el **principio de superposición**, que se expresa por medio del siguiente teorema.

Teorema 1.2. *Si $y_1(x)$, $y_2(x)$, ..., $y_m(x)$ son soluciones de una ecuación diferencial lineal homogénea, ecuación (1.59), la combinación lineal*

$$y = C_1 y_1 + C_2 y_2 + \cdots + C_m y_m$$

donde $C_1, C_2, \ldots, C_m$ son constantes, también es la solución de la ecuación (1.59).

Demostración. *Sean $y_1, y_2, \ldots, y_m$ soluciones de la ecuación (1.59), es decir,*

$$L^{(n)}[y_1] = 0, \quad L^{(n)}[y_2] = 0, \quad \ldots, \quad L^{(n)}[y_m] = 0$$

Es fácil ver que de cualquier solución $y(x)$ de la ecuación (1.59), multiplicada por cualquier constante C, resulta la solución de la misma ecuación:

$$L^{(n)}[Cy] = CL^{(n)}[y] = 0$$

Sea $y = C_1 y_1 + C_2 y_2 + \cdots + C_m y_m$. Entonces,

$$\begin{aligned} L^{(n)}[y] &= L^{(n)}[C_1 y_1 + C_2 y_2 + \cdots + C_m y_m] \\ &= L^{(n)}[C_1 y_1] + L^{(n)}[C_2 y_2] + \cdots + L^{(n)}[C_m y_m] \\ &= C_1 L^{(n)}[y_1] + C_2 L^{(n)}[y_2] + \cdots + C_m L^{(n)}[y_m] = 0 \end{aligned}$$

y queda demostrado el teorema. ■

Se dice que las funciones y_1, y_2, ... y_m son **linealmente independientes** sobre un intervalo $x_1 < x < x_2$ si para toda x de dicho intervalo, la igualdad

$$C_1 y_1(x) + C_2 y_2(x) + \cdots + C_m y_m(x) = 0 \tag{1.60}$$

se cumple sólo para las constantes $C_1 = C_2 = \cdots = C_m = 0$. En caso contrario las funciones se llaman **linealmente dependientes**.

Ejemplo 1.20. *Las funciones $y_1 = e^x$, $y_2 = e^{-x}$ son linealmente independientes. Para demostrarlo supóngase la relación*

$$C_1 e^x + C_2 e^{-x} = 0 \Longleftrightarrow C_1 y_1 + C_2 y_2 = 0$$

donde C_1 y C_2 son constantes. Al derivarla se obtiene

$$C_1 e^x - C_2 e^{-x} = 0 \Longleftrightarrow C_1 y_1' + C_2 y_2' = 0$$

Las dos ecuaciones forman el sistema homogéneo de ecuaciones lineales con respecto a C_1 y C_2. Resolviéndolo se tiene $C_1 = C_2 = 0$.

Ejemplo 1.21. *Las funciones* $y_1 = e^x$, $y_2 = e^{-x}$, $y_3 = 7e^x$ *son linealmente dependientes, ya que para* $C_1 = 7$, $C_2 = 0$, $C_3 = -1$ *se cumple la igualdad* $C_1e^x + C_2e^{-x} + 7C_3e^x = 0$ *para toda* x.

El concepto de independencia lineal de funciones tiene su importancia al aplicarlo a las soluciones de ecuaciones diferenciales lineales. El criterio de independencia lineal se establece en términos del **determinante de Wronski** o, simplemente, **wronskiano**:

$$W(y_1, y_2, \cdots, y_m) = \begin{vmatrix} y_1 & y_2 & \cdots & y_m \\ y_1' & y_2' & \cdots & y_m' \\ \vdots & \vdots & & \vdots \\ y_1^{(m-1)} & y_2^{(m-1)} & \cdots & y_m^{(m-1)} \end{vmatrix} \tag{1.61}$$

donde las funciones $y_1(x)$, $y_2(x)$, ..., $y_m(x)$ tienen derivadas continuas con respecto a x hasta el orden $(m-1)$.

Teorema 1.3. *(La condición necesaria y suficiente) Las funciones* $y_1(x)$, $y_2(x)$, ..., $y_m(x)$, *que tienen sus derivadas continuas hasta el orden* $(m-1)$, *son linealmente independientes sobre un intervalo* $x_1 < x < x_2$ *si y sólo si su wronskiano* $W(y_1, y_2, \ldots, y_m) \neq 0$ *en este intervalo. (Si* $W = 0$ *para un valor* $x = x_0$, *entonces* $W \equiv 0$ *sobre todo el intervalo).*

Demostración. *Supongamos que las funciones* $y_1(x)$, $y_2(x)$, ..., $y_m(x)$ *son linealmente dependientes, pues existe una combinación lineal tal que* $C_1y_1(x) + C_2y_2(x) + \cdots + C_my_m(x) = 0$. *Derivando a ésta* $(m-1)$ *veces consecutivamente se obtiene el sistema homogéneo de ecuaciones lineales con respecto a las constantes incógnitas* $C_1, C_2, \ldots, C_m$:

$$\begin{gathered} C_1y_1(x) + C_2y_2(x) + \cdots + C_my_m(x) = 0 \\ C_1y_1'(x) + C_2y_2'(x) + \cdots + C_my_m'(x) = 0 \\ \vdots \\ C_1y_1^{(m-1)}(x) + C_2y_2^{(m-1)}(x) + \cdots + C_my_m^{(m-1)}(x) = 0 \end{gathered}$$

Dicho sistema tiene una solución no trivial si y sólo si su determinante, que es precisamente el wronskiano, es igual a cero. Lo último contradice la condición del teorema de que $W(y_1, y_2, \ldots, y_m) \neq 0$. *Por tanto, la suposición de la dependencia lineal de las funciones* y_1, $y_2, \ldots, y_m$ *es errónea. Queda probado el teorema.* ■

Proposición 1.6. *La ecuación diferencial lineal homogénea de orden* n, *ecuación (1.59), tiene no más que* n *soluciones linealmente independientes. Si* $y_1(x)$, $y_2(x)$, ..., $y_n(x)$ *son* n *soluciones linealmente independientes de la ecuación (1.59), entonces, la primitiva de la ecuación (1.59) es*

$$y_0(x) = C_1y_1(x) + C_2y_2(x) + \cdots + C_ny_n(x) \tag{1.62}$$

donde $C_1, C_2, \ldots, C_n$ *son constantes arbitrarias.*

Un conjunto de n soluciones linealmente independientes $y_1(x)$, $y_2(x)$, ..., $y_n(x)$, se llama **base** o bien, **sistema fundamental de soluciones** de la ecuación (1.59).

Proposición 1.7. *Sea $y_p(x)$ cualquier solución particular de la ecuación diferencial lineal no homogénea, ecuación (1.56). Entonces,*

$$y(x) = C_1 y_1(x) + C_2 y_2(x) + \cdots + C_n y_n(x) + y_p(x) = y_0(x) + y_p(x) \tag{1.63}$$

es la primitiva de la ecuación (1.56). Es decir, la solución general de una ecuación diferencial lineal no homogénea es la suma de una solución particular cualquiera de la ecuación no homogénea y de la solución general de la ecuación diferencial lineal homogénea.

Aunque en general la primitiva de una ecuación diferencial no es necesariamente la solución completa de la ecuación, sin embargo, cuando la ecuación es lineal, la primitiva es su solución completa.

1.5.2. Ecuaciones lineales homogéneas con coeficientes constantes

Una ecuación diferencial lineal homogénea con coeficientes constantes $P_0 = 1$, $P_1, \ldots, P_n$ tiene la forma

$$y^{(n)} + P_1 y^{(n-1)} + \cdots + P_{n-1} y^{(1)} + P_n y = 0 \tag{1.64}$$

Mediante la notación $D \equiv \frac{d}{dx}$, la ecuación (1.64) toma la forma

$$\left(D^n + P_1 D^{n-1} + \cdots + P_{n-1} D + P_n\right)[y] = 0 \tag{1.65}$$

donde el operador diferencial $L^{(n)}(D) = D^n + P_1 D^{n-1} + \cdots + P_{n-1} D + P_n$ tiene la forma de un polinomio en la variable D. Según el teorema fundamental del álgebra, un polinomio de orden n tiene n raíces, sean $\lambda_1, \lambda_2, \ldots, \lambda_n$, algunas de ellas pueden ser múltiples. Esto implica que el polinomio puede ser factorizado, o bien presentado en la forma

$$\begin{aligned} L^{(n)}(D) &= D^n + P_1 D^{n-1} + \cdots + P_{n-1} D + P_n \\ &= (D - \lambda_1)(D - \lambda_2) \cdots (D - \lambda_n) \end{aligned}$$

Entonces la ecuación (1.65) toma la forma

$$(D - \lambda_1)(D - \lambda_2) \cdots (D - \lambda_n)[y] = 0 \tag{1.66}$$

Nótese que en el operador anterior el orden de los factores no tiene importancia, es decir, $(D - \lambda_k)(D - \lambda_m) = (D - \lambda_m)(D - \lambda_k)$.

Es fácil ver que la ecuación diferencial $(D - \lambda)[y] = 0$ tiene la solución $y = e^{\lambda x}$. Esto nos sugiere la solución de la ecuación (1.66), o bien, de la ecuación (1.65) en forma exponencial, $y = e^{\lambda x}$. Sustituyéndola en la ecuación (1.65), se tiene

$$e^{\lambda x}\left(\lambda^n + P_1\lambda^{n-1} + \cdots + P_{n-1}\lambda + P_n\right) = 0$$

que se cumple si

$$\lambda^n + P_1\lambda^{n-1} + \cdots + P_{n-1}\lambda + P_n = 0 \tag{1.67}$$

La última se conoce como **ecuación característica** de la ecuación (1.65); sus raíces o **valores característicos** $\lambda_1, \lambda_2, \ldots, \lambda_n$ permiten escribir la ecuación característica en la forma

$$(\lambda - \lambda_1)\ (\lambda - \lambda_2)\ \cdots\ (\lambda - \lambda_n) = 0$$

Del álgebra se sabe que la ecuación característica puede tener raíces de dos tipos:

Caso I. La ecuación característica tiene n raíces distintas $\lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \cdots \neq \lambda_n$. Entonces las n soluciones linealmente independientes

$$y_1 = e^{\lambda_1 x}, \quad y_2 = e^{\lambda_2 x}, \quad \ldots, \quad y_n = e^{\lambda_n x}$$

constituyen una base y la solución general de la ecuación (1.65), o bien de la ecuación (1.64) es

$$y_0 = C_1 e^{\lambda_1 x} + C_2 e^{\lambda_2 x} + \cdots + C_n e^{\lambda_n x} \tag{1.68}$$

Caso II. Algunos de los valores característicos son múltiples. Digamos que λ_1 es una raíz múltiple de grado k y las demás son distintas, $\lambda_1 \neq \lambda_{k+1} \neq \cdots \neq \lambda_n$. En este caso, las k soluciones linealmente independientes que corresponden al valor característico λ_1 son

$$y_1 = e^{\lambda_1 x}, \quad y_2 = x e^{\lambda_1 x}, \quad y_3 = x^2 e^{\lambda_1 x}, \quad \ldots, \quad y_k = x^{k-1} e^{\lambda_1 x}$$

que se comprueba notando que

$$L^{(n)}(D) = (D - \lambda_1)^k (D - \lambda_{k+1}) \cdots (D - \lambda_n)$$

y, además,

$$\begin{aligned}
(D - \lambda_1)\left[x^m e^{\lambda_1 x}\right] &= m x^{m-1} e^{\lambda_1 x}, \\
(D - \lambda_1)^2\left[x^m e^{\lambda_1 x}\right] &= m(m-1)\, x^{m-2} e^{\lambda_1 x} \\
&\ \ \vdots \\
(D - \lambda_1)^m\left[x^m e^{\lambda_1 x}\right] &= m! e^{\lambda_1 x} \\
(D - \lambda_1)^{m+1}\left[x^m e^{\lambda_1 x}\right] &= 0
\end{aligned}$$

De la última relación se sigue que

$$(D-\lambda_1)^k\left[x^m e^{\lambda_1 x}\right]=0 \quad \text{si} \quad m=0,1,\ldots,k-1$$

Soluciones linealmente independientes correspondientes a los otros valores característicos son

$$y_{k+1}=e^{\lambda_{k+1}x}, \quad \ldots, \quad y_n=e^{\lambda_n x}$$

y, por tanto, la solución general de la ecuación (1.65), o bien de la ecuación (1.64) es

$$y_0=C_1e^{\lambda_1 x}+C_2xe^{\lambda_1 x}+\cdots+C_kx^{k-1}e^{\lambda_1 x}+C_{k+1}e^{\lambda_{k+1}x}+\cdots+C_ne^{\lambda_n x} \quad (1.69)$$

Ejemplo 1.22. *En caso de una ecuación diferencial lineal de segundo orden con coeficientes constantes reales*

$$y''+2ay'+by=0$$

la ecuación característica es

$$\lambda^2+2a\lambda+b=0$$

y sus raíces son

$$\lambda_1=-a+\sqrt{a^2-b}, \quad \lambda_2=-a-\sqrt{a^2-b}$$

a) Raíces reales distintas, $\lambda_1\neq\lambda_2$: la solución general es

$$y_0=C_1e^{\lambda_1 x}+C_2e^{\lambda_2 x}$$

b) Raíces complejas conjugadas, $\lambda_1=-a+i\beta$, $\lambda_2=-a-i\beta$, donde $\beta=\sqrt{b-a^2}$. Para este caso,

$$y_1=e^{\lambda_1 x}=e^{-\alpha x}e^{i\beta x}, \qquad y_2=e^{\lambda_2 x}=e^{-\alpha x}e^{-i\beta x}$$

y tomando $C_2=C_1^*$ la solución general real es

$$y_0=C_1e^{\lambda_1 x}+C_1^*e^{\lambda_2 x}=Ae^{-\alpha x}\cos(\beta x+\phi)$$

Se queda como tarea el cálculo de las constantes arbitrarias A y ϕ en función de C_1.

c) Una raíz real doble, $\lambda=-a$, $(a^2=b)$: la solución general es

$$y_0=C_1e^{\lambda x}+C_2xe^{\lambda x}=C_1e^{-ax}+C_2xe^{-ax}$$

1.5.3. EDO (n,1) lineales no homogéneas

La solución general $y(x)$ de una ecuación diferencial lineal

$$L^{(n)}[y] = Q(x) \tag{1.70}$$

es la suma $y(x) = y_0 + y_p$ de una solución particular y_p de esta ecuación y de la solución general y_0 de la ecuación homogénea correspondiente. Se conocen dos métodos para obtener y_p. El **método de variación de parámetros** es completamente general. El **método de los coeficientes indeterminados**, siendo un método especial y más sencillo, sin embargo es práctico para aplicaciones en problemas de ingeniería.

Coeficientes indeterminados

Este método es aplicable a ecuaciones lineales con coeficientes constantes cuando $Q(x)$ tiene la forma

$$Q(x) = \left(Ae^{\alpha x} \pm Be^{\beta x}\right) P_n(x)$$

en donde A y B son algunas constantes y

$$P_n(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n$$

Para obtener una solución particular de la ecuación (1.70) se supone y_p de la forma semejante a la de $Q(x)$ y sus derivadas, conteniendo coeficientes desconocidos que se determinan sustituyendo y_p y sus derivadas en la ecuación (1.70). A continuación se presentan ejemplos ilustrativos.

Ejemplo 1.23. $y''' + ay = 3x$. *Suponiendo que* $y_p = Ax + B = A_1Q(x) + B_1Q'(x)$ *y sustituyendo en la ecuación diferencial, se tiene* $a(Ax + B) = 3x$, *de donde* $a(Ax) = 3x$, *pues* $A = \frac{3}{a}$; *también,* $aB = 0$, *y* $B = 0$. *Así,* $y_p = \frac{3}{a}x$.

Ejemplo 1.24. $y''' + ay = e^{bx}$. *Intentaremos* $y_p = Ae^{bx}$, *por lo que* $Ab^3e^{bx} + aAe^{bx} = e^{bx}$, *luego* $Ab^3 + Aa = 1$, *de donde* $A = \left(a + b^3\right)^{-1}$, *obteniéndose* $y_p = \left(a + b^3\right)^{-1} e^{bx}$.

Ejemplo 1.25. $y'' - 2y' + 3y = x^3 + \operatorname{sen} x$. *Suponemos que* $y_p = A_3x^3 + A_2x^2 + A_1x + A_0 + B\operatorname{sen} x + C\cos x$. *Sustituyendo, se obtiene*

$$\begin{aligned} x^3 + \operatorname{sen} x &= 6A_3x + 2A_2 - B\operatorname{sen} x - C\cos x - \\ &\quad 2\left(3A_3x^2 + 2A_2x + A_1 + B\cos x - C\operatorname{sen} x\right) + \\ &\quad 3\left(A_3x^3 + A_2x^2 + A_1x + A_0 + B\operatorname{sen} x + C\cos x\right) \end{aligned}$$

Igualando los coeficientes de $\cos x$ *y* $\operatorname{sen} x$, *y de potencias de* x, *se tiene*

$$\begin{aligned} -C - 2B + 3C &= 0, \quad -B + 2C + 3B = 1 \\ 3A_3 &= 1, \quad -6A_3 + 3A_2 = 0 \\ 6A_3 - 4A_2 + 3A_1 &= 0, \quad 2A_2 - 2A_1 + 3A_0 = 0 \end{aligned}$$

De donde, $A_3 = 1/3$, $A_2 = 2/3$, $A_1 = 2/9$, $A_0 = -8/27$, $B = C = 1/4$ *y, por lo tanto,*

$$y_p = \frac{1}{3}x^3 + \frac{2}{3}x^2 + \frac{2}{9}x - \frac{8}{27} + \frac{1}{4}\operatorname{sen} x + \frac{1}{4}\cos x$$

Solución por variación de parámetros

Para entender mejor la idea del método empezaremos con ecuaciones diferenciales lineales de segundo orden de la forma

$$y'' + A(x)\, y' + B(x)\, y = Q(x) \tag{1.71}$$

La ecuación diferencial homogénea correspondiente

$$y'' + A(x)\, y' + B(x)\, y = 0 \tag{1.72}$$

tiene una solución general $y_0(x)$ de la forma

$$y_0(x) = C_1 y_1(x) + C_2 y_2(x)$$

donde $y_1(x)$ y $y_2(x)$ son dos soluciones linealmente independientes de la ecuación (1.72), y C_1 y C_2 son dos constantes arbitrarias.

El método de variación de parámetros consiste en reemplazar las constantes C_1 y C_2 por las funciones $C_1(x)$ y $C_2(x)$ que deben determinarse de modo que la función resultante

$$y_p(x) = C_1(x)\, y_1(x) + C_2(x)\, y_2(x) \tag{1.73}$$

sea una solución particular de la ecuación (1.71). Derivando se tiene

$$y_p' = C_1 y_1' + C_2 y_2' + C_1' y_1 + C_2' y_2$$

La $y_p(x)$ contiene dos funciones $C_1(x)$ y $C_2(x)$, pero la condición de que y_p satisfaga la ecuación (1.71) impone sólo una relación entre ellas. Teniendo la libertad de imponer una segunda condición, suponemos que

$$C_1' y_1 + C_2' y_2 = 0 \tag{1.74}$$

Entonces,

$$y_p' = C_1 y_1' + C_2 y_2' \tag{1.75}$$

$$y_p'' = C_1 y_1'' + C_2 y_2'' + C_1' y_1' + C_2' y_2' \tag{1.76}$$

Sustituyendo (1.75) y (1.76) en la ecuación (1.71), se obtiene

$$\begin{aligned} &(C_1 y_1'' + C_2 y_2'') + C_1' y_1' + C_2' y_2' + A(x)(C_1 y_1' + C_2 y_2') + \\ &B(x)(C_1 y_1 + C_2 y_2) \\ = \;& Q(x) \end{aligned}$$

Como $y_1(x)$ y $y_2(x)$ son soluciones de la ecuación homogénea, ecuación (1.72), esta relación junto con la ecuación (1.74) se reduce al sistema algebraico de dos ecuaciones lineales con respecto a dos funciones incógnitas C_1' y C_2',

$$\begin{aligned} C_1'y_1 + C_2'y_2 &= 0 \\ C_1'y_1' + C_2'y_2' &= Q(x) \end{aligned}$$

en donde el determinante del sistema es el wronskiano de y_1 y y_2, $W(y_1, y_2) = y_1y_2' - y_1'y_2 \neq 0$. Por la regla de Cramer se halla

$$C_1' = -\frac{y_2Q}{W}, \quad C_2' = \frac{y_1Q}{W}$$

Integrando

$$C_1(x) = -\int \frac{y_2Q}{W}dx, \quad C_2(x) = \int \frac{y_1Q}{W}dx$$

y sustituyendo estas expresiones en la ecuación (1.73), se obtiene la solución particular deseada de la ecuación (1.71),

$$y_p(x) = -y_1(x)\int \frac{y_2Q}{W}dx + y_2(x)\int \frac{y_1Q}{W}dx \tag{1.77}$$

Si se dejan las constantes arbitrarias de integración que aparecen en la ecuación (1.77), entonces ésta se convierte en una solución general de la ecuación (1.71).

La generalización del método de variación de parámetros para una ecuación diferencial lineal no homogénea de orden n, ecuación (1.70), es bastante directa. La solución particular se busca en la forma

$$y_p(x) = C_1(x)y_1(x) + C_2(x)y_2(x) + \cdots + C_n(x)y_n(x)$$

donde $y_1(x), y_2, \ldots, y_n$ son soluciones fundamentales de la ecuación homogénea correspondiente. La $y_p(x)$ contiene n funciones $C_1(x)$, $C_2(x), \ldots, C_n(x)$ desconocidas, pero la condición de que y_p satisfaga la ecuación (1.70) impone sólo una relación entre ellas. Teniendo la libertad de imponer $(n-1)$ condiciones adicionales, suponemos que

$$\begin{gathered} C_1'y_1 + C_2'y_2 + \cdots + C_n'y_n = 0 \\ C_1'y_1' + C_2'y_2' + \cdots + C_n'y_n' = 0 \\ C_1'y_1'' + C_2'y_2'' + \cdots + C_n'y_n'' = 0 \\ \vdots \\ C_1'y_1^{(n-2)} + C_2'y_2^{(n-2)} + \cdots + C_n'y_n^{(n-2)} = 0 \end{gathered}$$

Junto con la relación

$$C_1'y_1^{(n-1)} + C_2'y_2^{(n-1)} + \cdots + C_n'y_n^{(n-1)} = Q(x)$$

que se obtiene al sustituir y_p en la ecuación (1.70), éstas constituyen un sistema algebraico de n ecuaciones lineales con respecto a C_1', C_2', ..., C_n', que es resuelto por la regla de Cramer, dado que su determinante es wronskiano $W(y_1, y_2, \ldots, y_n) \neq 0$. Luego, integrando se obtienen $C_1(x)$, $C_2(x)$, ..., $C_n(x)$ y finalmente $y_p(x)$.

Ejercicio. *Obtener las primitivas de las ecuaciones*

$$\begin{aligned}(D^2 - 6D + 9)y &= x^{-2}e^{3x} \\ (D^2 - 2D + 3)y &= x^3 + \operatorname{sen} x\end{aligned}$$

por variación de parámetros.

1.6. Problemas con valores frontera e iniciales

Cada ecuación diferencial ordinaria tiene sólo una variable independiente, la cual, a menudo, es el tiempo. En muchas aplicaciones se tiene interés en una solución particular $y(t)$ de una ecuación diferencial dada. La solución general $y(t; C_1, C_2, \ldots, C_n)$ de una ecuación diferencial ordinaria de orden n contiene un número n de constantes arbitrarias, C_1, C_2, ..., C_n. Por tanto, para definir una solución particular de esta ecuación hay que dar ciertos valores a estas constantes. En la práctica los valores de n constantes arbitrarias se determinan de n condiciones que se conocen como **condiciones iniciales**, dado que éstas prescriben la situación inicial a un problema físico en un instante, digamos $t = 0$. En muchas aplicaciones son n condiciones del tipo

$$\begin{aligned}y(t; C_1, C_2, \ldots, C_n)|_{t=0} &= K_0, \quad y'(t; C_1, C_2, \ldots, C_n)|_{t=0} = K_1, \ldots, \\ y^{(n-1)}(t; C_1, C_2, \ldots C_n)|_{t=0} &= K_{n-1}\end{aligned}$$

donde K_0, K_2, ..., K_{n-1} son números dados. Una ecuación diferencial ordinaria junto con sus condiciones iniciales se llama **problema con valor inicial**.

Las ecuaciones diferenciales parciales (en derivadas parciales) tienen por lo menos dos variables independientes. Por razones físicas, un conjunto de variables independientes suele separarse en una variable temporal, el tiempo t, y en un grupo de variables espaciales, por ejemplo x, y, z. Dicha diferencia entre las ecuaciones ordinarias y parciales da origen a dos tipos de **condiciones, iniciales** y **en la frontera**, para definir una solución particular de la ecuación diferencial parcial.

Una ecuación muy conocida es la ecuación de onda que describe la propagación de una perturbación (una onda) en un medio unidimensional, por ejemplo en una cuerda elástica:

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad t > 0, \quad 0 < x < L$$

En esta ecuación se supone que la cuerda de longitud L está ubicada sobre el eje x de un sistema de coordenadas con un extremo en el punto $x = 0$ y con el otro en $x = L$. La solución $u(x,t)$ describe la propagación de onda a partir del instante $t = 0$. Para definir una solución particular se requieren condiciones iniciales y condiciones en la frontera. Siendo una ecuación de segundo orden con respecto a la derivada temporal, hay que aplicar dos condiciones iniciales que son los valores de la solución y de su primera derivada temporal en todo el intervalo espacial:

$$u(x,0) = a_1(x), \quad \frac{\partial u}{\partial t}|_{t=0} = a_2(x), \quad 0 < x < L$$

donde $a_1(x)$ y $a_2(x)$ son funciones dadas en el intervalo $0 < x < L$. Dado que la ecuación de onda es de segundo orden con respecto a la derivada espacial, hay que satisfacer dos condiciones en la frontera, por ejemplo,

$$u(0,t) = b_1(t), \quad u(L,t) = b_2(t), \quad t > 0$$

que son valores de la onda en los extremos de la cuerda, donde $b_1(t)$ y $b_2(t)$ son funciones dadas de tiempo. Siendo siempre los dos para la ecuación de onda, condiciones de la frontera dependiendo de las condiciones físicas que los determinan, pueden tener una forma diferente a los escritos anteriormente. Por ejemplo, podrían ser una combinación lineal de la función y su primera derivada espacial en los dos extremos de la cuerda: $\alpha_1 u(0,t) + \beta_1 \frac{\partial u(x,t)}{\partial x}|_{x=0} = b_1(t), \quad \alpha_2 u(L,t) + \beta_2 \frac{\partial u(x,t)}{\partial x}|_{x=L} = b_2(t), \quad t > 0$. Una ecuación diferencial parcial junto con sus condiciones iniciales y en la frontera se llama **problema con condiciones frontera**.

1.7. Reducción de un sistema general a un sistema normal

Los teoremas de existencia y unicidad de soluciones se formulan, por lo general, en términos de sistemas normales de ecuaciones diferenciales. Un sistema general de EDO es muy común tanto en modelos matemáticos de ciencias naturales y exactas como en métodos numéricos. Dada la importancia de los sistemas de EDO, en esta sección se considera la reducción de una EDO o un sistema general de EDO a un sistema normal.

Un sistema de EDO en forma

$$u_i' = f_i(x, u_1, u_2, \ldots, u_n), \quad i = 1, 2, \ldots, n \tag{1.78}$$

se llama **sistema normal**. En este sistema x es la variable independiente, $u_1, u_2, \ldots, u_n$ son funciones incógnitas de la variable x, $u_i' = du_i/dx$, $f_1, f_2, \ldots, f_n$ son funciones dadas de las variables $x, u_1, u_2, \ldots, u_n$ en un dominio abierto del espacio de $(n+1)$ dimensiones. Se supone que las funciones $f_1, f_2, \ldots, f_n$ son continuas junto con sus derivadas parciales,

$$\frac{\partial f_i(x, u_1, u_2, \ldots, u_n)}{\partial u_j}, \quad i, j = 1, 2, \ldots, n$$

en el dominio de su definición. Un caso particular de sistemas normales es un sistema normal de EDO lineales:

$$u_i' = \sum_{j=1}^{n} a_{ij}(x)\, u_j + b_i(x), \quad i = 1, 2, \ldots, n \tag{1.79}$$

o en forma vectorial,

$$\frac{d\mathbf{u}}{dx} = A(x)\,\mathbf{u} + \mathbf{b}(x) \tag{1.80}$$

en donde la matriz $A(x) = [a_{ij}(x)]$ y el vector $\mathbf{b}(x)$ son funciones dadas de la variable independiente x.

En la práctica surgen ecuaciones y sistemas de EDO más generales que lo escrito en párrafos anteriores. En el caso de una sola función incógnita $y(x)$ de una variable independiente x, la ecuación diferencial de orden n puede escribirse en la forma

$$F\left(x, y, y', \ldots, y^{(n)}\right) = 0 \tag{1.81}$$

en donde F es una función dada de $(n+2)$ variables que está definida en un dominio abierto del espacio de sus variables. Una función $y = \varphi(x)$ definida en un intervalo $x_1 < x < x_2$, que tiene las derivadas hasta el orden n, es la solución de la ecuación (1.81) si al sustituirla en ésta se obtiene la identidad en la variable x en el intervalo $x_1 < x < x_2$.

En el caso de dos funciones incógnitas, $y(x)$ y $z(x)$, de una sola variable independiente x, que no son independientes, se consideran dos ecuaciones diferenciales que forman un sistema. Este sistema se puede escribir en la forma

$$\begin{aligned} F\left(x, y, y', \ldots, y^{(n)}, z, z', \ldots, z^{(m)}\right) = 0 \\ G\left(x, y, y', \ldots, y^{(n)}, z, z', \ldots, z^{(m)}\right) = 0 \end{aligned} \tag{1.82}$$

en donde F y G son funciones dadas, cada una $(n+m+3)$ variables, definidas en un dominio abierto. Si el orden de la derivada mayor de la función $y(x)$ es n y el orden de la derivada mayor de $z(x)$ es m, se dice que el sistema (1.82) es de orden n con respecto a $y(x)$ y de orden m con respecto a $z(x)$. El número $(n+m)$ es el orden del sistema (1.82). La solución de (1.82) es un par de funciones $y = \varphi(x)$ y $z = \psi(x)$, definidas en un intervalo $x_1 < x < x_2$, tales que al sustituirlas en (1.82) se obtienen dos identidades en x en el intervalo $x_1 < x < x_2$.

De modo similar se define un sistema de EDO de tres y más funciones incógnitas de una sola variable. Sean $y_1, y_2, \ldots, y_n$ funciones incógnitas y sea m_i el orden de derivada mayor de y_i, $i = 1, 2, \ldots, n$. Entonces, se dice que m_i es el orden del sistema con respecto a y_i y $m = m_1 + m_2 + \cdots + m_n$ es el orden del sistema. Por tanto, el sistema normal (1.78) tiene el orden n.

Si la ecuación (1.81) puede ser despejada con respecto a la derivada mayor, entonces esta ecuación se escribe en la forma

$$y^{(n)} = f\left(x, y, y', \ldots, y^{(n-1)}\right) \tag{1.83}$$

De manera similar, si las ecuaciones del sistema (1.82) pueden ser resueltas con respecto a las derivadas mayores, entonces este sistema se escribe en la forma

$$\begin{aligned} y^{(n)} &= f\left(x, y, y', \ldots, y^{(n-1)}, z, z', \ldots, z^{(m-1)}\right) \\ z^{(m)} &= g\left(x, y, y', \ldots, y^{(n-1)}, z, z', \ldots, z^{(m-1)}\right) \end{aligned} \tag{1.84}$$

Las ecuaciones (1.83) y (1.84) se llaman **resueltas con respecto a las derivadas mayores**. De manera similar se definen sistemas de ecuaciones con un número arbitrario de funciones incógnitas que son resueltas con respecto a las derivadas mayores. En particular, el sistema normal, ecuación (1.78), es resuelto con respecto a las derivadas mayores.

En lo que sigue se demuestra que cualquier sistema de EDO de orden n resuelto con respecto a las derivadas mayores es reducible a un sistema normal de orden n. Para empezar, mostramos que una ecuación diferencial de orden n se reduce a un sistema normal de orden n. La ecuación (1.83) es una ecuación diferencial de orden n resuelta con respecto a la derivada mayor, en donde $f\left(x, y, y', \ldots, y^{(n-1)}\right)$ es una función dada de $(n+1)$ variables en un dominio abierto. Se supone que tanto $f\left(x, y, y', \ldots, y^{(n-1)}\right)$ como sus derivadas parciales $\partial f/\partial y^{(k)}$, $k = 0, 1, \ldots, n-1$, son continuas en el dominio de definición de f. Con el fin de transformar la ecuación (1.83) en un sistema normal se introducen nuevas funciones incógnitas, $u_1, u_2, \ldots, u_n$, de la variable independiente x:

$$u_1 = y, \quad u_2 = y', \quad \ldots, \quad u_n = y^{(n-1)} \tag{1.85}$$

Se tiene que la ecuación (1.83) es equivalente al siguiente sistema normal:

$$\begin{aligned} u_1' &= u_2 \\ u_2' &= u_3 \\ &\vdots \\ u_{n-1}' &= u_n \\ u_n' &= f\left(x, u_1, u_2, \ldots, u_n\right) \end{aligned} \tag{1.86}$$

El método descrito en el párrafo anterior permite reducir a la forma normal cualquier sistema de EDO resuelto con respecto a las derivadas mayores. Con el fin de no complicar las expresiones, consideremos el sistema de cuarto orden constituido por dos ecuaciones de segundo orden:

$$\begin{aligned} y'' &= f\left(x, y, y', z, z'\right) \\ z'' &= g\left(x, y, y', z, z'\right) \end{aligned} \tag{1.87}$$

Dicho sistema se reduce a la forma normal por medio del cambio a las nuevas funciones incógnitas:

$$u_1 = y, \quad u_2 = y', \quad u_3 = z, \quad u_4 = z' \tag{1.88}$$

Finalmente, se tiene que el sistema (1.87) es equivalente al siguiente sistema normal:

$$\begin{aligned} u_1' &= u_2 \\ u_2' &= f\left(x, u_1, u_2, u_3, u_4\right) \\ u_3' &= u_4 \\ u_4' &= g\left(x, u_1, u_2, u_3, u_4\right) \end{aligned} \tag{1.89}$$

Nota bibliográfica: para ampliar los conocimientos sobre el tema de este capítulo, se puede consultar [2], [4], [5] y [11].

Capítulo 2

Matrices

2.1. Definición y origen de las matrices

Tanto en las matemáticas puras como en las aplicadas surgen múltiples situaciones en las que tenemos que tratar con arreglos rectangulares de números o funciones. Algunos de estos ejemplos son los siguientes.

Ejemplo 2.1. *Un número complejo $x + iy$, donde x y y son números reales e i es una unidad imaginaria, se define completamente por el par ordenado (x, y) de números reales, es decir, por el arreglo de 1×2 elementos.*

Ejemplo 2.2. *Las coordenadas cartesianas de un punto en el espacio tridimensional constituyen un triplete ordenado (x, y, z) de números reales x, y y z.*

Ejemplo 2.3. *La forma cuadrática*

$$ax^2 + by^2 + cz^2 + 2hxy + 2gxz + 2fyz$$

está completamente determinada por el arreglo simétrico de 3×3 de números

$$\begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix}$$

Ejemplo 2.4. *El sistema de ecuaciones lineales*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= c_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= c_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= c_3 \end{aligned}$$

está completamente determinado por el arreglo de 3×4 de coeficientes

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & c_1 \\ a_{21} & a_{22} & a_{23} & c_2 \\ a_{31} & a_{32} & a_{33} & c_3 \end{bmatrix}$$

dado que toda la información sobre las soluciones de este sistema de ecuaciones lineales se puede expresar en términos de propiedades del arreglo.

Ejemplo 2.5. *La transformación lineal*

$$\begin{aligned} u &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ v &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \end{aligned}$$

que mapea los puntos (x_1, x_2, x_3) del espacio tridimensional en los puntos (u, v) de un plano (espacio bidimensional), se determina completamente por el arreglo de 2×3 de números

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

Dicha transformación lineal se puede estudiar en términos del álgebra de arreglos rectangulares correspondientes.

La generalización de los ejemplos presentados anteriormente se da por medio del concepto de matriz.

Definición 2.1. *Una disposición (un arreglo) rectangular*

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \tag{2.1}$$

de $m \times n$ elementos tomados de un conjunto de elementos S y ordenados en m renglones y n columnas se llama **matriz** *de orden $m \times n$ sobre el conjunto S, si uno quiere subrayar la naturaleza de los elementos de la matriz A. Los elementos del conjunto S se llaman escalares. Si el conjunto S representa todos los números complejos o reales, las matrices definidas sobre estos conjuntos se llaman matrices complejas o reales, respectivamente.*

Cuando $m = n$, la matriz se llama **cuadrada** de orden n y se dice que los elementos $a_{11}, a_{22}, \ldots, a_{nn}$ constituyen la diagonal principal de A, mientras que los elementos $a_{n1}, a_{n-1,2}, \ldots, a_{1n}$ constituyen la diagonal secundaria. Se emplean paréntesis rectangulares o barras dobles para representar matrices, es decir, $A = [a_{ij}] = ||a_{ij}||$, donde los subíndices i y j del elemento a_{ij} de la matriz A identifican, respectivamente, el renglón y la columna en los que está situado a_{ij}. Casos particulares de matrices son el **vector columna** A, es decir, la matriz de orden $m \times 1$

$$A = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} \equiv \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}$$

y el **vector renglón** B, que es la matriz de orden $1 \times n$

$$B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \end{bmatrix}$$

Ejercicio. *Dada la matriz*

$$A = \begin{bmatrix} 1 & 5i & 7 \\ 3i & -1 & -6 \end{bmatrix}$$

a) ¿Cuáles son las dimensiones de A?; b) ¿a qué equivalen los elementos a_{12}, a_{21}, a_{22}?

Ejercicio. *Escribir en forma explícita la matriz $A = [a_{ij}]$ determinada por las siguientes condiciones: a) $m = 3$, $n = 2$, $a_{ij} = |2i - j|$; b) $m = n = 3$, $a_{ij} = |i - j|$.*

2.2. Operaciones básicas

En esta sección introducimos operaciones fundamentales sobre matrices, que son la adición de matrices, la multiplicación por un escalar, la multiplicación de matrices y la matriz inversa (de la izquierda, de la derecha) y la derivada de una matriz de funciones. Ante todo empezamos por la igualdad de matrices.

2.2.1. Igualdad de matrices

Definición 2.2. *Se dice que dos matrices $A = [a_{ij}]$ y $B = [b_{ij}]$ son iguales, $A = B$, si y sólo si tienen el mismo orden y todos sus elementos correspondientes son iguales, $a_{ij} = b_{ij}$.*

De esta definición se siguen cuatro propiedades de la **igualdad de matrices**:

1) Si A y B son dos matrices cualesquiera, se tiene que $A = B$ o $A \neq B$ (propiedad determinativa).

2) Si A es una matriz cualquiera, entonces $A = A$ (propiedad reflexiva).

3) Si $A = B$, entonces $B = A$ (propiedad simétrica).

4) Si $A = B$ y $B = C$, entonces $A = C$ (propiedad transitiva).

Nótese que cualquier relación entre pares de objetos matemáticos que posee estas propiedades se llama relación de equivalencia. La igualdad de dos matrices significa que éstas de hecho son idénticas y, por tanto, una matriz puede ser sustituida por cualquier matriz igual en las operaciones sobre matrices.

2.2.2. Adición de matrices

Definición 2.3. *La **suma**, $A + B$, de dos matrices $A = [a_{ij}]$ y $B = [b_{ij}]$ del mismo orden $m \times n$ es la matriz $C = [c_{ij}]$ de $m \times n$ con elementos $c_{ij} = a_{ij} + b_{ij}$.*

Se dice que dos matrices del mismo orden son conformables para la suma. Puesto que la suma de dos matrices $m \times n$ cualesquiera de nuevo es una matriz $m \times n$, se dice que el conjunto de todas las matrices $m \times n$ es cerrado con respecto a la adición.

Definición 2.4. *Una matriz cuyos elementos son todos ceros se llama **matriz cero** y se denota por O u $O_{(m,n)}$, cuando es necesario destacar el orden.*

La propiedad básica de la matriz $O_{(m,n)}$ es que para cualquier matriz A de orden $m \times n$, se cumple la igualdad

$$A_{(m,n)} + O_{(m,n)} = A_{(m,n)}$$

pues la matriz cero es un elemento idéntico con respecto a la adición.

Definición 2.5. *La* ***negativa*** *o* ***negación*** *de una matriz A, denominada por $-A$, es la matriz del mismo orden que satisfaga la igualdad*

$$A + (-A) = O$$

es decir, $-A$ es la matriz obtenida al multiplicar cada elemento de A por -1.

Dado que los elementos de las matrices son escalares, los cuales satisfacen la ley conmutativa y la asociativa de la adición, es fácil comprobar, basándonos en la definición de la suma de matrices, el siguiente teorema.

Teorema 2.1. *La adición de matrices es conmutativa y asociativa, esto es, si A, B y C son matrices del mismo orden, entonces*

$$A + B = B + A \tag{2.2}$$

$$A + (B + C) = (A + B) + C \equiv A + B + C \tag{2.3}$$

y además

$$A + O = A \tag{2.4}$$

$$A + (-A) = O \tag{2.5}$$

Definición 2.6. *El* ***producto*** *de una matriz $A = [a_{ij}]$ de orden $m \times n$ por un número (escalar) k, que se denota mediante kA, es la matriz $C = [c_{ij}]$ de orden $m \times n$ cuyos elementos son $c_{ij} = ka_{ij}$, es decir, que multiplicar una matriz por un número significa multiplicar todos los elementos de la matriz por este número.*

Basándose en las definiciones, se ve que para matrices cualesquiera de $m \times n$ y números cualesquiera, se tiene

$$k\,(A + B) = kA + kB \tag{2.6}$$

$$(k + q)\,A = kA + qA \tag{2.7}$$

$$k\,(qA) = (kq)\,A \tag{2.8}$$

$$1A = A \tag{2.9}$$

Nótese que $(-1)\,A = -A$, la negativa de A.

Las fórmulas, ecuaciones (2.2), (2.3), (2.4) y (2.5), así como el sistema de (2.6), (2.7), (2.8) y (2.9), determinan un espacio vectorial (véase capítulo 5).

Teorema 2.2. ***(Espacios vectoriales de matrices)*** *Todas las matrices reales (o complejas) del mismo orden $m \times n$ forman un espacio vectorial real (o complejo) de dimensión mn. Una base consta de todas las mn matrices $E(r,c) = [e_{ij}(r,c)]$ $(r = 1, 2, \ldots, m;\ c = 1, 2, \ldots, n)$, en donde $e_{ij}(r,c) = \delta_{ir}\delta_{jc}$, pues el elemento en el r-ésimo renglón y en la c-ésima columna es igual a 1, $e_{rc}(r,c) = 1$, y los demás elementos son ceros.*

Problema. *Describir la base del espacio lineal de matrices de a) 3×1; b) 2×2 y c) 2×3.*

Problema. *Por definición, una matriz positiva es una matriz real cuadrada con todos sus elementos positivos. ¿Forman un espacio vectorial las matrices positivas de $m \times m$?*

2.2.3. Transpuesta y conjugada de una matriz. Matrices (anti)simétrica y (anti)hermitiana

En esta sección se definen otras operaciones sencillas sobre matrices y se presentan ciertos tipos de matrices especiales que tienen importancia en la práctica.

Definición 2.7. *La **transpuesta** A^T de una matriz $A = [a_{ij}]$ de $m \times n$ es la matriz de $n \times m$ que se obtiene al intercambiar los renglones y las columnas en A, esto es, el i-ésimo renglón de A se convierte en la i-ésima columna de A^T: $a_{ij}^T = a_{ji}$.*

Definición 2.8. *La **transpuesta** o **conjugada hermitiana** A^+ de una matriz $A = [a_{ij}]$ de $m \times n$ es la matriz de $n \times m$ que se obtiene al tomar el complejo conjugado de todos los elementos en la transpuesta A^T, esto es, $a_{ij}^+ = a_{ji}^*$.*

Por ser conmutativas las operaciones de transpuesta y de conjugación compleja, la transpuesta hermitiana de una matriz A puede ser obtenida, primero, tomando la conjugada compleja A^* de A y, después, la transpuesta $(A^*)^T$, es decir, $A^+ = (A^T)^* = (A^*)^T$. Nótese que para las matrices reales la transpuesta y la transpuesta hermitiana son iguales.

Se queda de tarea para los estudiantes probar las siguientes propiedades:

$$\left(A^T\right)^T = A, \quad \left(A^+\right)^+ = A \tag{2.10}$$

$$(kA)^T = kA^T, \quad (kA)^+ = k^* A^+ \tag{2.11}$$

$$(A \pm B)^T = A^T \pm B^T, \quad (A \pm B)^+ = A^+ \pm B^+ \tag{2.12}$$

Ejemplo 2.6. *La transpuesta A^T y conjugada hermitiana A^+ de una matriz A son, respectivamente,*

$$A = \begin{bmatrix} -1 & 2i & 4 \\ 3 & 6 & i \end{bmatrix}, \quad A^T = \begin{bmatrix} -1 & 3 \\ 2i & 6 \\ 4 & i \end{bmatrix}, \quad A^+ = \begin{bmatrix} -1 & 3 \\ -2i & 6 \\ 4 & -i \end{bmatrix}$$

De este ejemplo es fácil ver que la transpuesta de una matriz columna es una matriz renglón, y viceversa.

Definición 2.9. *Se dice que una matriz $A = [a_{ij}]$ es* ***simétrica*** *si es igual a su transpuesta,*

$$A^T = A, \quad \textit{es decir,} \quad a_{ji} = a_{ij} \tag{2.13}$$

Se dice que una matriz $A = [a_{ij}]$ es ***antisimétrica*** *si*

$$A^T = -A, \quad \textit{es decir,} \quad a_{ji} = -a_{ij} \tag{2.14}$$

En el caso de una matriz A antisimétrica, para $i = j$ se tiene $a_{ii} = -a_{ii}$, lo cual implica que todos los elementos que se encuentran en la diagonal principal de una matriz antisimétrica son ceros.

Se comprueba fácilmente que, dada cualquier matriz cuadrada A, la matriz $A + A^T$ es simétrica, mientras que $A - A^T$ es antisimétrica. De aquí sigue que cualquier matriz cuadrada A puede escribirse de manera única como la suma de la matriz simétrica $\left(A + A^T\right)/2$ y la matriz antisimétrica $\left(A - A^T\right)/2$,

$$A = \left(A + A^T\right)/2 + \left(A - A^T\right)/2$$

Definición 2.10. *Se dice que una matriz $A = [a_{ij}]$ es* ***hermitiana*** *si es igual a su transpuesta hermitiana,*

$$A^+ = A, \quad \textit{es decir,} \quad a_{ji}^* = a_{ij} \tag{2.15}$$

Se dice que una matriz $A = [a_{ij}]$ es ***antihermitiana*** *si*

$$A^+ = -A, \quad \textit{es decir,} \quad a_{ji}^* = -a_{ij} \tag{2.16}$$

En el caso de una matriz A hermitiana, para $i = j$ se tiene $a_{ii}^* = a_{ii}$, lo cual implica que todos los elementos que se encuentran en la diagonal principal de una matriz hermitiana son números reales. En el caso de una matriz A antihermitiana, para $i = j$ se tiene $a_{ii}^* = -a_{ii}$, lo cual implica que todos los elementos que se encuentran en la diagonal principal de una matriz antihermitiana son números imaginarios o ceros.

Se comprueba fácilmente que, dada cualquier matriz cuadrada A, la matriz $A + A^+$ es hermitiana, mientras que $A - A^+$ es antihermitiana. De aquí se sigue que cualquier matriz cuadrada A puede escribirse de manera única como la suma de la matriz hermitiana $\left(A + A^+\right)/2$ y la matriz antihermitiana $\left(A - A^+\right)/2$,

$$A = \left(A + A^+\right)/2 + \left(A - A^+\right)/2$$

Ejemplo 2.7. *Considere las matrices*

$$\begin{aligned}
A &= \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}, \\
D &= \begin{bmatrix} 2 & 1+2i \\ 1-2i & -5 \end{bmatrix}, \quad \bar{D} = \begin{bmatrix} 2 & 1+2i \\ 1+2i & -5 \end{bmatrix}, \\
\tilde{D} &= \begin{bmatrix} 2i & 1+2i \\ -1+2i & -5i \end{bmatrix}
\end{aligned}$$

Entonces A es simétrica y hermitiana, B es antisimétrica y antihermitiana, C no es (anti)simétrica ni (anti)hermitiana, D es hermitiana pero no simétrica, $\bar{D}$ es simétrica pero no hermitiana y $\tilde{D}$ es antihermitiana.

2.2.4. Multiplicación de matrices

En relación con las transformaciones lineales, que son transformaciones de un sistema de coordenadas a otro o mapeos entre espacios lineales, surge la operación de **multiplicación** de dos matrices. Como se verá en capítulo 5, esta operación representa dos mapeos consecutivos.

Definición 2.11. *Sea $A = [a_{ij}]$ una matriz de $m \times n$ y $B = [b_{jk}]$ una matriz de $n \times p$. Entonces el* **producto** *AB (en este orden) se define como la matriz $C = [c_{ik}]$ de $m \times p$ cuyos elementos son*

$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk} \tag{2.17}$$

Nótese que para que el producto AB esté definido, el número de columnas en A debe ser igual al número de renglones en B. Se ve que el elemento c_{ik} es el producto interior (producto punto) del i-ésimo vector renglón de la matriz A por el k-ésimo vector columna de la matriz B.

Propiedad 1. Es fácil ver que, en general, la multiplicación de matrices no es conmutativa. Si A y B son matrices tales que tanto AB como BA están definidas, lo que implica que A es de $m \times n$ y B es de $n \times m$, entonces, en general,

$$AB \neq BA \tag{2.18}$$

Un ejemplo correspondiente de matrices cuadradas que comprueba esta propiedad es el siguiente:

$$\begin{aligned} A &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}, \\ AB &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \neq BA = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} \end{aligned}$$

Se dice que las matrices A y B son conmutativas si se cumple la igualdad $AB = BA$. La multiplicación de una matriz cuadrada A por sí misma es conmutativa, $AA = AA$. Esto nos permite definir la **potencia entera positiva** A^n de una matriz cuadrada A de la manera natural:

$$A^1 = A,\ A^2 = AA,\ A^3 = AA^2 = AAA,\ \ldots,\ A^n = AAA \cdots A\ (n \text{ factores}) \tag{2.19}$$

De aquí se sigue que $A^p A^q = A^{p+q} = A^q A^p$.

Propiedad 2. Es fácil verificar que

$$A_{(m,n)} O_{(n,p)} = O_{(m,p)}, \quad O_{(m,n)} A_{(n,p)} = O_{(m,p)} \tag{2.20}$$

Sin embargo, al contrario que con los números, tiene lugar la siguiente propiedad:

Propiedad 3. En general, no se cumple la ley de cancelación, es decir,

$$AB = O \quad \text{no implica que} \quad A = O \quad \text{o} \quad B = O \tag{2.21}$$

Un ejemplo correspondiente es el siguiente:

$$\begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}=\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Propiedad 4. La transposición de un producto es igual al producto de los factores transpuestos, tomados en orden inverso,

$$(AB)^T = B^T A^T \tag{2.22}$$

Para demostrar dicha propiedad suponemos que A es de $m \times n$ y B es de $n \times l$, de tal manera que la matriz AB esté definida y sea de $m \times l$, quedando $(AB)^T$ de $l \times m$. Luego, B^T es de $l \times n$, A^T es de $n \times m$ y, por lo tanto, el producto $B^T A^T$ está definido. Para el (ij)-ésimo elemento de la matriz $(AB)^T$ se tiene

$$\left\langle (AB)^T \right\rangle_{ij} = \langle AB \rangle_{ji} = \sum_{q=1}^{n} a_{jq} b_{qi} = \sum_{q=1}^{n} a_{qj}^T b_{iq}^T = \sum_{q=1}^{n} b_{iq}^T a_{qj}^T = \langle B^T A^T \rangle_{ij}$$

lo que completa la demostración.

De manera similar se comprueba la siguiente propiedad.

Propiedad 5. La conjugada hermitiana de un producto de dos matrices es igual al producto de los factores transpuestos hermitianos, tomados en orden inverso,

$$(AB)^+ = B^+ A^+ \tag{2.23}$$

Definición 2.12. *Una matriz diagonal en la que todos los elementos de la diagonal principal son iguales, se llama **matriz escalar**,*

$$E = [e_{ij}] = \begin{bmatrix} k & 0 & \cdots & 0 \\ 0 & k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & k \end{bmatrix}$$

en donde $e_{ij} = k\delta_{ij}$, δ_{ij} es el símbolo de Kronecker y k es un número cualquiera.

El nombre de la matriz E se debe a la siguiente propiedad:

$$A_{(m,n)} E_{(n,n)} = E_{(m,m)} A_{(m,n)} = k A_{(m,n)} \tag{2.24}$$

es decir, la multiplicación de una matriz rectangular A de $m \times n$ por una matriz escalar $E_{(n,n)}$ de la derecha o por una matriz escalar $E_{(m,m)}$ de la izquierda es igual a la multiplicación de A por un escalar k. Además, de la ecuación (2.24) se obtiene que E conmuta con cualquier matriz cuadrada del mismo orden. En el caso particular, $k = 1$, de la matriz escalar E se tiene la **matriz identidad** o **unidad** I, tal que

$$AI = IA = A \tag{2.25}$$

para cualquier matriz cuadrada A del mismo orden. Por definición $A^0 = I$.

Teorema 2.3. *(Propiedades de la multiplicación de matrices) La multiplicación de matrices es*

1) asociativa

$$A(BC) = (AB)C \tag{2.26}$$
$$(kA)B = k(AB) = A(kB) \tag{2.27}$$

2) distributiva con respecto a la adición de matrices

$$A(B+C) = AB + AC \tag{2.28}$$
$$(A+B)C = AC + BC \tag{2.29}$$

Demostración. Supongamos que A es de $n \times l$, B es de $l \times m$ y C es de $m \times p$. Entonces para el (ij)-ésimo elemento de la matriz $A(BC)$ se tiene

$$\begin{aligned}
\langle A(BC)\rangle_{ij} &= \sum_{q=1}^{l} a_{iq} \langle BC\rangle_{qj} = \sum_{q=1}^{l} a_{iq} \left(\sum_{s=1}^{m} b_{qs} c_{sj}\right) \\
&= \sum_{q=1}^{l}\sum_{s=1}^{m} a_{iq} b_{qs} c_{sj} = \sum_{s=1}^{m} \left(\sum_{q=1}^{l} a_{iq} b_{qs}\right) c_{sj} \\
&= \sum_{s=1}^{m} \langle AB\rangle_{is} c_{sj} = \langle (AB)C\rangle_{ij}
\end{aligned}$$

que completa la demostración de la ecuación (2.26). De manera similar o usando propiedades de la matriz escalar $E = kI$, se comprueba la ecuación (2.27), que se deja como ejercicio. La propiedad distributiva, ecuaciones (2.28) y (2.29), se demuestra fácilmente, y también se deja como ejercicio.

Problema. *Demuestre las igualdades* $(ABC)^T = C^T B^T A^T$ *y* $(ABC)^+ = C^+ B^+ A^+$; $ABCD = A(B(CD)) = (AB)(CD) = A((BC)D) = (A(BC))D$.

Problema. *¿En caso de que* $AB = AC$ *o* $BA = CA$ *se sigue necesariamente que* $B = C$*?*

Problema. *¿Cuándo es válida la igualdad* $(A+B)(A-B) = A^2 - B^2$*?*

2.3. Matrices inversas

2.3.1. Definición y propiedades

Dado el número a, su inverso recíproco x es el número tal que resuelva la ecuación $xa = ax = 1$. Por cierta analogía con los números y tomando en cuenta que, en general, el producto de matrices no es conmutativo, podemos introducir el concepto de la matriz inversa X de una matriz dada A, de tal manera que la inversa resuelva $XA = I$ y/o $AX = I$.

Definición 2.13. *Sea A una matriz rectangular de $m \times n$.*

*a) Cualquier matriz $L_{(n,m)}$ para la cual $L_{(n,m)}A_{(m,n)} = I_{(n,n)}$ se llama **inversa izquierda** de A.*

*b) Cualquier matriz $R_{(n,m)}$ para la cual $A_{(m,n)}R_{(n,m)} = I_{(m,m)}$ se llama **inversa derecha** de A.*

*c) Cualquier matriz $A^{-1}_{(n,m)}$ para la cual $A^{-1}_{(n,m)}A_{(m,n)} = I_{(n,n)}$ y $A_{(m,n)}A^{-1}_{(n,m)} = I_{(m,m)}$ se llama **inversa (bilateral)** de A.*

Si la matriz A es de $m \times n$, entonces cualquier inversa derecha R o izquierda L de A y, por lo tanto, la inversa bilateral A^{-1} deben ser de $n \times m$. Se puede determinar cuál de estos tipos de inversas tiene una matriz A mediante la solución de conjuntos de ecuaciones lineales correspondientes, en donde los elementos de las inversas están considerados como incógnitas. En caso de que A sea de $n \times n$, las inversas, si existen, son matrices cuadradas del mismo orden. Debe mencionarse que un análisis más a fondo de la existencia de inversas bilaterales nos llevará a la conclusión de que las inversas bilaterales existen sólo para las matrices cuadradas, tanto que en el inciso c de la definición anterior hay que suponer $m = n$.

Ejemplo 2.8. *Buscaremos una matriz inversa derecha R para la matriz*

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}$$

que requiere resolver la ecuación $AR = I$, esto es,

$$\begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x & z \\ y & w \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Multiplicando las matrices se tienen dos conjuntos de ecuaciones lineales:

$$\begin{array}{cc} x - y = 1 & z - w = 0 \\ x + 2y = 0 & z + 2w = 1 \end{array}$$

La solución de este sistema de ecuaciones es única: $x = 2/3$, $y = -1/3$, $z = 1/3$, $w = 1/3$. Por lo tanto, hay exactamente una matriz inversa derecha, que es

$$R = \begin{bmatrix} 2/3 & 1/3 \\ -1/3 & 1/3 \end{bmatrix}$$

Ejemplo 2.9. *Se busca una matriz inversa derecha R para la matriz*

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

que implica la resolución de la ecuación $AR = I$, esto es,

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} x & u \\ y & v \\ z & w \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

La última se reduce al sistema de ecuaciones lineales subdeterminadas, es decir, el número de incógnitas es mayor que el número de ecuaciones:

$$\begin{array}{ll} x-y+z=1 & u-v+w=0 \\ x+y+2z=0 & u+v+2w=1 \end{array}$$

Al resolverlas, se pueden dar valores arbitrarios para $z=\alpha$ y $w=\beta$. Resolviendo con respecto a las otras variables y considerando α y β como parámetros, finalmente se obtiene un número infinito de inversas derechas

$$R=\begin{bmatrix} (1-3\alpha)/2 & (1-3\beta)/2 \\ -(1+\alpha)/2 & (1-\beta)/2 \\ \alpha & \beta \end{bmatrix}$$

Ejemplo 2.10. *Se busca una inversa izquierda L de la matriz*

$$A=\begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix} \quad \text{en la forma} \quad L=\begin{bmatrix} x & y \\ z & w \end{bmatrix}$$

De la ecuación $LA=I$ se obtiene un sistema inconsistente de ecuaciones lineales,

$$\begin{array}{ll} x-2y=1 & z-2w=0 \\ -x+2y=0 & -z+2w=1 \end{array}$$

que no tiene solución. Por lo tanto, no existe la inversa izquierda para la matriz A dada.

Los ejemplos anteriores muestran la manera de encontrar matrices inversas y, además, el hecho de que las inversas pueden no existir o, al contrario, ser múltiples. El siguiente teorema aclara la existencia y unicidad de inversas bilaterales.

Teorema 2.4. ***(Inversas bilaterales)*** *Sea A una matriz. Si existen una inversa derecha R y una inversa izquierda L de A, entonces son iguales y son una inversa bilateral A^{-1}. Cualesquiera dos inversas bilaterales de A son idénticas.*

Demostración. *Sean R y L inversas derecha e izquierda de A, respectivamente. Entonces,*

$$R=IR=(LA)R=L(AR)=LI=L$$

de modo que $R=L\equiv A^{-1}$, y ésta es una inversa bilateral. Luego, supongamos que X y Y son dos inversas bilaterales de A. A partir de

$$X=IX=(YA)X=Y(AX)=YI=Y$$

se obtiene la unicidad. ■

La propiedad de tener la inversa bilateral es tan importante que las matrices cuadradas con tales inversas se distinguen con un nombre especial.

b) Suponemos que A tiene una inversa derecha R. Entonces, existe al menos una solución, que es $R\mathbf{b}$, a la ecuación $A\mathbf{x} = \mathbf{b}$.

c) Suponemos que A es no singular. Entonces existe exactamente una solución a $A\mathbf{x} = \mathbf{b}$, es decir, $A^{-1}\mathbf{b}$.

Demostración. *a) Si $\mathbf{x}_0$ es alguna solución a $A\mathbf{x} = \mathbf{b}$, entonces $\mathbf{b} = A\mathbf{x}_0$. Multiplicando la última ecuación por L se tiene*

$$L\mathbf{b} = L(A\mathbf{x}_0) = (LA)\mathbf{x}_0 = I\mathbf{x}_0 = \mathbf{x}_0$$

Pues, si existe una solución, ésta debe ser igual a $L\mathbf{b}$.

b) Se tiene que

$$A(R\mathbf{b}) = (AR)\mathbf{b} = I\mathbf{b} = \mathbf{b}$$

por lo tanto, $R\mathbf{b}$ es una solución.

c) Si A es no singular, existe la inversa bilateral A^{-1}. Ya que A^{-1} es una inversa derecha, $A^{-1}\mathbf{b}$ es una solución según el inciso anterior. Por otra parte, cualquier solución debe ser igual a $A^{-1}\mathbf{b}$ según el inciso a, ya que A^{-1} es una inversa izquierda. Además, la matriz A^{-1} es única. Entonces, la solución $A^{-1}\mathbf{b}$ es única también. Esto completa la demostración de la existencia y unicidad de la solución. ■

De lo anterior se ve cómo se pueden solucionar sistemas de ecuaciones lineales mediante matrices inversas. Pero resulta que cada matriz inversa a la vez se obtiene mediante la solución de sistemas de ecuaciones lineales. Pues, volvemos exactamente a problemas del mismo tipo que se querían resolver desde un principio. En la práctica, por lo general, es más eficiente resolver un sistema de ecuaciones lineales directamente (por ejemplo, usando el método de eliminación de Gauss o de Gauss-Jordan), que calcular primero A^{-1} y luego $A^{-1}\mathbf{b}$. Sin embargo, el concepto y representación de la solución como $\mathbf{x} = A^{-1}\mathbf{b}$ pueden ser muy poderosos para el entendimiento general de ciertos aspectos en muchos problemas. Proposiciones generales acerca de problemas de existencia y unicidad de solución de un sistema de ecuaciones lineales se dan mediante el concepto de rango de una matriz, y la información relevante se puede encontrar en varios libros, por ejemplo, véase [8], [11] y [16].

2.4. Matriz unitaria, ortogonal. Traza

Las matrices unitarias y ortogonales están estrechamente relacionadas con transformaciones entre las bases ortonormales en espacios vectoriales lineales (véase el capítulo 5). Dada la importancia de dichos conceptos, en esta sección se presenta la definición y algunas de las propiedades importantes de las matrices unitarias.

Sean $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ y $\{\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n\}$ dos bases en un espacio vectorial lineal L. Sea T la matriz de una transformación de coordenadas, $x' = Tx$ de un vector $\mathbf{x}$ que tiene coordenadas x' en la base $\{\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n\}$ y coordenadas x en la base $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$. Además, sea A la matriz que representa un mapeo

Teorema 2.7. *Sea A una matriz cuadrada de $n \times n$. El $\det A$ es igual a la suma de los productos de los elementos de cualquier renglón o de cualquier columna por sus cofactores correspondientes:*

$$\det A = \sum_{j=1}^{n} a_{1j}C_{1j} = \sum_{j=1}^{n} a_{rj}C_{rj} = \sum_{i=1}^{n} a_{ik}C_{ik} \tag{2.35}$$

Los propiedades importantes de los determinantes son las siguientes.

Corolario. *a) El determinante de una matriz y de su transpuesta son iguales*

$$\det A = \det A^T$$

b) Si cualquier renglón o columna de A se multiplica por un número c, quedando como resultado una $\tilde{A}$, entonces

$$\det \tilde{A} = c \det A$$

En particular, el determinante de una matriz con cualquier renglón o columna igual a cero es cero.

c) Para cualquier número c y una matriz A de $n \times n$, se tiene

$$\det cA = c^n \det A$$

d) Si A tiene dos renglones (columnas) iguales, entonces $\det A = 0$. De donde, con base en el inciso b se tiene que, si A tiene dos renglones (columnas) múltiplos, entonces $\det A = 0$.

e) Si A' se obtiene de A intercambiando dos renglones (o dos columnas), entonces

$$\det A' = (-1) \det A$$

f) Si A' se obtiene de A sumando a un renglón (una columna) un múltiplo de otro renglón (otra columna), entonces

$$\det A' = \det A$$

Demostración. *a) Desarrollar el $\det A$ con respecto a su primer renglón es lo mismo que desarrollar el $\det A'$ con respecto a su primera columna.*

b) Desarrollando el $\det \tilde{A}$ con respecto al renglón (a la columna) que fue multiplicado por un número c, se tiene $\det \tilde{A} = c \det A$.

c) Se comprueba fácilmente por inducción.

d) El enunciado es válido para cualquier matriz A de 2×2 con dos renglones iguales. Suponemos que es también válido para cualquier matriz A de $(n-1) \times (n-1)$. Al desarrollar una matriz A de $n \times n$ con respecto a un renglón que es diferente de los dos iguales se tiene que todos los productos son ceros según la suposición hecha con respecto a las matrices de $(n-1) \times (n-1)$. Entonces, el resultado es cierto por inducción.

Propiedad 2. La transpuesta A^T tiene los mismos eigenvalores que A, lo que sigue de $\det\left(A^T - \lambda I\right) = \det\left((A - \lambda I)^T\right) = \det(A - \lambda I)$.

Propiedad 3. Sean $\lambda_1, \lambda_2, \ldots, \lambda_n$ los eigenvalores de una matriz normal A de $n \times n$, $AA^+ = A^+A$. Se demuestra que $trA = \sum_{i=1}^{n} \lambda_i$.

Propiedad 4. La inversa bilateral A^{-1} tiene los eigenvalores $\tilde{\lambda}_1 = \lambda_1^{-1}, \tilde{\lambda}_2 = \lambda_2^{-1}, \ldots, \tilde{\lambda}_n = \lambda_n^{-1}$. Además, los eigenvectores de A^{-1} son los mismos que los de A. Para demostrar la propiedad, considérese la ecuación característica de la matriz A^{-1}:

$$\begin{aligned} 0 &= \det\left(A^{-1} - \tilde{\lambda} I\right) \\ &= \det\left(\tilde{\lambda} A^{-1}\left(\tilde{\lambda}^{-1} I - A\right)\right) = \det\left(\tilde{\lambda} A^{-1}\right)\det\left(\tilde{\lambda}^{-1} I - A\right) \end{aligned}$$

Dado que $\det\left(\tilde{\lambda} A^{-1}\right) \neq 0$, se tiene que $\det\left(A^{-1} - \tilde{\lambda} I\right) = 0$ si y sólo si $\det\left(\tilde{\lambda}^{-1} I - A\right) = 0$, es decir, $\tilde{\lambda}$ es un eigenvalor de A^{-1} si y sólo si $\tilde{\lambda}^{-1}$es un eigenvalor de A. Además, $\mathbf{x}$ es un eigenvector de A^{-1} asociado a $\tilde{\lambda}$ si y sólo si es el eigenvector de A asociado a $\lambda = \tilde{\lambda}^{-1}$: $A^{-1}\mathbf{x} = \tilde{\lambda}\mathbf{x} \Longleftrightarrow \tilde{\lambda}^{-1}\mathbf{x} = A\mathbf{x}$.

Propiedad 5. La matriz cA, en donde c es un número, tiene un eigenvalor $\tilde{\lambda} = c\lambda$ si y sólo si A tiene un eigenvalor λ. La demostración es la consecuencia inmediata de la igualdad: $0 = \det\left(cA - \tilde{\lambda} I\right) = c^n \det\left(A - \left(\tilde{\lambda}/c\right) I\right)$. Es obvio que cA y A tienen los mismos eigenvectores.

Propiedad 6. La matriz A^m, en donde m es un entero no negativo, tiene los eigenvalores $\tilde{\lambda}_i = \lambda_i^m$, y los eigenvectores de A asociados con un eigenvalor λ_i son eigenvectores de A^m asociados a $\tilde{\lambda}_i = \lambda_i^m$. La demostración es como sigue. Sea $\mathbf{x}_i$ un eigenvector de A asociado a λ_i, $A\mathbf{x}_i = \lambda_i \mathbf{x}_i$. Entonces, $A^m \mathbf{x}_i = A^{m-1}(\lambda_i \mathbf{x}_i) = ... = \lambda_i^m \mathbf{x}_i$.

Propiedad 7. Una vez que se han determinado los eigenvalores, es posible determinar los eigenvectores correspondientes a partir de la ecuación (2.40). Dado que el sistema es homogéneo, es obvio que, si $\mathbf{x}$ es un eigenvector de A, entonces $c\mathbf{x}$, en donde $c \neq 0$ es una constante cualquiera, también es un eigenvector de A correspondiente al mismo eigenvalor. En otras palabras, un eigenvector se determina hasta un factor constante.

Ejemplo 2.14. *La ecuación característica de la matriz*

$$A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

en donde a y b son reales y $b \neq 0$, es $(a - \lambda)^2 + b^2 = 0$. Las raíces de la ecuación son complejos conjugados, $\lambda_1 = a + ib$ y $\lambda_2 = a - ib$. Los eigenvectores asociados se obtienen de sistemas de ecuaciones

$$\begin{array}{ll} -ibx_1 + bx_2 = 0 & \quad ibx_1 + bx_2 = 0 \\ -bx_1 - ibx_2 = 0 & \quad -bx_1 + ibx_2 = 0 \end{array}$$

Teorema 2.14. *Una matriz A de $n \times n$ es normal si y sólo si es unitariamente semejante a una matriz diagonal,*

$$U^+AU = diag\,[\lambda_1, \lambda_2, \ldots, \lambda_n]\,, \quad U^+ = U^{-1} \tag{2.46}$$

Los elementos en la diagonal principal son, por supuesto, eigenvalores de A.

Del teorema anterior y del teorema de eigenvalores de una matriz hermitiana se sigue:

Teorema 2.15. *Una matriz A de $n \times n$ es hermitiana si y sólo si es unitariamente semejante a una matriz diagonal con elementos reales,*

$$U^+AU = diag\,[\lambda_1, \lambda_2, \ldots, \lambda_n]\,, \quad U^+ = U^{-1}, \quad \lambda_i = \lambda_i^* \quad (i = 1, 2, \ldots, n) \tag{2.47}$$

Teorema 2.16. *Una matriz A de $n \times n$ es unitaria si y sólo si es unitariamente semejante a una matriz diagonal con elementos que tienen el valor absoluto igual a uno en la diagonal principal,*

$$U^+AU = diag\,[\lambda_1, \lambda_2, \ldots, \lambda_n]\,, \quad U^+ = U^{-1}, \quad |\lambda_i| = 1 \quad (i = 1, 2, \ldots, n) \tag{2.48}$$

Problema. *Sea A una matriz cuadrada. Demuestre que λ es un eigenvalor de A si y sólo si λ^* es un eigenvalor de A^+. Sugerencia: considérese la ecuación característica* $\det(A - \lambda I) = (\det(A^+ - \lambda^* I))^* = 0$.

Problema. *Sea A una matriz normal, $AA^+ = A^+A$. Demuestre que si $\mathbf{x}$ es un eigenvector de A asociado con un eigenvalor λ, $A\mathbf{x} = \lambda\mathbf{x}$, y a la vez $\mathbf{x}$ es un eigenvector de A^+ asociado con un eigenvalor μ, $A^+\mathbf{x} = \mu\mathbf{x}$, entonces, $\mu = \lambda^*$.*

Problema. *Sean $\lambda_1, \lambda_2, \ldots, \lambda_n$ los eigenvalores de una matriz normal A de $n \times n$, $AA^+ = A^+A$. Demuestre que $tr A = \sum\limits_{i=1}^{n} \lambda_i$.*

2.8. Función de una matriz

Sean $A = [a_{ij}]$ una matriz cuadrada de $n \times n$ y $f(\lambda)$ una función de variable escalar λ. Es natural el deseo de extender el concepto de una función escalar a una función de variable matricial. Dado que anteriormente fue definida la potencia entera positiva A^n de una matriz, la solución de dicho problema es fácil en el caso de funciones que son polinomios de su variable λ; es decir, $f(\lambda) = a_0\lambda^m + a_1\lambda^{m-1} + \cdots + a_m$ tiene como su homólogo la siguiente función de matriz A, $f(A) = a_0A^m + a_1A^{m-1} + \cdots + a_mA^0$. Sin embargo, el problema general de función de una matriz es algo complicado y requiere conocimientos de la teoría espectral de matrices. Por esta razón, en esta sección las consideraciones serán un tanto reducidas. Una manera práctica de introducir el concepto de función de una matriz es como sigue.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

en donde $\mathbf{c}$ es un vector columna de constantes de integración y la integral de una matriz se entiende como la matriz de integrales de los elementos. Luego,

$$\mathbf{x}(t) = e^{tA}\mathbf{z}(t) = e^{tA}\left(\mathbf{c} + \int_{t_0}^{t} e^{-\tau A}\mathbf{f}(\tau)\, d\tau\right) = e^{tA}\mathbf{c} + \int_{t_0}^{t} e^{(t-\tau)A}\mathbf{f}(\tau)\, d\tau$$

Las condiciones iniciales para $\mathbf{x}(t)$ al instante inicial t_0 requieren que se cumpla la igualdad

$$\mathbf{x}_0 = \mathbf{x}(t_0) = e^{t_0 A}\mathbf{c}, \quad \text{es decir,} \quad \mathbf{c} = e^{-t_0 A}\mathbf{x}_0$$

Finalmente, la solución vectorial de la ecuación (2.60) es

$$\mathbf{x}(t) = e^{(t-t_0)A}\mathbf{x}_0 + \int_{t_0}^{t} e^{(t-\tau)A}\mathbf{f}(\tau)\, d\tau \tag{2.62}$$

2.10. Formas bilineales, cuadráticas y hermitianas

Las funciones cuadráticas de múltiples variables independientes son del siguiente nivel de complejidad después de las constantes y funciones lineales. Tales funciones surgen en distintas áreas de aplicación y los métodos matriciales permiten un estudio unificado de sus propiedades. Por otra parte, las funciones cuadráticas proporcionan conceptos matriciales importantes, especialmente en lo que se refiere a los sistemas de eigenvalores y eigenvectores. Las curvas llamadas secciones cónicas, que juegan un papel fundamental en geometría bidimensional, son un ejemplo conocido de las funciones cuadráticas. Muchos problemas de optimización y estabilidad de sistemas dinámicos llevan a la necesidad de estudiar el comportamiento de un sistema en vecindades de estados de equilibrio. Tales estudios se basan en una función de múltiples variables que puede llamarse función de costo o energía. Los valores extremos y el comportamiento de esta función en vecindades de los estados de equilibrio a menudo se aproximan por medio de funciones cuadráticas, desarrollando la función de costo en serie de Taylor. Cabe mencionar que las formas bilineales y cuadráticas hermitianas son de importancia especial para la construcción de espacios unitarios en Mecánica Cuántica.

Un polinomio homogéneo de segundo orden con respecto a las $2n$ variables x_i y y_j $(i, j = 1, 2, \ldots, n)$,

$$A(\mathbf{x}, \mathbf{y}) = \sum_{i,j=1}^{n} a_{ij} x_i^* y_j \tag{2.63}$$

se llama **forma bilineal** de las $2n$ variables x_i y y_j. Las $2n$ variables x_i y y_j

Esto conduce a la necesidad de la introducción de los números complejos. Si adjuntamos a los números reales un símbolo i tal que satisface por definición a la ecuación

$$i^2 = -1 \tag{3.1}$$

podemos construir los números complejos,

$$z = x + iy = \operatorname{Re} z + i \operatorname{Im} z \tag{3.2}$$

en donde x y y son números reales cualesquiera. A $\operatorname{Re} z = x$ se le llama **parte real** del número complejo z, e $\operatorname{Im} z = y$ recibe el nombre de **parte imaginaria** de z. Resulta que los números complejos satisfacen las leyes algebraicas de los números reales e incluyen a los números reales como un caso especial. Además, los números complejos permiten resolver no solamente la ecuación $x^2 + 1 = 0$, sino cualquier ecuación polinomial.

Se dice que dos números complejos $z_1 = x_1 + iy_1$ y $z_2 = x_2 + iy_2$ son iguales si y sólo si sus partes reales e imaginarias son respectivamente iguales, es decir, $x_1 = x_2$ y $y_1 = y_2$. Además, se admite que $iy = yi$. La adición y la multiplicación de números complejos se definen de tal manera que el álgebra de números complejos sea similar a la de los números reales.

La **suma (adición) de dos números complejos** $z_1 = x_1 + iy_1$ y $z_2 = x_2 + iy_2$ es

$$z_1 + z_2 = (x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2) \tag{3.3}$$

El producto (multiplicación) de dos números complejos z_1 y z_2 se define como

$$\begin{aligned} z_1 \cdot z_2 &= (x_1 + iy_1)(x_2 + iy_2) = x_1x_2 + iy_1x_2 + ix_1y_2 + i^2y_1y_2 \\ &= (x_1x_2 - y_1y_2) + i\,(x_1y_2 + y_1x_2) \end{aligned} \tag{3.4}$$

A partir de las leyes del álgebra para los números reales se obtienen las leyes para los números complejos. La **ley conmutativa** de la adición y la multiplicación:

$$z_1 + z_2 = z_2 + z_1, \quad z_1 \cdot z_2 = z_2 \cdot z_1 \tag{3.5}$$

La **ley asociativa** de la adición y la multiplicación:

$$(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3), \quad (z_1 \cdot z_2) \cdot z_3 = z_1 \cdot (z_2 \cdot z_3) \tag{3.6}$$

La **ley distributiva**:

$$z_1(z_2 + z_3) = z_1z_2 + z_1z_3 \tag{3.7}$$

Además, se define el número cero $\widehat{0}$ y la unidad $\widehat{1}$ tales que para cualquier número complejo z se cumplan las igualdades,

$$z + \widehat{0} = z, \quad z \cdot \widehat{1} = z \tag{3.8}$$

Es fácil ver que $\widehat{0} = 0 + i0 = 0$ y $\widehat{1} = 1 + i0 = 1$.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

3.2. Funciones complejas básicas

Una magnitud compleja $z = x + iy$, en la cual x y y están consideradas como variables reales, se llama **variable compleja**. Por función de una variable compleja z definida sobre un conjunto D se entiende una correspondencia (regla) que asigna a cada z en D un número complejo $w = u(x, y) + iv(x, y)$, en donde $u(x, y)$ y $v(x, y)$ son dos funciones reales, dependiendo cada una de las dos variables reales x y y. Entonces, se escribe $w = f(z)$. El conjunto D es el **dominio de definición** de $f(z)$. La función $w = f(z)$ se llama **univaluada** en una región D si a cada número complejo z de esta región le corresponde un valor único de w. Se dice que $w = f(z)$ es una función **multivaluada** si a cada z le corresponde más de un valor de w.

Ejemplo 3.1. *a) La función lineal $w = az$, es decir, $u = ax$ y $v = ay$, en donde a es una constante real, es univaluada en todo el plano complejo z.*

b) La función $w = z^2 = (x^2 - y^2) + i2xy$, es decir, $u = x^2 - y^2$ y $v = 2xy$, es también univaluada en todo el plano z.

c) La función $w = zz^*$, en donde $u = x^2 + y^2$ y $v = 0$ es univaluada.

d) Es fácil ver que la ecuación $w^2 = z$ tiene la solución $w = \sqrt{z}$ $= \sqrt{\rho}\left(\cos\frac{\varphi+2\pi k}{2} + i\,\mathrm{sen}\,\frac{\varphi+2\pi k}{2}\right)$, $k = 0, 1$, en donde la variable compleja es $z = \rho(\cos\varphi + i\,\mathrm{sen}\,\varphi)$. Por lo tanto, la función $w = \sqrt{z}$, que se define como la solución de $w^2 = z$, es doble valuada: $w_1 = \sqrt{\rho}\left(\cos\frac{\varphi}{2} + i\,\mathrm{sen}\,\frac{\varphi}{2}\right)$, $w_2 = \sqrt{\rho}(\cos\left(\frac{\varphi}{2} + \pi\right) + i\,\mathrm{sen}\left(\frac{\varphi}{2} + \pi\right))$.

En esta sección se consideran las funciones complejas básicas y sus propiedades. Al considerar funciones complejas, es necesario distinguir entre los puntos internos y los de frontera de un dominio. Se dice que un punto de dominio D es **punto interno** si existe un círculo de un radio no igual a cero, con el centro en este punto y tal que contiene en su interior únicamente los puntos del dominio D. Los puntos de la frontera de D no son interiores, dado que cualquier círculo alrededor de un punto de frontera necesariamente contiene tanto puntos de D como puntos que no pertenecen al dominio D. La frontera de un dominio D puede ser constituida por los puntos de D o los que no pertenecen a D. Un dominio constituido únicamente por puntos interiores se llama abierto. En este caso la frontera no pertenece a éste. Un ejemplo de **dominio abierto** es $|z| < R$. Cuando los puntos de frontera están incluidos en el dominio, éste se llama **cerrado**. Un ejemplo de región cerrada (de dominio cerrado) es $|z| \leq R$. Cuando cada punto de un dominio se encuentra a una distancia finita del centro, se dice que el **dominio es limitado** o **finito**. Todos los puntos de un dominio limitado se encuentran adentro de un círculo $|z| = R$, si el radio R es escogido suficientemente grande. El dominio que no es limitado se llama no limitado. Un ejemplo, $|z| \geq 1$.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Ejemplo 3.5. *Para calcular* $\operatorname{sen}(1-i)$ *podemos seguir dos caminos. Primero, según la definición,*

$$\begin{aligned}
\operatorname{sen}(1-i) &= \frac{e^{i(1-i)}-e^{-i(1-i)}}{2i}=\frac{e^{i+1}-e^{-i-1}}{2i} \\
&= \frac{1}{2i}\left[e^{1}\left(\cos 1+i\operatorname{sen} 1\right)-e^{-1}\left(\cos 1-i\operatorname{sen} 1\right)\right] \\
&= \frac{e^{1}-e^{-1}}{2i}\cos 1+\frac{e^{1}+e^{-1}}{2}\operatorname{sen} 1
\end{aligned}$$

Por otra parte, aplicando la fórmula trigonométrica, se tiene

$$\operatorname{sen}(1-i)=\operatorname{sen} 1\cos i-\cos 1\operatorname{sen} i$$

Sustituyendo

$$\cos i=\left(e^{-1}+e^{1}\right)/2 \quad y \quad \operatorname{sen} i=\left(e^{-1}-e^{1}\right)/2i$$

en la fórmula anterior, se tiene el resultado que obtuvimos de la definición.

Funciones hiperbólicas

Las funciones hiperbólicas de una variable compleja z se definen mediante las fórmulas:

$$\begin{aligned}
\cosh z &= \frac{1}{2}\left(e^{z}+e^{-z}\right), \quad \operatorname{senh} z=\frac{1}{2}\left(e^{z}-e^{-z}\right) && (3.37)\\
\tanh z &= \frac{\operatorname{senh} z}{\cosh z}, \quad \coth z=\frac{\cosh z}{\operatorname{senh} z} && (3.38)\\
\operatorname{sech} z &= \frac{1}{\cosh z}, \quad \operatorname{csch} z=\frac{1}{\operatorname{senh} z} && (3.39)
\end{aligned}$$

Éstas son univaluadas. De las ecuaciones (3.32) y (3.37) se obtiene que

$$\begin{aligned}
\cosh(iz) &= \cos z, \quad \operatorname{senh}(iz)=i\operatorname{sen} z && (3.40)\\
\cos(iz) &= \cosh z, \quad \operatorname{sen}(iz)=i\operatorname{senh} z && (3.41)
\end{aligned}$$

Logaritmo natural

El logaritmo de un número complejo z se define de manera similar que el logaritmo de una variable real, pues el logaritmo de base natural

$$w=\ln z \tag{3.42}$$

de un número complejo z es la solución de la ecuación

$$e^{w}=z \tag{3.43}$$

Puesto que $w=u+iv$ y $z=\rho\left(\cos\varphi+i\operatorname{sen}\varphi\right)$, de la ecuación (3.43) se tiene

$$e^{u}\left(\cos v+i\operatorname{sen} v\right)=\rho\left(\cos\varphi+i\operatorname{sen}\varphi\right)$$

Definición 3.1. *La función $f(z)$ se dice que es analítica (holomorfa) en el punto z_0 si existe la derivada $f'(z)$ en $z = z_0$ y en una vecindad del punto z_0. La función $f(z)$ es analítica (holomorfa) en un dominio D, si es diferenciable en todos los puntos de D. Los puntos del plano z, donde $f(z)$ no es analítica, se llaman puntos singulares de $f(z)$.*

El criterio básico de la analiticidad de una función compleja se da por el siguiente teorema que involucra la parte real e imaginaria de esta función.

Teorema 3.1. ***(Teorema de Cauchy-Riemann)*** *La función $f(z) = u(x,y) + iv(x,y)$ es analítica en un punto $z_0 = x_0 + iy_0$ si y sólo si en una vecindad de (x_0, y_0) las funciones $u(x,y)$ y $v(x,y)$ son continuas junto con sus primeras derivadas parciales $\partial u/\partial x$, $\partial v/\partial x$, $\partial u/\partial y$, $\partial v/\partial y$ y, además, se cumplen las ecuaciones de Cauchy-Riemann:*

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y} \tag{3.50}$$

Demostración. *Condiciones necesarias. Suponemos que $f(z)$ es analítica en z. Dado que $\Delta z = \Delta x + i\Delta y$ y $\Delta w = \Delta u(x,y) + i\Delta v(x,y)$, se tiene que el límite*

$$f'(z) = \lim_{\Delta z \to 0} \frac{\Delta u + i\Delta v}{\Delta z} = \lim_{\Delta x, \Delta y \to 0} \frac{\Delta u + i\Delta v}{\Delta x + i\Delta y}$$

es único. Para evaluar el límite, primero hacemos tender $\Delta z \to 0$ de la siguiente manera:

$$f'(z) = \lim_{\Delta x \to 0} \lim_{\Delta y \to 0} \frac{\Delta u + i\Delta v}{\Delta x + i\Delta y} = \lim_{\Delta x \to 0} \left(\frac{\hat{\Delta} u}{\Delta x} + i\frac{\hat{\Delta} v}{\Delta x} \right) = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}$$

en donde $\hat{\Delta} u = u(x + \Delta x, y) - u(x,y)$, $\hat{\Delta} v = v(x + \Delta x, y) - v(x,y)$. Luego, tendiendo primero $\Delta x \to 0$ y, después, $\Delta y \to 0$, se obtiene

$$f'(z) = \lim_{\Delta y \to 0} \lim_{\Delta x \to 0} \frac{\Delta u + i\Delta v}{\Delta x + i\Delta y} = \lim_{\Delta y \to 0} \left(\frac{\check{\Delta} u}{i\Delta y} + \frac{\check{\Delta} v}{\Delta y} \right) = \frac{\partial v}{\partial y} - i\frac{\partial u}{\partial y}$$

en donde $\check{\Delta} u = u(x, y + \Delta y) - u(x,y)$, $\check{\Delta} v = v(x, y + \Delta y) - v(x,y)$. Entonces, para que estos límites tengan los mismos valores en el punto z, deben satisfacerse las ecuaciones

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}$$

que reciben el nombre de ecuaciones de Cauchy-Riemann.

Condiciones suficientes. Suponemos que se cumplen las ecuaciones de Cauchy-Riemann y, además, $u(x,y)$ y $v(x,y)$ son diferenciables, es decir,

$$\Delta u = u(x + \Delta x, y + \Delta y) - u(x,y) = \frac{\partial u}{\partial x}\Delta x + \frac{\partial u}{\partial y}\Delta y + \epsilon_1 \Delta x + \epsilon_2 \Delta y$$

$$\Delta v = v(x + \Delta x, y + \Delta y) - v(x,y) = \frac{\partial v}{\partial x}\Delta x + \frac{\partial v}{\partial y}\Delta y + \epsilon_3 \Delta x + \epsilon_4 \Delta y$$

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Esto demuestra que la integral de línea, ecuación (3.56) existe y su valor es independiente de la partición de la curva C y los puntos intermedios ζ_k. La curva C recibe el nombre de **trayectoria de integración**.

Al aplicar la representación paramétrica, ecuación (3.55), la integral de línea toma la forma,

$$\begin{aligned}\int_C f(z)\,dz &= \int_C (udx - vdy) + i\int_C (udy + vdx) \\ &= \int_a^b (u\dot{x} - v\dot{y})\,dt + i\int_a^b (u\dot{y} + v\dot{x})\,dt \qquad (3.58)\end{aligned}$$

donde $u = u(x(t), y(t))$, $v = v(x(t), y(t))$ y el punto denota la derivada con respecto al parámetro t.

Ejemplo 3.15. *Con el fin de mostrar la aplicación de la fórmula (3.57), evaluamos la integral*

$$\int_C (z^*)^2\,dz$$

a lo largo de la línea recta que conecta los puntos $z = 0$ y $z = 1 + 2i$. Dado que $(z^*)^2 = x^2 - y^2 - i2xy$, al sustituir $u = x^2 - y^2$ y $v = -2xy$ en la ecuación (3.57), se tiene

$$\int_C (z^*)^2\,dz = \int_C \left(\left(x^2 - y^2\right)dx + 2xydy\right) + i\int_C \left(\left(x^2 - y^2\right)dy - 2xydx\right)$$

Al usar la representación cartesiana de C, $y = 2x$, la fórmula anterior se reduce a la evaluación de dos integrales definidas:

$$\int_C (z^*)^2\,dz = \int_0^1 5x^2dx + i\int_0^1 \left(-10x^2\right)dx = \frac{5}{3} - i\frac{10}{3}$$

Ejemplo 3.16. *La integral*

$$\oint_C \frac{dz}{z}$$

en donde C es el círculo unitario con centro en $z = 0$, es más fácil de evaluar usando la representación paramétrica, ecuación (3.58). La representación paramétrica de C es

$$z(t) = \cos t + i \operatorname{sen} t = e^{it}, \quad 0 \le t \le 2\pi$$

Por tanto, $dz = ie^{it}dt$ y, a partir de esto, se obtiene la integral

$$\oint_C \frac{dz}{z} = \int_0^{2\pi} \frac{ie^{it}dt}{e^{it}} = 2\pi i$$

Basándose en la definición de integral compleja, ecuación (3.56), es fácil obtener las propiedades siguientes.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

integral de $f(z)$ a lo largo de la curva exterior C es igual a la suma de las integrales por los contornos interiores $C_1, C_2, \ldots, C_n$ a condición de que todos los contornos se recorran en el mismo sentido,

$$\oint_C f(z)\,dz = \sum_{k=1}^{n} \oint_{C_k} f(z)\,dz \tag{3.67}$$

Antes de ver las aplicaciones específicas del teorema, se deduce un resultado importante que nos permite evaluar muchas integrales de una manera simple.

Teorema fundamental del cálculo integral

Sean $f(z)$ una función analítica en un dominio simplemente conexo D y C una curva que conecta dos puntos z_0 y z de este dominio. Considérese la integral

$$F(z) = \int_{z_0}^{z} f(\zeta)\,d\zeta \tag{3.68}$$

a lo largo de C, siendo z_0 cualquier punto fijo en D. Puesto que $f(z)$ es una función analítica en D, la integral no depende de la trayectoria de integración y está completamente determinada por los puntos z_0 y z. Por tanto, la integral (3.68) define una función $F(z)$. Se comprueba que ésta es una función analítica de la variable z en D y que

$$F'(z) = f(z) \tag{3.69}$$

evaluamos la razón,

$$\begin{aligned}
\frac{F(z+\Delta z) - F(z)}{\Delta z} &= \frac{1}{\Delta z}\left[\int_{z_0}^{z+\Delta z} f(\zeta)\,d\zeta - \int_{z_0}^{z} f(\zeta)\,d\zeta\right] \\
&= \frac{1}{\Delta z}\int_{z}^{z+\Delta z} f(\zeta)\,d\zeta \\
&= \frac{1}{\Delta z}\int_{z}^{z+\Delta z} \left(f(\zeta) - f(z) + f(z)\right)d\zeta \\
&= f(z) + \frac{1}{\Delta z}\int_{z}^{z+\Delta z} \left(f(\zeta) - f(z)\right)d\zeta
\end{aligned}$$

Resulta que

$$\lim_{\Delta z\to 0} \frac{1}{\Delta z}\int_{z}^{z+\Delta z} \left(f(\zeta) - f(z)\right)d\zeta = 0 \tag{3.70}$$

y, por consiguiente, es válida la ecuación (3.69). Para evaluar el límite anterior se usa la siguiente estimación del valor absoluto de la integral de línea:

$$\left|\int_C f(z)\,dz\right| \le \int_C |f(z)|\,|dz| \le ML \tag{3.71}$$

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Ejemplo 3.20. *Para el contorno $|z| = 2$, de la ecuación (3.78) se tiene*

$$\oint_C \frac{z^2 dz}{(z-i)^2} = 2\pi i\,(2z)\,|_{z=i} = -4\pi$$

3.3.5. Series de Taylor

En esta sección consideraremos la representación de funciones analíticas en series de potencias, que es de gran importancia en el análisis complejo. Recordaremos algunos de los conceptos básicos de la convergencia de series. Si una serie de potencias $\sum_{k=0}^{\infty} a_k z^k$ converge para $z = z_1$, ésta converge absoluta y uniformemente en cualquier disco circular $|z| \leq r$, donde $r < |z_1|$. El círculo de radio r, tal que la serie $\sum_{k=0}^{\infty} a_k z^k$ converge para toda z del disco $|z| < r$ y diverge para algunos puntos $|z| > r$, se llama **círculo de convergencia** y su radio r es el **radio de convergencia** de la serie. El radio de convergencia a menudo se determina mediante la prueba de la razón,

$$r = \lim_{k\to\infty} \left| \frac{a_{k-1}}{a_k} \right| \tag{3.79}$$

siempre y cuando este límite exista.

Ejemplo 3.21. *Para la serie*

$$\sum_{k=1}^{\infty} \frac{(-1)^k z^k}{k}$$

el radio de convergencia es $r = 1$, porque

$$r = \lim_{k\to\infty} \left| \frac{a_{k-1}}{a_k} \right| = \lim_{k\to\infty} \frac{k}{k-1} = 1$$

Ejemplo 3.22. *La serie*

$$\sum_{k=0}^{\infty} \frac{z^k}{k!}$$

converge en todo el plano complejo z, porque su radio de convergencia es infinito,

$$r = \lim_{k\to\infty} \left| \frac{a_{k-1}}{a_k} \right| = \lim_{k\to\infty} \frac{k!}{(k-1)!} = \infty$$

Suponiendo que $f(z)$ es una función analítica en $z = a$, ésta tiene derivadas de todos los órdenes $f^{(k)}(a)$, $k = 0, 1, 2, \ldots$ Por tanto, podemos asociar con $f(z)$ la serie de potencias

$$S(z) \equiv \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (z-a)^k$$

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

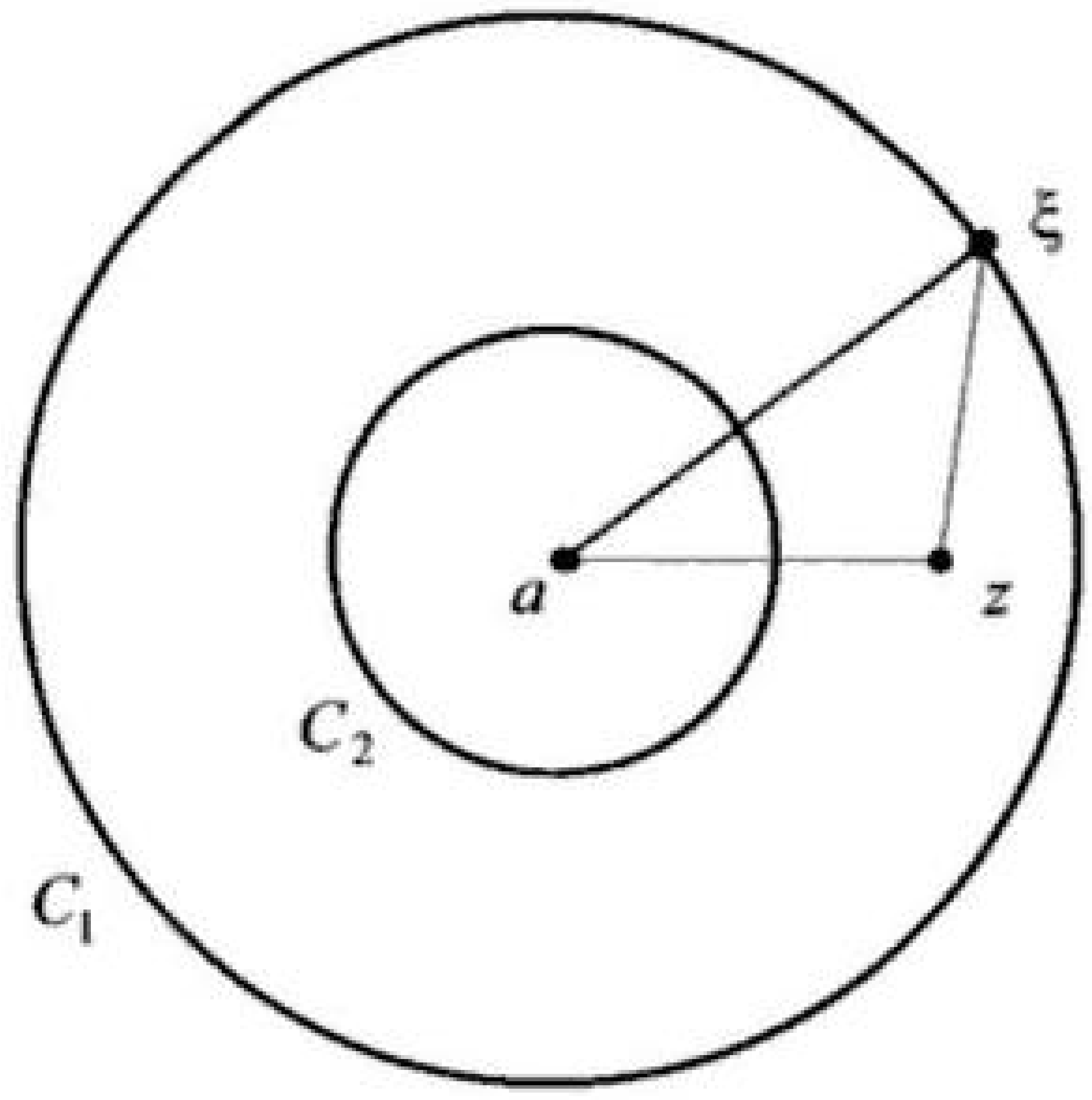

Figura 3.7. Dominio para el desarrollo en serie de Laurent

Teorema 3.6. *Una función $f(z)$ analítica en un anillo circular $R_2 \leq |z-a| \leq R_1$ puede ser representada en cada punto interior de éste por medio de la serie de Laurent*

$$f(z) = \sum_{k=0}^{\infty} a_k (z-a)^k + \sum_{k=1}^{\infty} \frac{a_{-k}}{(z-a)^k} \tag{3.84}$$

donde

$$a_k = \frac{1}{2\pi i} \oint_{C_1} \frac{f(\zeta)\, d\zeta}{(\zeta-a)^{k+1}}, \quad k = 0, 1, 2, \ldots \tag{3.85}$$

$$a_{-k} = \frac{1}{2\pi i} \oint_{C_2} \frac{f(\zeta)\, d\zeta}{(\zeta-a)^{-k+1}}, \quad k = 1, 2, \ldots \tag{3.86}$$

siendo C_1 y C_2 fronteras del anillo circular y tomándose cada una de las integrales en sentido contrario al de las manecillas del reloj.

Demostración. *Según la fórmula integral de Cauchy, ecuación (3.77), para cualquier punto z del interior del anillo $R_2 \leq |z-a| \leq R_1$ se tiene*

$$f(z) = \frac{1}{2\pi i} \oint_{C_1} \frac{f(\zeta)\, d\zeta}{\zeta - z} - \frac{1}{2\pi i} \oint_{C_2} \frac{f(\zeta)\, d\zeta}{\zeta - z} \tag{3.87}$$

donde se integra en sentido contrario al de las manecillas del reloj.

Evaluamos la integral por la circunferencia C_1. Para cualquier ζ de C_1 se cumple $|z-a| / |\zeta - a| < 1$ y, por tanto, usando la fórmula de progresión geométrica se obtiene

$$\frac{1}{\zeta - z} \equiv \frac{1}{\zeta - a} \frac{1}{1 - (z-a)/(\zeta - a)} = \frac{1}{\zeta - a} \sum_{k=0}^{\infty} \left(\frac{z-a}{\zeta - a} \right)^k$$

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

El coeficiente a_{-1} en la serie de Laurent de $f(z)$ en la vecindad de un punto singular aislado a juega un papel importante en la evaluación de las integrales de funciones analíticas. Este coeficiente recibe el nombre de **residuo** de $f(z)$ en el punto $z = a$ y se utiliza la notación

$$a_{-1}(a) = \operatorname{Res} f(z)|_{z=a}$$

Cuando la singularidad en $z = a$ es un polo de orden m, el residuo en este punto se puede calcular sin obtener la serie de Laurent. Multiplicando la ecuación (3.90) por $(z-a)^m$ se tiene

$$\begin{aligned} \phi(z) &= (z-a)^m f(z) \qquad (3.91) \\ &= a_{-m} + a_{-m+1}(z-a) + \cdots + a_{-1}(z-a)^{m-1} + a_0(z-a)^m + \cdots \end{aligned}$$

donde $a_{-m} \neq 0$. De la última ecuación se ve que el residuo de una función $f(z)$ en un polo de orden m se calcula mediante la fórmula

$$a_{-1} = \frac{1}{(m-1)!} \frac{d^{m-1}(z-a)^m f(z)}{dz^{m-1}} |_{z=a} \qquad (3.92)$$

Para un polo simple, de la ecuación (3.91) o (3.92) se tiene

$$a_{-1} = \lim_{z \to a} (z-a) f(z) \qquad (3.93)$$

De ésta se obtiene la fórmula más útil para calcular el residuo en un polo simple. Si $f(z)$ tiene un polo simple en $z = a$, ésta puede presentarse en forma

$$f(z) = \frac{p(z)}{q(z)}$$

donde $p(z)$ y $q(z)$ son analíticas en $z = a$, $p(a) \neq 0$ y $q(z)$ tiene un cero simple en $z = a$. Como consecuencia, $q(z)$ puede desarrollarse en una serie de Taylor de la forma

$$q(z) = (z-a) q'(a) + \frac{(z-a)^2}{2!} q''(a) + \cdots$$

De donde, por la ecuación (3.93),

$$a_{-1} = \lim_{z \to a} (z-a) f(z) = \lim_{z \to a} (z-a) \frac{p(z)}{q(z)} = \frac{p(a)}{q'(a)} \qquad (3.94)$$

Ejemplo 3.28. *La función*

$$f(z) = \frac{1+z}{z(2-z)}$$

tiene singularidades en los puntos $z = 0$ y $z = 2$. Notando que la función $\phi(z) = zf(z)$ ya no tiene singularidad en $z = 0$, concluimos que $z = 0$ es polo simple de $f(z)$ y, por tanto, el residuo en este punto es

$$a_{-1}(0) = \lim_{z \to 0} zf(z) = 1/2$$

Puesto que, por la ecuación (3.86) para $k = 1$,

$$\frac{1}{2\pi i}\oint_{C_k} f(z)\,dz = a_{-1}(a_k) = Res\,f(z)|_{z=a_k}$$

la fórmula (3.67) conduce a la (3.95) y el teorema queda demostrado. ■

A continuación presentamos algunos ejemplos de aplicación del teorema de residuos en relación con las integrales complejas y reales.

Ejemplo 3.34. *La función*

$$f(z) = \frac{1+z}{z(2-z)}$$

tiene dos polos simples en $z = 0$ y $z = 2$ con residuos correspondientes $a_{-1}(0) = 1/2$ y $a_{-1}(2) = -3/2$. Por lo tanto, la integral de esta función por una trayectoria simple cerrada C que: 1) encierre el punto $z = 0$ es igual a πi; 2) encierre el punto $z = 2$ es igual a $-3\pi i$; y 3) encierre los dos puntos a la vez es igual a $2\pi i\,(1/2 - 3/2) = -2\pi i$.

Ejemplo 3.35. *Para evaluar la integral*

$$\oint_C f(z)\,dz = \oint_C \frac{e^z}{z^2+1}dz$$

por el círculo $|z| = 2$, notamos que el integrando tiene dos polos simples en el interior de la trayectoria, $z = i$ y $z = -i$. Los residuos correspondientes son $a_{-1}(i) = e^i/2i$ y $a_{-1}(-i) = -e^{-i}/2i$. Por tanto, la integral es igual a $2\pi i\left(e^i/2i - e^{-i}/2i\right) = 2\pi i \operatorname{sen} 1$.

Ejemplo 3.36. *En la sección anterior encontramos que $f(z) = \cos(1/(z-1))$ tiene una singularidad esencial en el punto $z = 1$. Sin embargo, el residuo correspondiente es cero, $a_{-1}(1) = 0$. Por tanto, el valor de integral de esta función es cero para todo contorno C que no pasa por el punto $z = 1$. Si $z = 1$ se encuentra en C, entonces la integral es impropia y requiere otros métodos para evaluarla.*

Ejemplo 3.37. *La función $f(z) = 1/\left(z^3-1\right)^2$ tiene tres polos de orden dos en los puntos $z = 1$, $z = e^{i2\pi/3}$ y $z = e^{i4\pi/3}$. El contorno $|z-1| = 1$ contiene sólo el polo $z = 1$. Entonces, utilizando la ecuación (3.92), la integral de dicha función a lo largo del círculo $|z-1| = 1$ es igual a $2\pi i a_{-1}(1) = 2\pi i\,(-2/9) = -4\pi i/9$.*

El teorema de residuos proporciona también un método magnífico para evaluar ciertas clases de integrales reales complicadas. En primer lugar se consideran integrales del tipo

$$I = \int_0^{2\pi} R(\cos\theta, \operatorname{sen}\theta)\,d\theta \tag{3.96}$$

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Por lo tanto,

$$I_1 = -2\pi \operatorname{Im}\left(\frac{e^{-\omega k}}{2ik}\right) = \frac{\pi}{k}e^{-k\omega}, \quad I_2 = 2\pi \operatorname{Re}\left(\frac{e^{-\omega k}}{2ik}\right) = 0$$

Nótese que el integrando de I_2 es una función impar y el intervalo de integración es simétrico con respecto a cero. Por tanto, la integral I_2 debe ser igual a cero, que confirma nuestro resultado anterior.

Cabe mencionar que existen más clases de integrales reales que pueden evaluarse aplicando el teorema de residuos a integrales complejas apropiadas.

Nota bibliográfica: para ampliar los conocimientos sobre el tema de este capítulo, se puede consultar [1], [4], [6], [12], [15] y [17].

La recta que pasa por los dos puntos, (x_1, y_1) y (x_2, y_2), se representa por la ecuación

$$\frac{y - y_1}{y_2 - y_1} = \frac{x - x_1}{x_2 - x_1} \tag{4.14}$$

En coordenadas polares, $x = \rho \cos\varphi$ y $y = \operatorname{sen}\varphi$, por lo que, a partir de la ecuación (4.9) para una recta se tiene la ecuación

$$\rho\,(A\cos\varphi + B\operatorname{sen}\varphi) + C = 0 \tag{4.15}$$

Cuando la ecuación de una recta está dada en la forma general, ecuación (4.9), la relación de A, B y C con los parámetros de otras formas de ecuación de la recta es fácil de encontrar; esto se queda como ejercicio para los estudiantes.

4.1.3. Curvas de segundo orden en 2-D (secciones cónicas)

Ecuación general de segundo orden

Las curvas de segundo orden o **secciones cónicas** se representan por la ecuación general de segundo orden

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{13}x + 2a_{23}y + a_{33} = 0 \tag{4.16}$$

o, haciendo $a_{ij} = a_{ji}$ $(i, j = 1, 2, 3)$, se obtiene la ecuación equivalente

$$(a_{11}x + a_{12}y + a_{13})\,x + (a_{21}x + a_{22}y + a_{23})\,y + (a_{31}x + a_{32}y + a_{33}) = 0 \tag{4.17}$$

Resulta que las tres magnitudes

$$I = a_{11} + a_{22}, \quad D = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \tag{4.18}$$

y el signo de

$$\tilde{A} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} \tag{4.19}$$

son invariantes de las ecuaciones (4.16) y (4.17) con respecto a las transformaciones de traslación y rotación. Esos invariantes definen las propiedades de las secciones cónicas y no dependen de la posición de una sección cónica en el plano. Omitiendo la tabla completa, presentamos algunos ejemplos de clasificación de secciones cónicas.

Si $A \neq 0$, $D > 0$ y $A/I < 0$, la sección cónica es una elipse real (véase la figura 4.3).

Si $A \neq 0$ y $D < 0$, la sección cónica es una hipérbola (véase la figura 4.4).

Si $A \neq 0$, $D = 0$ y $\tilde{A} < 0$, la sección cónica correspondiente es una parábola (véase la figura 4.5).

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

tendiendo $M \to 0$, se obtiene la ecuación

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0, \quad y = \pm\frac{b}{a}x \tag{4.30}$$

de dos líneas rectas que se intersectan en el punto $(0,0)$.

De la ecuación (4.26) o la ecuación (4.27), tendiendo $b \to \infty$ se obtiene la ecuación de dos líneas rectas paralelas al eje y

$$x^2/a^2 = 1 \tag{4.31}$$

Si, a la vez, $a \to 0$, las dos rectas convergen a una sola recta, $x^2 = 0$.

Si $a \to \infty$ en la ecuación (4.26), se obtienen dos líneas rectas paralelas al eje x

$$y^2/b^2 = 1 \tag{4.32}$$

4.2. Superficies y curvas en 3-D

4.2.1. Representación analítica de superficies y curvas

Las coordenadas de cada punto $(x,y,z) = \mathbf{r}$ en una superficie continua en el espacio euclidiano de 3-D pueden representarse mediante un conjunto de tres ecuaciones paramétricas

$$x = x(u,v), \quad y = y(u,v), \quad z = z(u,v) \quad \text{o} \quad \mathbf{r} = \mathbf{r}(u,v) \tag{4.33}$$

donde u y v son parámetros reales apropiados. Las ecuaciones (4.33) se llaman **representación paramétrica** de una **superficie**. Una ecuación del tipo

$$\varphi(x,y,z) = 0 \quad \text{o} \quad z = f(x,y) \tag{4.34}$$

también representa una superficie. Una superficie puede tener más de una hoja; las hojas pueden ser unidas o no. Las superficies correspondientes a las ecuaciones

$$\varphi(x,y,z) = 0 \quad \text{y} \quad \lambda\varphi(x,y,z) = 0$$

son idénticas para cualquier constante $\lambda \neq 0$.

Una curva continua en el espacio euclidiano 3-D es un conjunto de puntos con coordenadas $(x,y,z) = \mathbf{r}$ que satisfacen un sistema de ecuaciones paramétricas

$$x = x(t), \quad y = y(t), \quad z = z(t) \quad \text{o} \quad \mathbf{r} = \mathbf{r}(t) \tag{4.35}$$

donde $x(t)$, $y(t)$, $z(t)$ son funciones continuas del parámetro real $t \in [t_1, t_2]$. De manera alternativa, la intersección de dos superficies también define una curva en 3-D, es decir, un conjunto de los puntos con coordenadas (x,y,z) que simultáneamente satisfacen dos ecuaciones

$$\varphi_1(x,y,z) = 0 \quad \text{y} \quad \varphi_2(x,y,z) = 0 \tag{4.36}$$

$\mathbf{a}$ y $\mathbf{b}$ pueden considerarse como la base de un sistema de coordenadas u y v en este plano. Entonces,

$$\mathbf{r} - \mathbf{r}_1 = u\mathbf{a} + v\mathbf{b} \quad \text{o} \quad \mathbf{r} = \mathbf{r}_1 + u\mathbf{a} + v\mathbf{b} \tag{4.46}$$

y ésta es la representación paramétrica de un plano.

Recta

La intersección de dos planos,

$$\mathbf{A}_1 \cdot \mathbf{r} + D_1 = 0 \quad \text{y} \quad \mathbf{A}_2 \cdot \mathbf{r} + D_2 = 0 \tag{4.47}$$

es una línea **recta**, si las normales $\mathbf{A}_1 = (A_1, B_1, C_1)$ y $\mathbf{A}_2 = (A_2, B_2, C_2)$ a cada superficie no son colineales, es decir, $\mathbf{A}_1 \times \mathbf{A}_2 \neq 0$. Cada línea recta puede representarse en la forma (4.47). La ecuación (4.47) representa una recta que pasa por el origen, si y sólo si $D_1 = D_2 = 0$.

La línea recta que tiene el vector director $\mathbf{a} = (a_x, a_y, a_z)$ y que pasa por el punto $\mathbf{r}_1 = (x_1, y_1, z_1)$ se describe, también, por una ecuación paramétrica

$$\mathbf{r} - \mathbf{r}_1 = \mathbf{a}t \tag{4.48}$$

donde t es una variable (parámetro) real. Aquí se puede tomar $\mathbf{a} = \mathbf{A}_1 \times \mathbf{A}_2$. En coordenadas cartesianas, la ecuación (4.48) toma la forma

$$x = x_1 + a_x t, \quad y = y_1 + a_y t, \quad z = z_1 + a_z t \tag{4.49}$$

Al excluir el parámetro t de las ecuaciones (4.49), se obtiene la ecuación de una recta en forma simétrica

$$\frac{x - x_1}{a_x} = \frac{y - y_1}{a_y} = \frac{z - z_1}{a_z} \tag{4.50}$$

4.2.3. Cuádricas

Ecuación general de segundo orden

Las **cuádricas** se describen por una ecuación general de segundo orden:

$$\begin{aligned} & a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz + \\ & 2a_{14}x + 2a_{24}y + 2a_{34}z + a_{44} \\ = \; & 0 \end{aligned} \tag{4.51}$$

o

$$\begin{array}{l} (a_{11}x + a_{12}y + a_{13}z + a_{14})\,x + \\ (a_{21}x + a_{22}y + a_{23}z + a_{24})\,y + \\ (a_{31}x + a_{32}y + a_{33}z + a_{34})\,z + \\ (a_{41}x + a_{42}y + a_{43}z + a_{44}) = 0 \end{array} \tag{4.52}$$

con $a_{ij} = a_{ji}$, $i, j = 1, 2, 3, 4$.

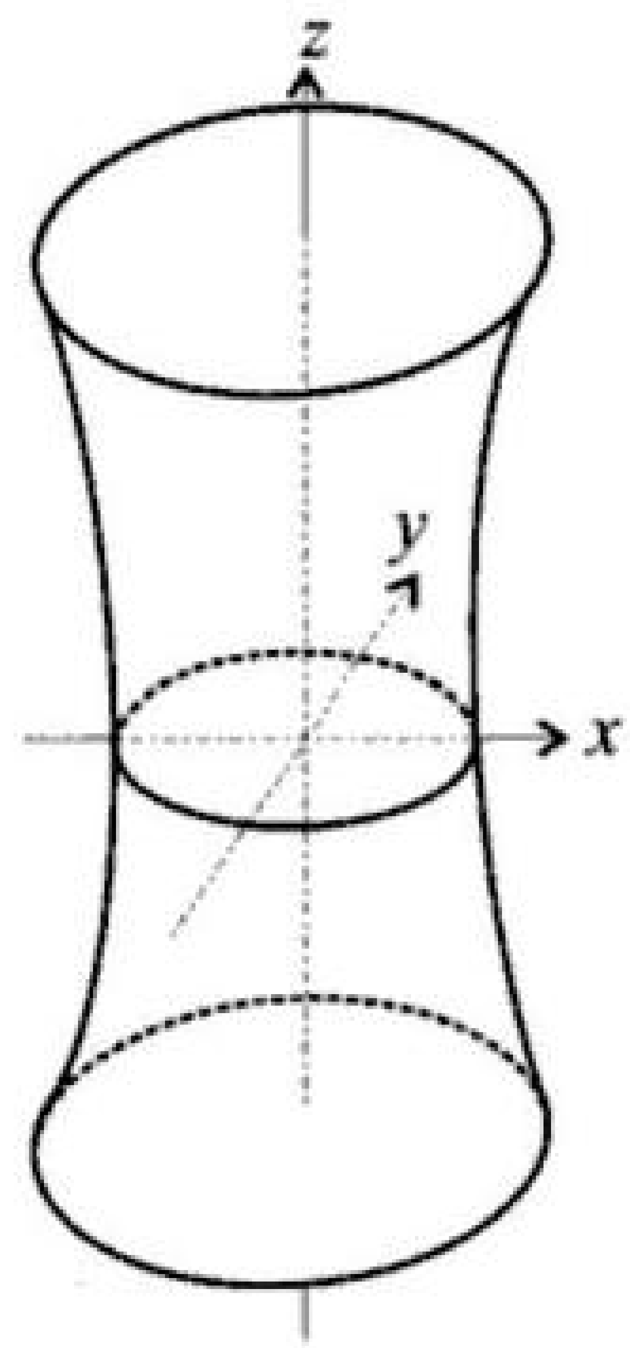

Figura 4.10. Hiperboloide de una capa

es también una cuádrica central, que en el caso de $a \neq b \neq c$ tiene tres planos y tres ejes principales (véase la figura 4.10). La intersección de una hiperboloide de una capa con el plano principal $x = 0$ y el $y = 0$ es una hipérbola; la intersección con el plano $z = 0$ es una elipse. En el caso de $a = b$ una hiperboloide de una capa se reduce a una superficie de revolución alrededor del eje z, y ésta ya tiene un número infinito de planos de simetría.

La **hiperboloide de dos capas**,

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 \tag{4.65}$$

es una cuádrica central de dos hojas; en el caso de $a \neq b \neq c$ tiene tres planos y tres ejes principales (véase la figura 4.11). La intersección de una hiperboloide de dos capas con el plano principal $z = 0$ y $y = 0$ es una hipérbola. No existe la intersección real con el plano $x = 0$. En el caso de $b = c$ una hiperboloide de dos capas se reduce a una superficie de revolución alrededor del eje x, y ésta ya tiene un número infinito de planos de simetría.

Una **paraboloide elíptica**,

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = z \tag{4.66}$$

es una cuádrica no central; en el caso de $a \neq b$ tiene dos planos y un solo eje principal (véase la figura 4.12). La intersección de una paraboloide elíptica con el plano principal $x = 0$ y $y = 0$ es una parábola. La intersección con el plano $z = const > 0$ es una elipse, no existe la intersección real con el plano $z = const < 0$. En el caso de $a = b$ una paraboloide elíptica se reduce a una superficie de revolución alrededor del eje z, y ésta ya tiene un número infinito de planos de simetría y un solo eje principal.

sucesión de longitudes de las líneas quebradas converge a un límite L, se dice que una **curva** C es **rectificable** y la longitud de C es igual a L.

Si una curva C no es simple pero consta de un número finito de curvas simples rectificables, la longitud de C se define como la suma de las longitudes de estas curvas.

Suponemos que una curva C está representada mediante una función vectorial continuamente diferenciable,

$$\mathbf{r} = \mathbf{r}(t) = x(t)\,\mathbf{i} + y(t)\,\mathbf{j} + z(t)\,\mathbf{k} \tag{4.69}$$

de un parámetro $a \leq t \leq b$. Dividimos el intervalo $a = t_0 < t_1 < \cdots < t_N = b$ en N subintervalos y unimos los pares de puntos consecutivos de C con segmentos de recta, obteniendo de esta manera una línea quebrada inscrita en C. El segmento k-ésimo $d\mathbf{r}_k$ de la línea quebrada inscrita en C es

$$\begin{aligned} d\mathbf{r}_k &= \mathbf{r}(t_k) - \mathbf{r}(t_{k-1}) \simeq dx(t_{k-1})\,\mathbf{i} + dy(t_{k-1})\,\mathbf{j} + dz(t_{k-1})\,\mathbf{k} \\ &= \left[\dot{x}(t_{k-1})\,\mathbf{i} + \dot{y}(t_{k-1})\,\mathbf{j} + \dot{z}(t_{k-1})\,\mathbf{k}\right]\Delta t_k \end{aligned} \tag{4.70}$$

donde $\Delta t_k = t_k - t_{k-1}$, $k = 1, \ldots, N$. El vector $d\mathbf{r}_k$ tiene la longitud

$$ds_k = (d\mathbf{r}_k \cdot d\mathbf{r}_k)^{1/2} \simeq \left[(\dot{x}(t_{k-1}))^2 + (\dot{y}(t_{k-1}))^2 + \left(\dot{z}(t_{k-1})^2\right)\right]^{1/2} \Delta t_k \tag{4.71}$$

La longitud L_N de la quebrada de N segmentos es la suma de las longitudes de dichos segmentos. La longitud de C es el límite de L_N cuando $N \to \infty$, es decir,

$$\begin{aligned} L &= \lim_{N\to\infty} L_N = \lim_{N\to\infty} \sum_{k=1}^{N} ds_k \\ &= \int_a^b \sqrt{\dot{\mathbf{r}} \cdot \dot{\mathbf{r}}}dt = \int_a^b \left[(\dot{x}(t))^2 + (\dot{y}(t))^2 + \left(\dot{z}(t)^2\right)\right]^{1/2} dt \end{aligned} \tag{4.72}$$

Reemplazando el límite superior fijo b en la ecuación (4.72) por un límite superior variable t, se obtiene la integral como función de t,

$$s(t) = \int_a^t \sqrt{\dot{\mathbf{r}} \cdot \dot{\mathbf{r}}}d\tau \tag{4.73}$$

Esta función es la **longitud de arco** de C. El uso de s como parámetro natural en la representación de una curva C simplifica diversas fórmulas. El sentido que corresponde a los valores crecientes de s se llama sentido positivo sobre C; de esta manera se define cierta orientación de la curva.

Un cambio de parametrización de C, digamos $s = s(t)$, reemplazará la ecuación (4.72) por la ecuación

$$L = \int_{s(a)}^{s(b)} \sqrt{\dot{\mathbf{r}}(s) \cdot \dot{\mathbf{r}}(s)}ds \tag{4.74}$$

que $\nabla\varphi$ es perpendicular a $d\mathbf{r}$ mientras que el vector $d\mathbf{r}$ une dos puntos P y Q que pertenecen a la superficie $\varphi(x,y,z)=const.$ Entonces, $\nabla\varphi$ es perpendicular a todas las tangentes a dicha superficie en el punto P, es decir, $\nabla\varphi$ es una normal a la superficie $\varphi(x,y,z)=const.$

Para un punto dado $P(x,y,z)$, el vector $\nabla\varphi$ es fijo. Por tanto, $d\varphi$ dependerá solamente de $d\mathbf{r}$. Sea $ds=|d\mathbf{r}|$, pues $d\mathbf{r}=\mathbf{u}\cdot ds$, donde $\mathbf{u}$ es el vector unitario en la dirección de $d\mathbf{r}$. Luego,

$$d\varphi = ds\,(\mathbf{u}\cdot\nabla\varphi) \quad \text{y} \quad \frac{d\varphi}{ds}=\mathbf{u}\cdot\nabla\varphi \tag{4.82}$$

La derivada $d\varphi/ds$ se llama **derivada direccional** en la dirección de $\mathbf{u}$. Si la dirección de $\mathbf{u}$ coincide con la del gradiente $\nabla\varphi$, entonces, la derivada direccional alcanza su máximo valor. Así, el gradiente nos indica la dirección del máximo cambio de la función φ.

Ejemplo 4.1. *Sea $\varphi(x,y,z)=r=|\mathbf{r}|=\sqrt{x^2+y^2+z^2}$. Por lo tanto, la superficie $\varphi(x,y,z)=const$ es una esfera. Dado que $\nabla\varphi$ es una normal a la superficie $\varphi(x,y,z)=const$, tenemos la igualdad*

$$\nabla r = k\mathbf{r}$$

donde k es un escalar por determinar. Ya que $r^2=\mathbf{r}\cdot\mathbf{r}$, se tiene que $d\left(r^2\right)=2rdr=2\mathbf{r}\cdot d\mathbf{r}$. Así que

$$d\varphi = dr = d\mathbf{r}\cdot\nabla r = kd\mathbf{r}\cdot\mathbf{r} = krdr$$

de donde $k=1/r$. Finalmente,

$$\nabla r = \frac{\mathbf{r}}{r} \tag{4.83}$$

es un vector unitario.

Ejemplo 4.2. *Demostraremos que*

$$\nabla f(u) = \frac{df(u)}{du}\nabla u, \quad \text{donde} \quad u=u(x,y,z) \tag{4.84}$$

Por definición,

$$\begin{aligned}
\nabla f(u) &= \frac{\partial f}{\partial x}\mathbf{i}+\frac{\partial f}{\partial y}\mathbf{j}+\frac{\partial f}{\partial z}\mathbf{k}\\
&= \frac{df}{du}\frac{\partial u}{\partial x}\mathbf{i}+\frac{df}{du}\frac{\partial u}{\partial y}\mathbf{j}+\frac{df}{du}\frac{\partial u}{\partial z}\mathbf{k}\\
&= \frac{df}{du}\nabla u
\end{aligned}$$

Operador diferencial ∇

Definimos el operador diferencial vectorial

$$\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial x} + \mathbf{k}\frac{\partial}{\partial x} \tag{4.85}$$

de tal manera que

$$\nabla\varphi = \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial x} + \mathbf{k}\frac{\partial}{\partial x}\right)\varphi = \mathbf{i}\frac{\partial\varphi}{\partial x} + \mathbf{j}\frac{\partial\varphi}{\partial x} + \mathbf{k}\frac{\partial\varphi}{\partial x}$$

El **operador** ∇, llamado **nabla**, tiene las características tanto de un vector como de un operador diferencial. Por lo tanto,

$$\begin{aligned} \nabla(uv) &= \mathbf{i}\frac{\partial(uv)}{\partial x} + \mathbf{j}\frac{\partial(uv)}{\partial x} + \mathbf{k}\frac{\partial(uv)}{\partial x} \\ &= u\nabla v + v\nabla u \end{aligned} \tag{4.86}$$

y

$$\nabla(\mathbf{a}\cdot\mathbf{r}) = \mathbf{a} \tag{4.87}$$

donde $\mathbf{a} = (a_x, a_y, a_z)$ es un vector constante.

4.6. Coordenadas curvilíneas ortogonales

A menudo los matemáticos, físicos o ingenieros se encuentran ante el problema de usar un sistema de coordenadas diferente del sistema cartesiano rectangular. La razón principal es que por la naturaleza de un problema físico o matemático a resolver, tanto el planteamiento como la resolución del problema resultan más simples en un sistema "natural" para el problema.

4.6.1. Coordenadas esféricas

Un problema físico con la simetría esférica es mucho más fácil de resolver en el sistema de coordenadas esféricas, que es natural para el problema. En este caso, la posición de un punto en el espacio físico se describe mediante las coordenadas esféricas r, θ, φ que se definen mediante el siguiente procedimiento (véase la figura 4.14). Notamos que la esfera $x^2 + y^2 + z^2 = r^2$, donde $r = |\mathbf{r}|$ es el módulo del vector de posición, el cono $z/(x^2+y^2+z^2)^{1/2} = \cos\theta$ y el plano $y/x = \tan\varphi$ pasan por el punto $P(r,\theta,\varphi)$. Por lo tanto, las relaciones

$$r = \sqrt{x^2+y^2+z^2} \tag{4.88}$$

$$\theta = \arccos\left(\frac{z}{\sqrt{x^2+y^2+z^2}}\right) \tag{4.89}$$

$$\varphi = \arctan\frac{y}{x} \tag{4.90}$$

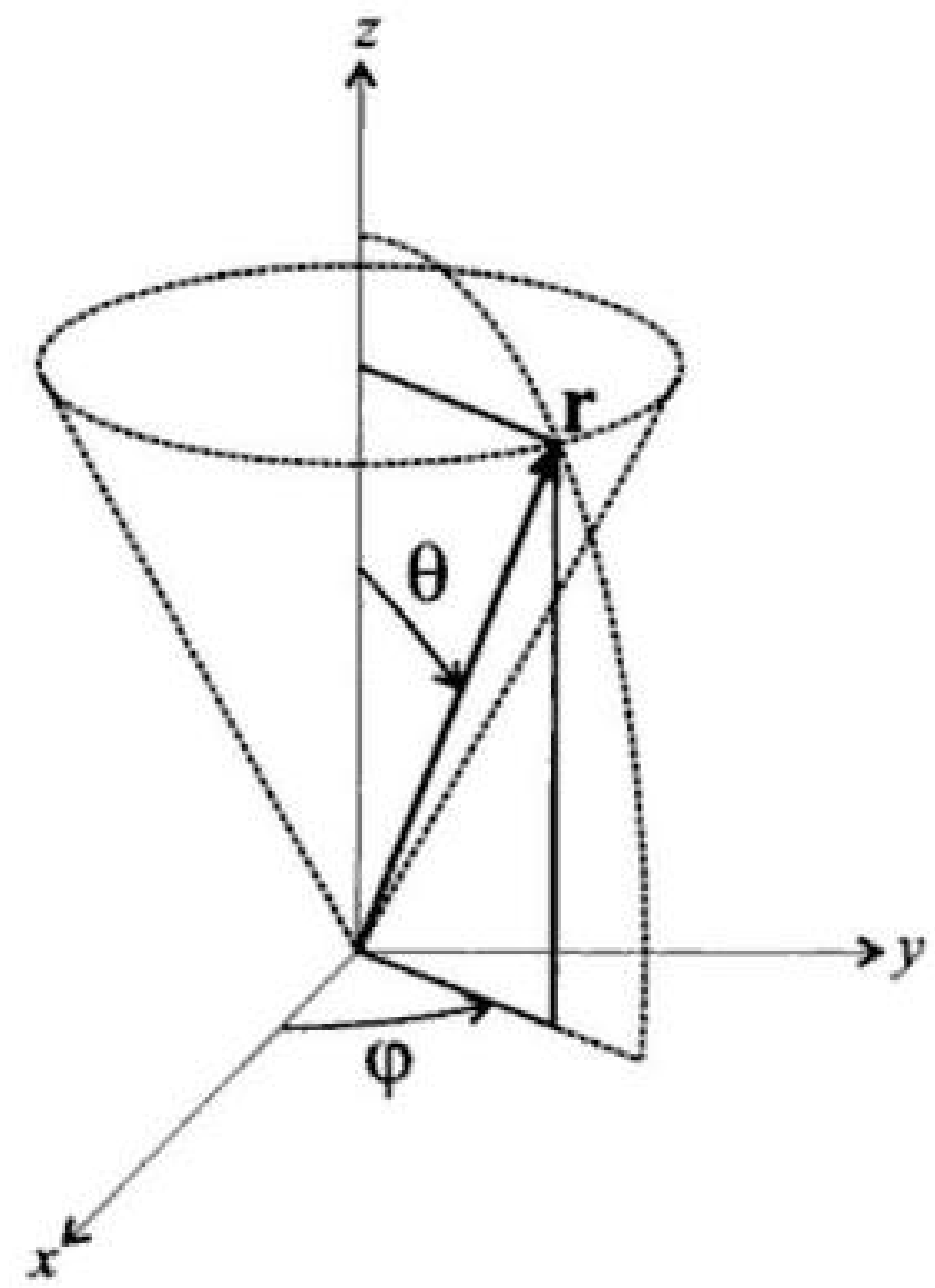

Figura 4.14. Coordenadas esféricas

pueden considerarse como la transformación del sistema de coordenadas cartesianas x, y, z al sistema de coordenadas esféricas r, θ, φ. A través de cada punto del espacio, salvo el origen, pasará una superficie del tipo $r = \sqrt{x^2 + y^2 + z^2} = c_1$, $\theta = \arccos(z/r) = c_2$ y $\varphi = \arctan(y/x) = c_3$, que son una esfera, un cono y un plano, respectivamente. Las coordenadas del punto P se determinarán por las constantes c_1, c_2, c_3.

La intersección de la esfera con el cono es una circunferencia de radio $r \operatorname{sen}\theta$ a una altura $z = r\cos\theta$ que tiene un vector tangente unitario $\mathbf{e}_\varphi$ en el punto P. Esta circunferencia se llama curva-φ ($r = c_1$, $\theta = c_2$) porque r y θ se mantienen fijos sobre ésta y sólo la coordenada φ se cambia al moverse a lo largo de la curva. La intersección de la esfera y el plano que pasa por el origen resulta en una circunferencia llamada curva-θ ($r = c_1$, $\varphi = c_3$), mientras que la intersección del cono con el plano resulta en la línea recta, curva-r ($\theta = c_2$, $\varphi = c_3$), que pasa por el origen y el punto P. Sean $\mathbf{e}_\theta$ y $\mathbf{e}_r$ vectores tangenciales unitarios en el punto P a las curvas -θ y -r, respectivamente. Es fácil ver que las tres superficies $r = c_1$, $\theta = c_2$ y $\varphi = c_3$ son perpendiculares entre sí en cada punto de intersección. Entonces, los tres vectores $\mathbf{e}_r$, $\mathbf{e}_\theta$, $\mathbf{e}_\varphi$ son perpendiculares entre sí y, por tanto, forman una base ortonormal en la vecindad del punto P. Dicha base es dependiente de cada punto P y, en este sentido, el sistema de coordenadas es local.

Dado que tanto $\mathbf{e}_\varphi$ como $\boldsymbol{\nabla}\varphi$ son perpendiculares a la superficie $\varphi = const$ en el punto P, el vector $\boldsymbol{\nabla}\varphi$ debe ser paralelo a $\mathbf{e}_\varphi$. Entonces, $\mathbf{e}_\varphi = h_3\boldsymbol{\nabla}\varphi$, en donde h_3 es el factor escalar de proporcionalidad entre $\mathbf{e}_\varphi$ y $\boldsymbol{\nabla}\varphi$. Si $d\mathbf{r}_3$ es un vector infinitesimal tangente a la curva-φ de longitud $ds_3 = |d\mathbf{r}_3|$, entonces se tiene

$$ds_3 = d\mathbf{r}_3 \cdot \mathbf{e}_\varphi = d\mathbf{r}_3 \cdot (h_3 \nabla \varphi) = h_3 d\varphi$$

ya que por la fórmula (4.80) $d\varphi = d\mathbf{r}_3 \cdot \nabla \varphi$. Se ve que h_3 representa el factor que, al haber multiplicado por el cambio diferencial $d\varphi$ de la coordenada φ, nos da la longitud del arco a lo largo de la curva-φ. Entonces, $h_3 = r \operatorname{sen} \theta$ y

$$\mathbf{e}_\varphi = h_3 \nabla \varphi = r \operatorname{sen} \theta \nabla \varphi \tag{4.91}$$

De manera similar se comprueba que

$$\mathbf{e}_\theta = h_2 \nabla \theta = r \nabla \theta \tag{4.92}$$

$$\mathbf{e}_r = h_1 \nabla r = \nabla r \tag{4.93}$$

Notamos que

$$\begin{aligned}
\mathbf{e}_r &= \mathbf{e}_\theta \times \mathbf{e}_\varphi = r^2 \operatorname{sen} \theta \, (\nabla \theta \times \nabla \varphi) \\
\mathbf{e}_\theta &= \mathbf{e}_\varphi \times \mathbf{e}_r = r \operatorname{sen} \theta \, (\nabla \theta \times \nabla r) \\
\mathbf{e}_\varphi &= \mathbf{e}_r \times \mathbf{e}_\theta = r \, (\nabla r \times \nabla \varphi)
\end{aligned}$$

Cada vector $\mathbf{f}$ en el punto P puede representarse en la base local como

$$\mathbf{f} = f_r \mathbf{e}_r + f_\theta \mathbf{e}_\theta + f_\varphi \mathbf{e}_\varphi = f_r \nabla r + r f_\theta \nabla \theta + f_\varphi r \operatorname{sen} \theta \nabla \varphi$$

en donde las coordenadas f_r, f_θ, f_φ pueden ser funciones de r, θ, φ. Además, la diferencial de volumen es como sigue:

$$\begin{aligned}
dV &= dxdydz = \Im\left(x, y, z / r, \theta, \varphi\right) dr d\theta d\varphi \\
&= d\mathbf{r}_1 \cdot d\mathbf{r}_2 \times d\mathbf{r}_3 = ds_1 ds_2 ds_3 \mathbf{e}_r \cdot \mathbf{e}_\theta \times \mathbf{e}_\varphi = ds_1 ds_2 ds_3 \\
&= h_1 dr h_2 d\theta h_3 d\varphi = r^2 \operatorname{sen} \theta dr d\theta d\varphi
\end{aligned} \tag{4.94}$$

en donde $\Im\left(x, y, z / r, \theta, \varphi\right)$ es el Jacobiano de la transformada $x, y, z \rightarrow r, \theta, \varphi$.

4.6.2. Coordenadas curvilíneas ortogonales: conceptos generales

Las coordenadas esféricas son un caso particular de los sistemas de coordenadas curvilíneas ortogonales. En esta sección se consideran sistemas de coordenadas curvilíneas ortogonales en general, con el fin de tener las representaciones del gradiente, la divergencia, el rotacional y el laplaciano en cualquiera de estos sistemas.

Sea

$$u_1 = u_1\left(x, y, z\right), \quad u_2 = u_2\left(x, y, z\right), \quad u_3 = u_3\left(x, y, z\right) \tag{4.95}$$

un cambio del sistema de coordenadas cartesiano x, y, z a un sistema u_1, u_2, u_3. Suponemos que el Jacobiano $\Im\left(u_1, u_2, u_3 / x, y, z\right) \neq 0$ en un dominio D de espacio y, por lo tanto, la transformación de coordenadas (4.95) es uno a uno en D. Las ecuaciones

$$u_1\left(x, y, z\right) = c_1, \quad u_2\left(x, y, z\right) = c_2, \quad u_3\left(x, y, z\right) = c_3 \tag{4.96}$$

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

En nuestros tiempos de "computación total", existe la creencia bastante común de que los problemas matemáticos aplicados en su mayoría absoluta se resuelven por computadora. Sin embargo, la complejidad de problemas a tratar en las ciencias exactas y naturales requiere la aplicación de conceptos y métodos de las Matemáticas cada vez más complicados y poderosos. Es por tanto, el objetivo principal de este libro de texto, que el estudiante se familiarice profundamente con los conceptos matemáticos fundamentales. El texto pretende abarcar los temas indispensables en matemáticas para los estudios de Física y áreas afines. El propósito de los *Fundamentos de Métodos Matemáticos para Física e Ingeniería* es dar a los estudiantes bases sólidas en métodos matemáticos, despertar su mente y enseñarles cierta cultura de estudios y aplicaciones de métodos matemáticos en su trabajo profesional.

ÁREA: MATEMÁTICAS
ISBN 968-18-6366-6
9 789681 863661

e-mail: limusa@noriega.com.mx
www.noriega.com.mx